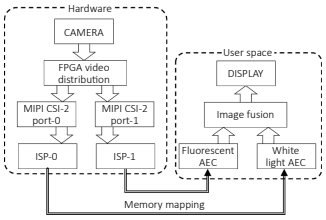
OEE | 光电工程 Opto-Electronic Engineering

Article

2026 年 6 月, 第 53 卷, 第 6 期

虚拟双相机架构下的腔镜实时荧光成像技术



邓奕城, 倪旭翔*, 王立强, 袁　波, 张基强, 陶子龙

浙江大学光电科学与工程学院, 浙江 杭州 310027

摘要： 本文针对荧光引导手术中腔镜复合成像的实时性与融合质量需求，提出一种基于虚拟双相机模式的近红外荧光与白光复合成像技术。通过基于 FPGA 的光源与相机精确同步控制及视频流硬件解复用架构，结合双模态联合自动曝光调节算法，实现异构图像流的实时分离与双模态图像曝光平衡。实验表明，该系统同步控制精度达帧级 (16.67 ms)，自动曝光收敛时间低于 0.8 s，所摄图像结构相似性 (SSIM) 为 0.9994，峰值信噪比 (PSNR) 为 46.23 dB。本方案在保持系统小型化、低成本的基础上，实现了接近双相机模式的成像性能，为术中实时影像导航提供了有效支持。

关键词： 腔镜；复合成像技术；医学图像处理；虚拟双相机；自动曝光控制

The English title and abstract appear at the end of the paper.

DOI: 10.12086/oee.2026.250341 | CSTR: 32245.14.oee.2026.250341 | 中图分类号：TN247; TB133 | 文献标识码：A

邓奕城，倪旭翔，王立强，等. 虚拟双相机架构下的腔镜实时荧光成像技术 [J]. 光电工程，2026，**53**(6): 250341

Deng Y C, Ni X X, Wang L Q, et al. Real-time fluorescence imaging technology for endoscopes under the virtual dual-camera architecture[J]. *Opto-Electron Eng*, 2026, **53**(6): 250341

1　引　言

荧光引导手术作为现代外科领域的重要技术革新，已在多个临床学科得到广泛应用。在神经外科领域，该技术用于胶质瘤、脑膜瘤等中枢神经系统肿瘤的术中定位与精确切除[1-2]。在头颈外科领域，内镜荧光引导手术为鼻腔鼻窦恶性肿瘤切除提供实时可视化辅助[3]。在肝胆外科领域，荧光引导技术应用于胆道解剖识别、胆漏检测、肝段划分及肿瘤定位，在腹腔镜手术中实时评估解剖结构和组织功能状态，并为吻合口血流状态评估提供实时信息[4-6]。在妇科肿瘤外科领域，该技术为卵巢癌和乳腺癌等恶性肿瘤的术中病灶识别提供了新的技术手段[7]。

该技术的原理是，通过向患者体内注射如吲哚菁绿 (indocyanine green, ICG) 等荧光造影剂，并在特定波长的激发光 (如 780 nm 的近红外光) 照射下，病灶或目标组织会发出可被特制相机捕捉的荧光信号，标识出病灶的轮廓和范围[8]。由于荧光图像缺乏色彩与细致解剖结构信息，临床应用中常将荧光图像与白光图像联合使用：白光提供解剖背景、组织质地和手术器械的直观视野，帮助外科医生定位边界和判断操作方向；荧光则补充病灶识别、血流灌注等信息，经过图像处理后高亮地融合显示在白光图像上[9]。

实现上述功能通常采用两种方法。1) 双相机方案，即在摄像系统中集成两路独立的光学成像单元，分别用于捕捉白光和荧光信号[9-11]。其优势在于两路信号并行采集，互不干扰，能够针对各自的光学特性进行独立优化：劣势是双相机及配套的光学组件 (分光棱镜，CMOS

收稿日期：2025-11-08；修回日期：2026-02-12；录用日期：2026-02-28；网络出版日期：2026-06-25

基金项目：北京市自然科学基金资助项目 (L252001)；国家重点研发计划 (2023YFF0720402)

*通信作者：倪旭翔，xuxiang@zju.edu.cn

版权所有©2026 中国科学院光电技术研究所

芯片) 使得摄像部件尺寸增大、结构复杂，增加了实现难度和制造成本，且两路成像光路存在物理上的装配偏差，导致输出的图像需要依赖软件算法进行配准。2) 单相机方案[12-13]。利用单一图像传感器，通过高速切换白光和激发光两种照明光源，分时交替采集白光帧和荧光帧。所有图像均由同一传感器在同一光路下采集，该方案从物理层面根除了空间配准问题，简化了摄像部件的光学与机械结构，更易于实现小型化、轻量化和低成本化。

然而，现有单相机方案存在以下问题：

1) 异构图像流的实时分离与处理：在分时模式下，单一传感器输出的是一个由白光帧和荧光帧交错组成的混合视频流。后端处理器首先需将两种图像数据实时准确地分发到不同的处理流水线中。传统的软件处理方式难以应对高带宽 (如 4 K 60 f/s 或更高) 视频流的分离任务，容易造成数据拥塞或处理延迟，无法满足手术对影像实时性的要求。

2) 曝光控制难题：白光图像与荧光图像的信号强度通常存在数个数量级的差异。荧光信号极其微弱，需要高增益和长曝光；而白光反而需要抑制过曝。在同一 CMOS 传感器上成像，所使用的曝光时长、模拟增益难以同时满足要求。因此设计一种智能的自动曝光控制算法，以在两种图像间取得最佳平衡，是提升单相机方案图像质量的关键。

因此，本文提出了一种基于虚拟双相机模式的荧光与白光分时成像全新方法，实现了一个集成高精度时序控制、高效数据分发以及智能化曝光策略的单相机腔镜复合成像系统，在保证系统低成本、小型化优势的同时，实现媲美传统双相机方案的实时成像性能与融合图像质量。

2 系统架构与实现

2.1 荧光腔镜摄像系统

本文提出的基于虚拟双相机模式的近红外荧光与白光复合成像系统，其整体架构如图 1(a) 所示。系统由双 LED 光源、FPGA 时序控制模块、CMOS 相机和嵌入式主机组成。FPGA 作为核心控制单元，负责生成精确的时序信号，控制近红外 LED 和白光 LED 的频闪时序，使其与 CMOS 相机的曝光窗口严格同步。两种 LED 光源在时间维度上交替发光，经由光学系统照射至目标组织，CMOS 相机同步采集对应的荧光图像和反射图像。FPGA 进一步将相机输出的混合视频流分发至两个独立的 ISP 端口，实现硬件级别的图像流分离，最终由主机接收并处理两路独立的视频信号。该系统通过时分复用照明和视频流硬件分发，实现了单传感器虚拟双相机的功能。

光源包括白光 LED 和近红外 LED 两个独立通道。白光 LED 的最大输出功率为 65 W，可提供覆盖可见光波段的连续光谱照明，满足组织表面形态和血管结构成像的需求。近红外荧光激发 LED 的中心波长为 780 nm，光谱半峰全宽 (FWHM) 为 30 nm，在 13.5 A 驱动电流条件下可输出 7.0 W 至 12.25 W 的辐射功率，能够充分激发 ICG 等近红外荧光造影剂产生荧光信号。780 nm 波

