

增强现实混合跟踪技术综述

罗斌¹ 王涌天² 沈浩¹ 吴志杰¹ 刘越²

摘要 混合跟踪是增强现实 (Augmented reality, AR) 领域最近二十年发展迅速的重要关键技术, 是增强现实姿态跟踪系统同时实现高精度和强鲁棒性的有效途径. 本文全面完整地论述了增强现实混合跟踪技术, 详细阐述了混合跟踪器所涉及的混合跟踪、标定和时间同步等重要方法, 介绍了不同类型混合跟踪器的当前应用现状, 然后探讨了混合跟踪技术的发展趋势及其存在的难题, 最后展望了混合跟踪技术的应用前景.

关键词 增强现实, 混合跟踪, 传感器融合, 标定, 时间同步

引用格式 罗斌, 王涌天, 沈浩, 吴志杰, 刘越. 增强现实混合跟踪技术综述. 自动化学报, 2013, 39(8): 1185–1201

DOI 10.3724/SP.J.1004.2013.01185

Overview of Hybrid Tracking in Augmented Reality

LUO Bin¹ WANG Yong-Tian² SHEN Hao¹ WU Zhi-Jie¹ LIU Yue²

Abstract Hybrid tracking is one of the most important key technologies having received rapid development over the past twenty years in the field of augmented reality (AR). It is an effective approach to realize head-pose tracking with high accuracy and strong robustness for AR application systems. This paper tries to provide a complete and comprehensive survey on hybrid tracking technology. Three important issues involved in hybrid trackers are described in detail, including hybrid tracking method, calibration, and time synchronization. The state-of-the-art applications of different hybrid trackers are introduced. The pending problems and development tendency of hybrid tracking technology are discussed, and finally the application prospect is reviewed.

Key words Augmented reality (AR), hybrid tracking, sensor fusion, calibration, time synchronization

Citation Luo Bin, Wang Yong-Tian, Shen Hao, Wu Zhi-Jie, Liu Yue. Overview of hybrid tracking in augmented reality. *Acta Automatica Sinica*, 2013, 39(8): 1185–1201

增强现实 (Augmented reality, AR) 是在用户观察到的真实自然环境中添加计算机生成的文字、3D 模型等信息的技术. 增强现实的基本原理是使用头部姿态跟踪器提供的姿态信息精确地对准虚拟的物体和摄像机摄取的视频图像或者通过透射方式观看到的自然景象, 当用户在转动或平移头部时, 用户还能够视频透射式或光学透射式头盔显示器 (Head-mounted display, HMD) 屏幕上观看到精确对准的虚实融合图像. 由于用户观察的真实场景与虚拟物体共存一个视觉感知空间, 从而增强

用户对真实环境的理解, 扩展与真实环境的交互能力, 为用户感知空间提供一个增强的、自然的交互接口. 1992 年 Caudell 等^[1] 在波音飞机维修项目中首次提出了 AR 概念. 之后, Milgram 等^[2] 使用一个横跨纯真实与纯虚拟之间的连续闭集对虚拟现实 (Virtual reality, VR)、增强现实、增强虚拟 (Augmented virtuality) 以及混合现实 (Mixed reality, MR) 进行清晰甄别, 将增强现实描述为更接近于纯真实空间之间的一段连续空间, 也是混合现实开集中的一部分连续空间, 如图 1 所示. 增强现实具有三个重要特征^[3]: 1) 虚实融合, 即虚拟物体与真实世界共存于用户观察视场内; 2) 实时交互, 即用户可与真实世界甚至虚拟物体进行实时的自然交互; 3) 三维注册, 即虚拟物体与真实世界精确地对准. 增强现实涉及到头部姿态跟踪、三维注册、显示设备、人机交互、系统评估等技术, 关于这些技术的最近发展与进步可参考文献 [4–5].

头部姿态跟踪是增强现实最重要的核心技术之一. 增强现实需要高精度鲁棒的头部姿态信息精确对准虚拟物体和真实世界. 当前基于硬件的头部姿态跟踪器包括机械、超声波、电磁跟踪器、全球定位系统 (Global positioning system, GPS)、数字罗

收稿日期 2011-11-22 录用日期 2012-10-25
Manuscript received November 22, 2011; accepted October 25, 2012

国家自然科学基金 (61072096), 国家科技重大专项基金 (2012ZX03002004), 中国工程物理研究院科学技术发展基金 (2010B0203023, 2012B0403068), 国家部委项目资助

Supported by National Natural Science Foundation of China (61072096), the Major National Science and Technology Projects (2012ZX03002004), the Science Technology Development Foundation of Chinese Academy of Engineering and Physics (2010B0203023, 2012B0403068), and the Project of National Ministries and Commissions

1. 中国工程物理研究院计算机应用研究所 绵阳 621900 2. 北京理工大学光电学院 北京 100081

1. Institute of Computer Application, Chinese Academy of Engineering and Physics, Mianyang 621900 2. School of Optoelectronics, Beijing Institute of Technology, Beijing 100081

盘、视觉摄像机以及惯性跟踪器等^[6],但是正如在第 1.1 节所阐述,它们自身都具有各自的缺点,在 AR 应用中都有各自的局限性.通常,室内 AR 系统使用视觉摄像机跟踪头部姿态可以获得高精度三维注册,但是由于室外环境的不易控制和可能利用的信息比较少使得视觉摄像机无法有效用于室外 AR 系统.室外 AR 系统头部跟踪的唯一有效途径是使用混合跟踪技术,通过融合不同类型跟踪器的测量数据获得高精度鲁棒的头部姿态.自上世纪 90 年代中期以来,混合跟踪技术已经改变了 AR 头部姿态跟踪的形态,并被广泛应用于室内外 AR 系统.

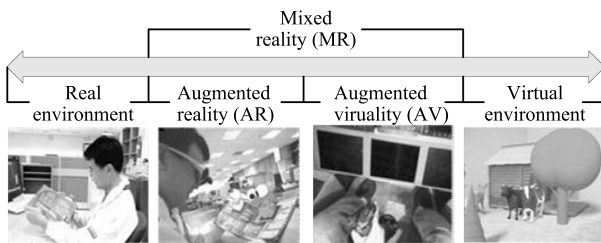


图 1 Milgram 的真实-虚拟连续闭集示意图
Fig. 1 Milgram's reality-virtuality continuum

AR 混合跟踪技术在国内还处于初始研究阶段.王涌天等^[7]在综述性文章中强调了混合跟踪技术在户外 AR 系统中应用的重要性;陈靖等^[8-9]使用视觉-惯性传感器用于运动结构恢复和室外 AR 系统;苏宏等^[10]和李薪宇等^[11]使用商业 VIS-1200 混合跟踪器和现有混合跟踪方法分别实现了一个基于 VIS-1200 跟踪器和视觉-惯性-无线传感混合跟踪的 AR 演示系统.罗斌等^[12-13]运用基于航迹融合 (Tracking fusion) 的多传感器融合算法实现了两个不同类型光学跟踪器的混合跟踪.杨浩等^[14]使用迭代无迹卡尔曼滤波器 (Unscented Kalman filter, UKF) 标定摄像机和惯性测量单元的相对姿态矩阵.

本文主要对国外最近二十年的 AR 混合跟踪技术进行综述性论述,包括混合跟踪技术所涉及的硬件跟踪器、混合跟踪方法、标定方法和时间同步法,介绍 AR 混合跟踪技术的应用现状,总结其发展趋势和存在难题,展望其应用前景.

1 混合跟踪问题描述

本节首先简要介绍 AR 领域中包括机械、电磁、超声波、GPS、视觉摄像机以及惯性传感器等在内的单一头部姿态跟踪器的原理及其具体特性,给出各类单一跟踪器的优缺点,总结混合跟踪器的好处,提出混合跟踪器所涉及的混合跟踪方法、标定和时间同步等三个重要关键问题.

1.1 单一姿态跟踪器

机械跟踪器使用一个机械机构测量用户头部的位置和方向,该机械机构一端与基座固联,另外一端与跟踪目标固联,在机械机构的关节位置安装光电编码器计算跟踪目标的位置.机械跟踪器的优点是精度高、更新快、无延迟,缺点是解码分辨率不高,存在运动空间限制.

电磁跟踪器一般由发射器、接收器、解码控制器构成,根据电磁场耦合原理测量位置和方向,利用接收器接收发射器发射的磁场信号,再传送给控制器进行计算处理.电磁跟踪器的优点是体积小、重量轻、更新率高、无延迟、价格低以及无遮挡,其缺点是工作空间小,容易受到金属、磁场等噪声影响.

超声波跟踪器使用多个超声波发射器作为发射源,根据发射源的超声波信号到达超声波接收器之间的时间差、相位差或声压差计算位置和方向.超声波跟踪器的优点是小巧、重量轻、无畸变,缺点是容易受到温度、湿度、压力、遮挡等影响,且更新率低.

GPS 主要由 24 个卫星和 12 个地面接收站构成,其基本原理是用户接收来自其中三个地球在轨卫星的信号,通过信号接收时间减去卫星信号发射之间的飞行时间 (Time of flight, TOF) 来确定 GPS 用户位置.当前,一般 GPS 系统的分辨率达到十米量级,微分 GPS 系统的分辨率可达到米量级.

惯性跟踪器也称为惯性测量单元 (Inertial measurement unit, IMU),包括惯性陀螺仪和微加速计,采用捷联式惯性导航原理计算方向和位置,其中惯性陀螺主要用于测量三轴旋转角速度或角加速度,并利用增量积分法获得跟踪器的方向信息;微加速计则主要用于测量位置加速度,再利用增量积分获得跟踪器位置信息.惯性跟踪器的优点是体积小、重量轻、免遮挡、工作范围不受限制、更新率高,但其内部漂浮随时间积累到一定程度也会导致跟踪精度下降.

视觉跟踪器使用摄像机跟踪目标,利用计算机视觉技术计算摄像机姿态.它包括“由内向外 (Inside-out, IO)”和“由外向内 (Outside-in, OI)”两类结构.IO 结构中的跟踪目标固定在环境,摄像机固定在头盔上且观察方向由内向外指向环境,而 OI 结构中的跟踪目标固定在头部,摄像机固定在环境中且观察方向由外向内指向头部.根据跟踪目标不同,也可分为标识 (Marker) 跟踪和免标识 (Markerless) 跟踪,前者使用 LED、标识板等作为摄像机跟踪目标,具有较高的计算效率和跟踪精度;后者使用边缘、平面、特征不变量等复杂自然特征作为摄像机跟踪目标,具有计算负载大、实时性差等特征;此外摄像机在遮挡条件下无法跟踪目标,在高

速运动条件下容易产生成像模糊而导致跟踪失败。

AR 应用也使用其他一些传感器定位用户, 比如使用数字罗盘、倾斜计等测量用户头部的二维方位和一维倾斜角, 使用激光测距仪测量用户的一维位置等。目前机械和电磁跟踪器在 AR 领域已较少使用, 超声波、GPS、视觉摄像机和惯性传感器是使用较多的姿态跟踪器, 并形成商业产品, 比如 InterSense 公司 InertialCube 系列惯性跟踪器, Xsense 公司 MT 系列惯性跟踪器, NDI 数字公司 Optotrak/Polaris 视觉跟踪器等。

