

增强现实系统的设计与实现

全红艳, 王长波, 林俊隽

QUAN Hong-yan, WANG Chang-bo, LIN Jun-juan

华东师范大学 软件学院, 上海 200062

Institute of Software Engineering, East China Normal University, Shanghai 200062, China

E-mail: hyquan@sei.ecnu.edu.cn

QUAN Hong-yan, WANG Chang-bo, LIN Jun-juan. Design and implementation of augmented reality system. Computer Engineering and Applications, 2009, 45(28): 164-167.

Abstract: A dynamic video tracking method based on planar structure is studied. By calculating the homography of the initial frame between the real and virtual object, homography of any frame between the real and virtual object can be obtained. The purpose of dynamic tracking can be achieved. On this basis this paper designs and implements dynamic tracking system based on planar structure and realizes dynamic tracking of the video. The augmented reality system has the characteristics of simplicity, convenience and efficiency.

Key words: augmented reality; dynamic tracking; planar structures; homography

摘要: 研究了一种基于平面特征的视频动态跟踪方法, 通过计算出初始帧图像的虚实坐标之间的单应矩阵, 可以计算出任意一帧的虚实坐标之间的单应矩阵, 可以实现动态跟踪的目的。在此基础上设计并实现了基于平面特征的增强现实动态跟踪系统, 实现了视频的动态跟踪功能。该增强现实跟踪系统具有实时性好、简单、方便的特点。

关键词: 增强现实; 动态跟踪; 平面结构; 单应矩阵

DOI: 10.3778/j.issn.1002-8331.2009.28.049 文章编号: 1002-8331(2009)28-0164-04 文献标识码: A 中图分类号: TP391

1 引言

随着计算机技术的快速发展, 增强现实技术得到了广泛的应用, 该技术的研究目前已经成为计算机视觉、计算机图形学等领域中研究的一个热点。增强现实技术是随着虚拟现实技术发展而出现的将真实场景同虚拟物体融合的一种技术手段, 增强现实系统的特点是虚实空间的一致性结合以及实时交互, 具体地包括 3D 物体的注册技术、显示技术、跟踪和定位技术等。该技术的研究主要应用在虚拟导游、虚拟产品浏览、虚拟房地产景观系统等。目前, 增强现实技术的研究中还存在着一些问题, 其技术的研究有着深远的意义。近些年研究人员利用该方法对于虚拟注册及动态跟踪技术进行了研究^[1-2]。Simon 等在 2000 年提出了一种基于平面特征的无缝跟踪方法^[3-5]。在他的研究中, 通过跟踪 3D 空间的参考平面的特征点, 计算初始平面单应矩阵, 从而实现了虚拟对象的嵌入。Yuan 等利用射影重建的方法^[6], 实现了虚拟场景的动态跟踪, 但存在的主要问题是射影重建的精度比较低, 从而影响了该方法应用。Lourakis 等利用线性优化的策略^[7], 通过对场景中某一虚拟平面的跟踪, 完成了对虚拟对象的动态跟踪功能, 该研究中存在的问题是非线

性优化导致算法计算复杂度大。目前在利用场景平面特征进行的研究中存在的主要问题是: (1) 利用 3D 模板进行研究, 给研究带来了不便; (2) 需要进行 3D 模板与 2D 图像之间的特征点匹配, 给研究带来了误差。

为了克服现有的基于平面场景特征研究中存在的这些问题, 研究了一种利用平面场景中的特征点, 计算出初始帧图像的虚实坐标之间的单应矩阵; 继而, 可以计算出任意一帧的虚实坐标之间的单应矩阵, 从而实现虚拟物体的注册和动态跟踪。利用了 VC++6.0 软件和 OpenGL 图形用户接口, 实现了基于视频的增强现实系统, 证明了该方法的有效性。

2 虚实坐标之间单应矩阵的计算方法

2.1 平面的单应矩阵

摄像机采用小孔成像模型的情况下^[8], 三维空间中的点 $M(X, Y, Z)^T$ 像的坐标为 $m(u, v)^T$, 那么:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P_w \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K[r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K[r_1 \ r_2 \ t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = H_w^0 \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (1)$$