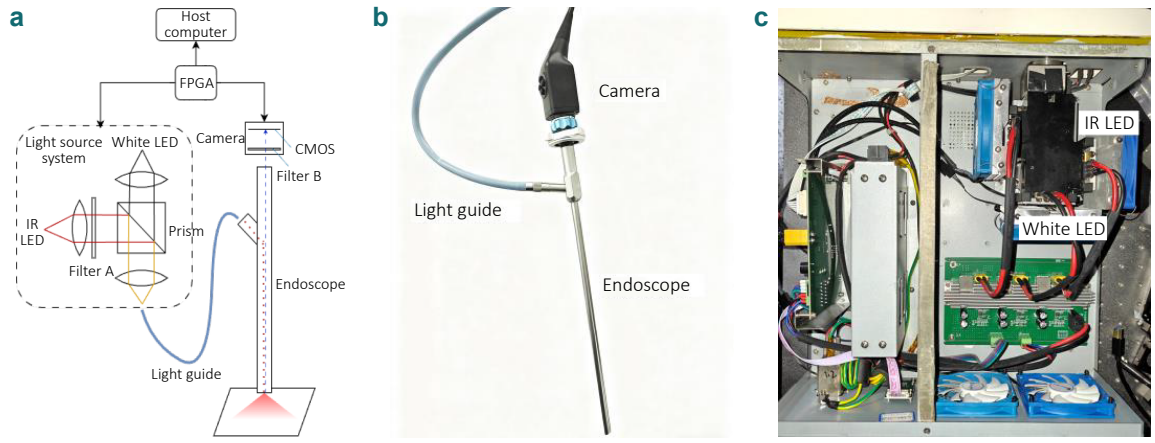


图 1 系统架构以及实物图。(a) 系统架构图；(b) 内窥镜体实物图；(c) 双 LED 光源实物图

Fig. 1 System architecture and physical diagrams. (a) System architecture diagram; (b) Physical diagram of the endoscope body; (c) Physical diagram of the dual-LED light source

长的选择与 ICG 荧光试剂的吸收峰相匹配，30 nm 的半峰全宽在保证激发效率的同时有效降低了光谱串扰。

相机采用 SONY IMX334 CMOS 图像传感器，分辨率为 3840 pixel×2160 pixel，单个像素尺寸为 2.0 μm×2.0 μm，默认以 60 f/s 的速率工作。该型号 CMOS 传感器在可见光区域 (400 nm 至 700 nm) 和 780 nm 以上的近红外荧光波段均保持良好的光电转换效率，能够为白光反射图像和近红外荧光图像提供充足的信噪比支持。CMOS 前装有一双带通滤光片，该滤光片的光谱特性与近红外 LED 的激发光谱对应，在 780 nm 激发光波长处截止深度达到 OD4 以上，有效阻隔激发光的反射干扰，同时在 LED 激发光谱之外的 820 nm 波段保持高透过率，确保 ICG 荧光信号有效传输。

FPGA 选用型号为 Elitsetek Ti60F225 芯片。图 2(a) 展示了 FPGA 硬件系统的电路原理图，包括主控芯片、电源转换模块、配置存储器及时钟电路等核心组成部分。图 2(b) 展示了 FPGA 硬件系统的实物照片，包括中心区域的 FPGA 核心模块、左侧的 MIPI 视频输入接口以及右侧的嵌入式主机核心板，各功能模块通过 PCB 板载走线连接，共同构成虚拟双相机模式的硬件处理平台。

嵌入式主机作为下游处理单元，负责接收及处理视频数据，实现了逻辑更加复杂的自定义算法，包括后续的自动增益控制、图像去噪、图像融合，以及显示等高级功能。

2.2 虚拟双相机系统

为解决单相机荧光腔镜系统光源同步控制、图像流实时分离的难题，本节提出并实现了一种“虚拟双相机”系统，在单一 CMOS 传感器与分时照明的硬件基础上，通过 FPGA 在硬件层面完成光源与相机的精确同步控制以及 MIPI 协议层的视频流解复用，从而在嵌入式主机中虚拟出两路并行的相机数据通道。这一设计使得单相机系统能够模拟双相机的功能，在保证图像质量的同时，显著降低了系统的复杂性和成本。

2.2.1 光源相机同步成像控制

采用分时成像策略在单个 CMOS 传感器上获取两种模式的图像，系统启动时，所有奇数帧采集白光图像，所有偶数帧采集荧光图像，以 60 f/s 的速率交替工作，实现了两种图像的分离。白光和荧光图像的先后次序不影响系统的功能，为便于说明，后文的奇数帧都代指白光图像。

实现这一策略的关键在于 FPGA 的精确同步控制设计。当相机被配置为从模式的时候，其曝光行为完全受控于外部输入的同步信号。FPGA 的内部逻辑负责生成驱动图像传感器的帧同步信号 (vertical sync, VS)，标志着当前帧曝光的结束和下一帧的开始。图 3 展示了系统光源与相机的同步控制时序。图中从上至下依次为相机曝光窗口、红外 LED 驱动信号、白光 LED 驱动信号和 VS 信号。VS 信号的下降沿作为帧切换的基准时刻，FPGA 在输出该下降沿时同步触发 LED 驱动信号的电平翻转。具体而言，当 VS 信号从高电平跳变至低电平时，若当前为奇数帧结束，则白光 LED 驱动信号由高电平切换至低电平 (关闭白光)，同时红外 LED 驱动信号由低电平切换至高电平 (开启红外光)，随后传感器在新的 VS 周期内执行曝光操作。反之，当偶数帧切换至奇数帧时，光源控制信号执行逆向切换。这种硬同步机制保

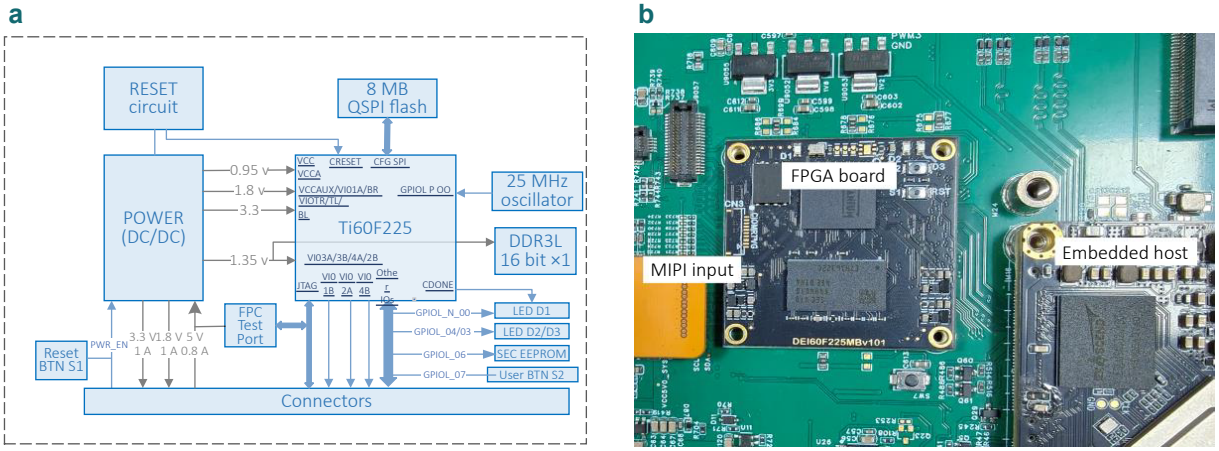


图 2 FPGA 原理与实物示意图。(a) FPGA 系统原理图；(b) FPGA 板卡实物图

Fig. 2 Schematic and physical illustration of FPGA. (a) Schematic diagram of the FPGA system; (b) Physical diagram of the FPGA board

