

基于模板跟踪的实时无标志点注册算法

林 惊 王涌天 刘 越 杨 珂

(北京理工大学信息科学技术学院, 北京 100081)

摘要 为了解决不依赖于标志点的增强现实系统的环境注册问题,提出了一种基于选定模板跟踪的无标志点注册算法。该算法使用已标定的单目摄像机进行场景捕捉,通过对运动目标上的选定模板进行实时跟踪,获得每帧场景图像中的对应特征,利用透射投影变换计算运动区域的单应性约束后,再结合摄像机内参数,完成对目标的3维姿态估计,从而得到场景的3维环境注册信息。实现了一种基于该算法的无标志点的增强现实系统。介绍了系统结构和 workflows,试验结果验证了理论的正确性和系统的鲁棒性。

关键词 增强现实 跟踪注册 无标志点 运动估计

中图分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)09-1812-08

Real-time Markerless Registration Algorithm for Augmented Reality Based on Template Tracking

LIN Liang, WANG Yong-tian, LIU Yue, YANG Ke

(School of Information Science and Technology, Beijing Institute of Technology, Beijing 100081)

Abstract Environment registration is a key issue for an augmented reality (AR) system. An efficient markerless registration algorithm is introduced in this paper. The correspondences between frames are obtained in the process of tracking planar target with a fixed template. The homograph of the target can be computed with the perspective projective transformation. Combined with the intrinsic parameters of the camera, pose estimation of the region can be completed resulting in 3D registration for the system. A markerless AR system based on this algorithm is designed. The system architecture and working flow are illustrated. Experimental results demonstrate the accuracy of the theoretical analysis and the robustness of the system.

Keywords augmented reality, markerless registration, virtual tracking, motion estimation

1 引言

增强现实(augmented reality, AR)是在虚拟现实(virtual reality, VR)技术的基础上发展起来的新兴研究领域,综合了光电成像、融合显示、多传感器、计算机图形、图像处理、计算机视觉等多门学科,是一种用计算机产生的附加信息对真实世界的景象增强或扩张的技术。AR系统的使用者可以在看到周围真实环境的同时看到计算机产生的增强信息,这种增强信息

可以在真实环境中与真实物体共存的虚拟物体,也可以是关于存在的真实物体的非几何信息^[1-4]。

AR系统能够利用融合显示、多传感器、计算机视觉等技术对真实场景和事物进行扩展和增强,因此,在系统构成上主要包括摄像与处理系统、注册定位系统、融合渲染系统与立体显示系统等几个关键部分^[5]。其中摄像与处理系统主要是通过图像采集设备获取真实场景并通过算法降低噪声;注册与定位系统通过硬件或者软件计算出观察者当前的位置和姿态,确定其视点和视线方向,建立观察坐标

基金项目:国家863计划项目(2006AA01Z339);国家自然科学基金委员会与微软亚洲研究院联合资助项目(60673198)

收稿日期:2006-06-22; **改回日期:**2007-05-11

第一作者简介:林 惊(1981~),男,北京理工大学信息科学技术学院博士研究生。主要研究领域为物体目标识别,场景理解,以及增强现实技术与应用。E-mail:linliang@bit.edu.cn

系;融合渲染系统在获得观察者的精确注册信息后将虚拟景物根据注册信息渲染,并无缝地结合到真实场景中;最后通过立体显示设备输出,使观察者“浸没”在增强后的场景中^[2,3]。增强现实系统的核心问题是如何实现完美的“虚实结合”,而注册与定位系统是其关键技术,在目前的研究中,可将注册技术分为基于标志点的注册方法与无标志点的注册方法两大类。

基于标志点的注册定位算法是当前最接近实际应用的增强现实注册技术^[6,7]。基于标志点的注册定位技术具有计算量小、执行速度快、不需要用户参与即可自动实现姿态测定等优势。但是标志点跟踪系统的精度有待提高,由于基于标志点的注册系统只能利用有限的 n 个点提供的信息,因此,系统需要每个点的高精度定位信息。而定位精度常常受到视频采集的白噪声、电路噪声和系统分辨率的限制,成为制约标志点算法精度的瓶颈,当若干个标志点被遮挡或者干扰时,系统稳定性更会大幅度下降。此外,在很多实际应用中,特别是户外条件下,不允许在场景中放置标志点。在这种情况下,必须使用无标志点的注册定位算法来完成虚实世界的配准任务^[4,8]。

本文提出了一种通过对场景中选定目标进行跟踪来实现环境注册的算法,并结合已标定摄像机的内参数来实现3维环境注册。新算法要求预先对场景进行初始化,先从选定区域内任意取出若干点,构成一个稀疏点集,作为跟踪目标,而初始时刻的跟踪目标同时作为跟踪阶段的参考模板。同时,还需计算初始时刻摄像机的透射投影矩阵。在实时环境注册阶段,通过计算当前场景图像与预存储模板的雅各比约束求解目标运动矢量的变化,在图像坐标下实现选定目标跟踪,从而得到摄像机运动中目标对应的图像坐标,即每帧图像的单应性关系,再利用已标定的摄像机内参数,结合初始时刻摄像机的透射投影矩阵,来估计摄像机的3维位置和姿态,完成3维环境注册。注册算法的实现分为初始化阶段、目标跟踪和相机姿态估计3部分。除初始化步骤费时较长,目标跟踪及姿态估计均可以实时完成。

