arXiv:2304.02002v1 [cs.RO] 4 Apr 2023

面向水下机器人遥操作的最优人机界面设计

Paulo Padrao、Jose Fuentes、Tero Kaarlela、Alfredo Bayuelo、Leonardo Bobadilla

摘要——**高效且直观的人机界面对于扩大操作员用户群，并推动精准农业、自动化建筑、康复和环境监测等关键领域的新应用至关重要。本文研究了动力系统遥操作的人机界面设计。所提出的框架旨在找到符合用户舒适度、效率、连续性和一致性等关键概念的最优界面。作为概念验证，我们提出了一种遥操作水下航行器的创新方法，可将人体运动转换为航行器控制指令。该方法无需潜水员在严酷的水下环境中作业，同时兼顾舒适度和通信限制。我们使用连接到智能手机的头戴式显示器，让人类受试者控制一台模拟遥控水下航行器，并开展了一项研究。此外，数值实验表明，最优转换往往是最直观、最自然的方式，也符合用户的预期。**

I. 引言

直观、稳健且高效的人机界面（HRI）设计能够扩大操作员用户群，并推动新一代应用在精准农业、自动化建筑、康复和环境监测等关键领域的发展 [12]。得益于近年来虚拟现实应用的发展，机器人遥操作的用户界面设计日益受到关注，因为遥操作系统的人类操作表现可能受到数据带宽、时延、帧率不足、注意力不集中以及其他界面相关因素的影响[4]。此外，海洋环境等时空变化剧烈、感知和通信面临挑战的环境，也给遥操作带来了额外挑战。

全球约有60%的人口居住在距离沿海或河口环境不足100公里的范围内 [6]。随着这些地区的人口密度和经济活动不断增长，沿海生态系统面临的压力也随之增大。因此，水下机器人在研究、监测和管理沿海与河口环境方面日益不可或缺。这些环境对许多应用至关重要，包括沿海保护 [31],珊瑚修复 [26],海洋相关现象的预测 [25],以及石油钻井平台维护 [28]。

以海岸保护与修复为例，水下机器人可以协助监测海洋保护区，确保遵守相关法规并发现非法活动。在珊瑚修复中，机器人可以帮助种植新的珊瑚，并清除受损珊瑚礁上的有害碎屑。同时，在石油钻井平台，机器人可以检查和维护管道，保障工作人员的安全并降低石油泄漏的风险。凭借收集高质量数据并在危险环境中执行任务的能力，水下机器人有望显著增进我们对这些关键环境的了解，并支持多种应用。

水下环境并不适合人类，若要在水面下工作，就需要潜水装备或遥操作水下机器人。本文提出了一种遥操作遥控水下航行器（ROV）的新概念。该概念打破了水下作业地点与用户之间的隔阂，使用户免受恶劣水下环境的影响。该方法使远程操作员能够通过身体和头部运动控制遥控水下航行器。此外，头戴式显示器（HMD）能够实现对遥控水下航行器的遥操作，并提供水下世界的视觉反馈。

本文提出的遥操作方法为人类操作者提供了一种直观、自然的遥控水下航行器高层控制接口。遥控水下航行器与远程操作员之间通过互联网以太网连接，使遥控水下航行器能够实现不受位置限制的远程操作。此外，该方案利用智能手机将远程操作员的头部运动转化为遥控水下航行器的控制指令，从而实现高层控制接口。该方案尤其适用于检查、焊接和救援等水下任务，可为相关行业带来益处。通过远程操作，远程操作员能够确保工作环境符合人体工学且安全，从而降低身体劳损或受伤的风险。此外，该方案可支持连续作业，无需频繁休息或将工作时间缩短为短时段，最终提高水下作业的生产力和效率。

本文的主要贡献如下：

- 我们构建了一个基于优化的人机界面设计框架，并使其符合用户舒适度和效率方面的约束。
- 我们提出了一种新颖的遥控水下航行器遥操作方案，通过捕捉人体运动并将其转换为遥控水下航行器的控制指令。
- 我们展示了该系统的原型，作为概念验证