证了每一帧的曝光都对应唯一且正确的照明光源，从根本上避免了成像端的时序错位。

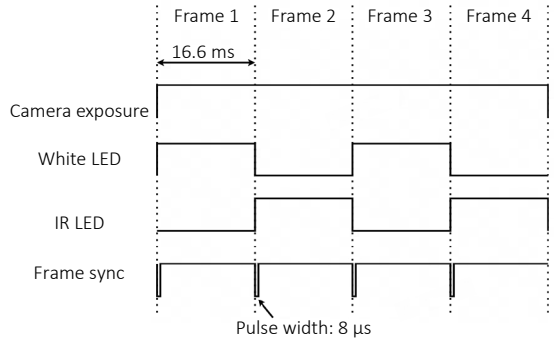


图 3 相机光源协同时序

Fig. 3 Camera-light source synchronization timing

2.2.2 视频信号实时解复用与分发系统

在单相机分时成像架构下，图像传感器输出的是由白光帧与荧光帧在时域上交错排列的混合视频流。如何将这一复合数据流高效、准确地分离并路由至不同处理通道，直接决定了系统的实时性与可靠性。传统的视频帧拆分方案通常采用软件处理方式，即通过主机 CPU 在应用层对视频流进行奇偶帧分离。然而，由于缺乏硬件级别的同步机制，在处理高带宽的 4K 60 f/s 视频流极易产生帧丢失和时序失配的问题。

图 4 展示了嵌入式主机的 MIPI CSI 视频传输链路，主要的性能瓶颈在于 4K 60 f/s 视频流数据的解复用步骤，本文将这一任务逻辑下沉到 MIPI 协议层，在硬件层面实现了异构视频的分离，即用 FPGA 处理器接收相机直出的视频流，再根据当前的光源状态发送当前帧的数据到不同的主机端口。这种设计实现了视频信号曝光和分离处理的硬同步机制，解决了用户空间的性能瓶颈。两路数据流经独立的 DMA 通道和缓冲区，避免交叉干扰，即使某一帧在传输层丢失或者阻塞，硬件分离逻辑仍基于真实帧序工作，上位机也总是能从不同的端口位置接收到正确的图像。

这种方案同时实现了 ISP 处理的解耦。ISP 的图像处理操作包括解马赛克、白平衡、色彩校正、色调映射、伽马校正、去噪等[14]。传统视频分离方式将异构视频流输入到了同一个 ISP，使用相同的参数进行处理，这种做法显然是不恰当的。具体来说，白光图像和荧光图像存在明显的特征差异，白光图像全光谱响应(400~700 nm)，色彩信息丰富，动态范围高，需要精确的色彩还原和光照校正。而荧光信号更加微弱，信噪比较低，需要精细的暗电流补偿以及更高的降噪力度。

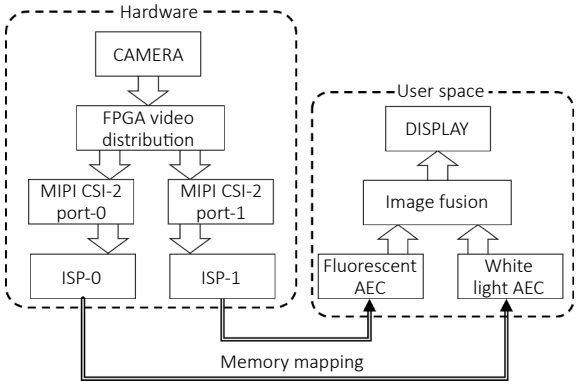


图 4 基于 FPGA 的硬件级视频流解复用方案

Fig. 4 FPGA-based hardware-level video stream demultiplexing solution

由于一个 ISP 在硬件链路上只能和一个 MIPI lvds 链路进行绑定，经过 FPGA 分发后，自然形成了两个 ISP 对同一视频流分别进行独立处理，分别实施不同的定制化图像处理参数。在本文方案中，按照 ISP 处理流程的先后顺序，主要对黑电平校正、Gamma 校正、时域降噪和边缘增强这四个算法进行了分模态的差异化参数配置。

黑电平校正通过获取传感器暗场输出基准，将有效像素值减去该暗电平以消除 CMOS 暗电流引起的灰雾状背景干扰。为防止系统将暗电流误识别为荧光信号导致假阳性结果，荧光通道采用了更严格的校正参数。

Gamma 校正通过施加非线性变换将传感器的线性光电响应映射至人眼的非线性感知特性，提升暗部亮度分辨率。近红外通道采用更高的 Gamma 值以增强荧光信号与背景的对比度，白光通道则采用标准值以保持视觉自然度。

时域降噪利用视频流的时间相关性融合多帧信息，在保持细节的同时显著提升信噪比。该算法的有效性依赖于 FPGA 视频分发架构所确保的帧间连续性：若不分离交替的双模态视频流，时域相关性将被破坏，无法正确实施降噪。

边缘增强通过高通滤波提取边缘细节并叠加至原图，本文设计中仅在白光通道的亮度分量上启用，以提升组织表面纹理和血管轮廓的清晰度。荧光通道关闭该算法以避免引入噪声和伪影。

上述关键算法的差异化配置充分体现了基于 FPGA 的视频分发架构的核心优势。虚拟双相机架构通过硬件

级的流分离，为两个模态提供独立的 ISP 处理通道，使得每个模态均能针对自身成像特性实施最优的算法参数配置，最终实现高质量的复合成像。

2.3 适用于虚拟双相机的自动调光算法

虚拟双相机模式的自动曝光控制 (auto exposure control, AEC) 需根据白光与近红外光照条件动态调整 CMOS 曝光参数，以获得亮度适中、层次丰富的双模态图像[15]。然而，两种模态的图像信号强度存在显著差异，给 AEC 算法的实现带来挑战。图 5 展示了常规单相机一次分时成像得到的白光/荧光图像对和灰度直方图对比，白光图像的 luma 均值为 126.07，而荧光图像仅为 3.92。因两种图像是从一块 CMOS 上依序采集的，其曝光时间和模拟增益全局共享。由此带来的问题是：若为提高荧光图像强度增加增益，将导致白光图像过曝；反之，为抑制白光过曝而降低增益，则会使荧光图像太弱。

针对这一矛盾，本文提出一种面向虚拟双相机的自动曝光控制算法。该算法以双路图像亮度统计值为反馈，通过联合调控白光光源强度、红外激发光强度、全局模拟增益及荧光图像数字增益，使双路图像亮度分别收敛至各自目标容差区间。值得注意的是，为避免破坏分时成像的硬同步时序，相机曝光时长保持固定。自动调光算法的流程图如图 6(a, b) 所示。算法分为两个阶段，阶段 1 只对白光帧进行收敛计算，阶段 2 只对荧光帧进行收敛计算。两阶段之间通过一帧时序延迟解耦，确保阶段 1 对模拟增益的调整能够在阶段 2 开始前生效并作用于荧光帧。

