

引用格式: DONG Yue, CHEN Yue-ting, FENG Hua-jun, *et al.* Scene Fusion System Based on Ambient Light Detection[J]. *Acta Photonica Sinica*, 2020, 49(1): 0110001
董月, 陈跃庭, 冯华君, 等. 基于环境光检测的场景融合系统[J]. 光子学报, 2019, 48(1): 0110001

基于环境光检测的场景融合系统

董月, 陈跃庭, 冯华君, 徐之海, 李奇

(浙江大学 现代光学仪器国家重点实验室, 杭州 310027)

摘 要: 基于深度相机和人眼检测模型实时返回观察者眼睛的位置坐标, 利用坐标转换模型计算合理的显示区域, 通过颜色传感器的颜色分量数据估计环境色温和亮度; 根据提前标定的显示参数映射表调整显示状态, 使其最接近环境光的色彩表现。搭建了场景融合实验系统, 系统分为位置监测相机、图像采集相机、颜色传感器、显示设备、处理器五部分。分别在色温为 6 354 K, 照度为 160 lx 的室内场景和色温为 6 197 K, 照度为 848 lx 的室外场景下进行融合实验。实验结果表明, 场景融合方案能够为不同位置的观察者调整显示画面, 并根据环境光信息改变显示参数, 融合效果优良, 单次执行仅需 283 ms。

关键词: 场景融合; 色温; 亮度; 小孔成像模型; 人眼检测

中图分类号: TP751

文献标识码: A

doi: 10.3788/gzxb20204901.0110001

Scene Fusion System Based on Ambient Light Detection

DONG Yue, CHEN Yue-ting, FENG Hua-jun, XU Zhi-hai, LI Qi

(State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China)

Abstract: The position coordinates of the observers' eyes were derived in real time based on the depth camera and human eye detection model. Reasonable display area was calculated based on coordinate transformation model. The ambient color temperature and brightness were estimated based on the color component data of the color sensor. Finally, the display state was adjusted to make it nearest to the color performance of ambient light based on the pre-calibrated display parameter mapping table. A scene fusion experimental system was built, which can be divided into five parts: position monitoring camera, image capture camera, color sensor, display device and processor. Fusion experiments were carried out in indoor scenes with color temperature of 6 354 K and illumination of 160 lx, and outdoor scenes with color temperature of 6 197 K and illumination of 848 lx. The experimental results show that this scene fusion scheme can adjust display area for the observers in different locations, and change display parameters according to ambient light information, then achieve excellent fusion effects. The single execution only takes 283 ms.

Key words: Scene fusion; Color temperature; Brightness; Pinhole model; Human eye detection

OCIS Codes: 100.2000; 110.0110; 130.6010; 150.0150

0 引言

场景融合是指: 对特定位置的观察者, 控制显示设备输出图像信号, 使其显示内容与当前场景匹配, 同时根据环境光源信息调节显示参数, 以还原当前场景的真实色彩表现。目前, 场景融合技术已被应用于多种商业场合。三星在 2018 年发布了新款的量子点 (Quantum Dot Light Emitting Diodes, QLED) 电视系列, 同时

基金项目: 国防科技创新特区资助项目, 中央高校基本科研业务费专项基金 (No.2018QN81002)

第一作者: 董月 (1995—), 男, 硕士研究生, 主要研究方向为光学成像与图像处理。Email: 21730058@zju.edu.cn

导师 (通讯作者): 冯华君 (1963—), 男, 教授, 博导, 主要研究方向为光学成像与图像处理。Email: fenghj@zju.edu.cn

收稿日期: 2019-07-05; 录用日期: 2019-09-27

<http://www.photon.ac.cn>

推出了名为环境模式(Ambient Mode, AM)的新显示方式.在该模式下,内置算法能够模拟融合图像,在屏幕上渲染出和墙壁同样的色彩和纹理.奔驰推出了隐形概念车,在车体一侧覆盖发光二极管(Light Emitting Diodes, LED),另一侧安装一台数码单反摄像机,将摄像机拍摄得到的影像实时呈现在LED上,达到“隐形”的视觉效果.

场景融合在实现细节上与增强现实(Augmented Reality, AR)技术接近.AR技术强调虚拟信息与真实世界的融合交互,首要任务就是根据用户视角再现真实的场景信息.YOSHIDA T^[1]提出使用头戴式投影仪和相机分别获取场景图像,并利用尺度不变特征变换(Scale-invariant Feature Transform, SIFT)特征提取方法匹配关键点,完成图像的单应性变换,避免追踪用户视点.HILL A^[2]假设景物深度固定且与显示设备平行,使用两台相机与视角变换方法,研制出第一个用户视角的视频AR系统.TOMOIOKA M^[3]使用有限数量的标记点标定相机图像到用户视角图像的单应性变换矩阵,提高了视角变换的精度.现有的增强现实方法更注重再现画面的位置准确性,而不关注场景的真实亮度、色彩信息.相比之下,本文提出的场景融合方法既能保证再现画面与场景内容的连续性,又能还原场景真实的亮度与色彩表现.

场景融合涉及人眼检测、经典成像模型推导、环境光源信息检测、亮度色温估计、显示设备参数控制等多项技术.其中,人眼检测算法的执行效率很大程度上决定了场景融合系统的视觉效果.传统的人眼检测算法一般分为人脸检测和人脸特征点标记两步,例如VJ检测器^[4].随着深度学习方法的普及,Cascade CNN^[5], DenseBox^[6], Face R-CNN^[7], SSH^[8]等卷积神经网络模型相继出现,检测效果提升显著.为尽可能简化网络结构,提高检测精度,现有模型通常采用多任务学习策略,共享网络的特征提取结构,分别输出人脸检测结果和人脸特征点标记结果.MTCNN模型^[9]是最经典的针对人脸的多任务学习框架,它将人脸区域检测和人脸关键点检测放在了一起,利用三个子网络级联的形式提升检测效果.HyperFace模型^[10]在一个网络中实现了人脸检测、特征点定位、姿态估计、性别判断多项工作.