新算法的鲁棒性优势体现在以下两点:

(1) 由于新算法将跟踪平面目标内的选定点集作为跟踪对象,利用其整体的灰度不变量,因此,相对于利用少量标志点极大地提高了注册精度和抗干扰性,甚至允许跟踪目标部分信息被遮挡。

(2) 由于新算法在摄像机采集的每帧图像中计

算目标灰度,同时记录灰度变化量,从而可以实时对模板进行校正,因此,很大程度上提高了注册算法在光照变化条件下的稳定性。

在通过本算法实现3维环境注册以后,可以很容易地将虚拟景物叠加到当前场景中来实现场景的增强。因此,可以实现高效稳定的无标志点的增强现实系统。

2 基于模板跟踪算法的无标志点注册

2.1 SSD平面跟踪算法

运动目标跟踪是计算机视觉研究的一个重要领域,其根本方法是在连续图像中匹配相似度,方法大致可以分为两类,一类是光流(optical flow)计算法,即通过匹配一些稀疏的点集,来计算每个像素的运动矢量^[7,9]。另一类是基于特征点提取的方法,即寻找每帧图像中的特征,如拐点、物体边缘等,再利用几何约束来最小化误差,实现跟踪^[7,10]。本文提出的注册算法是基于精确的平面目标跟踪结果上的,首先定义误差平方和(sum-of-squared differences, SSD)函数,在跟踪过程中最小化当前跟踪区域与预存模板的SSD函数,并通过透射投影运动模型求解跟踪目标的运动矢量。因此,本算法中实现目标跟踪的部分,属于第1类方法。

在摄像机移动过程中,视场每一个微弱的偏移都会引起跟踪目标内灰度的细微变化,在摄像机采样时间足够短的情况下,利用灰度的变化信息再通过迭代求解可以估计跟踪目标的运动参数。

假设 $I(\mathbf{X}, t)$ 表示图像上位置 $\mathbf{X} = (x, y)^T$ 在 t 时刻的亮度值,并且假设 $\nabla I(\mathbf{X}, t)$ 表示相应的空间灰度梯度。将初始时刻 t_0 的图像作为参考图像,同时在区域内抽样选定 N 个点构成稀疏点集 $\mathbf{G} = \{X_1, X_2, X_3, \dots, X_N\}$, 将点集灰度值定义为跟踪目标,因而,在参考图像上的点集 \mathbf{G} 的灰度集合也是参考模板。在实时跟踪过程中,摄像机和跟踪目标之间的相对运动造成了跟踪目标的扭曲或者畸变,因此首先必须定义一个运动模型 $f(\mathbf{X}; \boldsymbol{\mu})$ 来描述目标的运动变化,其中 $\boldsymbol{\mu}$ 为 n 维运动参数矢量,因此,有 $f(\mathbf{X}; 0) = \mathbf{X}, N > n$ 。此时,目标的跟踪问题抽象为在每一帧图像中通过灰度变化恢复运动矢量的问题。设定 $\hat{\boldsymbol{\mu}}(t)$ 为 t 时刻运动矢量的真实值,而 $\boldsymbol{\mu}(t)$ 为估算值。由以上约定可知,在初始时刻 t_0 , 有 $\hat{\boldsymbol{\mu}}(t) = \boldsymbol{\mu}(t) = 0$, 而在 $t > t_0$ 的任意一个时刻则有

$$I(\mathbf{X}, t_0) = I(f(\mathbf{X}, \hat{\boldsymbol{\mu}}(t)), t) \quad (1)$$

这就是一般形式的灰度不变量假设^[11]。因此,可以通过最小二乘方程来估计任意时刻 t 的运动参数:

$$O(\boldsymbol{\mu}) = \|(I(f(\mathbf{X}, \boldsymbol{\mu}), t) - I(\mathbf{X}, t_0))\|^2 \quad (2)$$

为了简化公式形式,用 $I(\boldsymbol{\mu}, t)$ 表示 $I(f(\mathbf{X}, \boldsymbol{\mu}), t)$, 表示 t 时刻经过运动矢量 $\boldsymbol{\mu}$ 变换后的目标灰度集合,且 t_0 时刻 $\boldsymbol{\mu} = 0$, 因此式(2)简化为

$$O(\boldsymbol{\mu}) = \|(I(\boldsymbol{\mu}, t) - I(0, t_0))\|^2 \quad (3)$$

在一般的情况下,式(3)是一个非凸目标函数,在缺少合适的初始位置的条件下,函数的求解需要通过代价巨大的全局优化途径来实现^[12,13]。在实时跟踪的情况下,能够在摄像机连续运动之前给出一个准确的跟踪目标位置的初始值,即 $\boldsymbol{\mu}(t_0)$, 而在