基金项目: 国家自然科学基金青年项目(the National Natural Science Foundation of China under Grant No.60603076); 国家高技术研究发展计划(863)(the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z316); 上海市科委自然科学基金项目(the Natural Science Foundation of Shanghai Municipality, China under Grant No.07ZR14035); 上海市青年科技启明星项目(Shanghai Rising-Star Program No.08QA14025)。

作者简介: 全红艳(1968-), 博士, 副教授, 研究方向为: 计算机视觉、虚拟现实、增强现实等; 王长波(1976-), 博士, 副教授, 研究方向为虚拟现实、计算机图形学等; 林俊隽, 硕士研究生, 研究方向为虚拟现实。

收稿日期: 2008-06-11 **修回日期:** 2008-09-08

其中 P_w 为 3×4 的投影矩阵, K 为摄像机的内部参数矩阵, $r_i (i=1, 2, 3)$ 及 t 为外部参数的列向量。 H_w^0 为视频第 0 帧的单应矩阵, 表示 3D 空间中 XOY 的 2D 平面的点与图像平面的点的对应关系。

设 H_w^i 表示视频第 i 帧的单应矩阵, 如果能够求得 H_w^i , 那么在视频的跟踪过程中, 根据 H_w^i 便可以得到该帧图像点的 3D 坐标和 2D 坐标关系, 便可以实现实时的跟踪过程。

设 H_w^{k+1} 表示相邻两帧第 $k+1$ 和 k 帧之间的单应矩阵, 则^[3]:

$$H_w^i = H_w^{i-1} H_w^{i-2} H_w^{i-3} \cdots H_w^1 H_w^0 \quad (2)$$

可见, 只要求得相邻两帧之间的单应矩阵以及 H_w^0 便可以得到 H_w^i , 从而可以实现对虚拟对象动态的跟踪。

2.2 虚拟空间坐标系的建立

世界坐标系的建立方法是, 以图像中的平面场景所在的平面作为世界坐标系的 XOY 平面, 垂直 XOY 平面且向上的方向为 Z 轴方向, 如图 1 所示。

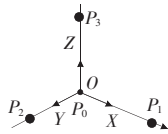


图 1 3D 空间的坐标系的示意图

现有的方法在研究时, 通过 3D 和 2D 点的对应来进行研究, 这样做的缺点是误差较大, 并且确定对应关系也给研究带来了一些问题。该文首先根据两幅图像, 取平面上的一条直线作为 3D 空间的 X 轴的像, 再根据摄像机的内部参数约束以及空间 3D 坐标系的 X, Y 轴的垂直约束关系, 可以求得 Y 轴的图像, 进而可以得到摄像机的外部参数, 即旋转矩阵及平移矩阵, 从而可以实现动态地嵌入及跟踪虚拟对象。

2.2.1 世界坐标系中 X 轴像的获取

设视频的初始两帧的图像为 I_0 和 I_1 , 其对应点的坐标假设为 $m(u, v)^T$ 及 $m'(u', v')^T$, 则有:

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = H_w^1 \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3)$$

其中 H_w^1 为该两帧图像之间的单应矩阵。一对对应点具有两个约束:

$$u'(h_{31}u + h_{32}v + h_{33}) = h_{11}u + h_{12}v + h_{13} \quad (4)$$

$$v'(h_{31}u + h_{32}v + h_{33}) = h_{21}u + h_{22}v + h_{23} \quad (5)$$

任意四对非共线的对应点就可以求得 H_w^1 。采用先提取两帧图像之间的多对对应点, 然后利用非线性投影误差最小进行优化:

$$\varepsilon = \sum d^2(m' - H_w^1 m) \quad (6)$$

其中 ε 为投影误差的平方和。 X 轴的像的获取算法步骤是:

步骤 1 在 I_0 和 I_1 图像上, 利用快速的点对应的方法(见 2.4 节), 求取它们之间的多对对应点;

步骤 2 设在 I_0 图像的平面区域求得的线段长度为 l 的两个对应点, 对于 I_0 图像的平面内任意两点 $p_0(u_{p_0}, v_{p_0})$ 和 $p_1(u_{p_1}, v_{p_1})$ 同时满足以下的条件:

(1) p_0 和 p_1 在平面 I_0 上;

(2) p_0 和 p_1 之间距离 $dis(p_0, p_1)$ 为 l , 即: $\{(p_0, p_1) | dis(p_0, p_1) = l\}$, 其中:

$$dis(p_0, p_1) = \sqrt{(u_{p_0} - u_{p_1})^2 + (v_{p_0} - v_{p_1})^2} \quad (7)$$

