

基于手势识别的文物数字化交互展示

张 月¹ 卫克晶¹ 闫宏斌²
ZHANG Yue WEI Kejing YAN Hongbin

摘 要

为解决传统文物展示方式互动性差、游客参与度不足的问题,文章研究并开发了一种基于手势识别的文物数字化交互展示系统。系统利用浏览器摄像头实时捕获用户手势,基于 MediaPipe 框架识别手部关键点并定义核心交互手势;通过 WebGL 的 Three.js 框架库在网页端加载并渲染三维文物模型,建立手势指令与模型旋转、缩放、切换控制之间的映射逻辑。系统采用分层状态管理与平滑插值算法,保证了在交互过程中系统的流畅与稳定。试验结果表明,该系统能准确识别预设手势,并实时驱动文物模型完成预期的缩放、旋转及模型切换等操作,证实了基于手势识别的自然交互在文物数字化展示中的可行性与有效性。

关键词

手势识别; 文物数字化; 三维交互展示; MediaPipe; Three.js

doi: 10.3969/j.issn.1672-9528.2026.03.046

0 引言

文物是连接过去与现在的桥梁,能够直观地传播历史文化知识^[1],借助文物展览可加深人们对文物历史、艺术和科学价值的理解^[2]。但传统的文物展示方式存在其局限性:实体文物展览受空间、时间等条件限制,大量文物难以向公众展示;二维图片或视频展示虽扩展了传播范围,但未展示文物的三维尺度和材质细节,使公众只能进行远距离、单向的观察,难以满足公众日益增长的互动性、探索性需求^[3]。

近年来,数字化与数字孪生技术的快速发展为文物展示与交流带来了新的机遇。通过高精度三维扫描、建模与渲染技术,文物可以被高保真、立体地复现为数字模型,形成虚拟“数字孪生体”^[4],观众能通过网络“零距离”接触、多角度观赏文物。数字化展示框架下交互方式的自然性与直观性是提升用户体验的关键^[5]。传统的鼠标、键盘或触摸屏交互方式对文物操作的真实感不足,对于基于文物的青少年文化教育而言,趣味性不足。若能实现基于手势识别的人与文物交互,用户仅通过自然的手部动作即可实现对虚拟文物的抓取、旋转、缩放等操作,观众从被动观察者转变为主动探索者,这是智慧文博的一个重要发展方向^[6]。因此,本系统将手势识别交互技术融入文物的数字化孪生展示中,以构建一种更具沉浸感、互动性的文物鉴赏模式。

1 文物交互展示功能需求

根据通过手势控制并观察文物的实际需求,本文所设计

文物交互展示系统的主要功能如图 1 所示。系统需具有初始化、实时手势识别和手势与文物交互功能。初始化功能中显示预先设置的文物模型、文物介绍文字界面或宣传画面。初始化完成后,该系统能够通过摄像头实时采集视频流、逐帧识别手势并赋予其含义。在手势识别的基础上,该系统能够根据不同的手势实现四种类型的交互:(1)能够展示当前文物的三维模型及其文字介绍、当前手势信息、手势控制进程等信息;(2)根据特定手势,文物模型可以特定速度旋转,展示文物不同周向位置的细节,实现 360°全覆盖;(3)根据特定手势,实现不同文物模型的稳定切换;(4)根据特定手势,实现文物模型的细节缩放,满足总体展示和细节观赏的需求。