$t > t_0$ 的任意一个时刻,目标区域的几何运动用 $\boldsymbol{\mu}(t)$ 来描述。在这种情况下,可以再次将跟踪问题简化为求解运动矢量的偏移 $\delta\boldsymbol{\mu}$, 即 $\boldsymbol{\mu}(t + \tau) = \boldsymbol{\mu}(t) + \delta\boldsymbol{\mu}$, τ 为时间变化量,因此式(3)可以转化为

$$O(\delta\boldsymbol{\mu}) = \|(I(\boldsymbol{\mu}(t) + \delta\boldsymbol{\mu}, t + \tau) - I(0, t_0))\|^2 \quad (4)$$

由于采用高采样率的摄像机,可以使 $\delta\boldsymbol{\mu}$ 足够小,为了将此非线性优化问题近似为线性处理,可以通过用泰勒级数展开 $I(\boldsymbol{\mu}(t) + \delta\boldsymbol{\mu}, t + \tau)$ 为

$$I(\boldsymbol{\mu} + \delta\boldsymbol{\mu}, t + \tau) = I(\boldsymbol{\mu}, t) + M(\boldsymbol{\mu}, t)\delta\boldsymbol{\mu} + \tau \times I_t(\boldsymbol{\mu}, t) + h. o. t \quad (5)$$

式中, $h. o. t$ 代表展开式的更高次项,而 $M(\boldsymbol{\mu}, t)$ 是图像的雅各比约束矩阵^[14], I_t 为任意时刻的图像,它仅与目标运动矢量和亮度变化相关,由于稀疏点集数量为 N , 且运动矢量是 n 维,所以雅各比约束 $M(\boldsymbol{\mu}, t)$ 为 $N \times n$ 维,可推导出其各个元素形式如下:

$$m_{ij} = I_{\mu_j}(f(X_i, \boldsymbol{\mu}), t) = \nabla_f I(f(X_i, \boldsymbol{\mu}), t)' f_{\mu_j}(X_i, \boldsymbol{\mu}) \quad (6)$$

式中, $\nabla_f I$ 是跟踪目标与运动模型相关的灰度梯度。

将展开后的 $I(\boldsymbol{\mu}(t) + \delta\boldsymbol{\mu}, t + \tau)$ 带入式(4)并忽略更高次项,可得

$$O(\delta\boldsymbol{\mu}) = \|I(\boldsymbol{\mu}, t) + M\delta\boldsymbol{\mu} + \tau \times I_t - I(0, t_0)\|^2 \quad (7)$$

在进一步近似后,令 $\tau \times I_t(\boldsymbol{\mu}, t) \approx I(\boldsymbol{\mu}, t + \tau) - I(\boldsymbol{\mu}, t)$, 可得

$$O(\delta\boldsymbol{\mu}) \approx \|I(\boldsymbol{\mu}, t + \tau) + M\delta\boldsymbol{\mu} - I(0, t_0)\|^2 \quad (8)$$

为得到 $\delta\boldsymbol{\mu}$ 的最小值,令式(8)右式为零,则有:

$$\delta\boldsymbol{\mu} = - (M^T M)^{-1} M^T [I(\boldsymbol{\mu}, t + \tau) - I(0, t_0)] \quad (9)$$

式(9)就是目标跟踪的基本模型,由此可知,只有在跟踪过程中计算雅各比约束 $M(\boldsymbol{\mu}, t)$, 才能计算目标的运动矢量的变化 $\delta\boldsymbol{\mu}$ 。

由式(6)可知,由于 $M(\boldsymbol{\mu}, t)$ 矩阵是与时间参数相关的,跟踪过程中对于每一帧都需要重新计算 $M(\boldsymbol{\mu}, t)$, 而计算代价是巨大的,因此必需重新分解式(6),将其中与时间无关的部分分解出来,在跟踪之前预先计算。

为了简化 $M(\boldsymbol{\mu}, t)$ 的形式,首先分解 $\nabla_f I$ 并将其用梯度公式展开,可得:

$$\nabla_x I(\mathbf{X}, t_0) = f_x(\mathbf{X}, \boldsymbol{\mu})' \nabla_f I(f(\mathbf{X}, \boldsymbol{\mu}), t) \quad (10)$$

代入式(6),在不考虑运动模型的情况下 ($M(\boldsymbol{\mu}, t)$ 对运动模型的约束在 2.2 节讨论)可以将 $M(\boldsymbol{\mu}, t)$ 整理为

$$M(\boldsymbol{\mu}) = \begin{pmatrix} \nabla_x I(X_1, \boldsymbol{\mu}_0)' f_x(X_1, \boldsymbol{\mu})^{-1} f_{\mu_j}(X_1, \boldsymbol{\mu})^{-1} \\ \nabla_x I(X_2, \boldsymbol{\mu}_0)' f_x(X_2, \boldsymbol{\mu})^{-1} f_{\mu_j}(X_2, \boldsymbol{\mu})^{-1} \\ \dots \\ \nabla_x I(X_N, \boldsymbol{\mu}_0)' f_x(X_N, \boldsymbol{\mu})^{-1} f_{\mu_j}(X_N, \boldsymbol{\mu})^{-1} \end{pmatrix} \quad (11)$$