本文构建了完整的场景融合流程,包括观察者位置监测、坐标变换、环境光源信息传感、亮度色温估计、显示参数映射及控制等步骤.实验结果表明,本文设计的系统及算法流程能够实时获取观察者位置和环境光源信息,结合坐标转换模型和显示参数映射表,实现场景融合.

1 系统工作流程

图1为场景融合系统示意图,系统由位置监测相机、图像采集相机、RGBC(Red-Green-Blue-Clear)环境光传感器、处理器、显示设备组成.位置监测相机选用一款深度相机,结合人脸检测算法返回观察者的位置及深度.图像采集相机用于实时获取场景图像,根据已知的观察者空间坐标和相机内参矩阵,计算合理的显示

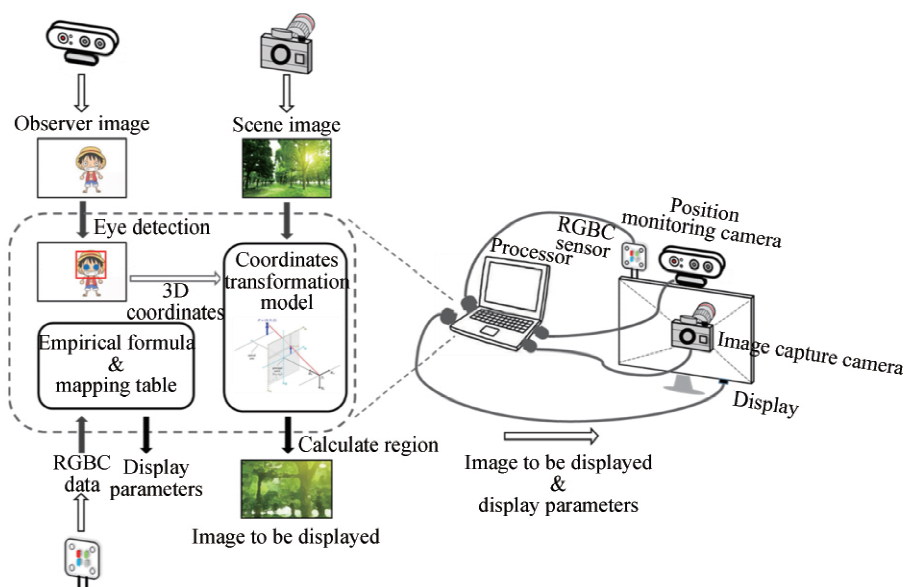


图1 场景融合系统示意

Fig.1 The diagram of scene fusion system

区域.RGBC 传感器可提取环境中不同波长的色光分量,用于估计当前色温与亮度.显示设备是系统中最重要的组成部分,用于最终实现场景融合.图 1 的虚线框中列举了系统的运行流程,位置监测相机和图像采集相机分别获取观察者图像和场景图像,在观察者图像中搜索人眼位置并返回其三维空间坐标,将人眼坐标输入坐标转换模型,计算场景图像中合理的显示区域.同时,RGBC 传感器获取各通道响应值,根据经验公式估计环境亮度与色温,最后根据显示设备的参数映射表选取合适的显示参数.处理器实时输出待显示图像和显示参数.显示设备接受显示参数,改变显示状态,并根据待显示图像实时变换显示内容,实现场景融合.

2 坐标转换模型

本文提出了场景融合系统的坐标转换模型,根据观察者与场景融合系统的相对位置关系,确定图像中的合理显示区域,以保证显示画面与当前场景准确匹配.

通用的相机成像模型由四部分组成:世界坐标系、相机坐标系、图像坐标系与像素坐标系.其中,世界坐标系与相机坐标系是三维坐标系,图像坐标系和像素坐标系是二维坐标系.四个坐标系中仅有像素坐标系是离散的,它的坐标值表示图像像素的行列索引.

相机坐标系下的物点 $[x, y, z]^T$ 和它在像素坐标系下的对应像素点 $[u, v]^T$ 存在如下关系

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}_{3 \times 3} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (1)$$

式中, $\mathbf{K}_{3 \times 3}$ 是当前相机的内参矩阵,与 CCD 像元尺寸和 CCD 分辨率相关.相机的装配误差,抖动、振荡等都会造成相机内参矩阵的变化,因此在实际场景中一般不使用内参矩阵的理论值.通常使用张正友标定法^[11]测量不同相机的实际内参矩阵.

式(1)将像素坐标 $[u, v]^T$ 写作齐次坐标形式 $[u, v, 1]^T$.在齐次坐标形式下, $[u, v, 1]^T$ 和 $z[u, v, 1]^T$ 表示同一个像素点.因此有如下结论:在同一相机坐标系下,空间坐标互成比例的两个物点 $[x, y, z]^T$ 和 $\alpha[x, y, z]^T$ 在成像过程中会被映射为像素坐标系下的同一个点,即对应同一个像素.

一般来说,场景融合系统搭建完毕后,显示器中心与位置监测相机、图像采集相机的相对偏移量保持不变.图 2 为坐标转换模型的示意图.观察者构成的相机坐标系记作 $O-X_o Y_o Z_o$,位置监测相机坐标系记作 $O-X_c Y_c Z_c$,图像采集相机坐标系记作 $O-X_c Y_c Z_c$.需要注意,由于位置监测相机面向观察者放置,故其相机坐标系的 X 轴、 Z 轴方向与其余两个坐标系的 X 轴、 Z 轴方向相反,三个坐标系的 Y 轴方向都相同.显示设备上的一点 Q 在观察者相机坐标系下的空间坐标为 $Q=[x_1, y_1, z_1]^T$,则距离为 z_2 的自然场景中空间坐标为 $Q' =$

$\begin{bmatrix} z_2 \\ z_1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}^T$ 的一点 Q' 与点 Q 对应,二者在

观察者的像素坐标系下被映射为同一个像素点.因

此,显示设备上的点 Q 应该显示自然场景中点 Q' 的内容.进一步地,需要计算自然场景中的点 Q' 在图像采集相机拍摄得到的图像中的对应点.

