

基于 Transformer-SSVEP 解码的脑控无人机系统设计与实现^{*}

李奥丹, 张 键, 兰晓红

(重庆师范大学 计算机与信息科学学院, 重庆 401331)

摘要:针对传统脑控无人机系统在实时性、控制稳定性和复杂环境适应性等方面存在的不足,设计并实现了一套基于 Transformer 类稳态视觉诱发电位(steady-state visual evoked potential, SSVEP)解码方法的脑控无人机闭环控制系统。该系统采用“脑电(electroencephalogram, EEG)采集—信号处理—指令映射—无人机控制”的4层架构,以 SSVEP 为控制范式,完成起飞、降落、前进、后退、左移、右移、上升和下降8类飞行动作的在线控制。在解码层,采用 Transformer 类 SSVEP 解码器对多通道 EEG 信号进行时空特征建模,再结合置信度阈值与连续一致性约束构建双重校验机制,以降低误触发并提高在线控制稳定性。为验证该方案的适用性,首先基于 BETA 和 Benchmark 这2个公开数据集开展离线对照实验,并与 CCA、FBCCA 及 EEGNet 等方法进行比较;随后组织12名真实受试者开展在线脑控飞行实验,从识别准确率、响应时间、信息传输率和误触发率等方面对系统性能进行评估。实验结果表明,所构建系统能够在在线闭环场景下稳定完成8类飞行指令控制,Transformer 类解码方法在离线和在线实验中均表现出较好的适用性,所设计的双重校验机制能够有效提高控制稳定性并降低误触发风险。本研究为 Transformer 类 SSVEP 解码方法在脑控无人机实时控制场景中的系统集成与工程应用提供了实验依据。

关键词:脑控无人机;稳态视觉诱发电位;Transformer;脑机接口;实时控制

中图分类号:TP391.4

文献标志码:A

文章编号:1672-6693(2026)04-0105-12

脑机接口(brain-computer interface, BCI)通过建立大脑与外部设备之间的直接通信通道,绕过传统神经肌肉传导路径,实现人机之间的替代性交互,在人机交互、康复医疗和智能控制等领域展现出重要应用价值^[1]。从系统组成看,典型 BCI 通常包括“信号采集—信号处理—控制命令转换”3个环节:首先采集脑电(electroencephalogram, EEG)等神经信号,然后完成预处理、特征提取与模式识别,最后将识别结果映射为外部设备可执行的控制指令,从而形成完整的闭环交互链路。

在设备控制类应用中,无人机具有自由度高、机动性强、应用场景丰富等特点,可以作为 BCI 能力的直观验证载体^[2]。早期关于脑控无人机的研究多采用运动想象(motor imagery, MI)范式实现二维控制,但 MI 通常依赖较长训练过程,且受个体差异影响明显,存在学习成本高、控制稳定性不足等问题,从而限制了 MI 工程化推广^[3]。相比之下,稳态视觉诱发电位(steady-state visual evoked potential, SSVEP)具有诱发稳定、无须长期训练、响应速度快和信息传输率高等优势,更适合无人机这类对实时性要求较高的控制任务,因此成为当前非侵入式 BCI 研究的重要方向之一^[4]。已有研究验证了 SSVEP 在实时控制中的可行性,例如基于 SSVEP 的多方向控制系统^[5-6],以及融合 SSVEP 与 MI 以扩展指令集、提升操作灵活性的混合控制模式等^[7]。

SSVEP 系统性能在很大程度上取决于频率识别算法。典型相关分析(canonical correlation analysis, CCA)因计算效率高、无须个体校准而被广泛应用于 SSVEP 频率识别, Lin 等人的工作推动了 CCA 在 BCI 领域的普及^[8];在此基础上,滤波器组典型相关分析(filter bank canonical correlation analysis, FBCCA)通过多子带分解与加权融合进一步提升了解码准确率与鲁棒性^[9]。然而,在脑控无人机等复杂在线应用场景中,传统相关分析方法仍存在若干局限:一方面,强闪烁刺激容易引发视觉疲劳,且控制界面与任务画面分离会导致用户视线频繁切换,增加注意负担并影响长时间交互体验;另一方面,脑电信号本身噪声强、个体差异大,同时无人机任务中还存在注意波动、眨眼伪迹和环境干扰等因素,传统方法对跨时间片、跨通道依赖关系的建模能力有限,容易

^{*} 收稿日期:2026-03-09 修回日期:2026-04-01 网络出版时间:2026-07-14T09:27

资助项目:国家级大学生创新训练计划项目(No. 202510637052)

第一作者简介:李奥丹,女,研究方向为教育机器人, E-mail: 1416662843@qq.com;通信作者:兰晓红,女,副教授, E-mail: 331512740@qq.com

网络出版地址: <https://link.cnki.net/urlid/50.1165.N.20260713.1927.006>

导致识别结果波动,从而影响控制稳定性与飞行安全。

近些年来,深度学习方法被逐步引入 SSVEP 解码,以它的增强特征表达能力和系统鲁棒性^[10-12]。相较于传统相关分析方法,深度模型能够直接从原始或轻度预处理后的脑电数据中学习判别性表示,在复杂场景下具有更强的适应能力。但在脑控无人机等实时应用场景中,SSVEP 解码仍面临 3 方面关键挑战:其一,在短时间窗条件下,刺激基频及谐波信息尚未充分稳定,传统方法对长时序依赖和非线性特征的建模能力不足,容易在快速决策时表现出性能下降;其二,多通道 EEG 信号不仅包含时间维度上的频率响应,还存在明显的空间耦合关系,而传统模板匹配方法及局部卷积方法对跨通道全局关联的利用仍不充分;其三,在线控制任务受个体差异、注意波动、眨眼伪迹及环境干扰影响较大,在跨被试和小样本条件下容易出现识别不稳定问题。因此,如何在保证实时性的同时提高解码鲁棒性和控制稳定性,仍然是脑控无人机研究中的关键问题。