在阶段 1，首先获取白光 luma 统计值 L_{w}，与目标 luma 值 L_{wt} 进行做差得到差值 $diff$。如果绝对差值在容差值内，则进入阶段 2 的计算，否则进入调光流程。系统对白光模态的输出控制量分别是白光 LED 的强度 i_{w} 和 CMOS 上的模拟增益 g_{a}。调节流程为当前 luma 低于目标值时，在不超过最大 g_{a} 上限的前提下优先提高 g_{a}，其次提高 i_{w}；当前 luma 高于目标值时，在不低于最小 i_{w} 下限的前提下优先降低 i_{w}，以维持较高 g_{a}。由于 g_{a} 是跨模态的共享参数，而荧光在绝大多数情况下是欠曝的一方，这一优先级策略能保证在避免白光过曝的前提下获得更清晰的荧光图像。g_{a}、i_{w} 的调节步长分别由 Δg 、Δi_{w} 表示，为实现快速收敛且避免超调振荡，采用基于误差比例的自适应步长策略：

$$\Delta i_{\mathrm{w}} = round\left(|diff| \times r_1\right), \tag{1}$$

$$\Delta g_{\mathrm{a}} = |diff| \times r_2 . \tag{2}$$

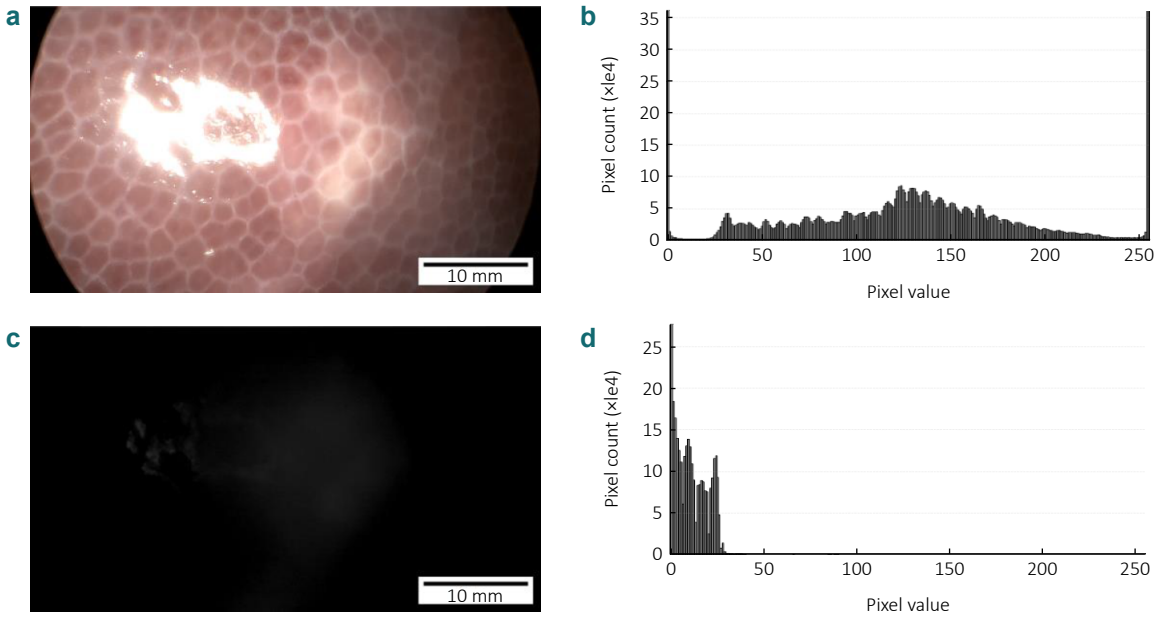


图 5 一次分时成像中，所得到的白光/荧光图像对。成像参数曝光时间 为 0.016 s，增益为 2.0×。(a) 白光照明图像；(b) 白光图像灰度直方图可视化分析结果；(c) 荧光照明图像；(d) 荧光图像直方图可视化分析结果

Fig. 5 White light/fluorescence image pairs obtained in the time-division imaging. Parameters: *expTime* = 0.016 s, *gain* = 2.0×. (a) White light illumination image; (b) Visualization analysis result of white light image grayscale histogram; (c) Fluorescence illumination image; (d) Visualization analysis result of fluorescence image histogram

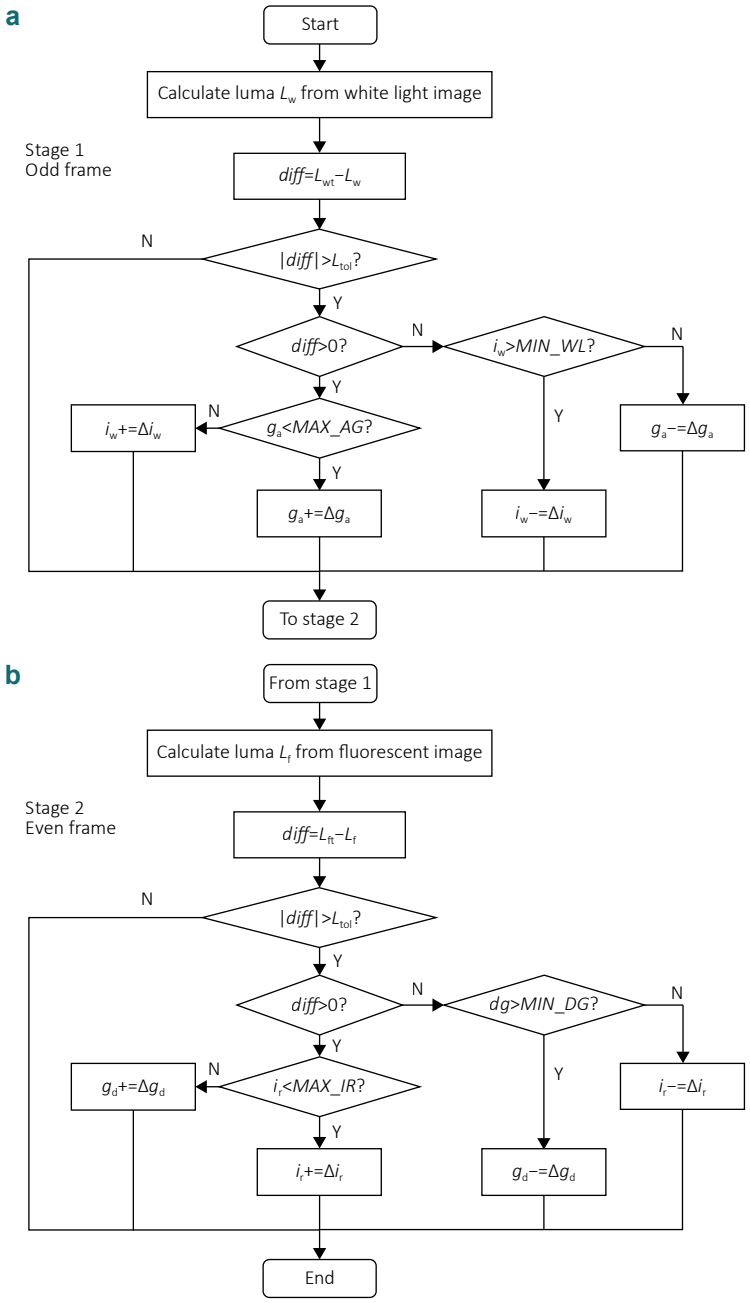


图 6 AEC 算法流程图。(a) 奇数帧 (白光图像)；(b) 偶数帧 (荧光图像)

Fig. 6 Flowchart of the AEC algorithm. (a) Odd frame (white light image); (b) Even frame (fluorescent image)

步长由绝对差值和缩放因子的乘积所得，缩放因子 r_1 和 r_2 为预设宏参数，用于平衡收敛速度与稳定性。由于光源只支持整数值作为调节的接口，所以需要对 Δi_w 的计算结果取整输出。

式 (1) 和 (2) 所得到的步长，还需要满足 *MAX_AG*，*MIN_AG*，*MAX_WL* 和 *MIN_WL* 所规定的上下界。*MIN_WL* 设置为光源允许输出的最小白光强度，*MAX_AG* 为在最小白光强度下，白光图像刚好不过曝所对应的 g_a 值，本方案设置为 8。*MIN_AG* 在本方案中设置为 4，能够给荧光图像一定程度的预增强，*MAX_WL* 则设置为此时白光图像刚好不过曝所对应的白光强度值。