已知图像采集相机坐标系 $O-X_c Y_c Z_c$ 与观测相机坐标系 $O-X_o Y_o Z_o$ 相互平行,且图像采集相机光心 O_c 在观测相机坐标系 $O-X_o Y_o Z_o$ 下的坐标为 $[\Delta x, \Delta y, \Delta z]^T$,则自然场景中的点 Q' 在图像采集相机坐标系

$O-X_c Y_c Z_c$ 下的坐标为 $\begin{bmatrix} z_2 \\ z_1 \end{bmatrix} \begin{bmatrix} x_1 - \Delta x \\ y_1 - \Delta y \\ z_1 - \Delta z \end{bmatrix}^T$,最终映射到图像采集相机拍摄得到图像中的一点 Q'' .

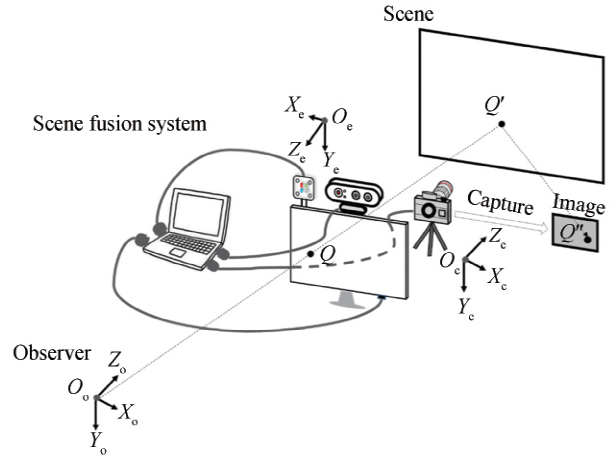


图 2 坐标转换模型示意

Fig.2 The diagram of coordinate transformation model

$$Q'' = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_e \begin{bmatrix} \frac{z_2}{z_1} x_1 - \Delta x \\ \frac{z_2}{z_1} y_1 - \Delta y \\ z_2 - \Delta z \end{bmatrix} \quad (2)$$

式(2)给出了齐次坐标形式的 Q'' 坐标,最终在场景图像中提取像素点 $[u, v]^T$,显示设备上的点 Q 应该显示场景图像中像素点 $[u, v]^T$ 的颜色及亮度.式(2)中 K_e 表示图像采集相机的内参矩阵.

已知对角的两顶点能确定唯一矩形.因此在实际的算法流程中,只选取显示设备的左上和右下两个顶点进行计算,即可确定图像中的合理显示区域.

3 人眼检测

第2节中给出了场景融合系统的坐标转换模型.该模型是建立在观察者位置已知的基础上的.如果将人眼也视作一台相机,那么晶状体中心就是该相机坐标系的光心.场景融合系统中的位置监测相机负责拍摄观察者图像,返回人眼坐标.为准确获取人眼相对于自适应场景变换和显示系统的空间坐标,选用一款深度相机作为位置监测相机.深度相机能够同时获取场景的三通道彩色图像和深度图像,使用MTCNN模型^[9]在彩色图像中搜索双眼位置.根据已知的双眼坐标,可在配准后的深度图像中提取双眼的深度数据.

以左眼为例,MTCNN模型给出左眼在彩色图像中的坐标 $[u_{\text{left}}, v_{\text{left}}]^T$,在配准后的深度图像的相同位置处可提取左眼深度 d_{left} .深度相机的深度测量范围是有限的,超出测量范围的像素点深度为0,即为无效点.为提高检测精度,在深度图像中,以像素点 $[u_{\text{left}}, v_{\text{left}}]^T$ 为中心,提取 11×11 大小的像素块,剔除无效点后,取像素块内的平均深度作为左眼深度 d_{left} .根据式(1),可得

$$\begin{bmatrix} u_{\text{left}} \\ v_{\text{left}} \\ 1 \end{bmatrix} = K_e \begin{bmatrix} x_{\text{left}} \\ y_{\text{left}} \\ d_{\text{left}} \end{bmatrix} \quad (3)$$

式中, K_e 是位置监测相机中可见光相机的内参矩阵.由式(3)可以计算左眼在位置监测相机坐标系下的空间坐标,其表达式为(为表达简便,式(3)中等号左边省去了 d_{left} ,在式(4)中不可忽略)

$$\begin{bmatrix} x_{\text{left}} \\ y_{\text{left}} \\ d_{\text{left}} \end{bmatrix} = K_e^{-1} \cdot \begin{bmatrix} u_{\text{left}} \\ v_{\text{left}} \\ 1 \end{bmatrix} \quad (4)$$

式中, K_e^{-1} 是 K_e 的逆矩阵, $[x_{\text{left}}, y_{\text{left}}, d_{\text{left}}]^T$ 是左眼在位置监测相机坐标系下的空间坐标.如此,系统可以实时获取人眼相对位置监测相机光心的三维空间坐标,由于位置监测相机和图像采集相机相对显示设备中心的三维空间坐标已知,坐标平移后即可获得显示设备顶点相对观察者的坐标,进一步计算得到目标区域顶点在图像中的位置和大小.