(3) p_0 和 p_1 之间所有像素点的非线性投影误差的平方和最小。

2.2.2 世界坐标系中 Y 轴像的获取

该文研究中思想方法是, 如果能够在 I_0 图像的平面内再求得 Y 轴的像 $p_0 p_2$, 就可以根据彼此垂直的 OX, OY 和 OZ 轴建立世界坐标系, 便可以实现虚实空间的影像。

设线段 $p_0 p_2$ 对应的三维空间的线段为 $P_0 P_2$, 对应的矢量 $P_0 P_2$ 作为世界坐标系的 Y 轴的方向, 设 $P_0 P_3$ 是世界坐标系的 Z 轴的方向上的矢量。用 p_0, p_1, p_2 和 p_3 分别表示 4 个点 (P_0, P_1, P_2 和 P_3) 在 I_0 平面上的坐标, 用 P_0, P_1, P_2 和 P_3 分别表示 4 个点在三维世界坐标系中坐标, 那么^[9]:

$$\lambda_0 p_0 = KRP_0 + Kt = Kt \quad (8)$$

$$\lambda_1 p_1 = KRP_1 + Kt \quad (9)$$

$$\lambda_2 p_2 = KRP_2 + Kt \quad (10)$$

$$\lambda_3 p_3 = KRP_3 + Kt \quad (11)$$

其中 $P_0 = (0, 0, 0, 1)^T$, K 为摄像机的内部参数矩阵, R 和 t 分别为摄像机外部参数矩阵, 它们的计算在下面进行讨论, 不失一般性设 $\lambda_0 = 1$, 把式(8)代入到式(9)、(10)和(11)中, 得:

$$K^{-1}(\lambda_1 p_1 - p_0) = RP_1 \quad (12)$$

$$K^{-1}(\lambda_2 p_2 - p_0) = RP_2 \quad (13)$$

$$K^{-1}(\lambda_3 p_3 - p_0) = RP_3 \quad (14)$$

再根据 $P_0 P_1, P_0 P_2$ 和 $P_0 P_3$ 的正交性, 得^[9]:

$$(\lambda_1 p_1 - p_0)^T K^{-1} K^{-1} (\lambda_2 p_2 - p_0) = 0 \quad (15)$$

其中

$$K^{-1} K^{-1} = \begin{bmatrix} f_u^{-2} & 0 & 0 \\ 0 & f_u^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

如果在 λ_1, λ_2 和 λ_3 计算后, 由式(12)、(13)和(14)可以得到旋转矩阵的列向量为^[9]:

$$r_1 = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \end{bmatrix} = K^{-1}(\lambda_1 p_1 - p_0) \quad (17)$$

$$r_2 = \begin{bmatrix} r_{21} \\ r_{22} \\ r_{23} \end{bmatrix} = K^{-1}(\lambda_2 p_2 - p_0) \quad (18)$$

这里研究了 p_2, λ_1 和 λ_2 的求取方法。在图 1 所示的世界坐标系中, 设在 OX 和 OY 轴方向上的矢量长度分别为 m 和 n , 那么根据式(18)便可以得到:

$$\|K^{-1}(\lambda_1 p_1 - p_0)\| = m \quad (19)$$

在摄像机标定的情况下, 给定 m 时, 可以得到 λ_1 , 再代入式(18), 便得到了旋转矩阵的列向量 r_1 , 因为 H_w^0 的旋转矩阵具有单位正交性, 所以 H_w^0 的第一列向量可以求得为:

$$\begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \end{bmatrix} = \begin{bmatrix} r_{11} \\ \|r\| \\ r_{12} \\ \|r\| \\ r_{13} \\ \|r\| \end{bmatrix} \quad (20)$$

在求得 $\lambda_1, p_0(u_0, v_0)$ 及 $p_1(u_1, v_1)$ 后, 根据式(18), 得 $\lambda_2 = fun_1(u_2, v_2)$, fun_1 表示 u_2, v_2 和 λ_2 的一个约束关系。同时根据 OX 轴与 OY 轴的正交性, 即根据式(15)便可以得到 u_2, v_2 和 λ_2 的另一个约束关系 $\lambda_2 = fun_2(u_2, v_2)$, 利用这两个约束中求得的 λ_2 的差值最小来求得 u_2, v_2 , 继而可以求得 λ_2 。同样根据式(18), 同时 H_w^0 的旋转矩阵具有单位正交性, 便可以求得 H_w^0 的第二列向量; 再根据式(8)便可以求得平移向量 t 。摄像机的外部参数求得后, 便可以实现虚拟对象的注册与动态跟踪。