Transformer 类模型凭借自注意力机制,在序列建模中能够更有效地捕获跨时间片长程依赖,并实现多通道脑电信号的全局空间信息融合,因此为复杂在线 SSVEP 场景下的解码提供了新的技术路径^[13]。需要指出的是,本研究并非提出新的 SSVEP 解码网络结构,而是面向脑控无人机实时闭环控制场景,选用 Transformer 类 SSVEP 解码方法作为核心识别模块,并围绕在线部署需求完成系统集成与控制决策设计。在此基础上构建了一套闭环脑控系统,并通过置信度阈值与连续一致性约束相结合的双重校验机制,降低在线控制过程中的误触发风险,提高飞行控制稳定性。

为验证所选解码方法在脑控无人机场景中的适用性,本研究采用“离线评估和在线验证”的实验思路:首先基于公开 SSVEP 数据集 BETA 和 Benchmark 开展离线实验^[14-15],对 Transformer 类解码方法与 CCA、FBCCA 等典型方法进行比较分析^[16];随后基于 12 名真实受试者开展在线飞行实验,从识别准确率、响应时间、信息传输率以及控制稳定性等维度评估系统性能。本研究的工作旨在为 Transformer 类 SSVEP 解码方法在脑控无人机实时闭环系统中的应用提供实验依据,并为非侵入式 BCI 在智能飞行控制中的工程化落地提供参考。

1 系统总体架构

本研究设计并实现了一套基于 Transformer-SSVEP 解码的端到端脑控无人机系统。为实现各功能模块的解耦与高效协同,系统采用分层模块化设计思想,构建了“EEG 采集—信号处理—指令映射—无人机控制”的 4 层处理架构,如图 1 所示。该架构通过层级间的高通量数据交互(TCP/IP 与 UDP 协议),形成了从视觉刺激诱发到实体控制执行的闭环回路。

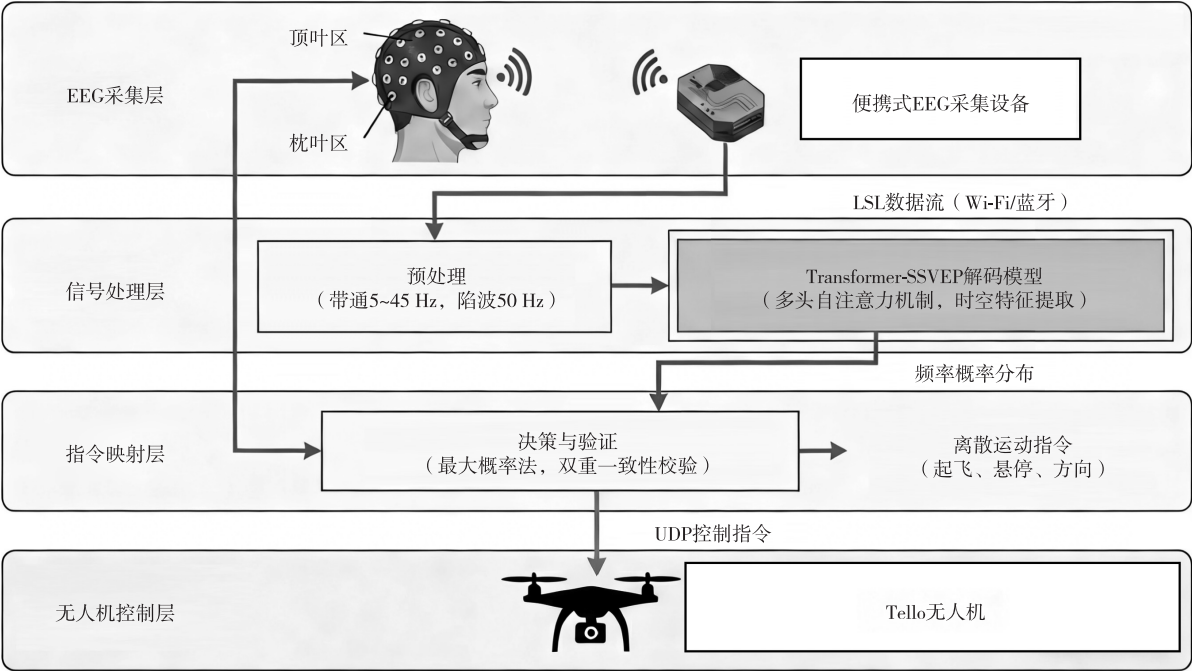


图 1 系统架构图
Fig. 1 System architecture diagram

EEG 脑电信号采集层负责高质量诱发脑电信号的获取与数字化传输。系统选用便携式多通道脑电采集设备,依据国际 10—20 系统标准^[17],将电极布局重点覆盖大脑后枕区 O1、O2、POz 及顶叶区域,以确保对稳态视觉诱发电位即 SSVEP 信号的有效捕捉。采集设备对模拟脑电信号进行放大、滤波及模数转换即 A/D 转换后,利用 Wi-Fi 或蓝牙通信链路,通过 lab streaming layer 即 LSL 流式协议^[18]将数字信号实时同步传输至上位机处理终端。

脑电信号处理层负责从噪声背景中提取特征并完成频率识别。接收到的原始 EEG 数据流首先经过 5~45 Hz 带通滤波与 50 Hz 工频陷波处理,以去除眼电漂移与工频干扰。随后,数据进入 Transformer-SSVEP 解码模型。该模型利用多头自注意力机制(multi-head self-attention, MHSA)并行挖掘脑电信号跨通道的空间相关性与跨时间片的时序依赖性,将时域信号映射为高维特征表示,并直接输出当前时间窗口对应的刺激频率概率分布。

指令映射层依据“最大概率原则”对处理层输出的频率概率进行判决,确定目标频率类别。为防止因注意力分散或眨眼等伪迹导致的误触发,本层引入了“双重一致性校验”策略,即要求连续 2 个时间窗口的识别结果一致且置信度高于设定阈值时,才确认为有效意图。确认后的意图依据预设的映射表被转换为离散的无人机运动指令。

无人机控制层负责指令的物理执行与人机交互反馈。上位机通过无线网络与无人机建立 UDP 通信连接,调用飞行控制接口将接收到的离散指令转化为具体的飞控命令,驱动无人机执行相应动作,实现全脑控的无人机飞行体验。