4 环境光检测

亮度和色温是最重要的光源特性,其测量值直接决定了融合系统的显示参数.专业级的亮度计和色温仪价格高昂、体积较大,且使用条件极为严苛,常要求对准光源方向测量.同时,手持式的亮度计和色温仪在没有人辅助的前提下,难以实现自动测量与即时通讯.因此,本文所提的场景融合方案选用RGBC颜色传感器获取环境中的光源信息.如图1, RGBC颜色传感器被固定在显示设备附近.

4.1 RGBC颜色传感器

RGBC颜色传感器共有Red、Green、Blue、Clear四个颜色通道,每个颜色通道都覆盖红外截止滤光片,仅接收可见光波段的光信号.Red、Green、Blue三个通道分别选用红色窄带滤光片、绿色窄带滤光片、蓝色窄带滤光片,结合同型号光电二极管,模拟人眼对不同波长色光的光谱响应.Clear通道除红外滤光片,不再单独设置滤光片,模拟人眼对复色光的光谱响应.

常用的 RGB 颜色传感器型号有 APDS9960、TCS34725、TCS3414CS 等。图 3 为 TCS34725 芯片的归一化光谱响应曲线。其中, Red 通道、Green 通道、Blue 通道的光谱响应呈现出单峰特性, 响应峰值分别位于 650 nm、550 nm、450 nm 附近, 基本符合 CIE 标准对三原色的波长定义。同时, Clear 通道的光谱响应呈现带通特性, 基本符合 380~780 nm 的可见光波段定义。由于滤光片及光电二极管的差异, 不同厂商出产的颜色传感器光谱响应曲线不同, 在峰值高度、半宽度、截止频率等方面存在明显差异。

不同通道的光电二极管接收光信号并转换为电信号, 经过时间积分与 A/D 转换, 实时输出传感器的 Red、Green、Blue、Clear 四通道响应值。

4.2 亮度估计

常用的光度量包括发光强度、光照度、光出射度、光亮度等^[12]。人眼的明暗感受与所处环境的光亮度线性相关。一般来说, 自然景物表面被光源照亮, 经过镜面反射或漫反射后被人眼接收。因此, 可以将自然场景视为二次光源。二次光源的光亮度由其所处平面的光照度决定。若主光源是余弦辐射体, 则二次光源的光亮度与光照度关系为

$$L = \frac{\rho E}{\pi} \quad (5)$$

式中, L 为光亮度, E 为光照度, ρ 为物体表面的反射率。

要获取自然场景的光亮度, 首先需要知道自然场景表面的光照度。在场景融合系统的应用场合中, 实验系统与真实场景往往存在较大间距。因此, 在没有外界帮助的前提下, 难以直接测量自然场景表面的光照度。另一方面, 阴天、雨天、多云天气时的环境光亮度基本由大气散射决定, 在无遮挡处的照度均匀性强。本算法使用显示设备表面的照度近似估计自然场景表面的照度, 从而计算得到二次光源亮度。

由图 2 可知, 尽管在 RGB 四个通道都设置了红外截止滤光片, 光谱响应曲线在红外波段仍呈现出一定程度的残余通过率, 且四个通道基本重合。因此, 首先估计红外波段的残余响应值

$$i_r = \frac{r + g + b - c}{2} \quad (6)$$

式中, r 、 g 、 b 、 c 分别为四通道初始响应值, i_r 表示红外波段残余响应值。然后, 将初始响应值分别减去 i_r , 结合经验系数计算 RGB 传感器表面照度:

$$E = [\alpha \quad \beta \quad \gamma] \cdot \begin{bmatrix} r - i_r \\ g - i_r \\ b - i_r \end{bmatrix} = [\alpha \quad \beta \quad \gamma] \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (7)$$

式中, $[\alpha \quad \beta \quad \gamma]$ 表示经验系数, 由传感器出厂商给出。

根据式(6)~(7)实时计算当前照度。将式(5)中的反射率 ρ 作为经验参数, 代入照度 E 估计当前拍摄场景的亮度。

4.3 色温估计

色温是指当前辐射源与某一温度黑体呈现的颜色相同时, 该黑体所处温度, 单位是开尔文(K)。实际情况中, 常用辐射源的光谱分布与黑体相差较远, 此时取该色品坐标距离黑体温度轨迹最接近处的黑体温度作为色温替代值, 称作相关色温。

一般来说, 色温不能直接测出, 而是根据光谱的相对光谱功率分布经计算得到。所以色温的测量实际上是测量光源的相对光谱功率分布, 其具体的测量方法可以分为光谱功率分布法和双色法两种^[13]。光谱功率分布法需要光谱辐射计来测量光源的光谱功率分布, 对设备的要求较高。相比之下, 双色法只需要测量某两个特定波长的相对光谱功率, 操作简便, 耗时较短, 应用广泛。

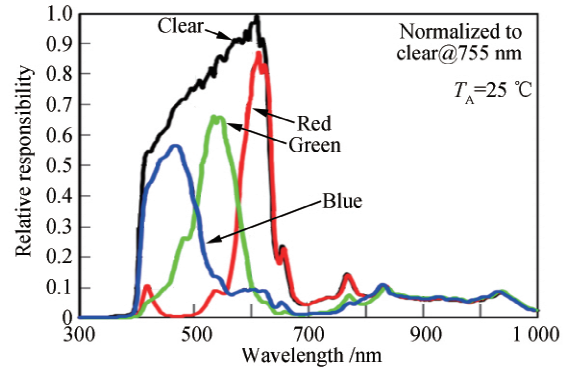


图 3 TCS34725 芯片的归一化光谱响应曲线
Fig.3 Normalized spectral response curve of TCS34725

一般的色温计测量蓝光对红光的比值,并由此确定被测光源的色温.此处,蓝光特指以 450 nm 为中心的窄波段,红光特指以 650 nm 为中心的窄波段.蓝-红比值与色温的关系曲线如图 4.由于图 4 中的曲线是单调的,而且可以近似看作线性,一般的 RGBC 传感器出厂商会以一次函数的形式给出色温与蓝红比的关系式,记为

$$T_c = k \frac{B}{R} + c \quad (8)$$

式中, T_c 为色温, $B = b - i_r$, $R = r - i_r$, 分别是蓝通道和红细胞响应值, k 和 c 表示一次函数的经验系数.