增强现实系统研究中关键技术之一就是虚实坐标系的对应问题, 这种对应主要是利用 H_w^0 进行映像。在上述建立的世界坐标系下, H_w^0 的计算步骤如下:

步骤 1 相邻两帧图像的快速点对应方法求取视频的起始两帧 I_0 和 I_1 的对应点;

步骤 2 在视频的第 0 帧 I_0 上平面区域中, 将投影误差的平方和 ε 最小作为约束(式(6)), 求取一段线段 $p\phi$;

步骤 3 计算摄像机的内部参数矩阵 K (下述方法);

步骤 4 根据摄像机的内部参数 K , 采用式(19)计算出比例因子 λ_1 ;

步骤 5 根据坐标轴的正交性, 可以求得世界坐标系中的 OY 轴上向量 P_0P_2 的像;

步骤 6 利用式(17)和(18)可以计算出旋转矩阵 R 的前两列的向量 r_1 和 r_2 , 再根据旋转矩阵 R 的正交性, 求得第三列向量 $r_3 = r_1 \times r_2$; 根据式(8)便可以求得平移矩阵 t ;

步骤 7 根据 $H_w^0 = K[r_1 \ r_2 \ r_3 \ t]$, 可以求得虚实坐标系对应的 H_w^0 。

该虚实坐标系的对应方法的主要特点是直接利用视频的初始两帧就可以求得 H_w^0 , 是一种简单、实用的方法。

2.3 摄像机内部参数 K 的计算

假设摄像机的内部参数矩阵的形式为:

$$K = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (21)$$

其中 f 为摄像机的焦距, (c_x, c_y) 为图像的主点坐标, 在研究中, 假设像素的扭曲为 0, 纵横比为 1, 主点坐标为图像的中心点时, 并根据旋转矩阵的性质, 平面单应矩阵 H 中的旋转向量 r_1 和 r_2 的内积为 0, 便可以求得摄像机的焦距 $f^{[10]}$:

$$f^2 = \frac{(h_{11}-c_x h_{31})(h_{12}-c_x h_{32}) + (h_{21}-c_y h_{31})(h_{22}-c_y h_{32})}{-h_{31} h_{32}} \quad (22)$$

2.4 相邻两帧图像的快速点对应方法

图像间的点对应问题一直是视觉研究中的一个主要问题, 视频图像相邻帧之间的视差比较小, 虚拟对象注册以及动态跟踪时需要一种图像帧之间快速的点对应的方法。

Zhang 提出了基于区域的图像点对应方法^[11], 其主要利用图像间的子窗口进行强度相关测试来确定对应点。该方法进行点对应的特点是准确。借助该方法的准确性, 研究了一种改进的视频的帧图像之间的快速准确的对应方法。

对于图像 I_0 中的角点, 设其坐标为 $pm(u, v)^T$, 并以该点为中心定义一个大小为 $(2n+1) \times (2m+1)$ 的相关窗口。选择第二幅图像 I_1 中的 $pm'(u', v')^T$, 利用式(7)计算 $dis(pm, pm')$, 求取其中满足 $dis(pm, pm') < d$ 的点, d 为一个较小的常数, 实验研究中取 d 为 10。新方法在确保准确的基础上, 大大提高了对应的速度。

3 视频的动态跟踪

动态跟踪过程的研究中, 采用动态地计算每帧图像的单应矩阵 H_w^j , 根据相邻帧之间的单应矩阵以及 H_w^0 , 利用式(2)可以计算出 H_w^j 。视频的动态跟踪过程是:

(1) 根据初始两帧的图像, 计算 H_w^0 ;

(2) 计算相邻帧之间的单应矩阵 H_j^{j+1} , 并计算每帧图像的单应矩阵 H_w^j ;