1.2 问题描述

混合跟踪是将来自不同类型跟踪器的输出姿态数据进行合并或融合的头部姿态跟踪技术, 采用混合跟踪技术的姿态跟踪器被称为混合跟踪器。与单一姿态跟踪器相比, 混合跟踪器虽然增加了 AR 系统复杂性, 但是其最大好处是保持高精度姿态跟踪时还增强了跟踪鲁棒性, 主要体现在四个方面: 1) 姿态自由度 (Degree of freedom, DOF) 组合或融合互补。不同 DOF 的跟踪器经过组合可以获得完整的 6DOF 姿态输出; 不同 DOF 属性的跟踪器经过融合可以分享各自 DOF 优点, 在输出 DOF 特性上达到优势互补。2) 避免跟踪不稳定性及跟踪失败。视觉跟踪器的不稳定性源于低频工作、遮挡以及复杂环境; 惯性跟踪器不稳定性源于内部漂浮; GPS 不稳定性来自于建筑遮挡等。通过某一特定跟踪类型的互补融合可以降低或者消除单一跟踪器的不稳定性, 提高跟踪鲁棒性。3) 扩展跟踪范围。可移动便携式 AR 导航、军事等应用需要通过增加视觉跟踪器标识板或利用 GPS 和惯性跟踪器等扩大跟踪范围以实现大范围跟踪。4) 姿态预测以降低系统注册延迟。AR 系统存在动态误差使得虚拟物体注册与视频图像摄取之间存在时间延迟。混合跟踪器可对头部姿态进行预测, 降低系统动态误差, 提高虚实物体重叠精度。

图 2 显示了一个典型的由三个不同类型跟踪器构成的混合跟踪器示意图, 三个跟踪器分别由跟踪器 1 (摄像机)、跟踪器 2 (惯性) 和跟踪器 3 (GPS) 表示并刚性连接, C、I、G 分别是三个跟踪器的局部坐标系, W 是场景世界坐标系, U 为摄像机的图像坐标系, S 为摄像机的跟踪目标。跟踪器输出数据可以是自身直接的测量数据, 也可以是间接计算得到的姿态数据, 比如视觉摄像机的直接数据是环境中的跟踪目标 S 在图像坐标系 U 中的像素坐标 s , 而间接得到的姿态数据是摄像机坐标系 C 相对世界坐标系 W 的坐标变换。由于各个跟踪器的测量数据所对应的参考坐标系各自不同, 因此必须首先标定各个跟踪器之间的恒定坐标变换, 然后将跟踪器测量

数据变换到一个统一的参考坐标系中才能够进行跟踪器测量数据的混合或融合; 而且因跟踪器的数据测量时间点不同, 进行数据融合时还要考虑跟踪器之间的时间同步问题。混合跟踪器一般由至少两个跟踪器构成, 混合跟踪器中跟踪器的类型越多, 所涉及的混合跟踪方法、相对姿态标定方法及时间同步方法就越复杂。可见, 混合跟踪方法、标定方法及时间同步方法是混合跟踪技术实现良好鲁棒性的关键, 是混合跟踪 AR 应用中需要重点解决的关键问题。

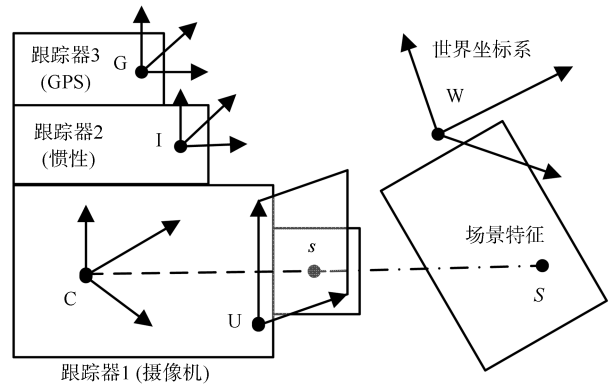


图 2 典型混合跟踪器及坐标系示意图

Fig. 2 Hybrid tracker and its coordinate systems

2 混合跟踪方法

将 Brooks 等^[15] 对定位跟踪的传感器融合分类类别应用于 AR 混合跟踪, 从多传感器数据融合的角度将 AR 混合跟踪方法分为互补式、竞争式和协作式传感器融合。如本节所阐述, 除 DOF 互补式组合和协作式融合外, 其他类型互补式融合和竞争式融合都需要使用线性卡尔曼滤波器 (Kalman filter, KF)^[16]、扩展卡尔曼滤波器 (Extended Kalman filter, EKF)^[17]、UKF^[18] 和粒子滤波器 (Particle filter)^[19] 等数学工具进行混合跟踪器的姿态数据融合。从传感器数据融合角度, 混合跟踪器被看作一个动态的线性或非线性的运动系统, 而混合跟踪器的姿态融合本质上是动态线性或者非线性系统的状态优化估计问题, 必须依赖上述数学工具进行求解, 可通过文献 [20-21] 进一步了解这些数学工具的理论知识和实际运用方法。本节重点阐述 AR 混合跟踪方法中所运用到的互补式、竞争式和协作式传感器融合, 揭示混合跟踪技术实现跟踪器优势互补的内部机理。

2.1 互补式融合

互补式传感器融合是组合或者融合相互独立的不同传感器的不完整测量数据, 以获得一个更完整的系统状态估计。对不同传感器不完整测量数据的不同融合处理方式导致四类不同的互补式融合, 包

括 DOF 组合或切换、单次单限制 (Single constrain at-a-time, SCAAT)、间接滤波 (Indirect filtering) 和航位推测法 (Dead reckoning)。

DOF 组合或者 DOF 切换是最简单的互补式融合。DOF 组合是将不同跟踪器的姿态 DOF 组合以获得完整的 6 DOF 姿态。比如在第一个集成 AR 与可移动计算的系统^[22]中,为实现市区环境 3D 可移动漫游,分别用 GPS 定位用户位置,用磁力计计算用户头部偏航,用双轴倾角计计算用户头部倾斜和滚转;在确定 GPS 位置条件下,使用室外轮廓视觉特征确定远距离注册的方向信息^[23];在用于城区战场训练的协作式 AR/VR 系统中,使用 GPS 和三轴数字罗盘实时提供位置和方向信息^[24]。DOF 组合的另一个变种是在 DOF 组合之前增加一个预测-校正过程^[25]或者滤波处理过程^[26]。DOF 切换是当某类型跟踪器跟踪头部姿态失败时,使用其他类型跟踪器继续进行跟踪。比如当视觉跟踪器的跟踪目标可见或部分遮挡时使用摄像机 6 DOF 姿态进行注册,而在跟踪目标全遮挡时使用 IMU 的 6 DOF 姿态进行注册^[27]。

SCAAT 是由 Welch 等^[28]提出来的互补式融合方法。SCAAT 解决了在传感器单个测量参数个数小于传感器状态估计参数个数条件下的退化解问题,其核心思想是单个传感器测量只提供用户状态部分信息并能够增量提高先验状态,从而在单个传感器测量被迭代用于合并一个新传感器的可测数据时获得低延迟和高精度的估计值。具体来说,当使用 EKF 实现基于 SCAAT 的头部姿态跟踪时,需要对所有跟踪器构建一个统一的过程模型,而对每个跟踪器单独构建一个测量模型,以协调处理跟踪器之间的测量不同步和不完整测量信息,如图 3 所示。只要有一个跟踪器测量可用,使用相应的测量变换矩阵和协方差矩阵进行状态校正以执行滤波器的预测-校正循环。典型例子是 Jiang^[29]在室外 AR 系统中使用 SCAAT 融合 GPS、陀螺仪、视觉跟踪器的测量数据。改进的 SCAAT 可以通过增加多个模型估计而自适应地在不同传感器模式之间进行切换^[30]来实现,也可通过使用粒子滤波器^[31]、基于李群代数描述的 EKF^[32-33]或误差状态滤波器^[34]等不同滤波器形式来实现。

间接滤波和航位推测法是源自于惯性导航领域并用于与其他测量系统集成的组合导航互补式融合方法,因而比较适用于惯性跟踪器与其他跟踪器的混合跟踪,其中惯性跟踪器被认为是主传感器,其他跟踪器被认为是辅助传感器。通常,间接滤波方法利用一个开环或闭环结构提高主传感器的跟踪精度。图 4 为开放式与封闭式间接滤波法结构示意图。在使用开环结构时,使用一个状态滤波器作用于主传

感器和辅助传感器之间的测量误差之差,状态滤波器的输出近似等于辅助传感器的负值,该负值与主传感器测量输出相作用近似消除了主传感器的误差,从而使主传感器获得误差优化的状态估计。与开环结构相比,闭环结构增加了一个反馈过程。通过此反馈过程,主传感器的误差估计被再次反馈回混合传感器以降低输出误差。在 AR 领域,间接滤波器的融合策略较早由 Fuchs^[35]提到并被 Welch^[36]深入探讨,之后被应用到 InterSense 公司的 IS-600 跟踪系统^[37-39]和其他应用^[40-41]。特别地,间接滤波法能够消减状态空间模型中惯性重力的作用以消除误差,获得平滑的位置信息^[40]。

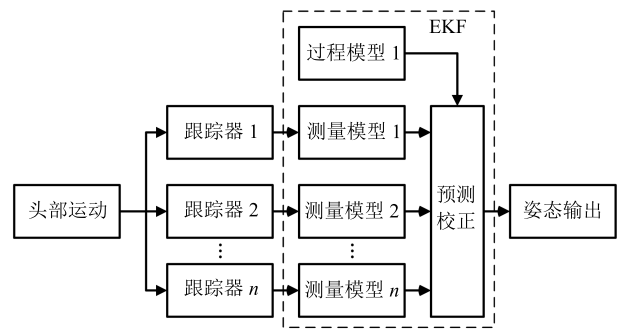


图 3 SCAAT 融合结构示意图
Fig. 3 SCAAT fusion configuration

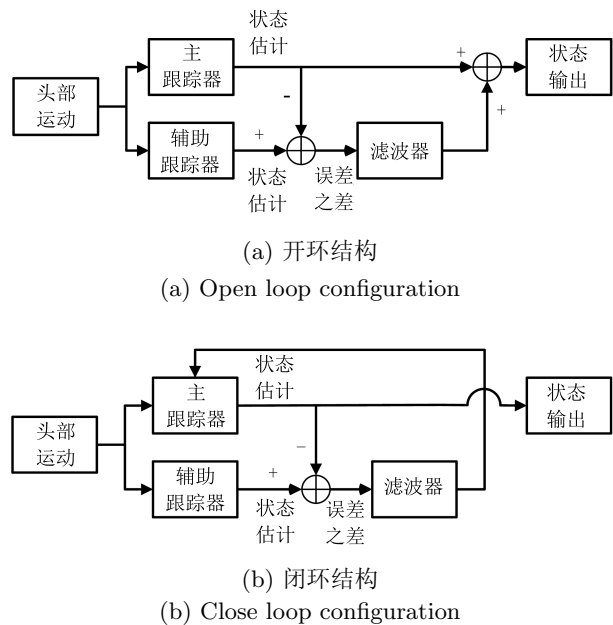


图 4 间接滤波融合结构示意图

Fig. 4 Indirect filtering fusion configuration

与间接滤波法不同,航向推测法所使用的滤波策略不直接作用于跟踪器测量误差,其关键之处是在滤波器融合结构中保留状态过程模型的输入信号,

且以惯性实际测量作为此输入信号, 利用滤波器内部的预测校正机制实现视觉测量对惯性测量的漂浮补偿, 如图 5 所示. 典型例子是在欧洲 MARTRIS 项目^[42-43]中, 使用加速计加速度和陀螺仪角速度之组合作为输入信号, 利用加速计/陀螺仪的直接测量参数构建过程状态的刚体动力学运动模型实现时滤波器的预测过程, 并利用视觉运动方程, 通过建立视觉摄像机图像测量与陀螺仪直接测量值之间的函数关系, 构建测量模型以实现惯性补偿校正. 此外, 这个输入信号也可以是加速计的线性加速度和角加速度^[44].

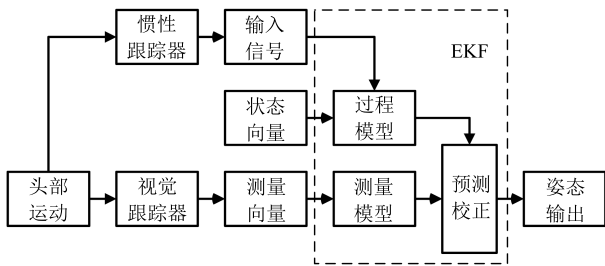


图 5 Dead reckoning fusion configuration
Fig. 5 Dead reckoning fusion configuration

上述四类互补式融合中, DOF 组合/切换最为简单, 其姿态数据耦合程度最为松散, 只需要先标定跟踪器之间的固定姿态, 再将跟踪器姿态变换到统一的参考坐标系中即可. SCAAT 融合不受某一跟踪器的跟踪失败影响, 而且能够耦合不同更新率的跟踪器. 间接滤波法侧重在通过对跟踪器测量之差进行优化估计来提高姿态精度, 航向推测法则重在补偿惯性漂浮, 但两者的缺点都是当辅助跟踪器(视觉或超声波)较长时间跟踪失败时主跟踪器(惯性)的姿态精度将降低, 甚至由于惯性漂浮的不断积累导致滤波器处于次优化甚至扩散状态.