阶段 2 是在阶段 1 所设定的 g_a 参数基础上，通过独立调节数字增益 g_d 与红外光源强度 i_r，进一步提升荧光图像强度。调节优先级设定为当前 luma 低于目标值时，优先提高 i_r，其次提高 g_d；当前 luma 高于目标值时，优先降低 g_d，其次降低 i_r。相比于模拟增益 g_a，数字增益

g_d 的噪声效应更为明显，该策略旨在最大限度抑制数字增益引入的额外噪声。

g_d、i_r 的步长由类似式 (1) 和 (2) 的公式取得，步长还需满足 *MAX_DG*，*MIN_DG*，*MAX_IR* 和 *MIN_IR* 所规定的上下界，*MAX_IR* 和 *MIN_IR* 被设置为红外光源输出的上下限，*MIN_DG* 设置为 1，考虑到最终的信噪比表现，本方案 MAX_DG 设置为 8。经过上述 AEC 算法调节后，系统能够有效且灵敏地在不同环境下输出清晰且一致的图像。图 7 展示了一例经自动调光后，将荧光区域融合到白光图像的效果，所采用的融合方法为像素级的加权融合。

3 实验设计与验证

为全面验证虚拟双相机成像方案的效果，我们从调光算法性能和图像质量表现两个维度设计了一套完整的实验方法。

调光算法重点考察其在实际应用中的实时响应能力和稳定性表现。采用收敛速度作为实时性指标，量化调光算法从初始状态到达稳态所需的帧数。同时，将本文提出的调光算法与主流的常规 AEC 算法进行对比实验，考察收敛过程中的稳定性。

在图像质量表现方面，选取 SSIM[16] 和 PSNR 作为评价指标。其中 SSIM 从亮度、对比度和结构量化融合图像与参考图像的相似程度，其计算公式如下：

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}, \tag{3}$$

式中：σ^2是图像的方差；σ_{xy}是两张图像的协方差；μ是图像的均值。计算中，C_1取 6.5，C_2取 58.5。

PSNR 基于均方误差 (mean square error，MSE) 进行定义，通过量化原始图像与处理后图像之间的像素级差异，评估图像的失真程度和信号保真度。假设有尺寸为 $m\times n$的参考图像 I 和待测图像 K，其 PSNR 计算公式如式 (4) 所示，*MAX* 为图 I 中像素的最大值。

$$\begin{aligned} MSE&=\frac{1}{mn}\sum_{i=1}^{m}\sum_{i=1}^{n}(I(i,j)-K(i,j))^2\\ PSNR&=10\log_{10}\left(\frac{MAX^2}{MSE}\right) \end{aligned}. \tag{4}$$

3.1 自动调光算法评价实验

为评估自动曝光控制算法的实时性能，本实验设计以曝光收敛时间为核心指标，即在极端初始条件下，系统亮度重新达到稳态所需的帧数。

在本方案的双模态自动曝光控制系统中，图像收敛是指系统通过 AEC 算法的迭代调节，使图像亮度逐渐逼近并稳定在目标亮度区间的动态过程。图像收敛的定量判定标准基于图像 luma 值与 AEC 算法目标区间的位置关系。AEC 算法为图像亮度预设了容忍区间，当图像的 luma 值进入该区间且算法不再输出增益调节量或光源功率控制量时，即认为图像已达到收敛状态。此时，图像的亮度参数不再随时间变化，系统进入稳态工作模式。

实验在 40 mm 工作距离下进行，以注射 0.1 mg/mL 浓度 ICG 荧光试剂的猪肝为拍摄对象。首先使 AEC 算法自然收敛至稳定状态，随后人为设定 CMOS 模拟增益与双 LED 光源强度至上下限，以模拟极端工况，并记录算法重新收敛的过程与时间。算法收敛过程分为两个阶段：第一阶段调节模拟增益使白光图像亮度收敛；第二阶段在模拟增益稳定的基础上，联合调节红外光源强度与数字增益，实现荧光图像亮度收敛。

实验设置 A、B 两组对照：A 组初始条件为模拟增益 32 倍、双 LEDl 亮度 100；B 组初始条件为模拟增益 1 倍、双 LED 亮度 0。收敛时间以帧数为单位，在 60 f/s

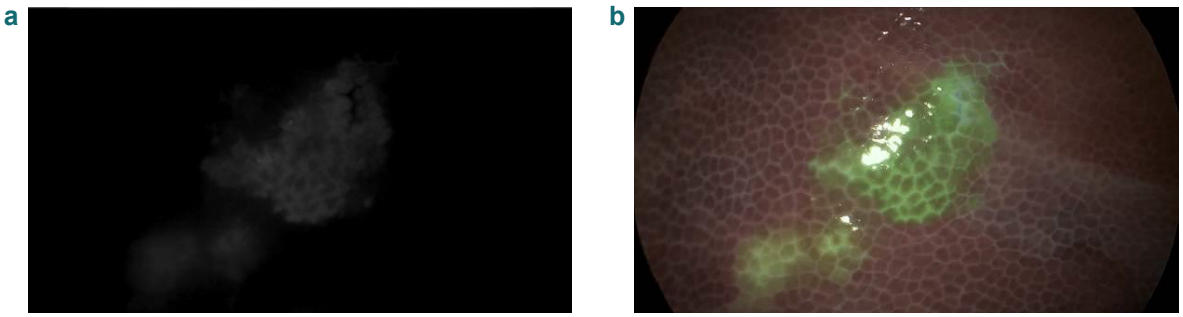


图 7 AEC 算法效果。(a) 荧光帧图像；(b) 融合帧图像

Fig. 7 Effect of the AEC algorithm. (a) Fluorescent frame image; (b) Fused frame image

系统中每帧对应 16 ms。白光收敛时间表现为 16 ms 的奇数倍，荧光收敛时间则为偶数倍。

表 1 说明 B 组的收敛时间通常大于 A 组，这是由于实验中面临的问题多是图像欠曝，稳态环境下光源强度和增益都需要从低值向高值增加更多步长，因此 B 组的调节幅度更长。另一方面，荧光图像的收敛时间长于白光图像，这是由于荧光需要模拟增益稳定后才能开始调整。实验数据表明，白光图像平均需要 23 帧收敛，荧光图像还需要再增加 15~20 帧达到收敛，最终 AEC 收敛时间不超过 0.8 s。

表 1　1~5 组实验算法收敛至目标亮度所需时间 (单位：秒)
Table 1　Convergence timetable of experiments 1–5 (Unit: s)

No.	White light image		Fluorescent image	
	Group A	Group B	Group A	Group B
1	0.304	0.432	0.544	0.704
2	0.240	0.368	0.512	0.672
3	0.400	0.528	0.640	0.768
4	0.272	0.336	0.544	0.608
5	0.464	0.304	0.672	0.640

为进一步直观展示算法的收敛过程，图 8 选取了 A 组初始条件下的代表性时间点图像。初始状态下，白光图像图 8(a) 呈现明显过曝特征，组织区域亮度饱和且伴随较多噪声；荧光图像图 8(b) 目标区域亮度饱和，同时背景噪声水平显著升高，非目标区域呈现假阳性信号。经 AEC 算法调控收敛后，白光图像图 8(c) 过曝得到有效抑制，噪声显著降低；荧光图像图 8(d) 背景干扰被充分抑制，目标信噪比明显改善。两种模态图像亮度均稳定于最优动态范围，成像质量显著提升。折线图 9(a) 定量地展示了收敛过程中 luma 值随时间的变化趋势。