由于将蓝红比与色温的关系近似看作线性函数会引入误差,于是结合另一种色温估计方法.McCamy 公式^[14]在 1991 年被提出,它能够根据光源的 CIE1931 色品坐标估计色温.色品坐标可由三刺激值计算得到

$$\begin{cases} x = \frac{X}{X+Y+Z} \\ y = \frac{Y}{X+Y+Z} \end{cases} \quad (9)$$

式中, X, Y, Z 为 CIE1931 三刺激值, (x, y) 为 CIE1931 色品坐标.

本文选用文献[14]中 CIE 标准照明体的计算结果,此时 McCamy 公式可表示为

$$\begin{cases} n = \frac{x - 0.3320}{0.1858 - y} \\ T_c = 449n^3 + 3525n^2 + 6823n + 5520 \end{cases} \quad (10)$$

由于 RGBC 传感器的光谱响应与 CIE1931 标准色度观察者的光谱响应存在差异,因此需要构建二者的变换矩阵,其表达式为

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathbf{M}_{3 \times 3} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (11)$$

式中, $\mathbf{M}_{3 \times 3}$ 即为变换矩阵.文献[15]针对 TCS3414CS 传感器,选取日光、荧光灯、卤素灯、白炽灯、LED 光源共计 5 类 18 种光源,测定其 RGB 响应值与 X, Y, Z 三刺激值,标定并得到 $\mathbf{M}_{3 \times 3}$ 变换矩阵.本文算法将沿用其标定结果.已知 RGBC 颜色传感器的响应值,代入式(9)~(11),计算得到第二个色温估计值.选取两个色温估计值的平均值作为最终的环境光色温.

5 显示设备控制

为增强场景融合的视觉效果,需要调节显示设备的亮度和色温,使其与环境光保持一致.市面上通用的显示屏一般都支持亮度和色温调节.确定系统要使用的显示设备型号后,首先构建一个显示参数映射表.当场景融合系统工作时,根据实时获取的环境亮度和色温,在映射表中选择对应的参数调节显示器状态.

5.1 显示参数映射表

显示器常用的可调参数包括色温、亮度、对比度等.以 Dell U2718qm 显示器为例,该显示器支持五档色温调节:5 000 K、6 500 K、7 500 K、9 300 K、10 000 K.首先设定最接近当前环境色温 T_c 的色温调节参数 CCT,其表达式为

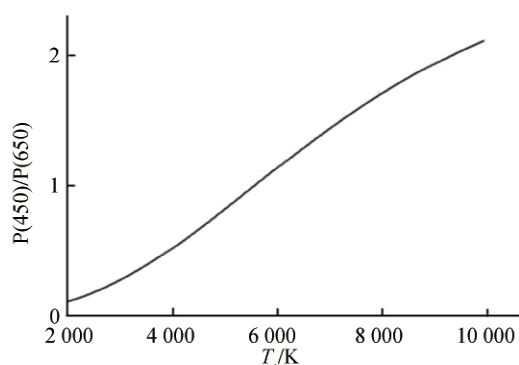


图 4 蓝-红响应值之比与色温的关系曲线
Fig.4 Relation curve between blue-red response ratio and color temperature

$$CCT = \begin{cases} 5\,000\text{ K}, & T_c \leq 5\,750\text{ K} \\ 6\,500\text{ K}, & 5\,750\text{ K} < T_c \leq 7\,000\text{ K} \\ 7\,500\text{ K}, & 7\,000\text{ K} < T_c \leq 8\,400\text{ K} \\ 9\,300\text{ K}, & 8\,400\text{ K} < T_c \leq 9\,650\text{ K} \\ 10\,000\text{ K}, & T_c \geq 9\,650\text{ K} \end{cases} \quad (12)$$

显示器设定的色温值不同,即使亮度值和对比度值完全相同,其实际辐射的光亮度也是不同的,这与显示器的驱动或增益控制有关.依次改变显示器的色温、亮度、对比度值,使用 MINOLTA 亮度计测量显示器实际辐射的光亮度,并记录在参数映射表中.对参数表做一次逆映射,即可由环境色温和光亮度确定合理的显示参数.表 1 为显示参数映射表示例.可知,若已将色温设定为 5 000 K,环境光亮度在 0~15 cd/m² 范围内变化,映射表总能给出一种显示参数设定方案.实物实验时,首先确定色温值,然后根据环境光亮度选择合适的亮度值和对比度值.

表 1 显示参数映射表示例
Table 1 An example of display parameter mapping table

Color temperature/K	Luminance/(cd·m ⁻²)	Brightness	Contrast
5 000	0	0	0
5 000	1	0	5
5 000	2	0	10
5 000	3	0	20
5 000	4	0	30
5 000	5	0	33
5 000	6	0	36
5 000	7	0	42
5 000	8	0	45
5 000	9	0	47
5 000	10	0	49
5 000	11	0	50
5 000	12	1	50
5 000	13	2	50
5 000	14	3	50
5 000	15	4	50

5.2 显示参数调节

在场景融合系统的工作过程中,无法手动调节显示器的色温、亮度、对比度.因此,考虑显示参数的程序控制方法.基于显示器数据通道命令接口(Display Data Channel Command Interface, DDC/CI)协议,用户可以通过程序接口来完成显示器的各项设置和色彩调节,而不必使用传统的屏幕菜单式调节方式(On-screen Display, OSD).市面上的大多数显示器都支持 DDC/CI 协议,这为自动参数调节带来了便利.