2.2 竞争式融合

如果多个传感器能够在相同的空间坐标系中给出一个不同测量类型的状态估计, 则所使用的传感器融合称为竞争式传感器融合. 竞争式融合通常利用传感器融合工具去处理多个传感器测量之间的冲突, 需要考虑由多个传感器测量引入的残余, 融合不同误差分布的传感器以最小化测量误差. 集中式融合(Centralized fusion)和航迹融合是两类重要的竞争式融合方法.

集中式融合方法适用于多个单一跟踪器在混合跟踪器相同坐标轴向测量不同参数的情况, 它采用一个紧密耦合结构, 利用一个优化滤波器融合不同跟踪器的测量数据, 其结构如图 6 所示. 此耦合结构处理传感器测量冲突的方式, 是扩展状态空间模型

中的状态向量, 使其包括所有跟踪器误差过程的状态分量和头部运动的动力学状态分量, 并将所有跟踪器的测量信息累积为一个测量状态作为滤波器测量输入, 再构建一个合适的测量模型以精确描述组合的测量状态和全状态向量之间的关系. 比如在文献^[45]中视觉传感器的四元数方向和陀螺仪的角速度构成了状态滤波器的测量状态向量, 而在时间更新方程中通过使用角速度和加速度的估计值在时间轴上向前插值以实现状态预测.

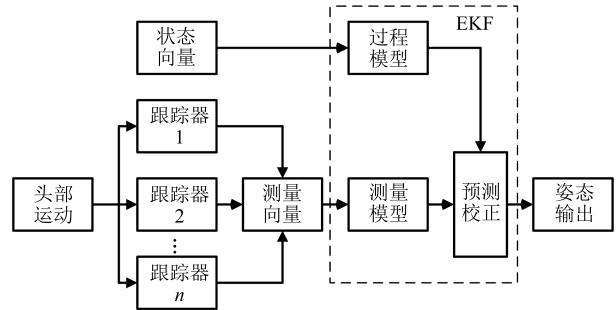


图 6 集中式滤波融合结构示意图
Fig. 6 Centralized filtering fusion configuration

航迹融合方法源自于多传感器融合环境中位置层级的航迹融合(Track-to-track fusion)问题. 多传感器融合环境中, 航迹(Track)定义为跟踪目标的运动轨迹. 可使用多个传感器对多个运动目标进行跟踪测量, 并且传感器对各个运动目标的跟踪测量值保持各自分开. 航迹融合就是当各个传感器的测量值被确定来自同一跟踪目标时如何将各个传感器的测量值融合在一起. Saha^[46]及其同事^[47]对航迹融合的理论算法进行了深入研究. 航迹融合方法的基本思想是为每个传感器分别构建一个滤波器, 然后在融合滤波器中使用合适的融合滤波方程融合每个滤波器的状态估计值, 从而获得误差最小化的状态输出, 其结构如图 7 所示. 航迹融合方法处理传感器测量冲突的关键在于融合滤波器. 由于头部运动噪声会通过每个跟踪器的姿态计算和姿态 EKF 滤波过程, 以姿态误差协方差矩阵的形式传播给融合滤波器, 因此融合滤波器如何耦合姿态误差协方差矩阵对混合跟踪器的最终姿态输出影响就显得非常重要. 当头部噪声引起的跟踪器姿态误差被假设为非相关时, 一般使用次优化的简化融合方程^[48], 甚至不使用平行滤波器而直接使用简化融合滤波器^[49]. 然而, 跟踪器的姿态误差被认为是互相关时, 则需要使用其他一些复杂的融合滤波方程以表述跟踪器测量误差的相关性对最终输出状态的作用.

集中式融合与航迹融合在测量冲突处理方式上的不同决定了两者的显著区别. 集中式融合需要所

有跟踪器都有测量才执行数据融合,但当某个跟踪器的测量不可获得时,整个滤波器就会崩溃,因而一般使用集中式融合来执行 AR 系统的头部姿态预测,降低 AR 系统的动态注册误差.航迹融合与 SCAAT 融合方式完全不同,但是具有与 SCAAT 相似的优点,它不受某一跟踪器的跟踪失败影响,而且能够融合不同 DOF 属性的跟踪器,使得混合跟踪器在姿态输出属性上达到优势互补.

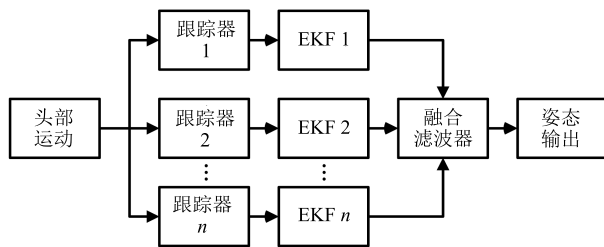


图 7 航迹融合结构示意图

Fig. 7 Tracking fusion configuration

2.3 协作式融合

协作式传感器融合建立在多传感器之间的紧密耦合基础上,此方法需要更多关于传感器的物理属性知识以建立不同参考坐标系中传感器参数之间的内在联系,分为简单协作式融合和复杂协作式融合.

简单协作式融合不直接使用传感器的原始测量数据,而是首先从直接测量数据计算出传感器的姿态,然后使用一个传感器的姿态数据去补偿另一个传感器的姿态数据,因而此方法相对简易而且计算负载小,在具体实践时能够实现惯性漂浮校正或视觉特征预测.比如使用超声波系统的输出姿态补偿惯性跟踪器的漂浮^[50];在视觉特征不可见时使用视觉跟踪器的方向补偿惯性漂浮,且在视觉特征可见时使用惯性方向稳定视觉跟踪^[51];使用辅助系统与视觉跟踪之间的测量差来校正辅助系统的姿态估计^[52];在立体摄像机视觉跟踪时,使用惯性方向和立体摄像机旋转中心计算得到的摄像机平移量预测视觉跟踪的标识板位置^[53].

相比之下,复杂协作式融合直接使用传感器原始测量数据而不是姿态输出,其基本思想是使用某一方式将传感器测量数据耦合进 AR 视觉注册的闭环框架^[54]中.第一种方式是减少视觉跟踪注册闭环框架中被估计的姿态矩阵参数个数.比如文献^[55]使用直线限制法,通过限制用户视点在俯视(Bird's eye)摄像机的视线上以减少 HMD 固定摄像机的视点估计参数个数,此直线限制法也适用于惯性与视觉混合^[56];或者通过将惯性跟踪器的 2D 偏转角嵌入到视觉跟踪的单应矩阵以减少单应矩阵中的估计参数个数^[57].第二种方式是在跟踪注册闭环框架中

实现如下的预测和校正过程:使用电磁^[58-59]、红外摄像机^[60]、惯性跟踪器^[61-62]等的直接测量数据预测图像特征点的运动范围,一旦探测到图像特征就将之用于姿态估计,而估计得到的姿态再一次被用于下一次特征预测行为.在实践中,需要某些诀窍以执行这种预测和校正行为.比如文献^[61-62]将惯性方向测量精妙地嵌入到视觉连续帧之间的单应矩阵,预测关键点匹配中的新图像特征点;文献^[63]使用前一帧的路标图像坐标计算陀螺仪的仰角漂浮更新值,然后嵌入此漂浮更新值到惯性陀螺仪视点的方向矩阵中以实现惯性漂浮补偿,而惯性方向测量用于预测当前帧的路标图像位置;文献^[64]则通过建立视觉运动方程,在摄像机连续运动时用惯性跟踪器旋转测量值计算当前视频帧的图像特征平移量,然后用惯性传感器和真实摄像机运动之间的特征速度之差来估计惯性漂浮值;文献^[65-66]使用陀螺仪旋转数据初始化视觉跟踪器在视频图像被处理前的姿态估计,修正边缘提取过程以避免运动模糊,并通过比较惯性传感器旋转预测值和视觉摄像机旋转测量值来校正系统偏差估计.

与其他融合方式相比,复杂协作式融合是耦合度最紧密的融合,适合于视觉跟踪器与其他类型跟踪器之间的姿态融合.与间接滤波和航向推测法不同,在复杂协作式融合中视觉跟踪器被认为是主传感器,其他类型跟踪器等被认为辅助传感器.辅助传感器通过提供给主传感器一定的 DOF 参数个数或者缩小主传感器成像特征的搜索范围,可提高主传感器执行视觉注册的计算效率,在主传感器图像模糊或跟踪失败时,还能使用辅助传感器进行跟踪注册,从而使混合传感器兼具计算高效率 and 强鲁棒特性.

3 标定方法

AR 混合跟踪系统中的标定对象包括各个跟踪器内部参数和各跟踪器之间恒定的坐标变换.混合跟踪器标定涉及到误差建模、标定算法、实验平台、标定步骤以及标定技巧.

3.1 跟踪器内部参数标定

跟踪器参数标定是指测量混合跟踪器中各个跟踪器的特定参数,因各个跟踪器自身属性不同而需采用不同的标定法.目前商用跟踪器一般都由制造商进行标定,比如 Xsense 公司和 Intersense 公司的系列惯性跟踪器等;而用户定制的跟踪器则需要由用户自己标定.对于 AR 中的视觉摄像机,其焦距、成像中心等内参数以及畸变参数确定了摄像机成像模型,通常采用 Zhang^[67]和 Tsai^[68]的标定方法进行标定.惯性跟踪器内部存在二次积分偏差以及电

路热效应等噪声,会导致惯性跟踪器在其内部参考坐标系上的测量输出产生偏差,其标定需要精心设计标定实验平台和标定步骤。比如文献 [69] 将铝制立方盒与惯性跟踪器刚性固定在标定台上,通过手动旋转立方体进行系列旋转以采集标定数据,再使用互补式 KF 估计惯性偏差、失准和尺度因子等参数;文献 [70] 使用 Optotrak 光学跟踪系统测量手动旋转的 IMU 角速度,并对加速计和陀螺仪分别构建数学模型,使用非线性最小二乘方法估算加速计和陀螺仪内部参数;此外文献 [71] 设计了一个转轴安装有解码器的钟摆,构建数学模型,通过钟摆运动方程、解码器读数和陀螺仪垂直方向的角速度来求解数学模型方程获得惯性偏差等参数;基于文献 [71] 提出的标定思路,文献 [72] 通过建立一次和二次多项式标定模型,使用一个与惯性跟踪器刚性固定的旋转支架,调节惯性跟踪器至不同方向,通过测量陀螺仪三轴角速度而不是垂直方向角速度来标定惯性偏差等参数。

3.2 混合跟踪器相对姿态标定

混合跟踪器的相对姿态标定是指测量混合跟踪器中各个跟踪器之间的平移和方向参数,相应的标定方法包括计算机视觉方法、视觉运动法、绝对朝向 (Absolute orientation) 法、手眼 (Hand-eye) 法、静态优化法和动态滤波法。

计算机视觉法是直接采用基于计算机视觉的姿态估计法在特定标定位置按步骤标定混合跟踪器的相对姿态。由于标定全过程仅仅使用视觉摄像机图像数据,此方法比较适合于标定不同类型光学跟踪器之间的姿态变换^[48, 73]。计算机视觉法利用 LED、标识板或棋盘板等作为摄像机跟踪目标,标定步骤相对简单,但是相应的标定精度严重依赖于所选择的标定位置,因而在远离标定位置之跟踪区域所获得的头部姿态跟踪误差相对较大。这主要归因于使用计算机视觉姿态估计算法时,沿摄像机光轴方向的平移估计与其他两个轴向相比,具有比较大的估计误差。