L. Bobadilla、P. Padrao 和 J. Fuentes 任职于美国佛罗里达州迈阿密市佛罗里达国际大学计算与信息科学学院，邮编 FL33199。bobadilla@cs.fiu.edu, ppadraol@fiu.edu, jfuen099@fiu.edu A. Bayuelo 任职于哥伦比亚波哥大哥伦比亚国立大学´，哥伦比亚。ajbayuelos@unal.edu.co T. Kaarlela 任职于芬兰伊利维耶斯卡 84100 森特里亚应用科学大学。tero.kaarlela@centria.fi

。

本文其余部分安排如下：第二节回顾该主题的既有研究。第三节阐述本文的两个研究问题。第四节介绍解决所提出问题的方法，第五节介绍作为概念验证的原型遥操作实现。第六节展示本研究的结果，第七节总结全文。

II. 相关工作

我们的工作探讨最优人机界面设计问题。人机界面的设计依赖于选择适当的期望属性，例如一致性、线性和连续性，以提升遥操作中的舒适度和自然感 [22]。我们的方法与 [13] 中提出的概念密切相关，因为我们研究了人类空间与机器人空间之间的映射，以及此类问题的数学表述。我们还与近期关于基于头戴式显示器的沉浸式遥操作接口 [11], 、以人类感知为优化目标的规划 [30], 以及利用最优控制遥操作机器人 [14]的研究有共同之处。

数十年来，研究人员一直在研究遥操作，以克服远程操作员与环境之间的障碍 [29], [19], [18]。障碍可能是墙壁等物理障碍，也可能是有毒或危险的环境，或深海环境等环境障碍。

近年来，水下远程操作研究主要聚焦于开发提升水下机器人的性能和能力的技术 [23]。当前水下远程操作应用的实例涵盖了为水下任务开发虚拟现实接口[7] 、设计水下仿人机器人 [32] ，以及用于水下航行器有限距离遥操作的可见光通信系统 [5]。

由于水下环境恶劣且多变，水下远程操作面临诸多挑战。其中一个主要挑战是通信带宽和信号质量有限，这可能导致时延和丢包。水的物理特性使得无线电波无法作为遥操作的无线通信方式。相反，研究人员已经研究并采用了声学通信和光通信 [27]。二者都能够实现水下无线通信，但分别存在带宽低或可用性受限的问题。声学调制解调器的带宽有限，只能实现低分辨率和低质量的视频传输 [1]。虽然光学调制解调器的带宽高于声学调制解调器，但其代价是发射器与接收器之间必须保持视距。要在遥控水下航行器与水面上方的发射器之间保持视距，需要复杂的跟踪电子设备和执行器 [17]。水下环境中缺乏视觉线索，使操作员难以准确感知机器人的位置和朝向。此外，水流和湍流的影响也会使机器人运动的精确控制变得困难。

III. 问题表述

本文研究基于视觉的水下航行器遥操作任务。参与其中的两个智能体——执行遥操作的人员和机器人本身——分别具有工作空间、动作空间和状态空间。设 $\mathcal{W}_o \subset \mathbb{R}^3$ 和 $\mathcal{W}_r \subset \mathbb{R}^3$ 分别为操作员和机器人的工作空间； $\mathcal{C}_o$ 和 $\mathcal{C}_r$ 为二者的构型空间。我们用 $\mathcal{U}$ 表示适用于机器人的控制量集合，用 $\mathcal{A}$ 表示人类操作者能够执行的动作集合。沿用 [13], 中的符号表示法，我们假设机器人的动力学由函数f 给出的关系 $\mathcal{C}_r \times \mathcal{U} \longrightarrow \mathcal{C}_r$所决定：

$$\dot{x} = f(x, u). \tag{1}$$

要建立遥操作系统，必须通过映射 g ：$\mathcal{C}_r \times \mathcal{A} \longrightarrow \mathcal{U}$ 将用户的动作映射到机器人的动作上 g ，使机器人受到用户所执行动作的影响