为验证本文提出的分阶段双模态自动曝光控制算法的性能优势，本节采用 ISP 自带的常规 AEC 算法作为对照对象，在相同实验条件下进行对比测试。实验条件具体设置与前文所述一致。

图 9(b) 展示了常规 AEC 算法的亮度收敛曲线。由图 9(a, b) 的对比可知，两种算法收敛速度相近，但控制精度存在显著差异。本文算法的白光亮度稳态值稳定收敛至目标值 50 luma 容差范围内，而常规 AEC 算法的白光亮度最终稳定在 85 luma 上下，存在明显过曝现象。在荧光图像方面，本文算法的荧光亮度曲线呈现平稳收敛趋势；而相比之下，常规 AEC 算法的亮度曲线在 0.25 s 处出现明显欠冲，随后经历振荡才缓慢回升至目标区间。表现在图像上是经历了一次明暗的闪烁，影响临床使用的观感。

上述实验差异揭示了两种算法在控制策略上的本质

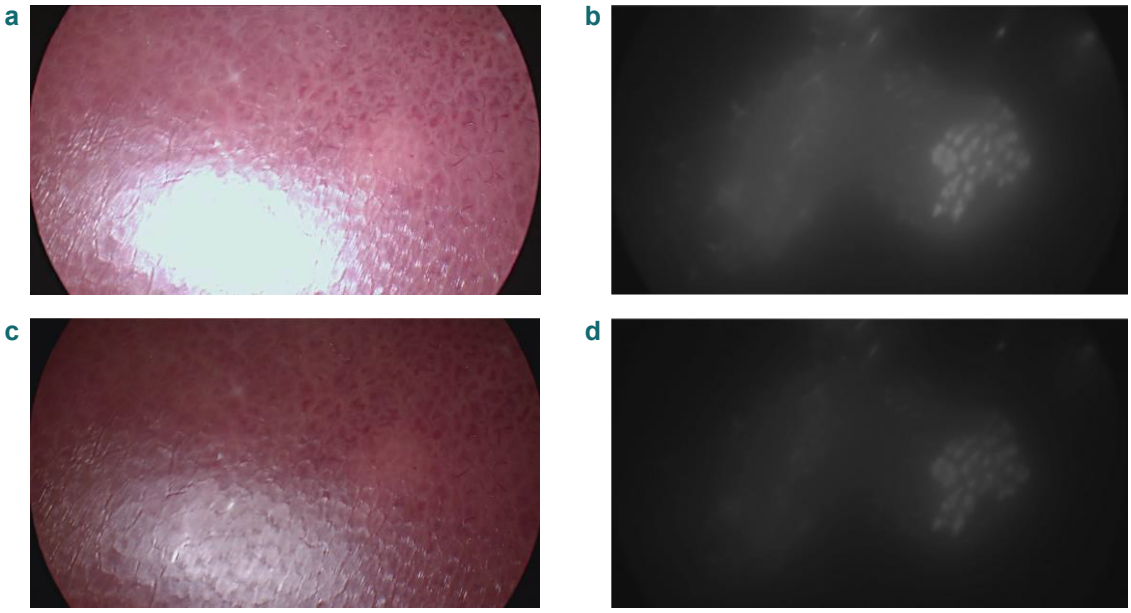


图 8　AEC 算法调节过程中代表性时间点图像。(a) 初始状态白光图像；(b) 初始状态荧光图像；(c) 白光达到收敛图像；(d) 荧光达到收敛图像

Fig. 8　Images at representative time points during AEC algorithm adjustment. (a) White-light image at initial state; (b) Fluorescence image at initial state; (c) White-light image when convergence is achieved; (d) Fluorescence image when convergence is achieved

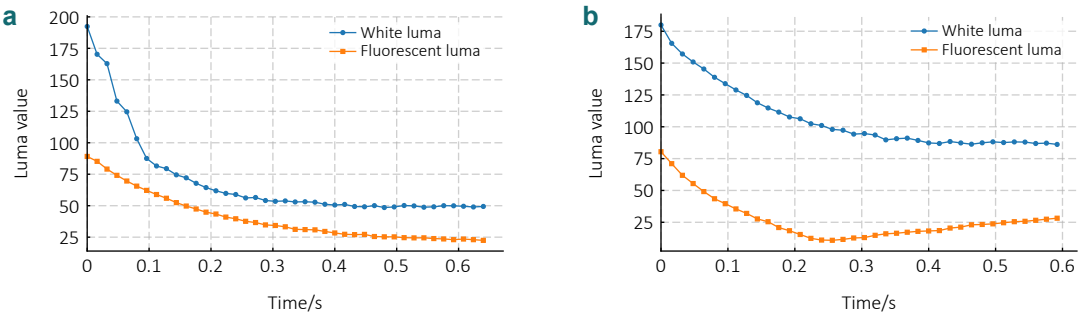


图 9 不同 AEC 算法调节过程中，图像 luma 随时间变化曲线。(a) 本研究提出的算法；(b) 常规 AEC 算法

Fig. 9 Curves of image luma over time during adjustment with different AEC algorithms. (a) The algorithm proposed in this study; (b) The conventional AEC algorithm

区别。常规 AEC 算法未实现白光与荧光模态的控制解耦，在每一帧的调节过程中，白光 AEC 和荧光 AEC 分别根据各自当前亮度与目标亮度的差值计算调节量，但这两个控制量最终同时作用到荧光模态图像上，导致控制量叠加，总调节量超过理论设计值，进而产生控制冲突，表现为欠冲后振荡收敛的不稳定现象。而本文算法通过分阶段控制策略，确保同一时刻只有一个算法输出控制量，实现了荧光图像的稳定收敛。

此外，常规 AEC 算法仅能输出相机端曝光参数，而无法控制光源亮度。当相机增益降至最小值时，若光源功率仍然过高，图像亮度将无法继续下降，这是白光稳态过曝现象的直接原因。本文算法集成的光源控制接口能够在增益调节基础上进一步调整 LED 驱动电流，通过协同控制确保图像精确收敛至目标区间，具备了更宽的动态调节范围。

3.2 图像质量评价实验

为客观评价图像质量，选取 SSIM 与 PSNR 作为评价指标。设定 10 mm、25 mm、40 mm 三个工作距离，覆盖内窥手术中的典型操作范围，验证系统在典型工作距离范围内的性能一致性。参考图像为在稳定单白光或单红外光源照明下采集的图像，测试图像则由本文提出的虚拟双相机方法在同一硬件平台上采集。这一方法能有效隔离硬件链路噪声与光学系统差异，模拟双相机系统的理想成像效果作为基准。在参考图像采集中，AEC 控制由 ISP 硬件模块完成；而在虚拟双相机模式下，则启用本文 2.3 节所提出的 AEC 算法。两种算法的目标亮度及容忍范围参数设置完全相同，并待其充分收敛后进行图像采集。SSIM 与 PSNR 的计算最终通过 Python 的 scikit-image 库实现，实验结果如表 2 所示。

从结果分析，随着工作距离的增加，测试图像相较于参考图像的失真程度总体呈上升趋势。白光图像在各工作距离下均能较好地保持与参考图像的一致性；而荧光图像的 SSIM 和 PSNR 指标则随距离增加出现显著衰减。这是由于荧光需要额外的数字增益以维持相同的亮度水平，例如，在单独成像的条件下，荧光收敛至目标亮度需要 64 倍的模拟增益，而分时成像的条件下，为了防止白光过曝，模拟增益上限被设定为 8 倍，另外还需要 8 倍的数字增益进行补偿，相较于模拟增益，数字增益会引入额外的量化噪声，从而导致 SSIM 和 PSNR 指标的下降。