视觉运动法利用摄像机在运动时跟踪目标在摄像机像平面上的运动信息。由于涉及到摄像机运动信息,此方法比较适用于视觉-惯性混合跟踪器之间的相对姿态标定^[64]。视觉运动法利用视觉运动原理,计算特定摄像机成像模型条件下视觉摄像机视频帧之间的图像特征 2D 速度,从而建立惯性陀螺仪角速度与图像特征 2D 运动速度之间的关系。通过多次旋转视觉-惯性跟踪器,多次采集相应陀螺仪角速度和摄像机图像像素坐标信息,求解视觉运动方程以获得混合跟踪器的相对姿态标定数据。

绝对朝向法利用 Horn^[74] 提出的绝对朝向估计算法,通过使用非共线的三个点计算两个不同坐标系之间的绝对方向闭解,其关键是根据特定的标定对象和条件比较技巧性地提炼测量数据以建立相应的绝对朝向方程。比如文献 [71] 通过让摄像机和惯性传感器在不同姿态时观测垂直方向,构建关于视觉-惯性相对方向的绝对朝向方程。该方程由两组测量数据求解,一组来自于惯性传感器在静态时刻测量重力方向的角速度值,另一组则是让视觉摄像机观察一组平行垂直边缘特征,在摄像机参考坐标系垂直方向测量得到的垂直消隐点信息。

手眼法源自于机器人领域,起初在于利用刚性固定在机器臂终端上的摄像机,采用 Park 等提出的算法^[75] 测量机器臂终端与机器人操作部件之间的相对变换。Baillot 等^[76] 首次将手眼法用于定位 AR 跟踪系统,之后 Klein 等^[32] 使用它计算边缘 IO 跟踪器和 LED OI 跟踪器之间的坐标变换, Aron 等^[61] 将之用于标定视觉-惯性相对方向。文献 [72] 通过分别计算视觉和惯性参考坐标系中的旋转测量之差建立手眼标定方程,其独特之处在于在任意方向上的测量数据都能够耦合进手眼方程中,而没有仅仅使用垂直方向测量之限制。基于文献 [71] 使用绝对朝向法所实现的相对方向标定,文献 [77] 进一步利用手眼法,通过使用摄像机观察固定的棋盘网格标定板以恢复摄像机外参数实现相对平移标定,而且其标定实验平台使用一个简单转盘,通过调节转盘转动视觉-惯性跟踪器,使惯性陀螺仪在几个姿态上达到零值以构建手眼标定方程。不过,文献 [71, 77] 所共同阐述的双步骤标定法需要频繁手工操作,使用手眼法标定平移获得的位置标定误差对方向标定误差非常敏感,这主要由于平移标定手眼方程包括了方向标定信息,使得方向标定误差被传播进入平移标定误差所致。

静态优化法是获得传感器静态测量值时,设计一个合理的代价函数,并最小化代价函数以获得一个误差优化的标定参数。文献 [56] 使用图像特征相对于被估计标定参数的雅克比矩阵构建一个非线性优化框架,在多个摄像机视点位置采集图像特征和陀螺仪方向以执行方向标定;文献 [78] 针对室外标定环境设计了一个基于惯性方向直接测量值和摄像机间接方向估计值的非线性误差函数,其所使用的姿态估计算法与文献 [71-72] 相似,但是由于在室外路道环境下很难测量特征位置,必须使用方向测量值代替 3D 位置测量值以修正误差函数。

动态滤波法将混合跟踪器标定认定为混合跟踪器随机运动条件下的滤波问题。动态滤波法主要用于视觉-惯性混合跟踪器标定,使用传感器融合工具实现,比如 EKF^[79-80]、迭代 EKF^[81] 和 UKF^[82]

等. 此方法的共同点是在测量惯性跟踪器姿态、速度和偏差之时使用棋盘标识板, 随机平移或旋转混合传感器动态采集标定数据, 自动标定混合跟踪器的 6 DOF 相对坐标变换. 采用此方法, IMU-视觉相对姿态标定被认为是一个灰盒问题, 通过 EKF 的革新参数设计一个预测器以构建误差代价函数^[79-80]. 此方法也可将视觉目标的结构信息包含在 UKF 的状态向量中, 通过恢复目标结构信息来实现全方向摄像机与惯性跟踪器之间的自动自标定^[82].

4 跟踪器同步

跟踪器同步是指不同跟踪器之间的时间匹配. 基本上, 跟踪器的采样周期是异步的, 每个跟踪器测量发生在不同时刻. 在实际应用中, 即使微小的时间匹配误差都能够削弱混合跟踪的潜在好处, 比如降低跟踪精度等. 一般可使用硬件和软件方式实现跟踪器时间同步. 最简单的硬件同步是为每个跟踪器测量设计一个增量计数器, 并使用外部信号强迫所设计的硬件在同一时刻触发所有跟踪器测量数据; 此外也可以直接使用一个中心定时器同时触发跟踪器测量^[43]. 如果每个跟踪器都有一个内置定时器提供采样时间戳, 则可以通过在线或者离线方式估计各个跟踪器测量发生时的相对时间差来实现时间同步. 在线同步可使用不同原理, 比如文献 [61-62] 利用摄像机图像和惯性跟踪器的突然运动, 使用“辨别视觉跟踪失败-计算传感器单应-重新初始化匹配”步骤处理跟踪器之间的时间偏差; 文献 [83] 利用跟踪器采样时间之间的互相关属性, 其基本思想是将其中一个传感器某时刻的测量数据变换为相隔一个离散时间偏移量的时刻点测量数据, 然后由变换后的传感器测量数据序列和另外一个传感器测量数据序列构建归一化的互相关函数作为相似度, 最大化相似度实现时间同步. 离线同步更多依赖于钟摆^[84]、可定制的标定架^[85]等平台; 文献 [85] 通过最小化估计的时间戳与采样到达时间之间的最小绝对差来实现离线同步. 当前, 考虑到实用性因素, 硬件方法是比较通用和方便的跟踪器同步手段.

5 混合跟踪器的应用现状

除了前述的传感器数据融合、相对姿态标定以及时间同步, 混合跟踪技术在实际应用中还必须根据特定的应用环境和需求对混合跟踪器组成进行合理选择与组合. 由于头部姿态跟踪器组合后的混合类别较多, 而且不同组合类别的混合跟踪器能够满足多样化应用需求. 依据当前调研文献, 本节将按照混合跟踪器类别描述混合跟踪技术的应用现状, 分别介绍光学混合跟踪器、视觉-惯性混合跟踪器以及其他混合跟踪器的特点以及相关应用.

5.1 光学混合跟踪器

光学混合跟踪器亦称为 Inside-outside-in 跟踪器, 由 IO 光学跟踪器与 OI 光学跟踪器构成. IO 跟踪器与 OI 跟踪器具有各自优缺点. 由于 IO 跟踪器摄像机的微小旋转可造成跟踪目标在图像平面上较大的偏移. 因而 IO 跟踪器有较高的方向分辨率; 由于使用多摄像机跟踪和三角化跟踪算法, OI 跟踪器具有比较高的位置分辨率和精度, 但因其方向估计严重依赖于位置估计, 导致方向估计误差对位置估计误差非常敏感. 光学混合跟踪器兼具 IO 跟踪器与 OI 跟踪器的优点, 具有比较好的位置和方向精度, 但因其大多利用 LED、标识板为跟踪目标且易受光照影响等原因, 主要被应用于室内环境, 比如医疗^[49,60]、桌面 AR^[32]、离线多人运动跟踪^[86]和座舱模拟器^[48]等. 光学混合跟踪器的共有特点是头部都固定有 IO 摄像机, 但按照特定的应用在视觉跟踪目标与 OI 摄像机的选择、摄像机与跟踪目标的位置摆放、融合方法等方面会有所不同. 比如文献 [49] 使用 Optotrak 跟踪器作为 OI 跟踪器, 并使用姿态协方差矩阵验证其结构设计合理性; 文献 [60] 将红外摄像机连接到 IGS 系统, 使用模板作为 IO 摄像机跟踪目标大大改善视频透射 AR 系统的图形重叠精度; 文献 [55] 使用俯视摄像机作为 OI 传感器, 运用直线限制法减小 IO 跟踪器姿态估计参数个数; 文献 [32] 使用边缘特征作为 IO 跟踪目标并使用基于李群的 EKF 框架进行姿态融合; 根据座舱环境及跟踪特点, 文献 [48] 精心设计单目 IO 和双目 OI 跟踪器的放置位置, 用以实现飞行员与座舱内所有设备之间的自然交互.

5.2 视觉-惯性混合跟踪器

早期的视觉-惯性混合器主要使用惯性跟踪器预测跟踪器头部姿态以减小系统延迟^[44-45], 之后视觉-惯性混合跟踪器因其能够利用视觉测量补偿惯性漂浮, 利用惯性测量预测视觉测量, 兼备视觉跟踪精度高和惯性跟踪擅于处理复杂场景和快速运动等优点而得到广泛应用. 不过, 视觉跟踪目标的不同选择导致相应视觉-惯性混合跟踪器的不同应用.

LED 标识点以及标识板等特征具有提取快速和计算负载小等优点, 这类视觉-惯性跟踪器主要用于室内环境. 对于室内局部区域跟踪应用, 基本的要求是使用合理的混合结构或融合方法实现高精度强鲁棒跟踪注册. 比如, 使用标识板立体图像和惯性方向构成立体视觉系统, 在立体摄像机平移时跟踪 6 DOF 姿态^[53]; 组合俯视摄像机与陀螺仪实现鲁棒注册^[56]; 用间接 KF 融合四个标识板的摄像机与三个陀螺仪^[87]; 惯性跟踪器直接用于解决标识板视觉跟踪器的遮挡^[27]; 甚至集成基于 LED 的光学混

合跟踪器与惯性跟踪器, 将算法嵌入 DSP 硬件以构建增强座舱系统^[88]等. 在室内导航应用中, 必须通过在导航路线不同位置安装和跟踪很多标识板来实现长距离跟踪, 比如 Studerstube 项目 SignPost 便携式系统在沿线布置 ARToolKit 标识板^[51]; 在室内工作区使用立体摄像机跟踪固定在墙上的多个红外标识点^[25]等. 对于大范围室内跟踪, 则需配置大量紧密排列的标识板, 比如 VIS-1200 系统^[89-90]. 对于其他应用, 可集成惯性与不同的视觉跟踪器获得不同用途, 比如用 LED Optotrak 跟踪系统辅助跟踪工业工具^[91], 用四个摄像机辅助跟踪手持移动设备^[40], 甚至使用以激光二级管为跟踪目标的六个摄像机辅助跟踪全封闭 VR 显示系统的 6 DOF 绝对姿态^[92].

拐角点特征的主要优点是在较大距离范围内的探测鲁棒性和摄像机较大方向姿态角条件下的探测可靠性, 因而使用拐角点的视觉-惯性跟踪器, 主要用于在室内外便携式或可移动 AR 系统中提供快速、精确和稳定可靠的输出属性^[93]. 在改进的 TOWNEAR 便携式室外 AR 系统中, 使用拐角点标识板, 结合基于视觉的漂浮补偿算法解决了高精度陀螺仪的仰角漂浮问题^[63]. 在室内 3D 跟踪条件下, 拐角点与惯性跟踪器组合, 借助于粒子滤波器融合框架可获得增强的跟踪鲁棒性和可靠性^[94].