2 视觉刺激与指令编码设计

2.1 视觉刺激界面设计

为兼顾诱发效率与用户视觉舒适度,本研究设计了一种基于透明度调制的低对比度 SSVEP 视觉刺激范式^[19]。系统在屏幕周围均匀布置 8 个刺激模块,分别对应无人机的不同飞行动作,如图 2 所示。

刺激信号的生成采用正弦波对刺激块的透明度即 Alpha 通道进行调制^[2]。设第 i 个刺激模块的闪烁频率为 f_i ,则该刺激模块在时刻 t 的透明度 $T_i(t)$ 定义为:

$$T_i(t) = \frac{1}{2} [\sin(2\pi f_i t + \varphi_i) + 1],$$

其中: φ_i 为初始相位,该设计使得刺激块在“全透明”与“半透明”之间平滑切换,既保证了频率特征的稳定呈现,又实现了刺激层与任务背景的无缝融合,有效降低了强视觉冲击对受试者的干扰。

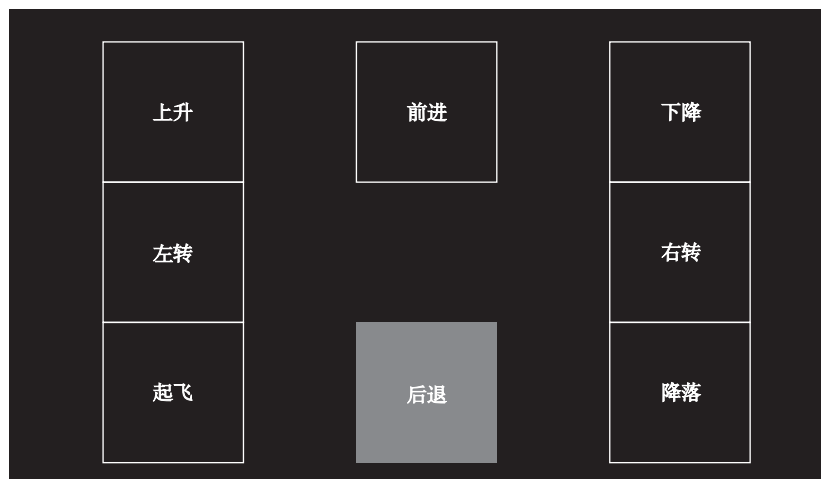


图 2 视觉刺激界面

Fig. 2 Visual stimulus interface

2.2 频率分配与指令编码

系统构建了“频率—指令”映射表,将 8 个数值范围为 8~15 Hz,步长为 1 Hz 的刺激频率分别映射至无人机的离散控制指令。具体编码方案如表 1 所示。为确保飞行安全,引入防误触机制:仅当解码结果满足置信度阈值且连续 2 帧识别一致时,才触发无人机运动指令,否则无人机保持当前运动状态或自动悬停。

表 1 频率-指令映射表
Tab. 1 Frequency-instruction mapping table

频率/Hz	动作指令
8	起飞
9	降落
10	前进
11	后退
12	左移
13	右移
14	上升
15	下降

3 Transformer-SSVEP 解码方法

针对传统的相关方法在分析中存在短时间窗输入条件下识别稳定性不足、对多通道空间耦合关系利用有限以及跨被试场景下易出现性能波动等问题,本研究在脑控无人机系统中引入 Transformer-SSVEP 解码模块^[20]。该类模型能够利用自注意力机制联合建模 EEG 信号在时间维度上的长程依赖关系和通道维度上的全局空间关联,从而提升复杂在线场景下的特征表达能力。与此同时,考虑到脑控无人机任务中在线样本有限、控制稳定性要求较高,本研究进一步结合公开数据集离线训练与在线双重校验机制,以增强系统在动态干扰和小样本条件下的适用性与鲁棒性。

3.1 数据预处理与序列嵌入实现

为提高脑电信号的信噪比并保留 SSVEP 的有效成分,系统首先对采集到的原始 EEG 信号进行预处理。具体而言,采用 5~45 Hz 的四阶巴特沃斯带通滤波器,以保留各刺激频率的基频及主要谐波成分,同时抑制低频漂移和高频肌电噪声;此外,为削弱工频干扰,进一步叠加 50 Hz 陷波滤波器。经滤波处理后的 EEG 信号用于后续离线训练与在线解码。

为保证方法设计与实验设置的一致性,本研究在离线训练阶段与在线推理阶段统一采用长度为 2 s 的 EEG 片段作为模型输入。对于采样频率为 500 Hz 的数据,每个输入样本包含 1 000 个采样点。在离线训练阶段,基于公开数据集构建训练样本,并从刺激开始后 0.5 s 处截取长度为 2 s 的有效信号,以减弱视觉诱发初始过渡段对分类性能的影响;在在线推理阶段,为兼顾识别稳定性与系统实时性,采用滑动窗口策略对连续 EEG 数据进行处理,窗口长度设为 2 s,步长设为 0.25 s,即系统每隔 0.25 s 更新 1 次输入数据,并基于最近 2 s 的 EEG 片段完成 1 次分类判定,从而形成持续更新的决策序列。

需要说明的是,BETA 和 Benchmark 数据集原始数据均为 64 通道 EEG 信号,而本研究在线脑控无人机系统采用 8 通道便携式采集设备。为保证离线训练与在线部署在输入维度和空间布局上的一致性,本研究从公开数据集中选取与在线系统覆盖区域相对应的 8 个后枕-顶叶通道构建训练样本,分别为 O1、Oz、O2、PO3、POz、PO4、P3 和 P4。在线阶段则直接使用 Brduino 八通道脑电仪采集的实时信号作为模型输入。通过该通道选择策略,可使离线模型训练与在线系统部署保持一致的输入形式。

为将连续 EEG 信号适配到 Transformer 网络中,本研究对输入样本进行序列化嵌入处理。设单次输入 EEG 片段表示为:

$$X \in R^{C \times T},$$

其中: C 表示通道数, T 表示采样点数。本研究中: $C=8, T=1\ 000$ 。沿时间维度将输入信号划分为 N 个不重叠 Patch,每个 Patch 的长度设为 250 个采样点,因此每个输入样本共划分为 4 个 Patch,即 $N=4$ 。考虑到 SSVEP 信号具有明显的频率响应特性,为增强频域特征的可分性,对每个 Patch 进行快速傅里叶变换(FFT)处理^[21],并提取其幅值谱特征。随后,通过线性投影层将各 Patch 的特征映射为 128 维嵌入向量,并加入可学习的位置编