(3) 根据当前 H_w^j , 求虚拟物体的二维投影, 生成虚实融合的场景。

4 增强现实系统的设计与实现

4.1 系统的结构

该文设计并实现了基于平面结构的增强现实系统。系统开发的硬件环境是 Intel Pentium、内存 512 M 的 PC 机; 软件环境是 VC++6.0 和 OpenGL。整个系统分为预处理和实时跟踪两部分。利用 Sony7051 摄像机进行了拍摄, 得到了视频。

预处理部分的功能主要是完成初始帧虚实之间的单应矩阵 H_w^0 以及相邻帧间的单应矩阵 H_k^{k+1} 的计算, 并将预处理的数据记录到文本文件中。 H_w^0 的计算是根据上述第 2 章中介绍的方法进行的。首先提取出相邻帧图像的对应点, 然后再进行计算; H_k^{k+1} 的计算根据式(2)进行。文本文件中包括相邻帧间的单应矩阵 H_k^{k+1} 信息。

在研究过程中, 拍摄的视频是 avi 格式, 在预处理时, 首先进行了读取视频文件的操作, 以位图形式将视频读取出来。在完成过程中, 定义了一个 avi 视频读取的类 CAvi。

CAvi 类的数据信息包括位图图像的高度、宽度、位图数据成员 pData、位图的文件头信息、avi 的流信息; CAvi 类的动态成员包括读 avi 文件的第 mFrame 帧的成员函数 AviRead, 其形式为 BOOL AviRead(int mFrame)。此外, 还有 CAvi 类的初始化的成员函数 CAviCreate, 该函数中主要完成位图大小的初始化以及位图信息头的初始化工作。avi 视频的读取的过程, 如图 2 所示。

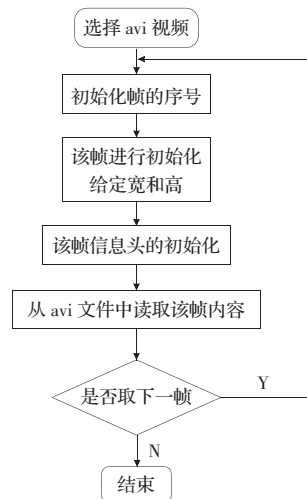


图 2 avi 文件读取的流程

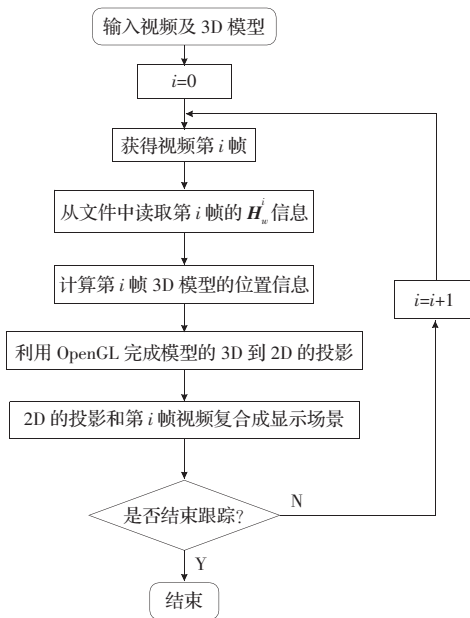


图3 实时跟踪的流程图

实时跟踪部分是系统的核心部分,主要完成的功能是增强现实场景的跟踪与实时绘制。

实时跟踪部分的主要任务是实时计算虚拟物体的三维空间的位置,并嵌入到真实场景中。实时跟踪部分的系统流程如图3所示。

4.2 X轴像的提取结果

图4是利用该方法对两帧图像的第一帧进行平面区域的线段提取的结果,其中标志黑色的线段是提取的 p_0p_1 线段的结果。设线段 p_0p_1 对应的三维空间的线段为 P_0P_1 ,把 P_0 作为世界坐标系的原点 O , P_0P_1 对应的矢量 P_0P_1 作为世界坐标系的X轴的正方向进行研究。

4.3 动态的视频跟踪的结果

利用该文的基于视频的动态跟踪算法,计算出视频的相应帧的单应矩阵,并实现了对虚拟物体的动态跟踪的功能,图5给出了实验中第4、7、10及13帧的动态跟踪结果,实验结果表明了该文基于视频的动态跟踪算法的正确性和有效性。