$$\dot{x} = f(x, g(x, a)). \tag{2}$$

假设给定集合 $\mathcal{A}$、$\mathcal{U}$ 和函数 f，我们的问题是依据诸如连续性、一致性 和可达性 等原则构建 g ，如 [13]中所述并在下文重述。然而，我们可能面临这样的挑战：用户完全沉浸在遥操作系统中，以致无法全面、准确地了解机器人的构型状态。此外，遥操作系统提供的信息可能难以让用户快速、准确地解读，从而影响最佳性能。尽管如此，函数 g 假设机器人构型状态的信息是完美的。因此，我们需要映射 $\psi : \mathcal{C}_o \longrightarrow \mathcal{C}_r$；该映射将在机器人构型状态与用户构型状态之间进行双向转换。事实上，我们假设 ψ 是一个微分同胚。在这种情况下，我们将 z 定义为 $\psi(z) = x$ ，并且（2）变为

$$\dot{z} = [D\psi(z)]^{-1} \cdot f(\psi(z), g(\psi(z), a)). \tag{3}$$

这种表述使我们能够从操作机器人者的角度来考虑这一问题。在这种情境下，我们假设机器人能够付出远超操作员的努力。这一考量使我们能够定义专门针对个人问题的泛函，而这些问题可能无法通过函数 ψ 得到保留。如前所述，为了实现真实且舒适的遥操作体验，映射 g 应满足若干关键条件 [13],，其中我们认为有四项最为重要。首先是 一致性，即保留机器人与操作员共有的属性。具体而言，这意味着对称性：如果操作员和机器人的动作相对于某一特定轴对称，映射 f 就应尽可能保持这种对称性。其次是 连续性，即将操作员的动作映射为与之密切相关的机器人动作。在这种情况下，可能还需要限制 f 的导数

以防机器人因感知运动方面的因素而出现不一致的移动。第三是 线性，它为智能体遥操作机器人提供了一种直观的方式。在这种情况下，可以合理地预期，如果输入加倍，操作员也会期待机器人的输出大致加倍。最后，可达性 和 完备性 使用户能够操作机器人到达给定的期望状态。设 $\mathcal{F}_o \subseteq \mathcal{C}_o$ 和 $\mathcal{F}_r \subseteq \mathcal{C}_r$ 分别为操作员和机器人的可行配置子集。如果存在一个控制函数$u(t)$，使状态 $x' \in \mathcal{F}_r$ 从 $x \in \mathcal{F}_r$ 到达 x'，则称状态 $x' \in \mathcal{F}_r$ 可从 $x \in \mathcal{F}_r$ 经 u 到达。类似地，如果存在一个控制函数 $a(t)$，使状态 x 到达 x'，则称状态$x' \in \mathcal{F}_o$ 可从 $x \in \mathcal{F}_o$ 经 a 到达。因此，如果所有 u-可达对 $(x, x') \in \mathcal{F}_r$ 都有一个等价的 a- 可达对 $(x, x') \in \mathcal{F}_o$，则称函数 g 为 完备。

一般而言，该问题可表述为寻找一个函数 g，使其最小化针对特定问题设计的某个泛函 $J(g)$。通常，该函数具有如下形式

$$J(g) = \int_0^T L(x, g, Dg)dt, \tag{4}$$

其中，L 是代价函数，D^n 表示 g 的 n^{th} 阶导数。有时，如果有需要，可以同时求出 a 和 g，因此该优化问题可能包含两个变量；除非另有说明，我们将假设两者都需要。自然地，最优控制问题中，$g \in C^1(\mathcal{C}_r \times \mathcal{A},\ \mathcal{U})$ 和 $a \in C^1(\mathcal{A})$ 是约束问题的解

$$\begin{aligned} \min_{g,a}\ & \int_0^T L(x, g, Dg)dt \\ \text{s.t. }\ & \dot{x} = f(x, g(x, a)) \\ & x(0) = x_{initial},\ x(T) = x_{final}. \end{aligned} \tag{5}$$