其中, $f_x(X_i, \boldsymbol{\mu})$ 是运动模型在方位矢量上的偏导数,而 $f_{\mu_j}(X_i, \boldsymbol{\mu})$ 是运动模型在运动矢量上的偏导数。其中 $\nabla_x I(X_i, \boldsymbol{\mu}_0)'$ 与时间无关,即其值在跟踪过程中恒定不变, $f_{\mu_j}(X_i, \boldsymbol{\mu})$ 由于与运动矢量 $\boldsymbol{\mu}(t)$ 相关,因此,其值随时间而改变的,而 $f_x(X_i, \boldsymbol{\mu})$ 则可以进一步分解,于是假设 $f_x(X_i, \boldsymbol{\mu})$ 中与时间无关部分为 $\Gamma(X_i)$, 式(11)可以写作:

$$M(\boldsymbol{\mu}) = \begin{pmatrix} \nabla_x I(X_1, \boldsymbol{\mu}_0)' \Gamma(X_1) \\ \nabla_x I(X_2, \boldsymbol{\mu}_0)' \Gamma(X_2) \\ \dots \\ \nabla_x I(X_N, \boldsymbol{\mu}_0)' \Gamma(X_N) \end{pmatrix} \boldsymbol{\Sigma}(\boldsymbol{\mu}) = M_0 \boldsymbol{\Sigma}(\boldsymbol{\mu}) \quad (12)$$

式中, M_0 是与时间无关部分,描述的就是跟踪目标的先验信息,可以在跟踪之前的初始化过程中进行计算。而 $\boldsymbol{\Sigma}(\boldsymbol{\mu})$ 明显依赖于运动矢量 $\boldsymbol{\mu}$, 因此,要在跟踪过程中实时计算。将分解并重新整理后的 $M(\boldsymbol{\mu}, t)$ 带入目标跟踪的基本模型即式(9)中可得到:

$$\delta\boldsymbol{\mu} = - \sum^{-1} (M_0^T M_0)^{-1} M_0^T [I(f(\mathbf{X}, \boldsymbol{\mu}), t_n) - I(\mathbf{X}, t_0)] \quad (13)$$

综上所述,目标跟踪算法流程可以分为初始化计算和实时跟踪两个阶段,具体流程如图 1 所示。其中

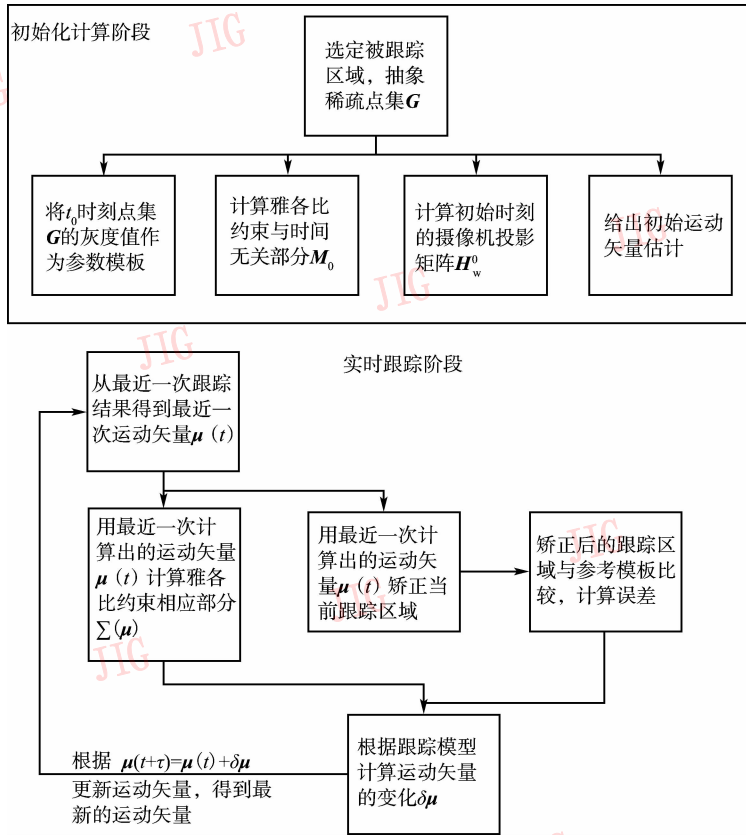


图 1 目标跟踪算法流程图

Fig. 1 Working flow of target tracking algorithm

在初始化计算阶段,还需要计算初始时刻的摄像机透射投影矩阵,其目的是为了求解摄像机运动外参数。

2.2 目标跟踪的透射投影运动模型

要跟踪目标的运动,首先要定义目标的运动模型,如平移、旋转、仿射变换等,由于研究目标是求解环境的 3 维注册信息,因此在跟踪过程中,应该将运动模型定义为透射投影变换运动^[15,16]。

设 $(u, v)^t$ 是点 X 的笛卡尔坐标,而 $(x_c, y_c, z_c)^t$ 是相应的摄像机投影坐标,如图 2 所示:

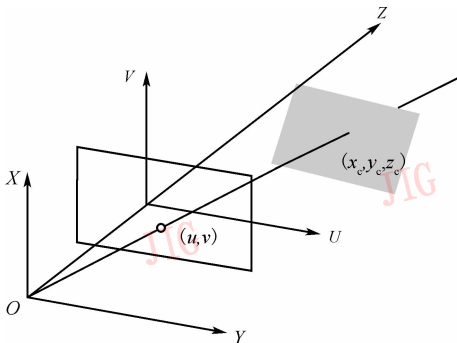


图 2 跟踪目标平面在摄像机坐标系中的透射投影模型

Fig.2 Projection model of the tracked target in camera coordinate

其关系如下

$$X_h = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \rightarrow X = \begin{pmatrix} x_c/z_c \\ y_c/z_c \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \quad \forall X \in I(X) \quad (14)$$

由于约定跟踪目标在一个平面内,运动模型描述的是跟踪目标的 2 维运动变换^[16],即

$$f_X(X, \mu) = PX_h = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & 1 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \quad (15)$$

式中, P 为投影矩阵。此时运动矢量 μ 可以定义为 $\mu = (a, b, c, d, e, f, g)^t$

进一步将雅各比约束根据透射投影变换运动模型展开,分别可求出:

$$\nabla_X I(X, \mu)^t = \left(\frac{\partial I}{\partial u}, \frac{\partial I}{\partial v}, - \left(u \frac{\partial I}{\partial u} + v \frac{\partial I}{\partial v} \right) \right) \quad (16)$$

$$f_X(X, \mu)^{-1} = P^{-1} \quad (17)$$

$$f_\mu(X, \mu) = \begin{pmatrix} x_c & 0 & 0 & y_c & 0 & 0 & z_c & 0 \\ 0 & x_c & 0 & 0 & y_c & 0 & 0 & z_c \\ 0 & 0 & x_c & 0 & 0 & y_c & 0 & 0 \end{pmatrix} \quad (18)$$

并且可以计算得

$$f_X(\mathbf{X}, \boldsymbol{\mu})^{-1} f_\mu(\mathbf{X}, \boldsymbol{\mu}) = (\mathbf{R}\mathbf{P}^{-1} | \mathbf{S}\mathbf{P}^{-1} | \mathbf{T}\mathbf{P}_{1,2}^{-1}) = \Gamma(\mathbf{X}) \boldsymbol{\Sigma}(\boldsymbol{\mu}) \quad (19)$$

其中, $\mathbf{P}_{1,2}^{-1}$ 是矩阵 \mathbf{P}^{-1} 的第 1 列与第 2 列, $\mathbf{R}, \mathbf{S}, \mathbf{T}$ 为摄像机相对于场景的旋转, 缩放平移矩阵, 由于根据约定 $\Gamma(\mathbf{X}_i)$ 是与时间无关部分, 因此, 可以得到

$$\Gamma(\mathbf{X}) = (x_c I_{3 \times 3} | y_c I_{3 \times 3} | z_c I_{3 \times 3}) \quad (20)$$

$$\boldsymbol{\Sigma}(\boldsymbol{\mu}) = \begin{pmatrix} \mathbf{P}^{-1} & 0 & 0 \\ 0 & \mathbf{P}^{-1} & 0 \\ 0 & 0 & \mathbf{P}_{1,2}^{-1} \end{pmatrix} \quad (21)$$

由式(21)可知, 矩阵 $\boldsymbol{\Sigma}(\boldsymbol{\mu})$ 是一个 9×8 的不可逆矩阵, 代入式(13)就可以得到基于透射投影变换运动模型的目标跟踪最终模型如下:

$$\delta \boldsymbol{\mu} = -(\boldsymbol{\Sigma}^t \mathbf{M}_0^t \mathbf{M}_0 \boldsymbol{\Sigma})^{-1} \boldsymbol{\Sigma}^t \mathbf{M}_0^t [I(f(\mathbf{X}, \boldsymbol{\mu}), t_n) - I(\mathbf{X}, t_0)] \quad (22)$$

2.3 3 维位姿估计

通过 2.1 节和 2.2 节的理论推导和分析可知, 根据跟踪目标的运动可以计算出跟踪目标, 从而得出所有的点在每一帧中对应关系如下:

$$(X_{1 \dots N}, t) = \mathbf{H}_0^n (X_{1 \dots N}, t_0) \quad (23)$$

式中, \mathbf{H}_0^n 表示 t_n 时刻点集与 t_0 时刻点集的对应关系, 即为单应性矩阵。摄像机透射投影模型与跟踪目标的关系如图 3 所示。为了简化透射投影方程, 将场景坐标系定义在跟踪目标上, 即跟踪平面为 Z 平面。

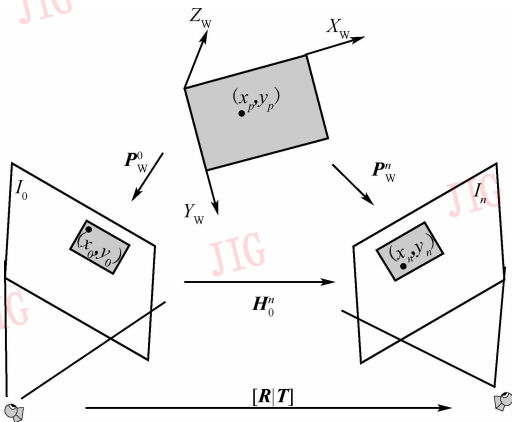


图 3 跟踪过程中的目标投影和摄像机运动关系

Fig. 3 The target projection and the moving camera in tracking process

用 (x_p, y_p) 表示跟踪目标上的任意点的坐标, (x_0, y_0) 表示在 t_0 时刻其投影到摄像机平面上坐标, (x_n, y_n) 在 t_n 时刻其投影到摄像机平面上坐标, 由图 3 可知,

$$\begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix} = \mathbf{P}_w^0 \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} \quad (24)$$