线或边缘特征比标识板较复杂. 边缘特征的可伸长结构使得其在部分遮挡时还可较长时间保持在摄像机观察视场内, 但是边缘特征提取需要较大的计算负载. 在特征匹配时需要一个鲁棒的特征对应机制. Rehbinder 等^[95-96]提出了一个理论框架, 从位置估计中分解出方向估计, 使用线特征和惯性信息估计刚体姿态. 这个框架之后被 Wu 等^[97]扩展到线性相关的线特征条件, 但是两者都没有进行实验验证. 在具体实践中, 线特征和惯性信息的组合可以跟踪头部快速旋转运动^[65]; 边缘跟踪器也可使用纹理 3D 模型代替边缘模型, 用基于表面的边缘指示器获得基于尺度的自动细节裁减和匹配改进^[33]; 在室外建筑区域, 也可利用建筑直线特征与惯性组合处理头部快速旋转, 在 CAD 模型中离线定义直线, 使用启发式直线选择法避免错误直线对应^[98].

平面、纹理、3D 模型甚至关键点特征也能够集成到视觉-惯性跟踪器, 主要利用惯性跟踪器辅助处理室内外纷繁复杂的场景, 增强跟踪的鲁棒性和实时性, 用于室内外 AR 应用. 文献 [64, 99] 通过使用陀螺仪提供视觉跟踪器相邻视频帧之间的相对方位来提高自然特征跟踪的计算效率, 但是仅仅实现 3 DOF 方向混合跟踪; 之后文献 [100] 使用双通道运动滤波器, 充分利用视觉跟踪器低频稳定和惯性跟踪器高频跟踪的特点, 进一步扩展到 6 DOF 自然特

征跟踪. 在欧洲 MATRIS 项目^[42-43, 101-106]中, 使用纹理、2D/3D 模型和光流等构建实时视觉-惯性跟踪系统, 用于增强电视转播和保护湮灭的文化遗迹. 在其他应用方面, 摄像机平面信息和加速计 2D 倾斜旋转能够在笔型 6 DOF 人机接口中增强姿态估计鲁棒性^[57], 连续视频帧之间的关键点与 XSense MT9 惯性跟踪器能够处理电子商务应用中的头部突然运动^[61-62], 甚至可以组合 IMU、两个前向摄像机和两个后向摄像机用于室内外大范围内的军事训练和游戏应用^[34].

5.3 其他混合跟踪器

除了光学混合跟踪器和视觉-惯性跟踪器, 其他混合跟踪器还包括惯性-超声波、惯性-数字罗盘、视觉-电磁、惯性-GPS、数字罗盘-GPS、视觉-GPS 以及由至少三类不同跟踪器构成的复合混合跟踪器等.

电磁跟踪器能实现用惯性跟踪器提高视觉跟踪注册精度和鲁棒性的相同功能. 与惯性跟踪器一样, 电磁跟踪器能够预测彩色解码圆形标识板的图像特征^[58], 也能够预测边缘在立体摄像机图像平面上的特征^[59]. 不过由于受金属环境影响导致误差较大, 电磁跟踪器在后续并不被广泛应用于对姿态精度要求比较高的 AR 应用. 此外, 可组合测距仪和视觉摄像机, 构成移动传感器平台应用于室内大范围跟踪, 利用测距系统提供相对于地面的相对位置和方向以改善标识板跟踪和自然特征跟踪条件下摄像机的跟踪属性^[107].

超声波、GPS、罗盘等也能够实现用视觉跟踪器补偿惯性累积漂浮的相同功能. 在 InterSense 公司的系列产品中, 3 DOF IS-300 系统通过测量重力倾斜信息补偿陀螺仪在倾斜和滚转方向上的漂浮, 使用地磁罗盘补偿陀螺仪在偏航方向上的漂浮, 6 DOF IS-600 跟踪系统使用超声波飞行感知系统校正惯性漂浮^[38], 而 IS-900 跟踪系统则使用星群分布的超声波无线电应答信标校正惯性漂浮^[37]. 这类漂浮补偿策略可用于不同应用环境, 比如在便携式 AR 系统中使用 GPS 补偿 IS-300 惯性/磁力计^[108]; 在自然环境中使用罗盘和倾斜计补偿三个陀螺仪, 以稳定与用户运动相关的虚拟物体^[109-110]; 在可移动陆地运输工具的惯性导航系统中使用电磁罗盘补偿由陀螺仪、两个加速计、两个倾斜计等构成的惯性跟踪器^[111].

目前, 比较受关注的混合跟踪器是复合型跟踪器. 由于跟踪器组成个数不低于三个, 因而复合型跟踪器需要更复杂的融合策略融合跟踪器测量, 但是它能够集成各个跟踪器优点, 被广泛用于复杂的室外应用. Hagure 等^[112]较早利用惯性、激光和雷达

定位器以及立体视觉系统测量自动化农业运输工具的位置. 之后, 视觉-惯性-GPS 复合型跟踪器因其具有处理快速运动和复杂场景以及扩展室外跟踪范围等优点, 在路径导航^[113]、便携式 GroundCam 系统^[41]、可移动定位^[52, 114]和手持式室外 AR^[26]等领域得到广泛应用. 比如, 视觉-惯性-GPS 组合能够为便携式 Ground-Cam 系统提供高频和高分辨率相对位置信息^[41]; 由 GPS 和惯性跟踪器构成的辅助定位子系统被用于初始化视觉定位系统, 并在视觉定位系统跟踪失败时用于估计用户姿态^[52]; GPS 和视觉跟踪器通过多传感器融合框架用于改善 IMU 内的瞬时振荡效应^[26]. 最近, 视觉-罗盘-GPS 混合跟踪器也被用于城区定位服务, 利用建筑拐角点的视觉信息提高 GPS 位置测量和数字罗盘方向信息的精度^[115].

6 发展趋势

本节主要在前述基础上总结 AR 混合跟踪技术中传感器融合方法和标定方法的发展趋势.

6.1 融合方法

多传感器融合在 AR 混合跟踪方法中有重要作用, 各种不同的比较成熟的融合策略被运用于 AR 混合跟踪. 除了 DOF 组合/切换和协作式融合, AR 混合跟踪的其他融合方法都依赖于传感器融合工具, 传感器融合工具的内部机制通过建立合适的过程模型和测量模型, 有效地将不同类型跟踪器的测量数据耦合到统一的状态系统中, 而且不同过程模型、不同测量模型甚至不同滤波组合结构形成了不同的传感器融合方法. 可见, AR 混合跟踪方法当前的总体发展趋势是使用传感器融合工具对不同类型跟踪器测量数据进行融合, 而下一步需要发展的方向就是新的传感器工具或融合策略在混合跟踪器中的运用, 主要有以下几点:

1) 噪声互相关假设条件下的航迹融合. 文献^[48]仅仅研究了头部噪声相互独立假设条件下的航迹融合算法融合两个光学跟踪器姿态数据, 而头部噪声互相关假设条件下的航迹融合算法及其在其他类型混合跟踪器中的应用至今还未有文献报道. 在航迹融合算法框架中, 单一跟踪器的跟踪对象是头部运动, 而头部的随机运动噪声会同时在两个跟踪器的姿态计算过程中进行传播, 会导致头部运动对两个跟踪器输出的姿态误差都有贡献且相互关联, 相互关联的跟踪器姿态误差对混合跟踪器的姿态输出有很大影响.

2) 多速率滤波器融合 (Multi-rate filter fusion). 多速率滤波^[116]主要用于信号处理领域, 它可以将低频采集信号和高速采集信号融合获得高带宽的

信号输出, 因此可以将这一信号融合策略应用到低采样率摄像机和高采样率惯性传感器之间的数据融合, 使得视觉-惯性混合跟踪器在最快采样率条件下实现传统的预测和校正. 比如文献^[96]利用多速率滤波结构估计摄像机相对于惯性跟踪器的方向矩阵; 文献^[117]使用多速率 EKF 和多速率 UKF 融合视觉-惯性测量数据实现快速的自我运动 (Egomotion) 估计, 但是至今还未有文献报道多速率滤波在 AR 混合跟踪系统中的应用.

3) 粒子滤波器融合. 当前 AR 混合跟踪方法大多利用 KF、EKF 和 UKF 等传感器融合工具进行姿态融合, 通常假设滤波器的过程模型和测量模型是线性的或者是一级泰勒线性近似的, 即将头部运动按照线性或线性近似的运动系统来处理, 但是当头部无规律非线性运动时将导致滤波器产生次优化、扩散甚至崩溃的问题. 粒子滤波器是通过寻找一组在状态空间传播的随机样本对概率密度函数进行近似. 以样本均值代替积分运算, 可获得状态最小方差分布的过程, 比较适用于非线性和非高斯分布的运动系统, 能够处理用户头部无规律的非线性运动, 但是大量的工作也还需要进一步扩展, 包括不同类型粒子滤波器在混合跟踪器中的运用, 在跟踪精度、鲁棒性和计算复杂性等方面对粒子滤波器与其他滤波器进行实验比较等.

6.2 标定方法

由于混合跟踪器中单个跟踪器的标定主要由制造商或采用众所周知的方法进行标定, 因此混合跟踪器的相对姿态标定是受关注重点. 混合跟踪器的相对姿态的标定方法与具体的标定对象、实验平台等条件紧密相关, 其发展趋势主要有以下几点:

1) 6 DOF 标定. 目前在光学混合跟踪器的 AR 系统中, 使用计算机视觉法、手眼法标定 6 DOF 相对姿态. 但是在视觉-惯性混合跟踪器的 AR 应用中, 通常假设惯性跟踪器的中心与视觉摄像机的中心是重合的, 在相对姿态标定过程中忽略了相对平移参数. 这种假设适用于 AR 系统远距离的场景注册, 但是对于近距离场景注册, 这个假设会给系统的跟踪注册带来一定误差, 从而可能限制混合跟踪器的应用潜力, 因此在实际应用中标定 6 DOF 姿态是很重要的.

2) 高精度标定. 由于混合跟踪器的相对姿态矩阵将用于混合跟踪器姿态融合, 其标定误差对混合跟踪器输出姿态精度有影响. 从标定算法角度考虑, 相对姿态标定本质上是求解线性或非线性方程组的优化估计问题, 需要多次采集单个传感器的测量数据, 测量数据量越多, 标定误差越小, 但是多次采集数据会导致人操作繁累. 与其他标定方法相比, 动态

滤波法由于使用传感器融合工具, 6 DOF 相对姿态参数以及跟踪器内部偏差等都能够被设计为滤波器的状态向量, 并通过滤波器内部的协方差矩阵优化估计机制获得误差最小化的标定参数估计值。

3) 自动操作标定. 计算机视觉法、视觉运动法、手眼法、绝对朝向法以及静态优化法在实际应用中, 都需要构建复杂的标定实验平台, 设计复杂标定操作步骤和手工采集测量数据. 与之相反, 动态滤波法在标定视觉-惯性混合跟踪器时, 仅仅使用摄像机标定板, 不需要其他复杂标定平台, 不需要繁琐的手工操作, 不需要在每次标定时重建标定平台. 进行动态滤波法标定实验时, 使混合传感器随机运动, 让其自动采集任意姿态状态下的测量数据, 并将测量数据送入滤波器优化估计标定参数。

7 存在难题

目前, 混合跟踪技术在混合跟踪方法、标定方法和时间同步法等方面已取得快速进展, 但依然存在一些难题需要进一步解决, 以满足更实用和更普遍的增强现实应用需要. 这些难题体现在以下几个方面:

1) 在混合跟踪方法方面, 存在难点是普适跟踪 (Ubiquitous tracking) 环境下任意传感器的信息融合. 目前 AR 混合跟踪方法研究基本上局限在某一类融合方式及其传感器融合工具的不同方式运用, 但是由于混合跟踪器自身类型多和融合算法多样性等原因, 迄今为止还缺乏一个能够包容各类融合方法的通用混合跟踪框架及开发平台. 文献 [118] 提出的“普适跟踪”概念, 将混合跟踪推动到更高级别, 在这个级别需要利用各类传感器尽可能获取自然环境中的所有可利用信息, 并能够自动和动态地进行融合所有可利用信息. 普适跟踪可推动通用混合跟踪框架及其开发平台的研究发展, 促进 AR 混合跟踪系统的开发更快速、简易和方便。