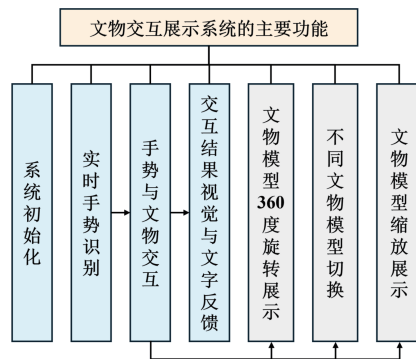


图 1 本文所设计文物交互展示系统的主要功能

2 交互系统设计与实现

2.1 系统架构

根据上述功能需求,本文拟定的文物交互系统总体架构如图 2 所示。总体采用分层次、模块化设计思路,将系统分

1. 南京理工大学紫金学院 江苏南京 210023

2. 南京理工大学 江苏南京 210094

为游客层、GUI 交互界面层、算法层、模型库和传感器层，各层与模块职责清晰、耦合度低，在保证系统可靠运行的同时，方便开展后续功能扩展。

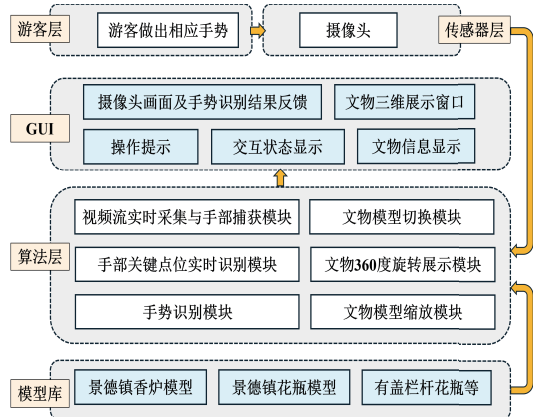


图2 基于手势识别的文物数字化交互展示系统总体架构

游客层和传感器层构成本系统的物理空间，其余各层构成本项目的虚拟空间。游客的手势动作由摄像头采集并传输至算法层，该层的视频实时采集与手部捕获模块用于获取视频图像，以捕获不同时刻用户的手势照片；该照片被传输至手部关键点实时识别模块，基于关键点信息，实现手势定义与识别；根据不同的手势，算法层可操作三维模型，实现模型切换、缩放和旋转等展示功能。模型库则负责管理 3D 文物模型，可通过对真实文物扫描等方式获取模型、扩充模型库。游客通过 GUI 界面查看系统信息、手势操作反馈信息并欣赏文物，本文所设计的 GUI 界面主要包括摄像头画面及手势识别结果反馈、操作提示、交互状态显示、文物三维展示窗口、文物信息显示等界面，实际可根据需要灵活拓展。

2.2 文物模型及其立体展示

本文用于系统测试的文物模型如图 3 所示，来源于史密森尼学会 3D 数字化平台。由于原始模型尺寸较小，为提高文物纹理的清晰度、便于游客观赏，本文针对性地对模型做了放大处理，分别将放大 40 倍和 60 倍的模型设定为基准模型和最大模型，作为模型缩放控制的极限值。



图3 本文用于交互展示的代表性文物模型

本文基于 WebGL 的 Three.js 框架库实现文物模型的三维展示^[7]，程序逻辑及其所依赖的关键函数如图 4 所示。主要包括 Three.js 场景初始化、WebGL 渲染器创建、相机实例创建、光照条件设置、三维模型加载、模型运动更新与实时渲染等关键环节^[8]。