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \mathbf{P}_w^n \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} \quad (25)$$

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \mathbf{H}_0^n \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix} \quad (26)$$

式中, 在 t_0 时刻的投影矩阵 \mathbf{P}_w^0 可以在初始化阶段根据式(24)计算, 理论上只要从选定的点集中取 4 个点计算就可以求解投影矩阵, 为了提高精度, 可以适当增加点数。

视觉透射投影方程^[17]如下:

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \lambda K [\mathbf{R} | \mathbf{T}] \begin{pmatrix} x_p \\ y_p \\ z_p \\ 1 \end{pmatrix} \quad (27)$$

其中, λ 为缩放因子, K 为摄像机内参数, $[\mathbf{R} | \mathbf{T}]$ 为摄像机相对于场景的旋转、平移矩阵, 即摄像机外参数, 其中, $\mathbf{R} = [r_1 r_2 r_3]$ 为旋转矩阵。由于场景坐标定义在跟踪目标上, 因此 $[\mathbf{R} | \mathbf{T}]$ 就是跟踪目标在摄像机坐标系中的位置和姿态, 即为所求的跟踪目标的 3 维位姿。由于跟踪目标平面为 Z 平面, 所以 $z_p = 0$, 再将式(24)、(25)、(26)带入视觉透射投影方程, 即式(27)可得:

$$\mathbf{H}_0^n = \mathbf{P}_w^n (\mathbf{P}_w^0)^{-1} = \lambda K [r_1 r_2 | \mathbf{T}] (\mathbf{P}_w^0)^{-1} \quad (28)$$

而由于 \mathbf{R} 作为旋转矩阵的特性, 即 $|r_1| = |r_2| = |r_3|$ 且 r_1, r_2, r_3 互相垂直, 因此, 可以已知 r_1, r_2 计算出 r_3 。

如前所述, \mathbf{H}_0^n 可以根据目标跟踪计算得出, 因此求解式(28)就可以计算出跟踪目标平面的位姿 $[\mathbf{R} | \mathbf{T}]$ 。

3 无标志点的 AR 系统结构

如上文所述, AR 系统在构成上主要有几个关键部分: 摄像与处理系统、注册定位系统、融合渲染系统与立体显示系统。围绕本文提出的基于实时目标跟踪的 3 维环境注册算法, 可以构建完整的无标志点的 AR 系统。其中采用已标定的高速采集单目摄像机作为场景采集设备, 并将摄像机安置在使用

者的头部,在通过本算法进行精确注册的基础上,使用高效图形渲染引擎 OGRE 进行图像融合和渲染,

最后输出到头盔显示设备(HMD),系统结构如图4所示:

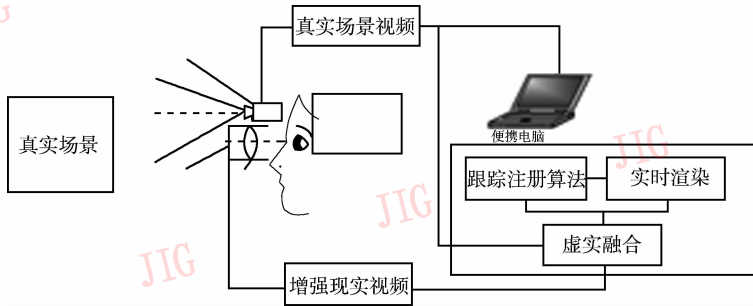


图4 增强现实系统结构

Fig. 4 The architecture of the AR system

系统工作流程如下:

(1) 系统预处理

①建立景物的虚拟3维模型;

②对摄像机进行标定,得到摄像机内参数 K ,并且存储;

③在场景中选定一个平面跟踪区域作为跟踪目标,并存储为跟踪模板;

④初始化计算相关参数,并且存储。

(2) 根据算法注册并渲染融合

①系统工作时,通过摄像机跟踪选定目标,并且实时计算出跟踪目标的3维位姿;

②渲染引擎得到跟踪目标的3维位姿进行渲染后,再对虚拟场景进行渲染,得到虚拟图像。

③将虚拟图像叠加到摄像机采集的实时视频流中得到增强后的场景。

(3)将虚实结合的增强场景输出到HMD中以实现增强现实。

用本算法对选定区域进行跟踪。试验过程如下,首先初始化场景,选定跟踪目标,计算相关参数,并存储参考模板,如图5所示,图6(a)、6(b)、6(c)、6(d)为跟踪过程,为了验证算法的抗噪声能力,人为加入了遮挡物对目标进行干扰,加入了强烈的光照变化测试其鲁棒性,如图6(c)所示。在本试验中,图像采用 320×240 大小的标准输入图像,在跟踪区域中,抽样120个稀疏点构成跟踪目标。