码以保留序列的时序位置信息。进一步地,在输入序列首部拼接1个分类标记[CLS]Token,最终形成模型输入序列:

$$Z_0 \in R^{(N+1) \times 128},$$

通过上述处理,原始 EEG 信号被统一转换为适用于 Transformer 编码器输入的序列表示,为后续时空特征建模奠定了基础。

3.2 解码网络架构部署

本系统的解码核心采用标准 Transformer 编码器(encoder)架构^[13],如图3所示。网络由2层堆叠的编码块组成,旨在并行建模 EEG 信号中跨通道的空间相关性与跨时间片的长时序依赖关系。每个编码块均包含多头自注意力和前馈网络(feed-forward network,FFN)2个子模块,并结合层归一化(layer normalization,LN)与残差连接实现稳定训练。

网络的前向计算流程如下:在第 l 层中,输入序列 Z_{l-1} 首先经过层归一化,再输入多头自注意力模块;随后通过残差连接与层归一化进入前馈网络。该计算过程可表示为:

$$Z'_l = \text{MHSA}(\text{LN}(Z_{l-1})) + Z_{l-1},$$

$$Z_l = \text{FFN}(\text{LN}(Z'_l)) + Z'_l,$$

其中:多头自注意力模块的头数设为8,嵌入维度设为128,前馈网络隐藏层维度设为256,激活函数采用GELU,并在注意力模块和前馈网络后设置0.1的Dropout以减轻过拟合。经过2层编码块后,系统提取首部对应的特征向量,并通过带有64个隐藏单元的单隐藏层多层感知机即MLP分类头映射为8类刺激频率对应的类别概率,最终通过Softmax函数输出各类别的置信度。

在模型训练阶段,本研究采用离线训练、在线部署的实现方式。离线训练数据来自BETA和Benchmark 2个公开SSVEP数据集^[14-15],并基于筛选后的8个后枕-顶叶通道构建训练样本。跨被试评估采用leave-one-subject-out协议,即每轮实验选择1名受试者作为测试集,其余受试者作为训练集;训练集中进一步按8:2的比例划分训练子集与验证子集,用于模型优化与超参数选择。训练过程中采用交叉熵损失函数对网络参数进行优化,优化器选用Adam,初始学习率设为 1×10^{-3} ,批量大小设为64,训练轮数设为100,权重衰减系数设为 1×10^{-4} 。当验证集性能连续10个轮回无明显提升时,提前终止训练。本研究未采用额外的数据增强策略。离线训练完成后,将性能最优的模型参数部署至上位机在线解码模块,以支持后续脑控无人机实时闭环控制。