5 结论

对基于视频的增强现实技术进行了研究,平面场景是增强现实的实际应用中较为常见的情形,对平面场景应用背景下的技术开展了一些有益的研究,主要贡献如下:(1)研究了一种基于平面结构的虚拟物体新的注册方法;(2)给出了一种应用视频初始帧图像解决虚实坐标系的对应的策略,有效地解决了虚实坐标系的对应问题;(3)对视频的动态跟踪进行了研究,给出了视频动态跟踪算法,设计并实现了基于平面特征的增强现实系统,从而证明了该方法的有效性。

该文给出的算法不仅仅适用于室内近距离的平面场景,对于室外远距离的场景的情形,在采用无穷远点的单应矩阵的情形下,同样是可以使用的;另外,对于摄像机做纯旋转的情形时,相邻两帧图像间的映射同样可以用单应矩阵来解决,所以该文提出的方法是仍然适用的。

总之,该文方法有效地解决了增强现实技术中虚拟物体的注册和动态跟踪问题,克服了现有方法的3D模板(或参考图

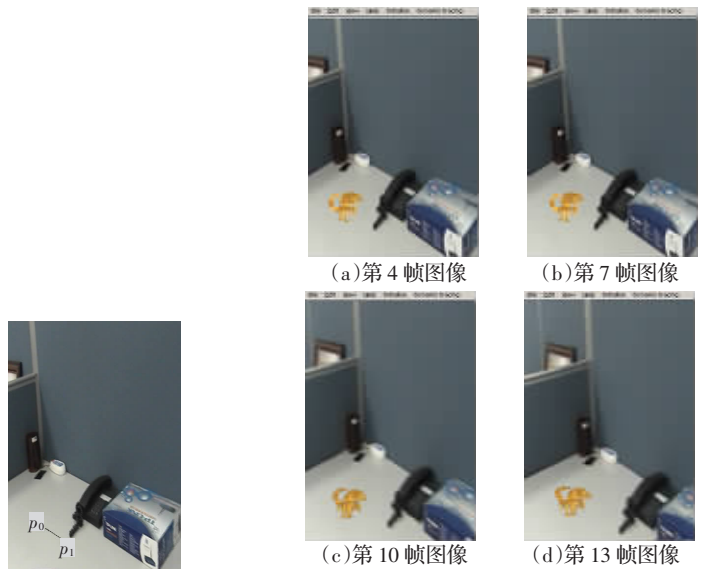
图4 在 I_0 图像上提取的X轴的像

图5 动态跟踪的结果

像)与2D图像之间的特征点匹配问题,它的研究必将推动增强现实技术研究的进一步发展。

参考文献:

- [1] Pressigout M, Marchand E. Hybrid tracking algorithms for planar and non-planar structures subject to illumination changes [C]// IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR, 2006.
- [2] Lourakis M, Halkidis S, Orphanoudakis S. Matching disparate views of planar surfaces using projective invariants [C]// British Machine Vision Conference, 2000, 6: 673-683.
- [3] Simon G, Fitzgibbon A, Zisserman A. Markerless tracking using planar structures in the scene [C]// Proceedings of the ISAR, 2000: 120-128.
- [4] Simon G, Berger M. Reconstructing while registering: A novel approach for markerless augmented reality [C]// Proc of the ISMAR, 2002: 285-294.
- [5] Simon G, Berger M. Real time registration known or recovered multi-planar structures: Application to AR [C]// British Machine Vision Conference (BMVC), 2002: 567-576.
- [6] Yuan M L, Ong S K. Registration based on projective reconstruction technique for augmented reality systems [J]. IEEE Visualization and Computer Graphics, 2005, 11(3): 254-264.
- [7] Lourakis M, Argyros A. Vision-based camera motion recovery for augmented reality [C]// International Conference on Computer Graphics, 2004: 569-576.
- [8] 马颂德, 张正友. 计算机视觉——计算理论与算法基础 [M]. 北京: 科学出版社, 1998: 52-56.
- [9] Chu Song C, Wen Ten H. Composition of 3D graphic objects and panoramas [C]// Proceedings of International Conference on Artificial Reality and Tele-existence, ICAT2000, 2000: 207-214.
- [10] Yuko U, Hideo S. Vision-based registration for augmented reality with integration of arbitrary multiple planes [C]// Roli F, Vitulano S ICIAP 2005, 2005: 155-162.
- [11] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry [J]. Artificial Intelligence Journal, 1995, 78: 87-119.