Windows 开发人员中心给出了监视器配置应用程序编程接口(Monitor Configuration Application Programming Interface, MCAPI),它基于 DDC/CI 协议,可发送指令给显示器.借助 MCAPI,最终实现了显示参数自动调节.

6 实验及结果

6.1 设备选型

位置监测相机选用 ORBBEC Astra 相机,该相机分辨率为 640×480,帧率为 30 fps,可同时提供彩色图像和深度图像.Astra 的深度测量范围约为 0.5~5 m,其在 1 m 距离处的深度测量误差可以控制在 3 mm 以内.

图像采集相机选用面阵相机 Mars5470S-18um/uc,该相机分辨率为 5 472×3 648,帧率为 30 fps.为该相机配备一款 6 mm 焦距, $F1.8\sim 16$,视场角为 $95.4^\circ \times 78.4^\circ \times 107.5^\circ$,最大畸变小于 2.3% 的广角镜头.

RGBC 颜色传感器选用 TCS34725.TCS34725 支持 I2C 总线通讯,故使用 Arduino UNO 开发板连接该

传感器读取响应值,Arduino 再经串口连接,将环境光信息实时返回给处理器.

显示设备选用 Dell U2718qm LCD 显示器,屏幕尺寸为 27 英寸,最大亮度为 350 cd/m^2 .处理器的 CPU 主频为 3.40 GHz,内存为 8 G.

6.2 实验系统搭建

图 5 给出了实验系统的侧视图.将位置监测相机固定在显示设备的正上方,将图像采集相机固定在显示设备的正后方.显示设备和位置监测相机面向观察者放置,图像采集相机面向拍摄场景放置.RGBC 颜色传感器与 Arduino 相连,同样面向拍摄场景固定,实时获取环境光信息.图 5 中所有的仪器设备都与处理器相连,有序执行算法流程.

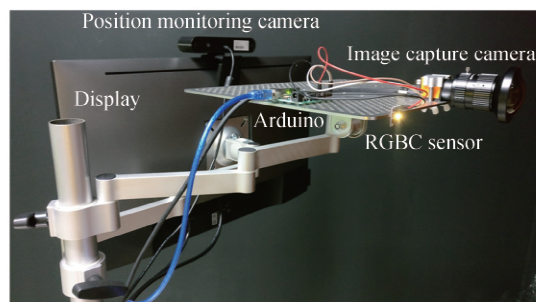
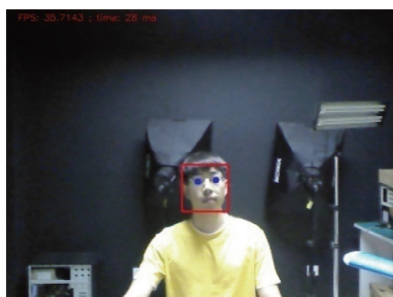


图 5 实验系统侧视图

Fig.5 Side view of our experimental system

6.3 人眼检测结果

第 2 节提到,使用深度相机作为位置监测相机,在拍摄得到的彩色图像中使用 MTCNN 检测算法搜索人眼位置,然后在深度图像中获取人眼深度.图 6 给出了一组人眼检测结果,图 6(a)为位置监测相机拍摄的彩色图,图 6(b)是它对应的深度图.为便于提取人眼深度,提前配准彩色图和深度图.在图 6(a)中,人脸用红色方框标识,人眼用蓝色点标识.图 6(b)为单通道图像,灰度越高表示距离越远,纯黑区域即为视场外部分.使用 OpenBLAS 开源计算库和 OpenCV 开源视觉库实现 MTCNN 检测算法,可达到 35 fps 的检测速度,同时保证精确的人眼定位效果.



(a) Eyes detection results in color image



(b) Corresponding depth image of Fig.6(a)

图 6 人眼检测结果

Fig.6 Human eyes detection results

6.4 场景融合实验

分别对室内和室外场景进行了融合实验.经验参数选取如下:式(7)中 $[\alpha \quad \beta \quad \gamma] = [0.136, 1, -0.444]$;

式(8)中 $k=3810$, $c=1391$;式(11)中 $M_{3 \times 3} = \begin{bmatrix} 0.2469 & 0.2707 & -0.1473 \\ 0.1735 & 0.3443 & -0.1143 \\ -0.1646 & 0.2210 & 0.3310 \end{bmatrix}$.场景融合系统按照指定算法

流程进入工作状态后,识别到显示设备前方的人脸,变换显示区域并调节显示参数.然后在人眼的相同位置处使用 Canon 5D Mark II 相机记录场景融合结果.采用这种方式,获得了图 7 和图 8 的实验结果.

图 7 是在室内记录的第一组场景融合结果,物距为 3.6 m,测量得色温为 6354 K,照度为 160 lx.查找显示参数映射表,设定色温为 6500 K,亮度为 23,对比度为 50.图 7(a)~(f)是系统完全工作时获得的实验结果,每一张图像对应的观察者位置不一致,但系统都表现出优秀的匹配效果,显示区域基本准确.同时,显示画面的亮度色温都与实际场景非常接近.

图 8 是阴雨天气记录的室外场景融合结果,物距为 15 m,测量得色温为 6197 K,照度为 848 lx.查找显示参数映射表,设定色温为 6500 K,亮度为 100,对比度为 100.图 8(a)和(b)是关闭显示参数调节模块获得的实验结果,图 8(a)明显色调偏冷,图 8(b)明显色调偏暖.图 8(c)~(f)是每一张图像都呈现出优秀的视觉效果,显示区域基本准确.同时,显示画面的亮度色温都与实际场景非常接近.