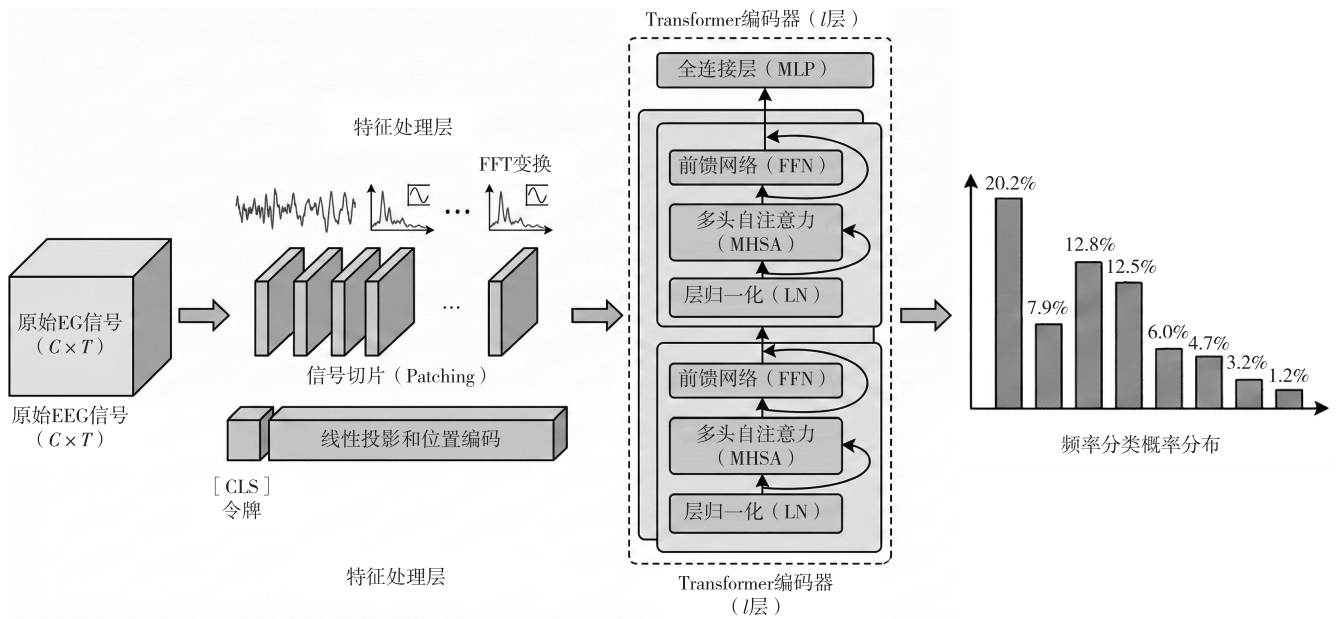


图3 Transformer-SSVEP 架构图

Fig. 3 Transformer-SSVEP architecture

本研究所采用Transformer-SSVEP解码模型的输入参数、网络结构参数、训练超参数及在线决策相关参数汇总如表2所示。

表 2 Transformer-SSVEP 模型关键参数设置

Tab. 2 Key parameter settings for Transformer-SSVEP model

参数类别	参数名称	取值
输入设置	EEG 通道数	8
输入设置	采样频率	500 Hz
输入设置	输入时间窗长度/s	2 s
输入设置	单样本采样点数/个	1 000
输入设置	滑动步长/s	0.25
输入设置	选取通道	O1、Oz、O2、PO3、POz、PO4、P3、P4
序列嵌入	Patch 长度/个	250
序列嵌入	Patch 数量/个	4
序列嵌入	嵌入维度/维	128
网络结构	Transformer 编码器层数/层	2
网络结构	注意力头数	8
网络结构	FFN 隐藏层维度	256
网络结构	激活函数	GELU
网络结构	Dropout	0.1
分类头	MLP 隐藏单元数	64
分类头	输出类别数	8
训练设置	优化器	Adam
训练设置	损失函数	交叉熵损失
训练设置	批量大小	64
训练设置	训练轮数/轮	100
训练设置	训练：验证划分	8：2
训练设置	提前停止策略	验证集连续 10 个轮回无提升则停止
训练设置	初始学习率	1×10^{-3}
训练设置	权重衰减系数	1×10^{-4}
在线决策	连续一致性窗口数/个	2
在线决策	动作冷却时间/ms	500

3.3 在线决策与稳健控制策略

针对在线环境下 EEG 信号波动、注意力短时偏移以及瞬时伪迹干扰可能引发的误动作等问题,本研究在解码输出层基础上进一步设计了适用于无人机控制场景的“概率—时序”双重校验机制,如图 4 所示。该机制从分类置信度和时序一致性这 2 个层面对识别结果进行约束,以提高在线闭环控制的稳定性与安全性。

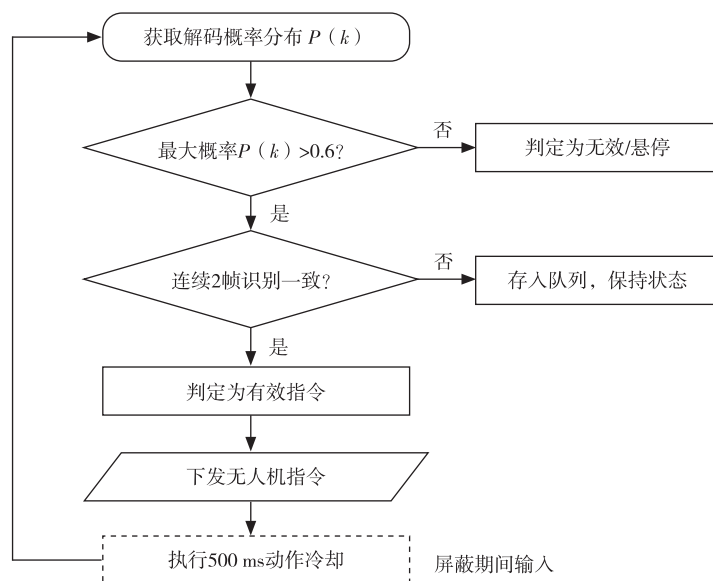
具体而言,设当前时刻模型输出的类别概率向量为 p ,当最大置信度满足:

$$\max (p) \geq \theta$$

时,认为当前窗口的识别结果具备候选指令的可信度。本文中阈值设定为 $\theta=0.6$ 。在此基础上,系统进一步引入时序一致性约束:仅当连续 2 个滑动窗口内的识别结果保持一致时,才将该结果判定为有效控制指令并下发至无人机执行;否则维持等待或自动悬停状态,不触发飞行动作。通过该策略,可以有效抑制由单窗口误判所引起的控制抖动。

为避免指令连续重复触发,系统在每次有效指令执行后进一步设置动作冷却时间 $\Delta t=500\text{ ms}$ 。在冷却期间,系统暂不响应新的重复输入,从而防止由于持续注视同一刺激块而导致的多次连续触发。结合置信度约束、

时序一致性约束和动作冷却机制,系统形成了完整的稳健控制策略,在保证响应效率的同时,有效降低了误触发率,提高了无人机飞行姿态切换的平滑性与安全性。



注: k 代表获得的脑电信号的类别, $P(k)$ 指获得该类别脑电信号的概率。

图4 双重校验机制图

Fig. 4 Double verification mechanism diagram

4 实验设计

4.1 实验平台与评价指标

为验证所构建脑控无人机的实际控制性能,本研究从离线算法验证和在线闭环控制验证这2个层面开展实验。系统沿用前文提出的“EEG采集—信号处理—指令映射—无人机控制”4层闭环架构。脑电采集设备采用Brduino八通道脑电仪,采样率设为500 Hz,如图5所示;刺激呈现设备采用24英寸LCD显示器,刷新频率为120 Hz,分辨率为 $1\,920 \times 1\,080$;控制对象为Tello EDU四旋翼无人机,如图6所示,通过Wi-Fi与上位机建立UDP通信连接,实现飞控指令实时下发。系统设置8、9、10、11、12、13、14和15 Hz共8个刺激频率,步长为1 Hz,分别对应起飞、降落、前进、后退、左移、右移、上升和下降8类飞行动作。为保证飞行安全,系统在在线决策阶段采用“最大置信度阈值和连续2个没去窗口结果一致”的双重校验策略,即仅当2项条件同时满足时才触发无人机动作。



图5 脑电采集设备

Fig. 5 EEG acquisition device

图6 Tello 无人机

Fig. 6 Tello UAV

在线实验共招募 12 名健康受试者参与测试。每名受试者在完成系统说明和操作熟悉后,进行多轮单指令测试与组合导航测试。评价指标包括识别准确率(accuracy)、F1 分数(F1-score)、平均响应时间(response time)、信息传输率(information transmission rate, ITR)和误触发率(false trigger rate, FTR)。其中, ITR 按下式计算:

$$ITR = \left[\log_2 X + P \log_2 P + (1-P) \log_2 \left(\frac{1-P}{X-1} \right) \right] \times \frac{60}{Y},$$