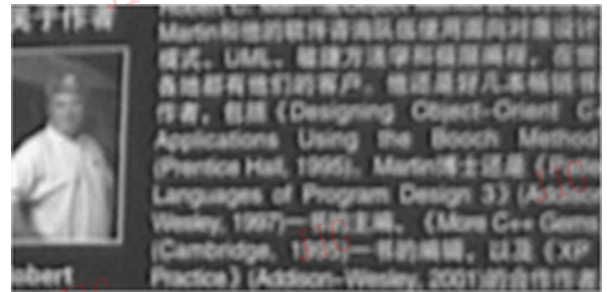


图5 跟踪目标模板

Fig. 5 The target template

4 实验与分析

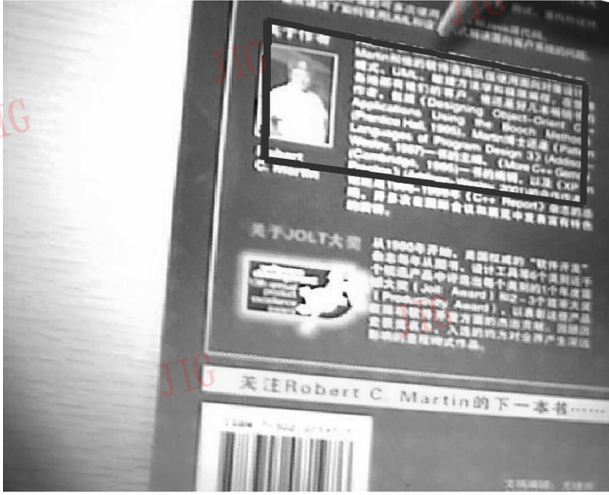
目前已经在普通图形工作站上实现基于模板跟踪的环境注册算法。针对本文所开发的算法的试验分为两个阶段,首先对某一具有复杂背景的目标进行跟踪并在场景中标出目标,同时记录跟踪目标与参考模板的残差来验证跟踪算法的准确性;其次根据环境注册算法的计算结果,在摄像机实时采集的场景中叠加虚拟信息,并使真实环境和虚拟物体无缝融合,实现增强现实的效果。

在第一阶段控制一个具有复杂背景的物体在环境中自由运动,通过手持摄像机采集场景图像,再采

为定量分析算法的准确性,将跟踪结果矫正以后与预存模板进行比较,从而得到跟踪的误差分析曲线,如图7所示。

由实验结果和误差曲线可知,无论在有无遮挡的干扰下,目标跟踪的精度都是亚像素级的,即在人类视觉的可视范围内基本没有误差。

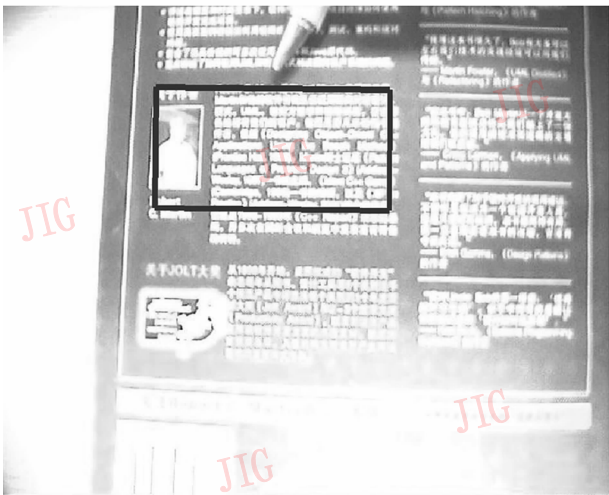
跟踪完成之后,可以根据计算出的环境3维注册信息将虚拟物体渲染并叠加到真实场景之上,实现增强现实系统。实验效果如图8所示,图8(a)和8(b)分别是增强视频中的两帧图像。由于跟踪精度高,环境注册精度也足够高,因此虚拟物体能够自然地与真实场景相融合。



(a)实时跟踪中的第37帧图像



(b)实时跟踪中的第53帧图像



(c)实时跟踪中的第69帧图像

图 6 实时跟踪结果
Fig. 6 The real-time tracking results

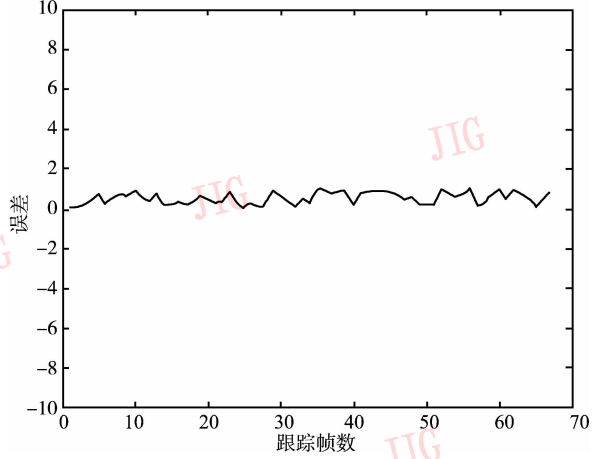


图 7 模板跟踪的误差分析曲线

Fig. 7 The curve for error analysis in template tracking



(a)