2) 在混合跟踪器的相对姿态标定方面, 难点在于完全自动化标定. 与其他标定方法相比, 动态滤波法在标定参数自由度个数、精度、操作复杂度等方面都有很大优势, 但是也还很难实现完全自动化的标定操作. 完全自动化标定要求不使用任何自制的标定板, 在标定混合传感器相对姿态时还要即时标定摄像机内外参数; 特别在室外环境条件下, 由于受到复杂环境、光照等各种因素影响, 要实现全自动化标定就更为困难。

3) 在跟踪器同步方面, 难点在于混合跟踪器的在线同步法. 硬件同步法简单且精度高, 离线同步法需设计标定平台但其同步效果显著, 但是在线同步法可靠性差. 其原因在于, 在线同步法需要考虑头部跟踪全程中的各种病态输入条件, 比如传感器静止

不动、高速运动或进行周期性运动等. 文献 [83] 提出的互相关同步法在普通运动情况下没有显示出较好的同步效果, 在存在病态输入条件下更会导致错误计算, 无法实现有效同步. 在线同步法必须首先确定当前的数据输入是否有足够的信息量确保可靠的时间同步. 可见, 要实现满足各种病态输入条件的跟踪器在线同步是非常困难的。

8 应用前景

上世纪 90 年代中期, 单个姿态跟踪器已经不能解决 AR 应用中因遮挡、剧烈运动变化、光照以及复杂环境等导致的头部姿态跟踪失败问题. 混合跟踪成为解决这一跟踪失败问题的唯一途径. 那时虽然提出了一些新颖的融合方法, 但是所构建的应用于室内外环境条件下的混合跟踪器原型, 在跟踪稳定性和硬件集成等方面离实用性都还有一段距离. 在本世纪初, 在其他类型混合跟踪器发展缓慢时, 光学混合跟踪器和视觉-惯性跟踪器却倍受关注且至今已经发展到比较成熟的技术水平, 复合混合跟踪器正成为最近几年的研究重点. 随着惯性跟踪器使用微电机技术在小型化和稳定性方面取得进步, InterSense 公司和 Xsense 公司的系列商业混合跟踪器陆续面世, 同时伴随当前计算机性能的快速提高, 混合跟踪器在不同领域得到更实际应用. 除了传统的工业、军事^[119-121]、医疗手术、文化遗迹保护、飞行模拟器、室内导航、电视转播等应用, 混合跟踪器在如下几个领域具有广泛应用前景。

1) 核工业. 目前在核电站维修^[122] 和辐射材料手套箱加工^[123] 的 AR 应用中仅仅使用标识板进行场景注册. 然而, 核工业环境由于材料成本高, 对人体健康有辐射威胁, 操作环境具有较高危险性, 需要使用视觉-惯性等混合跟踪器跟踪用户头部高速运动, 帮助人员及时处理危险紧急事件, 降低辐射威胁等。

2) 机器人. AR 与机器人结合可以实现远程呈现、远程操作等, 为医用机器人和工业机器人等提供更自然的交互接口. 混合跟踪技术可帮助机器人 AR 系统提供路径导航和定位等, 比如视觉-惯性传感器可以实时为系统创建地图和实现定位^[124], 摄像机与激光测距仪组合可为远程人脑机器人系统提供导航^[125] 等。

3) 互动娱乐. AR 与多媒体数字技术结合, 可以为数字文化产业注入新元素, 而混合跟踪器可以帮助 AR 技术实现互动娱乐游戏等. 典型例子是微软公司开发的 Kinect 体感设备, 综合使用体感摄像机、激光测距仪等跟踪用户 3D 体感动作, 实现互动娱乐应用^[126]. 此外, 通过在可移动智能通信设备上安装微型视觉-惯性-GPS 跟踪器, 混合跟踪技术

还能够将互动娱乐文化移植到移动通信设备上,实现移动式互动娱乐,为用户提供与环境栩栩如生的交互感知,改变普通人的文化娱乐方式。

4) 普适计算 (Ubiquitous computing). 普适计算利用用户所在环境中的所有跟踪传感器,将所有跟踪传感器的信息进行融合,获得周围环境的情景意识,满足普通人日常工作和衣食住行方面的需要。麻省理工学院开发的“第六感知”交互式 AR 系统^[127],利用摄像机已经实现了超市商品查询、电话拨号等方面的普适生活应用。可以想象,若使用混合跟踪技术将这一交互概念与 GPS 以及其他跟踪传感器相混合,结合 Google 地理信息系统和未来的各类物联网知识信息系统,能够为任意地理位置上的用户提供无处不在的生活信息,从而有可能在将来革命性地改变普通人在日常工作和衣食住行等方面的生活方式。

9 总结

混合跟踪是增强现实领域最近二十年发展迅速的重要关键技术,有效解决了增强现实系统既要求高精度跟踪又要求强鲁棒性的技术难题。本文主要对 AR 混合跟踪技术进行了综述性论述,其论述要点总结如下:

1) 构成混合跟踪器的硬件跟踪器各具优缺点,主要用于特定增强现实应用,而混合跟踪器能够兼具硬件跟踪器各自优势达到优势互补,实现这种优势互补依赖于混合跟踪技术所涉及的混合跟踪方法、标定方法和时间同步法。

2) 混合跟踪方法是多传感器融合在 AR 姿态跟踪技术中的典型应用,其数学基础为多传感器融合的 KF、EKF、UKF 以及粒子滤波器等数学工具。混合跟踪方法分为互补式、竞争式和协作式传感器融合,面向不同应用有多种形式,且各具优缺点,其发展趋势是使用传感器融合进行姿态融合,未来发展方向包括航迹融合、多速率滤波融合、粒子滤波器融合等,未来发展的难点是普适跟踪条件下任意传感器的数据融合。

3) 混合跟踪器的标定方法经历绝对朝向、手眼标定到动态滤波法的发展过程,发展方向是使用传感器融合工具进行动态标定以及实现自动化标定操作,而未来发展的难点是混合跟踪器的全自动标定,特别是在室外环境下的全自动标定。

4) 混合跟踪器中比较成熟和稳定的时间同步法是硬件同步法,其他离线或在线的同步法还处于探索研究阶段,未来发展的难点是在线同步法。

5) 混合跟踪器面向不同应用具有不同混合类型,其中光学混合、超声波-惯性和视觉-惯性混合跟踪器的发展技术水平已比较成熟,而未来发展重

点在于视觉-惯性-GPS 等复合跟踪器。除了传统的工业、医疗、军事、文化遗迹保护、导航等领域,混合跟踪器还能够为核工业、机器人、互动娱乐以及普适计算等室内外 AR 应用注入新活力。

综上,混合跟踪技术在增强现实领域发挥非常重要的作用。没有混合跟踪技术,增强现实将不得不局限在有限的应用空间范围内。可以确信,在现有发展基础上将会有更多新颖的混合跟踪方法、标定方法以及时间同步法被应用于混合跟踪,先进的混合跟踪技术将帮助增强现实涉入普通人的日常生活,进而可能改变人们的生活形态。

References

- 1 Caudell T P, Mizell D W. Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Proceedings of the 25th Hawaii International Conference on System Sciences. Hawaii: IEEE, 1992. 659-669
- 2 Milgram P, Furnio K. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 1994, **E77-D**(12): 1321-1329
- 3 Azuma R, Bailiot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 2001, **21**(6): 34-47
- 4 Feng Z, Duh H B L, Billingham M. Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR. In: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. Cambridge, UK: IEEE, 2008. 193-202
- 5 Carmigniani J, Furht B, Anisetti M, Ceravolo P, Damiani E, Ivkovic M. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 2011, **51**(1): 341-377
- 6 Rolland J P, Bailiot Y, Goon A A. A Survey of Tracking Technology for Virtual Environments, Technical Report, University of Central Florida, Orlando FL, USA, 1999
- 7 Wang Yong-Tian, Liu Yue, Hu Xiao-Ming. Study on key technique and application of outdoor AR system. *Journal of System Simulation*, 2003, **15**(3): 329-333
(王涌天, 刘越, 胡晓明. 户外增强现实系统关键技术及其应用的研究. 系统仿真学报, 2003, **15**(3): 329-333)
- 8 Chen Jing, Wang Yong-Tian, Liu Yue, Axel P. Real-time structure and motion by fusion of inertial and vision data for mobile AR system. *Journal of Beijing Institute of Technology*, 2006, **15**(4): 431-436
- 9 Chen Jing, Wang Yong-Tian, Liu Yue, Liu Wei, Guo Jun-Wei, Lin Jing-Dun. Hybrid tracking for outdoor augmented reality system. *Journal of Computer-Aided Design and Computer Graphics*, 2010, **22**(2): 204-209
(陈靖, 王涌天, 刘越, 刘伟, 郭俊伟, 林精敦. 适用于户外增强现实系统的混合跟踪定位算法. 计算机辅助设计与图形学学报, 2010, **22**(2): 204-209)
- 10 Su Hong, Kang Bo. Implementation of hybrid tracking technique in augmented reality system. *Computer Engineering*, 2009, **35**(4): 268-270
(苏宏, 康波. 混合跟踪技术在增强现实系统中的应用. 计算机工程, 2009, **35**(4): 268-270)