式中: $X=8$ 为指令类别数, P 为平均识别准确率, Y 为单次决策总耗时。需要说明的是, 在线系统中的 Y 不仅包括 EEG 时间窗长度, 还包括连续一致性确认引入的额外判决时间以及数据传输和飞控执行时延, 因此本研究采用端到端响应时间参与 ITR 计算, 更能反映真实在线控制效率。为检验不同方法之间差异的统计学意义, 本研究对 12 名受试者的总体结果采用配对统计检验; 当数据满足正态分布时使用配对 t 检验, 反之则采用 Wilcoxon 符号秩检验, 显著性水平设为 $p < 0.05$ 。

4.2 离线对照实验

为验证所选 Transformer 类解码器在 SSVEP 频率识别中的适用性, 本研究首先在 2 个公开的数据集上开展离线实验。离线训练数据来自 BETA 和 Benchmark 2 个公开 SSVEP 数据集。其中, BETA 数据集包含 70 名受试者, Benchmark 数据集包含 35 名受试者, 二者均为 64 通道 EEG 数据。参考统一设置, 本研究对 2 个数据集均采用 2 s EEG 片段作为模型输入, 并从刺激开始后 0.5 s 处截取有效信号; 跨被试评估采用 leave-one-subject-out (LOSO) 协议, 即每次选取 1 名受试者作为测试对象, 其余受试者作为训练集。所有方法均采用相同预处理流程和相同窗口设置, 以保证公平性。

4.3 在线脑控飞行主实验

在离线验证的基础上, 本研究进一步组织 12 名真实受试者开展在线脑控飞行实验。实验过程中, 受试者佩戴脑电采集设备, 在室内安全环境下通过注视不同频率的视觉刺激块控制无人机完成起飞、降落、前进、后退、左移、右移、上升和下降等动作, 如图 7 所示。系统采用在线闭环控制方式运行, 即将实时采集的 EEG 数据经预处理后输入 Transformer 类解码器进行识别, 并在双重校验机制确认后映射为对应飞控指令, 最终下发至 Tello EDU 无人机执行。



图 7 在线飞行测试场景

Fig. 7 Online flight test scenario

4.4 抗干扰实验

为考察系统在实际应用环境中的鲁棒性, 本研究进一步设计了抗干扰实验, 并设置正常条件与轻度干扰条件 2 种测试场景。正常条件下, 受试者在相对安静环境中保持稳定注视; 轻度干扰条件下, 允许受试者出现少量眨眼、轻微头动以及短暂注意分散, 以模拟真实脑控飞行过程中可能出现的动态干扰因素。考虑到在线闭环控制中 FBCCA 具有较强代表性, 本文选取 FBCCA 与 Transformer 2 种方法进行比较。评价指标包括识别准确率、平均响应时间和误触发率。

5 实验结果

5.1 离线对照实验结果

为体现对比的代表性,将典型相关分析方法 CCA^[8]、改进的滤波器组方法 FBCCA^[9]、轻量级深度学习模型 EEGNet^[10]以及本研究所采用的 Transformer 类解码器进行比较。图 8 给出了离线跨被试识别结果。从结果可以看出,CCA 作为经典模板相关方法,具有实现简单和无须个体校准的优点,但在跨被试条件下对个体差异较为敏感,因此整体性能相对较低;FBCCA 通过多子带分解与加权融合提高了对谐波信息的利用率,在 2 个数据集上的识别准确率和 F1 分数均明显优于 CCA;EEGNet 作为典型轻量级深度学习模型,进一步提升了分类性能,说明数据驱动特征学习方法在复杂 SSVEP 模式建模中具有一定优势;相比之下,Transformer 类解码器在 BETA 和 Benchmark 数据集上均取得了最高的平均准确率和 F1 分数,表明 Transformer 类解码器在跨时间片依赖建模和多通道空间信息融合方面具有更强的表征能力。

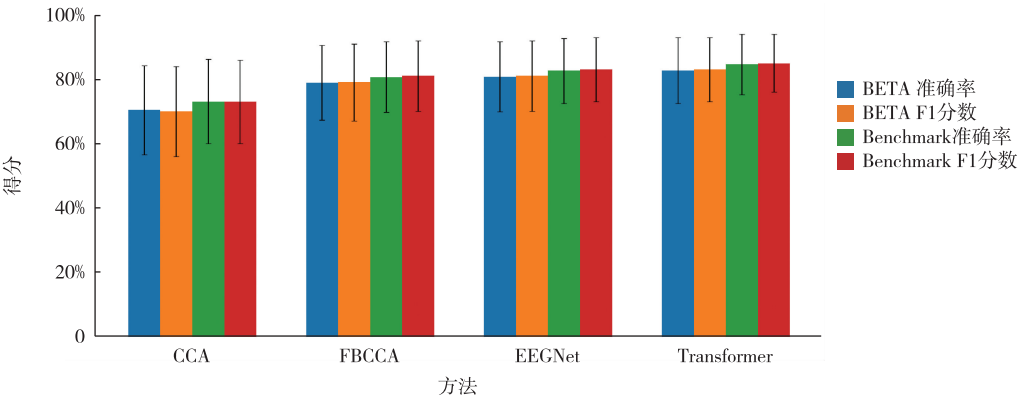


图 8 离线跨被试对照结果图
Fig. 8 Offline cross-subject control analysis

从相对提升幅度看,Transformer 在 BETA 数据集上的平均准确率较 FBCCA 提升约 3.8%、较 EEGNet 提升约 1.9%;在 Benchmark 数据集上较 FBCCA 提升约 3.9%、较 EEGNet 提升约 2.0%。总体来看,Transformer 在 2 个数据集上均取得了最高的平均准确率和 F1 分数,表现出更优且更稳定的识别性能。

5.2 在线脑控飞行主实验结果

表 3 为 8 类飞行指令在在线实验中的识别结果。可以看出,系统总体平均识别率达到 92%左右,各指令均保持了较高的识别水平。其中,起飞对应的边缘频率 8 Hz 和下降对应的边缘频率 15 Hz 识别率略低,而中间频段对应指令的识别效果更为稳定。识别结果以 12 名受试者在线实验结果的“平均值±标准差”形式给出,以反映系统性能的整体水平及实验波动情况。

表 3 在线飞行实验中各指令识别结果
Tab. 3 Instruction recognition results in online flight experiments