表 2 不同工作距离下 SSIM 和 PSNR 的表现

Table 2 SSIM and PSNR under different working distances

Distance/mm	White light image		Fluorescent image	
	SSIM	PSNR/dB	SSIM	PSNR/dB
10	0.9994	46.23	0.9893	42.25
25	0.9932	45.15	0.9645	38.31
40	0.9836	40.80	0.9224	31.46

图 10 展示了三个工作距离下的成像融合结果。在 10 mm 距离下，双模态成像质量均良好，组织纹理清晰，荧光信号明亮。在 25 mm 距离下，双模态图像仍保持稳定的成像效果，亮度适中，目标区域易于识别。在 40 mm 距离下，白光图像和荧光图像的亮度略有下降，但 AEC 算法通过自动调节增益和光源功率，确保图像亮度仍处于目标区间内，成像质量满足临床诊断要求。

4 结 论

本文针对荧光引导手术中对实时、高质量医学影像的需求，提出了一种基于虚拟双相机的荧光与白光分时

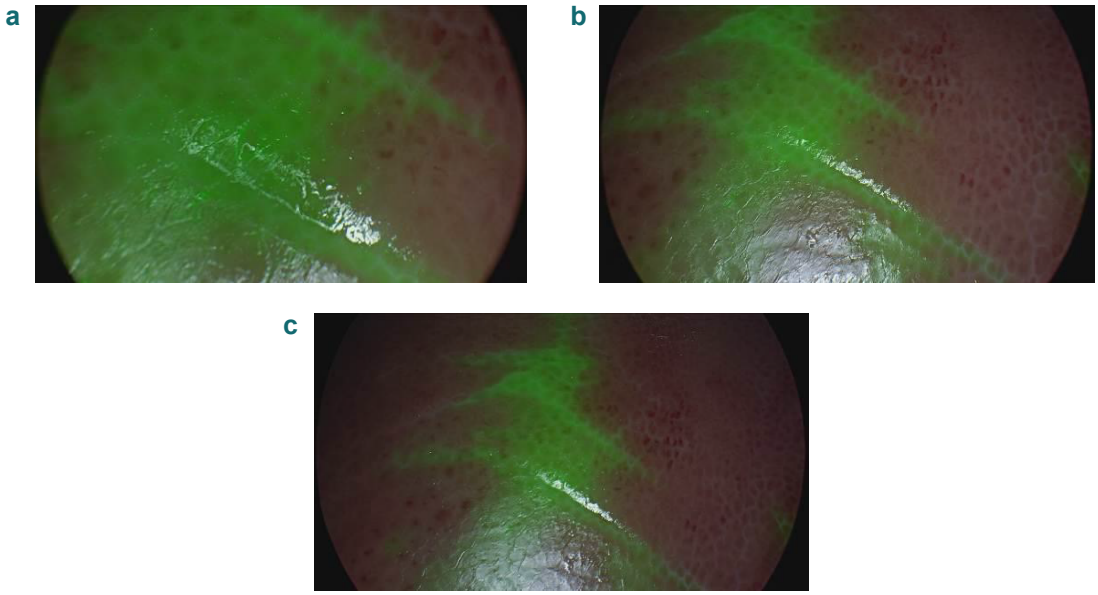


图 10 不同工作距离下荧光内窥镜融合成像结果。(a-c) 分别为 10 mm、25 mm、40 mm 工作距离图像

Fig. 10 Fluorescence endoscopic fusion imaging results at different working distances. Images at working distances (a) 10 mm, (b) 25 mm, and (c) 40 mm

成像、处理与融合的创新方法。采用硬件级同步控制实现了光源切换与传感器曝光的精确耦合；设计了视频流硬件解复用架构，在 FPGA 层面实现了异构视频流的实时分离与路由，避免了高帧率下的软件处理瓶颈；提出了适用于双模态图像的 AEC 算法，通过联合调控光源强度和增益参数，实现了白光与荧光图像的独立优化与快速收敛 (0.8 s 内)。实验表明，该方案在保持单相机系统小型化、低成本、免配准等优势的同时，白光图像 SSIM 保持在 0.98 以上，荧光图像 SSIM 最高可达 0.9893，展现出优异的成像性能。

作者贡献

邓奕城：研究资料的收集、分析，论文写作；倪旭翔，王立强，袁波：指导论文选题、论文修改；张基强，陶子龙：硬件链路及 FPGA 功能设计。

利益冲突

所有作者声明无利益冲突

参考文献

1. Rodriguez B, Brown C S, Colan J A, et al. Fluorescence-guided surgery for Gliomas: past, present, and future[J]. *Cancers*, 2025, **17**(11): 1837.
2. Cho S S, Salinas R, Lee J Y K. Indocyanine-green for fluorescence-guided surgery of brain tumors: evidence, techniques, and practical experience[J]. *Front Surg*, 2019, **6**: 11.
3. Hart Z P, Nishio N, Krishnan G, et al. Endoscopic fluorescence-guided surgery for sinonasal cancer using an antibody-dye conjugate[J]. *Laryngoscope*, 2020, **130**(12): 2811−2817.
4. Lwin T M, Hoffman R M, Bouvet M. Fluorescence-guided hepatobiliary surgery with long and short wavelength fluorophores[J]. *Hepatobiliary Surg Nutr*, 2020, **9**(5): 61539−61639.
5. Verbeek F P R, van der Vorst J R, Schaafsma B E, et al. Image-guided hepatopancreatobiliary surgery using near-infrared fluorescent light[J]. *J Hepato-Biliary-Pancreatic Sci*, 2012, **19**(6): 626−637.
6. Fujimoto H, Kinoshita M, Ahn C, et al. Real-time blood flow assessment using ICG fluorescence imaging during hepatobiliary and pancreatic surgery with consideration of vascular reconstruction[J]. *Cancers*, 2025, **17**(5): 868.
7. Tummers Q R J G, Hoogstins C E S, Gaarenstroom K N, et al. Intraoperative imaging of Folate receptor alpha positive ovarian and breast cancer using the tumor specific agent EC17[J]. *Oncotarget*, 2016, **7**(22): 32144−32155.
8. Nagaya T, Nakamura Y A, Choyke P L, et al. Fluorescence-guided surgery[J]. *Front Oncol*, 2017, **7**: 314.
9. Zhejiang Healnoc Technology Co., Ltd. Bifluorescence image processing method and endoscope imaging system: 118787297B[P]. 2025-03-25.
浙江华诺康科技有限公司. 双荧光图像处理方法和内窥镜成像系统: 118787297B[P]. 2025-03-25.
10. Guangdong OptoMedic Technologies, Inc. 3D fluorescent endoscope, imaging method, device and debugging method: 117398043B[P]. 2024-07-02.
广东欧谱曼迪科技股份有限公司. 一种 3D 荧光内窥镜、成像方法、装置及调试方法: 117398043B[P]. 2024-07-02.
11. Olympus Corporation. Fluorescence endoscopy system: 103687529B[P]. 2016-05-18.
奥林巴斯株式会社. 荧光内窥镜装置: 103687529B[P]. 2016-05-18.