(b)

图 8 基于模板跟踪实现增强现实
Fig. 8 The augmented reality based on template tracking

5 结 论

论文提出了一种高效的基于模板跟踪的环境注册算法,实现了无标志点的3维环境实时注册。由于该算法可以实时实现并且准确性高,在该算法的基础上可以实现高效稳定的无标志点的AR系统。

由于本文算法在跟踪目标的过程中使用选定点集的灰度信息进行计算求解,因此,相对单个或者少量标志点的跟踪,有较大的稳定性优势,对光照变化的抗干扰能力较强。目前,正在进行的研究是对环境光照和光源建模,并且在跟踪过程中计算光照模型的参数,从而进一步提高其鲁棒性。此外还应该进一步研究存在遮挡和干扰物体时的情况,并在这些情况下判断跟踪结果的可信度并适当调整迭代计算次数。而本文算法的明显缺点就是比较依赖于环境给定的初始值,在增强现实的系统中,可以利用场景中的一些特定信息来给定初值,或者利用一些高精度的定位设备(如差分GPS)来给定初值。这些问题都是值得深入分析和探讨的,将在下一步的工作中进一步研究。

参考文献 (References)

- Chen Jing, Shi Qi, Wang Yong-tian. The application of augmented reality[J]. Computer Engineering and Applications, 2001, 37(21): 55~57. [陈靖,施琦,王涌天. 增强现实技术及其应用[J]. 计算机工程与应用, 2001, 37(21): 55~57.]
- Lin Liang, Liu Yue, Wang Yong-tian. Outdoor registration algorithm based on image match[J]. IEE Proceedings on Vision, Image & Signal Processing, 2007, 153(1): 57~62.
- Lin Liang, Wang Yong-tian, Liu Yue, *et al*, Outdoor Registration Method Based on FFT [J]. Journal of image and graphics, 2005, 19(10): 1146~1151. [林惊,王涌天,刘越等. 基于FFT的户外增强现实的环境注册算法[J]. 中国图象图形学报, 2005, 19(10): 1146~1151.]
- Wang Yong-tian, Lin Liang, Liu Yue. Outdoor augmented reality and its application in digital reconstruction of yuanmingyuan[J]. Bulletin of National Natural Science Foundation of China, 2006, 20(2): 76~81. [王涌天,林惊,刘越. 亦真亦幻的户外增强现实系统——圆明园的数字重建[J]. 中国科学基金, 2006, 20(2): 76~81.]
- Bajura M, Henry F, Ohbuchi R. Merging virtual reality with the real world: Seeing ultrasound imagery within the patient [J]. Computer Graphics, 1992, 26(2): 203~210.
- Nakazawa Y, Nakano S, Komatu T, *et al*. A system for composition of real moving images and CG images based on image feature points[J]. The Journal of the Institute of Image Information and Television Engineers, 1997, 51(7): 1086~1095.
- Shiping Z, Yang G. Non-contact 3D coordinates measurement of cross-cutting feature points on the surface of large-scale workpiece based on machine vision method [A]. In: Proceedings of IEEE Instrumentation and Measurement Technology Conference [C]. Ottawa, Canada, 2005:1255~1259.
- Zhou Ya, Yan Da-yuan, Wang Yong-tian, *et al*. A registration method in augmented reality[J]. Journal of Image and Graphics, 2000, 5(5): 430~433. [周雅,闫达远,王涌天等. 一种增强现实系统的3维注册方法[J]. 中国图象图形学报, 2000, 5(5): 430~433.]
- Black M J, Jepson A D. Eigentracking: robust matching and tracking of articulated objects using a view-based representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 26(1): 63~84.
- Hutchinson S, Hager G D, Corke P. A tutorial introduction to visual servo control [J]. IEEE Transactions Robotics and Automation, 1996, 12(5): 651~670.
- Stauffer C, Kih T. Automated multi-camera planar tracking correspondence modeling[A]. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. Madison Wisconsin, USA, 2003, 1:259~266.
- Betke M, Makris N. Fast object recognition in noisy images using simulated annealing[A]. In: Proceedings of the Fifth International Conference on Computer Vision [C]. Washington, DC. USA, 1995: 523~530.
- Rehg J, Kanade T. Visual tracking of high DOF articulated structures: An application to human hand tracking [A]. In: Proceedings of European Conference on Computer Vision [C]. Stockholm, Sweden, 1994, B: 35~46.
- Darrell T, Essa I A, Pentland A P. Task-specific gesture analysis in real-time using interpolated views [J]. IEEE Transactions Pattern Analysis and Machine Intelligence, 1996, 18(12): 1236~1242.
- Gleicher M. Projective registration with difference decomposition [A]. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. San Juan, Puerto Rico, 1997: 331~337.
- Murase H, Nayar S K. Visual learning and recognition of 3D object from appearance [J]. International Journal of Computer Vision, 1995, 18(14): 5~24.
- Simon G, Fitzgibbon A, Zisserman A. Markerless tracking using planar structures in the scene [A]. In: Proceedings of International Symposium on Augmented Reality [C], Alamos, CA, USA, 2000: 111~119.