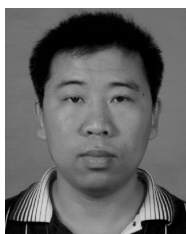
- 11 Li Xin-Yu, Chen Dong-Yi. Design and implementation of augmented reality system based on hybrid tracking. *Journal of Computer Applications*, 2009, **29**(10): 2852–2854, 2858 (李新宇, 陈东义. 基于混合跟踪的增强现实系统设计与实现. 计算机应用, 2009, **29**(10): 2852–2854, 2858)
- 12 Luo Bin. On Multi-Sensor Fusion Based Pose Tracking System and Calibration Technology [Ph.D. dissertation], Beijing Institute of Technology, China, 2010 (罗斌. 多传感器融合的姿态跟踪系统及其标定技术研究 [博士学位论文], 北京理工大学, 中国, 2010)
- 13 Luo Bin, Wang Yong-Tian, Liu Yue. Multi-sensor data fusion for optical tracking of head pose. *Acta Automatica Sinica*, 2010, **36**(9): 1239–1249 (罗斌, 王涌天, 刘越. 光学头部姿态跟踪的多传感器数据融合研究. 自动化学报, 2010, **36**(9): 1239–1249)
- 14 Yang Hao, Zhang Feng, Ye Jun-Tao. A camera-IMU relative pose calibration method. *Robot*, 2011, **33**(4): 419–426 (杨浩, 张峰, 叶军涛. 摄像机和惯性测量单元的相对位姿标定方法. 机器人, 2011, **33**(4): 419–426)
- 15 Brooks R R, Iyengar S S. Real-time distributed sensor fusion for time-critical sensor readings. *Optical Engineering*, 1997, **36**(3): 767–779
- 16 Kalman R E. A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 1960, (82: Series D): 34–45
- 17 Welch G, Bishop G. An introduction to the Kalman filter [Online], available: http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf, August 5, 2012
- 18 Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 2004, **92**(3): 401–422
- 19 Schön T B, Törnqvist D, Gustafsson F. Fast particle filters for multi-rate sensors [Online], available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.138.8290&rep=rep1&type=pdf>, August 10, 2012
- 20 Bar-Shalom Y, Li X R, Kirubarajan T. *Estimation with Applications to Tracking and Navigation*. New York: Wiley-Interscience Publication, 2001
- 21 Gelb A, Kasper J F, Nash R A, Price C F, Sutherland A A. *Applied Optimal Estimation*. Massachusetts, US: MIT Press Publication, 2001
- 22 Feiner S, Macintyre B, Hollerer T, Webster A. A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. In: Proceedings of the 1st International Symposium on Wearable Computers. Cambridge, Massachusetts, USA: IEEE, 1997. 74–81
- 23 Behringer R. Registration for outdoor augmented reality applications using computer vision techniques and hybrid sensors. In: Proceedings of the 1999 IEEE Virtual Reality. Houston, Texas, USA: IEEE, 1999. 244–251
- 24 Piekarski W, Gunther B, Thomas B. Integrating virtual and augmented realities in an outdoor application. In: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. San Francisco, CA, USA: IEEE, 1999. 45–54
- 25 Maeda M, Ogawa T, Kiyokawa K, Takemura H. Tracking of user position and orientation by stereo measurement of infrared markers and orientation sensing. In: Proceedings of the 8th International Symposium on Wearable Computers. Arlington, VA, USA: IEEE, 2004. 77–84
- 26 Schall G, Wagner D, Reitmayr G, Taichmann E, Wieser M, Schmalstieg D, Hofmann-Wellenhof B. Global pose estimation using multi-sensor fusion for outdoor augmented reality. In: Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality. Orlando, Florida, USA: IEEE, 2009. 153–162
- 27 Maida M, Ababsa F, Mallem M. Vision-inertial tracking system for robust fiducials registration in augmented reality. In: Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Multimedia Signal and Vision Processing. Nashville, TN, USA: IEEE, 2009. 83–90
- 28 Welch G, Bishop G. SCAAT: incremental tracking with incomplete information. In: Proceedings of the 24th annual conference on SIGGRAPH 1997. Los Angeles, CA, USA: ACM Press/Addison-Wesley Publication, 1997. 333–344
- 29 Jiang B. Robust Hybrid Tracking for Outdoor Augmented Reality [Ph.D. dissertation], University of Southern California, USA, 2004
- 30 Lin C, Nguyen K, Hoff B, Vincent T. An adaptive estimator for registration in augmented reality. In: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. San Francisco, CA, USA: IEEE, 1999. 23–32
- 31 Ababsa F E, Mallem M. Inertial and vision head tracker sensor fusion using a particle filter for augmented reality systems. In: Proceedings of the 2004 International Symposium on Circuits and Systems. Vancouver, British Columbia, Canada: IEEE, 2004. 861–864
- 32 Klein G, Drummond T. Sensor fusion and occlusion refinement for tablet-based AR. In: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington, VA: IEEE, 2004. 38–47
- 33 Reitmayr G, Drummond T W. Going out: robust model-based tracking for outdoor augmented reality. In: Proceedings of the 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality. Santa Barbara, CA, USA: IEEE, 2006. 109–118
- 34 Oskiper T, Chiu H P, Zhu Z W, Samarasekera S, Kumar R. Stable vision-aided navigation for large-area augmented reality. In: Proceedings of the 2001 IEEE Virtual Reality Conference. Singapore: IEEE, 2001. 63–70
- 35 Fuchs E M. Inertial Head-tracking [Master dissertation], MIT, Cambridge, USA, 1993
- 36 Welch G F. Hybrid Self-tracker: An Inertial/Optical Hybrid Three-dimensional Tracking System, Technical Report, University of North Carolina at Chapel Hill, 1995
- 37 Foxlin E, Harrington M, Pfeifer G. Constellation: a wide-range wireless motion-tracking system for augmented reality and virtual set applications. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. Orlando, FL, USA: ACM Press, 1998. 371–378
- 38 Foxlin E M, Harrington M, Altshuler Y. Miniature 6-DOF inertial system for tracking HMDs. In: Proceedings of the 1998 SPIE Helmet and Head-Mounted Displays III. Orlando, FL, USA: SPIE, 1998. 214–228

- 39 Foxlin E. Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter. In: Proceedings of the 1996 IEEE Virtual Reality Annual International Symposium. Santa Clara, California, USA: IEEE, 1996. 185–194, 267
- 40 Parnian N, Golnaraghi F. Integration of a multi-camera vision system and strapdown inertial navigation system (SDINS) with a modified Kalman filter. *Sensors*, 2010, **10**(6): 5378–25394
- 41 Diverdi S, Hollerer T. Heads up and camera down: A vision-based tracking modality for mobile mixed reality. *IEEE Transactions on Visualization and Computer Graphics*, 2008, **14**(3): 500–512
- 42 Schön T B, Gustafsson F. Integrated navigation of cameras for augmented reality. In: Proceedings of the 16th IFAC World Congress. Prague, Czech Republic: IEEE, 2005. 1–6
- 43 Hol J D, Schön T B, Luinge H, Slycke P J, Gustafsson F. Robust real-time tracking by fusing measurements from inertial and vision sensors. *Journal of Real-Time Image Processing*, 2008, **2**: 149–160
- 44 Yokokohji Y, Sugawara Y, Yoshikawa T. Accurate image overlay on video see-through HMDs using vision and accelerometers. In: Proceedings of the 2000 IEEE Virtual Reality. New Brunswick, New Jersey, USA: IEEE, 2000. 247–254
- 45 Azuma R, Bishop G. Improving static and dynamic registration in an optical see-through HMD. In: Proceedings of the 21st International Conference on Computer Graphics and Interactive Techniques. Orlando: ACM Press, 1994. 197–204
- 46 Saha R K. Effect of common process noise on two-sensor track fusion. *Journal of Guidance, Control, and Dynamics*, 1996, **19**(4): 829–835
- 47 Chang K C, Saha R K, Bar-Shalom Y. On optimal track-to-track fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 1997, **33**(4): 1271–1276
- 48 Luo B, Wang Y T, Liu Y. Sensor fusion based head pose tracking for lightweight flight cockpit systems. *Multimedia Tools and Applications*, 2011, **52**(1): 235–255
- 49 Hoff W A. Fusion of data from head-mounted and fixed sensors. In: Proceedings of the 1st International Workshop on Augmented Reality. San Francisco, CA, USA: IEEE, 1998. 167–182
- 50 Newman J, Ingram D, Hopper A. Augmented reality in a wide area sentient environment. In: Proceedings of the 2001 IEEE and ACM International Symposium on Augmented Reality. New York, NY, USA: IEEE, 2001. 77–86
- 51 Kalkusch M, Lidy T, Knapp N, Reitmayr G, Kaufmann H, Schmalstieg D. Structured visual markers for indoor pathfinding. In: Proceedings of the 1st IEEE International Workshop on Augmented Reality Toolkit. Darmstadt, Germany: IEEE, 2002. 8–15
- 52 Zendjebil I M, Ababsa F, Didier J Y, Mallem M. Hybrid localization system for mobile outdoor augmented reality applications. In: Proceedings of the 1st Workshops on Image Processing Theory, Tools and Applications. Sousse, Tunisia, France: IEEE, 2008. 1–6
- 53 Kanbara M, Fujii H, Takemura H, Yokoya N. A stereo vision-based augmented reality system with an inertial sensor. In: Proceedings of the 2000 IEEE and ACM International Symposium on Augmented Reality. Munich, Germany: IEEE, 2000. 97–100
- 54 Bajura M, Neumann U. Dynamic registration correction in video-based augmented reality systems. *IEEE Computer Graphics and Applications*, 1995, **15**(5): 52–60
- 55 Satoh K, Uchiyama S, Yamamoto H, Tamura H. Robust vision-based registration utilizing bird's-eye view with user's view. In: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality. Tokyo, Japan: IEEE, 2003. 46–55
- 56 Satoh K, Uchiyama S, Yamamoto H. A head tracking method using bird's-eye view camera and gyroscope. In: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington, VA, USA: IEEE, 2004. 202–211
- 57 Grimm M, Grigat R R. Real-time hybrid pose estimation from vision and inertial data. In: Proceedings of the 1st Canadian Conference on Computer and Robot Vision. Canada: IEEE, 2004. 480–486
- 58 State A, Hirota G, Chen D T, Garrett W F, Livingston M A. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New Orleans, LA, USA: ACM Press, 1996. 429–438
- 59 Auer T, Pinz A. Building a hybrid tracking system: integration of optical and magnetic tracking. In: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. San Francisco, CA, USA: IEEE, 1999. 13–22
- 60 Fischer J, Eichler M, Bartz D, Straper W. Model-based hybrid tracking for medical augmented reality. In: Proceedings of the 12th Eurographics Conference on Virtual Environments. Lisbon, Portugal: IEEE, 2006. 71–80
- 61 Aron M, Simon G, Berger M O. Handling uncertain sensor data in vision-based camera tracking. In: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington, USA: IEEE, 2004. 58–67
- 62 Aron M, Simon G, Berger M O. Use of inertial sensors to support video tracking. *Computer Animation and Virtual Worlds*, 2007, **18**(1): 57–68
- 63 Satoh K, Anabuki M, Yamamoto H, Tamura H. A hybrid registration method for outdoor augmented reality. In: Proceedings of the 2001 IEEE and ACM International Symposium on Augmented Reality. New York, NY, USA: IEEE, 2001. 67–76
- 64 You S, Neumann U, Azuma R. Orientation tracking for outdoor augmented reality registration. *IEEE Computer Graphics and Applications*, 1999, **19**(6): 36–42
- 65 Klein G S W, Drummond T W. Tightly integrated sensor fusion for robust visual tracking. *Image and Vision Computing*, 2004, **22**(10): 769–776
- 66 Klein G, Drummond T. Robust visual tracking for non-instrumental augmented reality. In: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality. Tokyo, Japan: IEEE, 2003. 113–122

- 67 Zhang Z Y. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330–1334
- 68 Tsai R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 1987, **3**(4): 323–344
- 69 Foxlin E, Naimark L. Miniaturization, calibration & accuracy evaluation of a hybrid self-tracker. In: Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality. Tokyo, Japan: IEEE, 2003. 151–160
- 70 Kim A, Golnaraghi M F. Initial calibration of an inertial measurement unit using an optical position tracking system. In: Proceedings of the 2004 Position Location and Navigation Symposium. Monterey, California, USA: IEEE, 2004. 96–101
- 71 Alves J, Lobo J, Dias J. Camera-inertial sensor modeling and alignment for visual navigation. *Machine Intelligence and Robotic Control*, 2003, **5**(3): 103–111
- 72 Lang P, Pinz A. Calibration of hybrid vision/inertial tracking systems [Online], available: <http://www.emt.tugraz.at/tracking/Publications/Lang2005.pdf>, August 3, 2012
- 73 Bin L, Wang Y T, Liu Y. Sensor fusion for vision-based indoor head pose tracking. In: Proceedings of the 5th International Conference on Image and Graphics. Xi'an, China: IEEE, 2009. 677–682
- 74 Horn B K P. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 1987, **4**(4): 629–642
- 75 Park F C, Martin B J. Robot sensor calibration: solving $AX = XB$ on the Euclidean group. *IEEE Transactions on Robotics and Automation*, 1994, **10**(5): 717–721
- 76 Baillet Y, Julier S J, Brown D, Livingston M A. A tracker alignment framework for augmented reality. In: Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality. Tokyo, Japan: IEEE, 2003. 142–150
- 77 Lobo J, Dias J. Relative pose calibration between visual and inertial sensors. *The International Journal of Robotics Research*, 2007, **26**(6): 561–575
- 78 Randeniya D I B, Gunaratne M, Sarkar S, Nazef A. Calibration of inertial and vision systems as a prelude to multi-sensor fusion. *Transportation Research Part C: Emerging Technologies*, 2008, **16**: 255–274
- 79 Hol J D, Schon T B, Gustafsson F. Modeling and calibration of inertial and vision sensors. *The International Journal of Robotics Research*, 2010, **29**(2–3): 231–244
- 80 Hol J D, Schon T B, Gustafsson F. A new algorithm for calibrating a combined camera and IMU sensor unit. In: Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision. Hanoi, Vietnam: IEEE, 2008. 1857–1862
- 81 Mirzaei F M, Roumeliotis S I. A Kalman filter-based algorithm for IMU-camera calibration: observability analysis and performance evaluation. *IEEE Transactions on Robotics*, 2008, **24**(5): 1143–1156
- 82 Kelly J, Sukhatme G S. Self-calibration of inertial and omnidirectional visual sensors for navigation and mapping. In: Proceedings of the 2010 IEEE International Conference on Robotics and Automation. Anshorage, Alaska, USA: IEEE, 2010. 1–6
- 83 Huber M, Schlegel M, Klinker G. Temporal calibration in multisensor tracking setups. In: Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality. Orlando, Florida, USA: IEEE, 2009. 195–196
- 84 Liang J D, Shaw C, Green M. On temporal-spatial realism in the virtual reality environment. In: Proceedings of the 4th annual ACM symposium on User Interface Software and Technology. New York, NY, USA: ACM Press, 1991. 19–25
- 85 Kantonen T. Sensor synchronization for AR applications. In: Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality. Seoul, Korea: IEEE, 2010. 245–246
- 86 Yamazoe H, Utsumi A, Tetsutani N, Yachida M. Vision-based human motion tracking using head-mounted cameras and fixed cameras. *Electronics and Communications in Japan (Part II: Electronics)*, 2007, **90**(2): 40–53
- 87 Young M C, Young S S, Sang K. Pose estimation from landmark-based vision and inertial sensors. In: Proceedings of the 2006 International Joint Conference on SICE-ICASE. Bexco, Busan, Korea: IEEE, 2006. 1668–1671
- 88 Foxlin E, Altshuler Y, Naimark L, Harrington M. Flight-Tracker: a novel optical/inertial tracker for cockpit enhanced vision. In: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington, USA: IEEE, 2004. 212–221
- 89 Foxlin E, Naimark L. VIS-Tracker: a wearable vision-inertial self-tracker. In: Proceedings of the 2003 IEEE Virtual Reality Conference. Los Angeles, CA: IEEE, 2003. 199–206
- 90 Wormell D, Foxlin E, Katzman P. Advanced inertial-optical tracking system for wide area mixed and augmented reality systems [Online], available: <http://www.intersense.com/pages/44/127/>, August 3, 2012
- 91 Parnian N, Golnaraghi M F. Integration of vision and inertial sensors for industrial tools tracking. *Sensor Review*, 2007, **27**(2): 132–141
- 92 Hogue A, Jenkin M R, Allison R S. An optical-inertial tracking system for fully-enclosed VR displays. In: Proceedings of the 1st Canadian Conference on Computer and Robot Vision. Canada: IEEE, 2004. 22–29
- 93 Ribo M, Lang P, Ganster H, Brandner M, Stock C, Pinz A. Hybrid tracking for outdoor augmented reality applications. *IEEE Computer Graphics and Applications*, 2002, **22**(6): 54–63
- 94 Ababsa F, Mallem M. Hybrid three-dimensional camera pose estimation using particle filter sensor fusion. *Advanced Robotics*, 2007, **21**(1–2): 165–181
- 95 Rehbinder H, Ghosh B K. Multi-rate fusion of visual and inertial data. In: Proceedings of the 2001 International Conference on Multisensor Fusion and Integration for Intelligent Systems. Kongresshaus Baden-Baden, Germany: IEEE, 2001. 97–102