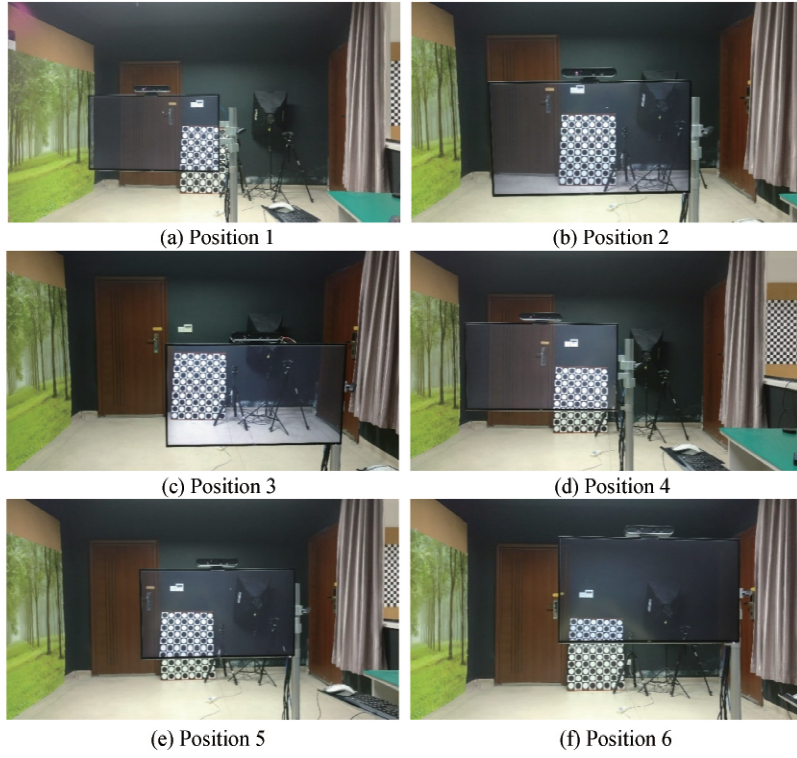


图7 场景融合实验(室内)
Fig.7 Scene fusion experiments (indoor)

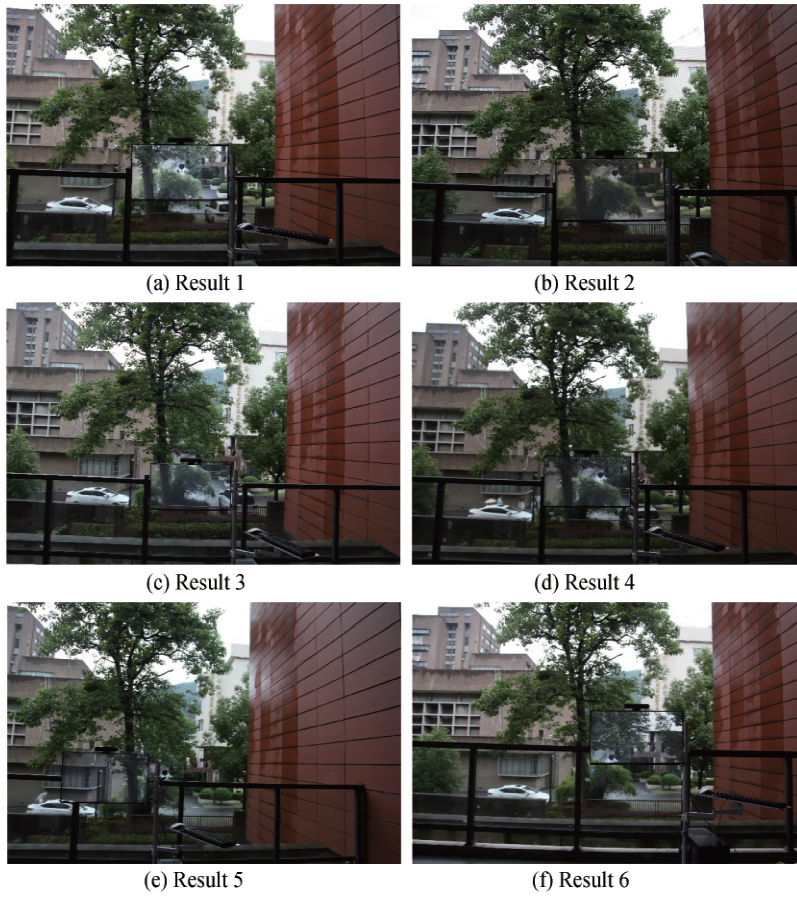


图8 场景融合实验(室外)
Fig.8 Scene fusion experiments (outdoor)

6.5 时间效率

本项目程序使用 C++ 语言编写,基本的编程环境为 Windows10+Visual Studio2015.处理器的 CPU 主频为 3.40 GHz,内存为 8 G.

算法的执行时间如表 2.初始化过程包括相机初始化、检测模型初始化和显示参数控制几部分,总耗时 4 906 ms.初始化过程仅执行一次.循环部分主要负责观察者位置检测和显示区域计算,单次执行耗时 283 ms,基本满足实时变换要求.