飞行指令	频率/Hz	识别准确率
起飞	8	89.5%±6.7%
降落	9	90.7%±6.1%
前进	10	93.4%±4.8%
后退	11	94.1%±4.6%
左移	12	94.6%±4.2%
右移	13	92.8%±4.9%
上升	14	91.8%±5.2%
下降	15	90.3%±5.7%
总体		92.2%±4.6%

表 4 为 Transformer 与 FBCCA 在在线闭环飞行场景中的总体性能比较。由于 CCA 在实时闭环控制中的整体性能通常低于 FBCCA,因此本研究在在线实验中选取 FBCCA 作为主要对照方法。结果表明,Transformer 在平均识别率、F1 分数、平均响应时间、信息传输率和误触发率等指标上均优于 FBCCA。这说明,在在线脑控无人机闭环控制场景中,Transformer 类解码器能够在保证较高识别准确率的同时,还能够兼顾控制效率与系统的稳定性。

表 4 在线飞行总体性能比较
Tab. 4 Comparison of overall performance of online flight simulations

方法	识别准确率	F1 分数	平均响应时间/s	信息传输率/(bit · min ⁻¹)	误触发率/(次 · min ⁻¹)
FBCCA	87.1%±6.1%	0.87±0.06	3.12±0.43	40.2±5.8	0.13±0.06
Transformer	92.2%±4.6%	0.92±0.05	2.85±0.37	50.1±5.2	0.06±0.03

综合表 3 和表 4 可以看出,所选 Transformer 类解码器在在线脑控无人机闭环控制中表现出较好的稳定性和实时性。它的优势不仅体现在更高的平均识别准确率上,还体现在更短的响应时间和更低的误触发率上。这说明,通过将 Transformer 类解码方法与双重校验机制结合,能够在实际飞控场景中兼顾控制效率与安全性。

5.3 抗干扰实验结果

表 5 为 FBCCA 和 Transformer 这 2 种方法在正常条件和轻度干扰条件下的在线性能比较结果。可以看出,正常条件下 Transformer 在识别准确率、响应时间和误触发率这 3 个指标上均优于 FBCCA;在轻度干扰条件下,2 种方法的性能均出现不同程度下降,但 Transformer 的下降幅度相对较小。具体而言,在轻度干扰条件下,Transformer 的平均识别准确率仍保持在 88%以上,平均响应时间约为 3.03 s,误触发率约为 0.11 次 · min⁻¹;相比之下,FBCCA 的识别准确率下降更为明显,响应时间进一步延长,误触发率也明显升高。

上述结果表明,Transformer 类解码器在存在动态噪声、注意波动和轻微运动伪迹干扰的情况下,仍能够保持较好的识别性能和控制稳定性。与此同时,本研究采用的双重校验机制在抑制误触发方面发挥了积极作用,使系统在复杂在线场景下仍具备较好的鲁棒性。综合来看,该抗干扰实验进一步验证了本研究方案在非理想条件下的适用性,为脑控无人机的实际部署提供了实验支持。

表 5 正常条件与轻度干扰条件下的在线性能比较
Tab. 5 Comparison of online performance under normal conditions and mild disturbance conditions

条件	方法	识别准确率	响应时间/s	误触发率/(次 · min ⁻¹)
正常条件	FBCCA	87.1%±6.1%	3.12±0.43	0.13±0.06
正常条件	Transformer	92.2%±4.6%	2.85±0.37	0.06±0.03
轻度干扰	FBCCA	81.6%±7.0%	3.36±0.47	0.22±0.08
轻度干扰	Transformer	88.5%±5.4%	3.03±0.40	0.11±0.05

5.4 实验结论

本研究通过离线对照实验、在线飞行主实验和抗干扰实验构建了较完整的性能验证体系。离线实验表明,所选 Transformer 类解码器在公开数据集上具有较好的跨被试识别能力;在线飞行实验验证了该方法在真实脑控无人机场景中的实时控制性能;抗干扰实验进一步说明了该方案在动态环境下的鲁棒性。结合 4 层闭环系统架构与双重校验机制,研究结果表明,Transformer 类 SSVEP 解码方法能够较好地适配脑控无人机应用需求,并在保证实时性的同时提升系统控制稳定性与安全性。

6 结束语

本研究面向脑控无人机实时控制场景,设计并实现了一套基于 Transformer 类 SSVEP 解码方法的闭环控制系统。系统采用“EEG 采集—信号处理—指令映射—无人机控制”的 4 层架构,并通过置信度阈值与连续一致性约束构建在线双重校验机制,以提升飞行控制稳定性并降低误触发风险。结合 BETA 与 Benchmark 数据集

上的离线对照实验以及 12 名真实受试者参与的在线飞行实验,结果表明:所采用的 Transformer 类解码器在跨被试离线识别中表现出较好的性能,在实际脑控无人机场景中能够实现稳定的 8 类飞行指令控制,并在识别准确率、响应时间、信息传输率和误触发率等指标上体现出较好的综合性能。抗干扰实验进一步说明,该方案在存在轻度动态干扰的条件下仍保持了较好的鲁棒性。总体来看,本研究从解码方法选取、闭环系统构建到在线性能验证,形成了较完整的脑控无人机实现方案,为 Transformer 类 SSVEP 解码方法在实时外设控制中的系统集成与工程应用提供了实验参考。

参考文献:

- [1] Wolpaw J R, Birbaumer N, McFarland D J, et al. Brain-computer interfaces for communication and control[J]. *Clinical Neurophysiology*, 2002, 113 (6): 767-791.
- [2] Chiuzbaian A, Jakobsen J, Puthusserypady S. Mind controlled drone: an innovative multiclass SSVEP based brain computer interface[C]//2019 7th International Winter Conference on Brain-Computer Interface (BCI). Piscataway: IEEE, 2019: 1-5.
- [3] Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication[J]. *Proceedings of the IEEE*, 2001, 89 (7): 1123-1134.
- [4] 潘洁, 高小榕, 高上凯. 同时诱发电位频率与相位特征的脑电研究[J]. *清华大学学报 (自然科学版)*, 2011, 51 (2): 250-254.
Pan Jie, Gao Xiaorong, Gao Shang kai. EEG study on the frequency and phase features of evoked potentials[J]. *Journal of Tsinghua University (Science and Technology)*, 2011, 51 (2): 250-254.
- [5] 伏云发, 郭衍龙, 李松, 等. 基于 SSVEP 直接脑控机器人方向和速度研究[J]. *自动化学报*, 2016, 42 (11): 1630-1640.
Fu Yunfa, Guo Yanlong, Li Song, et al. Direct-brain-controlled robot direction and speed based on SSVEP brain computer interaction[J]. *Acta Automatica Sinica*, 2016, 42 (11): 1630-1640.
- [6] 徐光华, 张锋, 王晶, 等. 面向智能轮椅脑机导航的高频组合编码稳态视觉诱发电位技术研究[J]. *机械工程学报*, 2013, 49 (6): 21-29.
Xu Guanghua, Zhang Feng, Wang Jing, et al. Research on key technology on time series combination coding-based high-frequency SSVEP in intelligent wheelchair BCI navigation[J]. *Journal of Mechanical Engineering*, 2013, 49 (6): 21-29.
- [7] 迟新一, 崔红岩, 陈小刚. 结合稳态视觉诱发电位的多模态脑机接口研究进展[J]. *中国生物医学工程学报*, 2022, 41 (2): 204-213.
Chi Xinyi, Cui Hongyan, Chen Xiaogang. Advances in multi-modal brain-computer interface combined with steady-state visual evoked potential[J]. *Chinese Journal of Biomedical Engineering*, 2022, 41 (2): 204-213.
- [8] Lin Z L, Zhang C S, Wu W, et al. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs [J]. *IEEE Transactions on Biomedical Engineering*, 2006, 53 (12): 2610-2614.
- [9] Chen X G, Wang Y J, Gao S K, et al. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface[J]. *Journal of Neural Engineering*, 2015, 12 (4): 046008.
- [10] Lawhern V J, Solon A J, Waytowich N R, et al. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces[J]. *Journal of Neural Engineering*, 2018, 15 (5): 056013.
- [11] Xu D C, Tang F Z, Li Y P, et al. An analysis of deep learning models in SSVEP-based BCI: a survey[J]. *Brain Sciences*, 2023, 13 (3): 483.
- [12] Guney O B, Oblokulov M, Ozkan H. A deep neural network for SSVEP-based brain-computer interfaces[J]. *IEEE Transactions on Biomedical Engineering*, 2022, 69 (2): 932-944.
- [13] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. 2017, 30: 5998-6008.
- [14] Liu B C, Huang X S, Wang Y J, et al. BETA: a large benchmark database toward SSVEP-BCI application[J]. *Frontiers in Neuroscience*, 2020, 14: 627.
- [15] Wang Y J, Chen X G, Gao X R, et al. A benchmark dataset for SSVEP-based brain-computer interfaces[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017, 25 (10): 1746-1752.
- [16] 高诺, 翟文文, 杨玉娜. MSI 和 CCA 算法对稳态视觉诱发电位信号分类的比较研究[J]. *信号处理*, 2018, 34 (8): 984-990.

- Gao Nuo, Zhai Wenwen, Yang Yuna. A comparison of MSI and CCA for SSVEP classification[J]. *Journal of Signal Processing*, 2018, 34 (8): 984-990.
- [17] Klem G H, Lüders H O, Jasper H H, et al. The ten-twenty electrode system of the international federation[J]. *Electroencephalography and Clinical Neurophysiology. Supplement*, 1999, 52: 3-6.
- [18] Kothe C, Shirazi S Y, Stenner T, et al. The lab streaming layer for synchronized multimodal recording[J]. *Imaging Neuroscience*, 2025, 3: IMAG-a. 136.
- [19] Peirce J W. PsychoPy: psychophysics software in Python[J]. *Journal of Neuroscience Methods*, 2007, 162 (1/2): 8-13.
- [20] Chen Y R, Cao L. Fliter bank convolution neural network based on transformer for SSVEP classification[C]//2025 4th International Symposium on Computer Applications and Information Technology (ISCAIT). Piscataway: IEEE, 2025: 69-72.
- [21] Cooley J W, Tukey J W. An algorithm for the machine calculation of complex Fourier series[J]. *Mathematics of Computation*, 1965, 19 (90): 297-301.

Design and Implementation of a Brain-Controlled UAV System Based on Transformer-SSVEP Decoding

LI Aodan, ZHANG Jian, LAN Xiaohong

(College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: This paper addresses the limitations of traditional brain-controlled UAV systems in real-time performance, control stability, and adaptability to complex environments by designing and implementing a closed-loop brain-controlled UAV system based on a Transformer-based SSVEP decoding method. The system adopts a four-layer architecture consisting of electroencephalogram (EEG) acquisition, signal processing, command mapping, and UAV control. Steady-state visual evoked potentials (SSVEPs) are used as the control paradigm to achieve online control of eight types of flight actions: takeoff, landing, forward, backward, leftward, rightward, ascent, and descent. In the decoding layer, a Transformer-based SSVEP decoder is employed to model the spatiotemporal features of multi-channel EEG signals. In addition, a dual-verification mechanism combining a confidence threshold and a consecutive consistency constraint is introduced to reduce false triggers and improve the stability of online control. To validate the applicability of the proposed approach, offline comparative experiments were first conducted on the two public datasets, BETA and Benchmark, and comparisons were made with methods such as CCA, FBCCA, and EEGNet. Subsequently, online brain-controlled flight experiments were carried out with 12 real participants, and the system performance was evaluated in terms of recognition accuracy, response time, information transfer rate, and false trigger rate. The experimental results demonstrate that the proposed system can stably accomplish online closed-loop control of eight flight commands in a real-time scenario. The Transformer-based decoding method shows good applicability in both offline and online experiments, and the designed dual-verification mechanism effectively improves control stability and reduces the risk of false triggers. This study provides experimental evidence for the system integration and engineering application of Transformer-based SSVEP decoding methods in real-time brain-controlled UAV scenarios.

Keywords: brain-controlled drone; steady-state visual evoked potential (SSVEP); Transformer; brain-computer interface (BCI); real-time control

(责任编辑 陈新颖)