- 96 Rehbinder H, Ghosh B K. Pose estimation using line-based dynamic vision and inertial sensors. *IEEE Transactions on Automatic Control*, 2003, **48**(2): 186–199
- 97 Wu Y X, Hu D W, Wu M P, Wu X P, Wu T. Observability analysis of rotation estimation by fusing inertial and line-based visual information: a revisit. *Automatica*, 2006, **42**(10): 1809–1812
- 98 Jiang B, Neumann U, Suya Y. A robust hybrid tracking system for outdoor augmented reality. In: Proceedings of the 2004 IEEE Virtual Reality. Chicago, IL, USA: IEEE, 2004. 3–275
- 99 Suya Y, Neumann U, Azuma R. Hybrid inertial and vision tracking for augmented reality registration. In: Proceedings of the 1999 IEEE Virtual Reality. Houston, Texas, USA: IEEE, 1999. 260–267
- 100 You S, Neumann U. Fusion of vision and gyro tracking for robust augmented reality registration. In: Proceedings of the 2001 IEEE Virtual Reality. Yokohama, Japan: IEEE, 2001. 71–78
- 101 Hol J, Slycke P, Schön T, Gustafsson F. 2D-3D model correspondence for camera pose estimation using sensor fusion [Online], available: <http://www.xsens.com>, August 3, 2012
- 102 Bleser G, Wohlleber C, Becker M, Stricker D. Fast and stable tracking for AR fusing video and inertial sensor data. In: Proceedings of the 2006 International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision. Plzen, Czech Republic: IEEE, 2006. 109–115
- 103 Hol J D, Schon T B, Gustafsson F, Slycke P. Sensor fusion for augmented reality. In: Proceedings of the 9th International Conference on Information Fusion. Florence, Italy: IEEE, 2006. 1–6
- 104 Chandaria J, Thomas G A, Stricker D. The MATRIS project: real-time markerless camera tracking for augmented reality and broadcast applications. *Journal of Real-Time Image Processing*, 2007, **2**(2): 69–79
- 105 Bleser G, Stricker D. Advanced tracking through efficient image processing and visual-inertial sensor fusion. In: Proceedings of the 2008 IEEE Virtual Reality Conference. Reno, Nevada, USA: IEEE, 2008. 137–144
- 106 Bleser G, Hendebly G. Using optical flow as lightweight SLAM alternative. In: Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality. Orlando, Florida, USA: IEEE, 2009. 175–176
- 107 Waechter C, Huber M, Keitler P, Schlegel M, Pustka D, Klinker G. A multi-sensor platform for wide-area tracking. In: Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality. Seoul, Korea: IEEE, 2010. 275–276
- 108 Hollerer T, Feiner S, Pavlik J. Situated documentaries: embedding multimedia presentations in the real world. In: Proceedings of the 3rd International Symposium on Wearable Computers. San Francisco, California, USA: IEEE, 1999. 79–86
- 109 Azuma R, Weon Lee J, Jiang B, Park J, You S, Neumann U. Tracking in unprepared environments for augmented reality systems. *Computers and Graphics*, 1999, **23**(6): 787–793
- 110 Azuma R, Hoff B, Neely H III, Sarfaty R. A motion-stabilized outdoor augmented reality system. In: Proceedings of the 1999 IEEE Virtual Reality. Houston, Texas, USA: IEEE, 1999. 252–259
- 111 Lobo J, Lucas P, Dias J, de Almeida A T. Inertial navigation system for mobile land vehicles. In: Proceedings of the 1995 IEEE International Symposium on Industrial Electronics. Athens, Greece: IEEE, 1995. 843–848
- 112 Hagure T, Marchant J A, Tillett N D. Ground based sensing systems for autonomous agricultural vehicles. *Computers and Electronics in Agriculture*, 2000, **25**(1): 11–28
- 113 Hu Z, Uchimura K. Fusion of vision, GPS and 3D gyro data in solving camera registration problem for direct visual navigation. *International Journal of ITS Research*, 2006, **4**(1): 3–12
- 114 Karlekar J, Zhou S Z, Nakayama Y, Lu W Q, Loh Z C, Hii D. Model-based localization and drift-free user tracking for outdoor augmented reality. In: Proceedings of the 2010 IEEE International Conference on Multimedia and Expo. Singapore: IEEE, 2010. 1178–1183
- 115 Park J. 3DOF tracking accuracy improvement for outdoor augmented reality. In: Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality. Seoul, Korea: IEEE, 2010. 263–264
- 116 Milic L, Saramaki T, Bregovic R. Multirate filters: an overview. In: Proceedings of the 2006 IEEE Asia Pacific Conference on Circuits and Systems. Singapore: IEEE, 2006. 912–9150
- 117 Armesto L, Tornero J, Vincze M. Fast ego-motion estimation with multi-rate fusion of inertial and vision. *International Journal of Robotics Research*, 2007, **26**(6): 577–589
- 118 Newman J, Wagner M, Bauer M, MacWilliams A, Pintaric T, Beyer D, Pustka D, Strasser F, Schmalstieg D, Klinker G. Ubiquitous tracking for augmented reality. In: Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington, USA: IEEE, 2004. 192–201
- 119 Piekarski W, Thomas B. ARQuake: the outdoor augmented reality gaming system. *Communications of the ACM*, 2002, **45**(1): 36–38
- 120 Höllerer T, Feiner S, Terauchi T, Rashid G, Hallaway D. Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers and Graphics*, 1999, **23**(6): 779–785
- 121 Julier S, Baillot Y, Lanzagorta M, Brown D, Rosenblum L. BARS: Battlefield augmented reality system. In: Proceedings of the 2000 NATO Symposium on Information Processing Techniques for Military Systems. Istanbul, Turkey: IEEE, 2000. 9–11
- 122 Ishii H, Bian Z, Fujino H, Sekiyama T, Nakai T, Okamoto A, Shimoda H, Izumi M, Kanehira Y, Morishita Y. Augmented reality applications for nuclear power plant maintenance work. In: Proceedings of the 2007 International Symposium on Symbiotic Nuclear Power Systems for 21st Century. Tsuruga, Jaolan, Japan, 2007. 262–268
- 123 Reinhart W, Fursch A. A camera-based support system with augmented reality functions for manual tasks in radioactive production environments. *Production Engineering*, 2008, **2**(2): 139–147

- 124 Jones E S, Soatto S. Visual-inertial navigation, mapping and localization: a scalable real-time causal approach. *The International Journal of Robotics Research*, 2011, **30**(4): 407–430
- 125 Escolano C, Antelis J, Minguez J. Human brain-teleoperated robot between remote places. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan: IEEE, 2009. 4430–4437
- 126 Kinect for windows [Online], available: <http://www.microsoft.com/en-us/kinectforwindows/>, August 3, 2012
- 127 Mistry P, Maes P. SixthSense: a wearable gestural interface. In: Proceedings of the 2009 SIGGRAPH Asia'09. Ketch, Yokohama, Japan: ACM, 2009. 85–92



罗斌 中国工程物理研究院计算机应用研究所高级工程师, 博士。1997 年获北京理工大学光电工程系学士学位, 2000 年获中国工程物理研究院硕士学位, 2010 年获北京理工大学光电学院博士学位。主要研究方向为虚拟现实和增强现实技术。

E-mail: luobin1827@bit.edu.cn

(**LUO Bin** Senior engineer and Ph. D. at the Institute of Computer Application, China Academy of Engineering and Physics (CAEP). He received his bachelor degree from Beijing Institute of Technology (BIT) in 1997, master degree from the Graduate Faculty of CAEP in 2000, and Ph. D. degree from BIT in 2010, respectively. His research interest covers virtual reality and augmented reality.)



王涌天 北京理工大学光电学院教授。1981 年于天津大学获得学士学位, 1986 年获英国 Reading 大学博士学位。主要研究方向为光学系统设计, 虚拟现实和增强现实, 医学图像处理与手术导航。本文通信作者。E-mail: wyt@bit.edu.cn

(**WANG Yong-Tian** Professor at the School of Optoelectronics, Beijing Institute of Technology. He received his bachelor degree from

Tianjin University in 1981 and Ph. D. degree from Reading University, England in 1986, respectively. His research interest covers optical design, virtual reality, augmented reality, medical image processing, and operation navigation. Corresponding author of this paper.)



沈浩 中国工程物理研究院计算机应用研究所研究员。主要研究方向为计算机数值模拟与仿真。

E-mail: shenh@ica.caep.cn

(**SHEN Hao** Professor at the Institute of Computer Application, China Academy of Engineering and Physics. His research interest covers computer numerical simulation and emulation.)



吴志杰 中国工程物理研究院计算机应用研究所研究员。主要研究方向为计算机软件应用。E-mail: wuzj@ica.caep.cn

(**WU Zhi-Jie** Professor at the Institute of Computer Application, China Academy of Engineering and Physics. His main research interest is computer software application.)



刘越 北京理工大学光电学院教授。1995 年获吉林大学硕士学位, 2000 年获吉林大学博士学位。主要研究方向为光电信号处理, 计算机视觉, 虚拟现实和增强现实。E-mail: liuyue@bit.edu.cn

(**LIU Yue** Professor at the School of Optics and Electronics, Beijing Institute of Technology. He received his

master degree from Jilin University of Technology in 1995 and Ph. D. degree from Jilin University in 2000, respectively. His research interest covers optoelectronics signal processing, computer vision, virtual reality, and augmented reality.)