表 2 程序执行时间
Table 2 Program execution time

Process	Procedure	Time/ms	Total/ms
Initialize (once)	Initialize the image capture camera	545	4 906
	Initialize the position monitoring camera	828	
	Initialize MTCNN model	3 453	
	Get luminance and color temperature, adjust display	80	
Adaptive scene transformation (loop)	Get one frame of scene image, convert to Mat format	71	283
	Get one frame of observer image, convert to Mat format	135	
	MTCNN human eye detection	35	
	Extract eyes' depth, calculate eyes' 3D coordinates	6	
	Calculate reasonable display area	16	
	Crop image and display	20	

6.6 讨论

我国一年四季的平均色温约在 8 000~9 500 K 之间,纬度越高色温越低,北美地区的平均色温约在 5 000~6 500 K 之间.大多数在售显示器的色温调节范围能够满足要求.然而,自然环境的亮度变化范围跨度极大,地表所见的太阳表面亮度可达 10^{10} cd/m² 量级,晴朗的白天天空亮度约为 3 000 cd/m².普通的 LCD、QLED、PDP 显示器最大亮度均不超过 1 000 cd/m²,这就导致晴朗天气时的场景融合效果不佳,显示画面整体偏暗.室外显示常用 LED 显示器,峰值亮度高,但同时像素点也较大,显示颗粒感明显.因此,新型高亮显示设备的出现可以促进本文提出的场景融合研究.使用高亮度、高动态范围的新型显示设备,在本文研究的基础上适当改进当前方案,即可确保场景融合系统满足大多数光照条件下的应用需求,进一步提升算法的鲁棒性和目视效果.

另一方面,MTCNN 模型能检测到的最小人脸为 12 pixel×12 pixel,位置监测相机的图像分辨率越高,人眼检测精度也就越高.本文使用的 Astra RGBD 相机能够在 60°视场范围内、3 000 mm 深度范围内精确定位人眼.使用更大视场、更大分辨率、更大深度范围的深度相机作为位置监测相机,系统就能够准确识别更大立体空间内的观察者,并为他们营造场景融合效果.

再者,本文提出的场景融合系统为了节约成本、减轻系统重量,弃用成像式亮度计、色温仪等高精度检测仪器,选用 RGBC 颜色传感器,结合经验公式估计环境色温与亮度.如果经费充裕、系统承重足够,不妨考虑使用高精度环境光检测仪器.

需要注意,本文提出的场景融合方案主要探究单目视觉的透视关系,目前仍然没有比较理想的针对双目视觉的解决方案.同时,本文方案唯一的未知参数是物距,通常采用独立手段测量后(如:激光测距仪)作为算法输入使用.如何解决双目视觉的深度感知对场景融合视觉效果的影响,如何添加自动测距模块到场景融合系统,将是下一阶段的研究重点.

7 结论

本文提出了完整的场景融合算法,包括观察者位置监测、坐标变换、环境光源信息传感、亮度色温估计、显示参数映射及控制等步骤.实验结果表明,本文提出的场景融合方案能够为不同位置的观察者调整显示画面,并结合环境信息改变显示参数,获得优良的融合效果.场景融合在实现细节上与增强现实技术接近.本文提出的场景融合方法既能保证再现画面与场景内容的连续性,又能还原场景真实的亮度与色彩表现,在以增强用户沉浸感为目的的商业场合中有很强的应用价值.本文提出的场景融合方案仍有较大的改进空间,由于

显示设备亮度调节范围的限制,本实验系统在晴朗天气下显示画面整体偏暗,融合效果不佳,可以考虑在室外场景下使用高亮度 LED 显示器.另一方面,本文提出的场景融合方案主要探究单目视觉的透视关系,目前仍然没有比较理想的针对双目视觉的解决方案.最后,本方案中的物距通常采用独立手段测量后的结果(如:激光测距仪)作为算法输入使用,并没有实现完全意义上的全自动场景融合.各模块设备更新、双目视觉、自动测距将是下一步的研究重点.

参考文献

- [1] YOSHIDA T, KUROKI S, NII H, *et al.* ARScope[C]. Conference on Computer Graphics, 2008.
- [2] HILL A, SCHIEFER J, WILSON J, *et al.* Virtual transparency: Introducing parallax view into video see-through AR [C]. IEEE International Symposium on Mixed & Augmented Reality, IEEE, 2012.
- [3] TOMIOKA M, IKEDA S, SATO K. Approximated user-perspective rendering in tablet-based augmented reality[C]. IEEE International Symposium on Mixed & Augmented Reality, IEEE, 2013.
- [4] VIOLA P, JONES M J. Robust real-time face detection[J]. *International Journal of Computer Vision*, 2004, **57**(2): 137-154.
- [5] LI Hao-xiang, LIN Zhe, SHEN Xiao-hui, *et al.* A convolutional neural network cascade for face detection[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2015.
- [6] HUANG Li-chao, YANG Yi, DENG Ya-feng, *et al.* DenseBox: Unifying landmark localization with end to end object detection[J]. *Computer Science*, 2015.
- [7] JIANG Huai-zu, LEARNED M E. Face detection with the faster R-CNN [C]. IEEE International Conference on Automatic Face & Gesture Recognition, IEEE, 2017.
- [8] NAJIBI M, SAMANGOUEI P, CHELLAPPA R, *et al.* SSH: Single stage headless face detector[J]. 2017, **1**(1):4885-4894.
- [9] ZHANG Kai-peng, ZHANG Zhan-peng, LI Zhi-feng, *et al.* Joint face detection and alignment using multitask cascaded convolutional networks[J]. *IEEE Signal Processing Letters*, 2016, **23**(10): 1499-1503.
- [10] RANJAN R, PATEL V M, CHELLAPPA R. HyperFace: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, **41**(1): 121-135.
- [11] ZHANG Zheng-you. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330-1334.
- [12] 李晓彤,岑兆丰. 几何光学·像差·光学设计[M]. 浙江:浙江大学出版社,2003.
- [13] 徐海松. 颜色信息工程[M]. 浙江:浙江大学出版社,2005.
- [14] MCCAMY C. Correlated color temperature as an explicit function of chromaticity coordinates[J]. *Color Research & Application*, 1992, **17**(2): 142-144.
- [15] VALENCIA J, GIRALDO F, BONILLA J. Calibration method for correlated color temperature (CCT) measurement using RGB color sensors [C]. IEEE 2013 XVIII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA), IEEE, 2013.