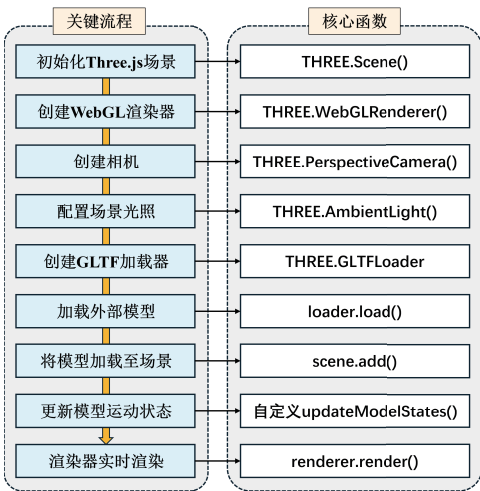


图4 文物模型展示程序逻辑及其所依赖的关键函数

2.3 手势定义及其识别

如图 5 所示，本系统通过浏览器调用摄像头获取视频流，并基于 MediaPipe 框架对视频帧进行实时手势识别^[9]。

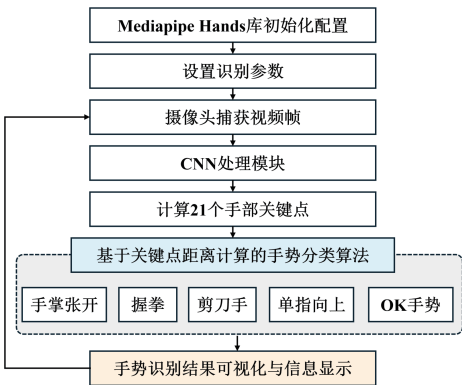


图5 手部关键点识别、手势定义及其识别算法流程

在初始化阶段，通过 MediaPipe Hands 设置识别参数，视频流经逐帧捕获后，被传递至 MediaPipe 的 WASM 模块进行处理，该模块首先对图像进行预处理，随后调用内置的 TFLite 轻量化 CNN 网络，该模型加载了预训练的神经网络权重，可实时识别并返回表征手部结构的 21 个关键点坐标。基于这些关键点之间的空间位置关系，通过几何规则判定手势类型。若所有指尖关键点均远离手腕，则识别为“手掌张开”手势；若所有指尖关键点均靠近手腕，则判定为“握拳”手势；当仅食指指尖远离手腕而其余指尖靠近时，识别为“单指向上”手势；若食指与中指指尖远离手腕而其他指尖靠近，则判定为“剪刀手”手势；若拇指和食指距离很近，而中指、无名指、小指都距离手腕较远，则判定为“OK”手势。识别完成后，在界面中实时更新反馈。系统共支持手掌张开、握拳、剪刀手、单指向上及 OK 五种预设手势的识别与反馈，对未识别或未知手势则显示“？”图标与“等待识别”文字提示。反馈信息通过直接操作 DOM 元素实现即时更新，形成图形与文本相结合的双重反馈机制。



图9 手势控制文物模型切换功能的试验验证



图10 手势控制文物模型缩放功能的试验验证



图11 手势控制文物模型旋转功能的试验验证

3 结语

本文针对文物数字化展示中对自然、直观交互方式的需求，设计并实现了一套基于手势识别的文物数字化交互展示系统。研究通过整合 MediaPipe 手势识别技术与 Three.js 三维渲染引擎，成功地将用户的手部动作转化为对虚拟文物模型的控制，实现了模型的旋转、缩放与切换等交互功能。系统测试结果表明，所提出的技术方案能够准确、稳定地响应手势指令，各项交互功能均达到了预期设计目标。根据实际需求，可在本系统框架下开展模型数量、控制功能、展示功能等的拓展开发，进一步增强了系统的实用性。

参考文献:

[1] 段天璟,曹文哲.文物的有效保护与利用:张忠培先生的文化遗产保护论[J].东南文化,2024(3):6-12.

[2] 袁晓梅.博物馆文物藏品数字化管理及其作用[J].办公室业务,2021(14):167-168.

[3] 郝石东.基于三维文物模型的数字化编辑系统的设计与实现[D].北京:北京邮电大学,2024.

[4] 王伟.实景三维模型在文物保护与修复中的应用探究[J].测绘与空间地理信息,2025,48(S1):181-183.

[5] 宾淑冰.新时代背景下文物陈列展示与教育传播的创新路径[J].文物鉴定与鉴赏,2025,19:92-95.

[6] LAVIOLA J J, KRUIJFF E, BOWMAN D, et al. 3D user interfaces: theory and practice[M]. Boston: Addison-Wesley Professional, 2017.

[7] 韩义.Web3D及Web三维可视化新发展:以WebGL和O3D为例[J].科技广场,2010,5:81-86.

[8] 乔斯·德克森.Three.js权威指南:在网页上创建3D图形和动画的方法与实践[M].北京:机械工业出版社,2025.

[9] 宋浩翔,张旭,沈宏晔,等.基于MediaPipe的手语识别系统[J].物联网技术,2025,15(10):4-6.

[10] POUPYREV I, BILLINGHURST M, WEGHORST S, et al. The go-go interaction technique: non-linear mapping for direct manipulation in VR[C]//UIST'96: Proceedings of the 9th annual ACM symposium on User interface software and technology. New York: ACM, 1996:79-80.

【作者简介】

张月(1989—),女,河北石家庄人,硕士,助教,研究方向:物联网技术。

卫克晶(1978—),女,山西运城人,硕士,副教授,研究方向:物联网技术。

闫宏斌(1988—),男,江苏南京人,博士,副教授,研究方向:深度学习。

(收稿日期:2026-01-10 修回日期:2026-03-12)