12. Zhejiang University, Zhejiang Zhike Lishang Medical Technology Co., Ltd. Endoscope fluorescence and visible light image fusion method: 115330651B[P]. 2025-04-29.
浙江大学, 浙江之科立上医疗科技有限公司. 一种内窥镜荧光与可见光图像融合方法: 115330651B[P]. 2025-04-29.
13. Stryker European Operations Limited. Imaging systems and methods for displaying fluorescence and visible images: 20210274131[P]. 2021-09-02.
14. Ramanath R, Snyder W E, Yoo Y, et al. Color image processing pipeline[J]. *IEEE Signal Process Mag*, 2005, **22**(1): 34−43.
15. Bernacki J. Automatic exposure algorithms for digital photography[J]. *Multimedia Tools Appl*, 2020, **79**(19): 12751−12776.
16. Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2004, **13**(4): 600−612.

作者简介



邓奕城 (2001-)，男，硕士研究生，主要研究医学图像处理方向。

E-mail：22330019@zju.edu.cn



【通信作者】倪旭翔 (1970-)，男，博士，副教授，主要研究方向为光电检测技术，数字信号处理技术。

E-mail：xuxiang@zju.edu.cn

Real-time fluorescence imaging technology for endoscopes under the virtual dual-camera architecture

Deng Yicheng, Ni Xuxiang*, Wang Liqiang, Yuan Bo, Zhang Jiqiang, Tao Zilong

College of Optical Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China

Abstract

Objective Fluorescence-guided surgery (FGS) has emerged as a critical technique in modern medical procedures, utilizing fluorophores such as indocyanine green (ICG) to visualize anatomical structures and pathological tissues that are invisible to the naked eye. While effective, the clinical utility of fluorescence imaging relies heavily on its integration with white-light (WL) imaging to provide necessary anatomical context. Conventional dual-camera systems, which employ beam splitters and separate sensors, offer high performance but suffer from bulky form factors, high costs, and physical registration errors. Conversely, single-camera solutions that utilize time-division multiplexing often face significant technical bottlenecks. These include data congestion when processing high-bandwidth 4K video streams via software, and, more critically, the inability to balance exposure between high-dynamic-range white-light scenes and weak fluorescence signals using a single sensor. The primary objective of this study was to overcome these limitations by developing a "virtual dual-camera" architecture. This system aimed to achieve the performance and independent image processing capabilities of a physical dual-camera setup within a single-sensor hardware framework, specifically targeting real-time performance, hardware-level synchronization, and optimized automatic exposure control (AEC) for composite endoscopic imaging.

Methods The proposed system integrated a custom hardware platform with a novel control architecture. The optical front-end utilized a high-resolution CMOS image sensor (Sony IMX334) paired with a dual-bandpass filter capable of blocking excitation light (780 nm) while transmitting visible light and fluorescence signals (above 820 nm). Illumination was provided by a synchronized dual-LED source delivering high-power white light and 780 nm near-infrared excitation.

The core innovation lay in the field-programmable gate array (FPGA)-based control logic. The FPGA functioned as the central timing master, generating precise frame synchronization signals. It coordinated the CMOS exposure window with the alternating strobe of white and near-infrared LEDs, achieving frame-level synchronization accuracy. To resolve the bottleneck of software-based frame separation, the system implemented a hardware-level video stream demultiplexing architecture. The FPGA intercepted the raw MIPI CSI-2 video stream from the sensor. Based on the current lighting state (odd frames for white light, even frames for fluorescence), the FPGA routed the image data to two independent direct memory access (DMA) channels and image signal processor (ISP) pipelines. This hardware separation

Foundation item: Beijing Natural Science Foundation (L252001), and National Key Research and Development Program of China (2023YFF0720402)
*Correspondence: Ni X X, E-mail: xuxiang@zju.edu.cn

allowed for the decoupling of image processing parameters; the white-light channel utilized standard color correction and gamma settings, while the fluorescence channel employed aggressive black level correction, higher gamma values for contrast enhancement, and temporal noise reduction, which relied on the frame continuity preserved by the hardware splitting.

Furthermore, a dual-modal joint automatic exposure control (AEC) algorithm was developed to address the drastic intensity disparity between the two imaging modes. The algorithm operated in two distinct phases separated by a single frame delay. In Phase 1, the system calculated the luminance of the white-light frame and adjusted the white LED intensity and the sensor's global analog gain. Crucially, a constraint strategy prioritized the prevention of white-light overexposure. In Phase 2, utilizing the analog gain fixed during Phase 1, the system optimized the fluorescence image brightness by independently adjusting the near-infrared excitation power and applying digital gain. This decoupled approach ensured that the shared analog gain did not compromise the white-light image quality while still allowing sufficient amplification for the weaker fluorescence signal.

Results and Discussions Experimental validation was conducted using a porcine liver model injected with ICG, imaged at varying working distances (10 mm, 25 mm, and 40 mm) to simulate clinical conditions. The hardware synchronization mechanism achieved a control precision of 16.67 ms, strictly aligning the exposure windows with the respective light sources and eliminating channel crosstalk.

The proposed AEC algorithm demonstrated superior stability and speed compared to conventional ISP-based exposure algorithms. Quantitative analysis of convergence time revealed that the system recovered to a stable brightness state within 0.8 seconds (approximately 23 frames for white light and an additional 15–20 frames for fluorescence) even under extreme lighting changes. In contrast, standard AEC algorithms exhibited significant oscillation and "flashing" artifacts due to the coupled feedback loop between the alternating bright and dark frames. The proposed method maintained the white-light luminance within a tight tolerance (50 luma) without the overshoot observed in conventional methods.

Image quality was assessed using structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) metrics. At a standard working distance of 10 mm, the system achieved an SSIM of 0.9994 and a PSNR of 46.23 dB for white-light images, indicating negligible information loss compared to reference images. Fluorescence images at the same distance yielded an SSIM of 0.9893 and a PSNR of 42.25 dB. As the working distance increased to 40 mm, the fluorescence image quality naturally degraded due to the reliance on higher digital gain to compensate for signal attenuation, resulting in an SSIM of 0.9224. However, the system successfully maintained diagnostic visibility without introducing significant noise artifacts or false positives, a benefit attributed to the independent ISP pipeline tuning enabled by the virtual dual-camera architecture. Visual inspection confirmed that the fused images retained clear anatomical textures from the white-light channel and distinct, high-contrast tumor margins from the fluorescence channel.

Conclusions The study successfully implemented a virtual dual-camera imaging system that resolves the fundamental conflicts inherent in single-sensor fluorescence endoscopy. By shifting the frame synchronization and video demultiplexing tasks to the FPGA hardware layer, the design eliminated the latency and bandwidth limitations associated with software processing, enabling smooth 4K real-time imaging. The novel two-stage automatic exposure control algorithm effectively decoupled the brightness regulation of white-light and fluorescence modes, allowing for optimal dynamic range utilization for both anatomy and pathology without physical filter switching or dual-sensor alignment.

The experimental results confirmed that this architecture delivers image quality comparable to dual-camera systems, with high structural similarity and signal-to-noise ratios, while retaining the cost and size advantages of a single-chip solution. The fast convergence time of the exposure algorithm ensures seamless transitions during intraoperative maneuvering. Consequently, this technology provides a robust, low-cost, and high-performance solution for fluorescence-guided surgery, offering surgeons reliable real-time navigation aids with precise anatomical and functional information fusion. The proposed architecture establishes a viable technical pathway for the next generation of miniaturized, multi-modal endoscopic platforms.

Keywords: endoscope; composite imaging technology; medical image processing; virtual dual-camera; automatic exposure control

Deng Y C, Ni X X, Wang L Q, et al. Real-time fluorescence imaging technology for endoscopes under the virtual dual-camera architecture[J]. *Opto-Electron Eng*, 2026, **53**(6): 250341; DOI: 10.12086/oee.2026.250341

Opto-Electronic Journals Group
www.oejournal.org

扫描二维码，获取 PDF 全文