我们强调，$x = x(t)$ 和 $a = a(t)$ 都具有时间依赖性；为简化 ˆ 符号表示，我们在式（5）中省略了这些依赖关系。同样，可以定义泛函 J，并将式（3）纳入考虑，以明确体现对操作员所执行轨迹的依赖关系。

无论如何，求解该问题需要求解 Jacobi-Bellman方程 [13], ，其中涉及求解偏微分方程以获得最优函数。此外，这一过程通常颇具挑战性，有些解可能不符合前述原则，或者普通人难以执行。因此，我们将研究一些具有直观设计且便于普通人操作的特例。

问题1：转换构型空间

给定操作员的配置空间 $\mathcal{C}_o$ 以及操作员能够执行的动作集合 $\mathcal{A}$，计算映射 g，将操作员的配置空间和动作空间分别转换为机器人的动作空间和配置空间 $\mathcal{C}_r$ 与 $\mathcal{U}$。

水下通信受到各种限制，可能影响数据的可靠性和传输速度，例如衰减、噪声、干扰、延迟和功耗。此外，水下通信的可用带宽有限，因此难以快速传输大量数据 [33]。由于人类天生能够从模糊或像素化等低质量图像中重建几何结构和含义，而水下环境又存在严苛的数据通信约束，我们研究如何转换观测空间，以便为人机交互实现高效的数据交换。设 $\mathcal{Y}_r$ 和 $\mathcal{Y}_o$分别为机器人和操作员的观测空间。我们希望计算出相应的映射，使传输的信息量尽可能少，同时又足够有意义，让操作员能够有效且高效地操控机器人。这对于不涉及高计算量任务的实时场景尤其有用。

通过这种方式，我们定义一个待优化的泛函，以便找到合适的映射。设 $h\colon\ \mathcal{Y}_r \longrightarrow \mathcal{Y}_o$ 为该映射。因此，**问题1**可以使用另一个代价函数 K 进行表述，并且需要在合适的函数空间 $\mathcal{H} = C^n$（例如 $\mathcal{C}_r \times \mathcal{A},\ \mathcal{U}$）中找到一个函数 $h \in \mathcal{H}$，使得

$$\max_h K(h, Dh, D^2h, \ldots, D^nh). \tag{6}$$

如前所述，需要满足某些特定属性，才能使信息有用且不会产生额外失真。因此，我们借用了之前介绍的理想函数 g 的属性，以及 [13], 的连续性和线性；不过，由于环境几何形状在某种程度上得以保留，可以放宽线性假设，以保留仿射变换。由于我们还面临与通信相关的问题，因此可以将 h 视为时间相关函数 $h\colon\ \mathcal{Y}_r \times [0, T] \longrightarrow \mathcal{Y}_o$，该函数具有与通信阈值 $c\ (t)$ 相关的时间约束。因此，可以将（6）更新为

$$\begin{aligned} \max_h\ & K(h, Dh, D^2h, \ldots, D^nh) \\ \text{s.t }\ & I(h(y, t)) \le c(t) \text{ for each } y \in \mathcal{Y}_r. \end{aligned} \tag{7}$$

其中 $I : \mathcal{Y}_o \longrightarrow \mathbb{R}$ 计算在给定时刻发送给操作员的信息 t_o。

问题2：转换观测空间

给定机器人 $\mathcal{Y}_r$ 的观测空间，计算映射 $h\colon\ \mathcal{Y}_r \longrightarrow \mathcal{Y}_o$，以向操作员提供足够的信息，从而有效地远程控制水下机器人的运动。

IV. 方法

A. 问题1：转换配置空间

为解决将操作员空间映射到机器人空间这一首要问题，我们考察这样一种情形：人类操作者通过头部运动和身体运动向机器人发送指令。人头的活动范围取决于多个因素，例如年龄、性别、健康状况和个体解剖差异。平均而言，成年人可以将头部向左右两侧各旋转至多 90 度，并上下俯仰约 45 度，因此

总活动范围约为 180 度。上述运动由智能手机的内置传感器捕捉，包括固定在操作员所佩戴潜水面罩上的惯性测量单元（IMU）和气压计。随后，操作员的运动会被转换为指令，以实现对水下机器人的遥操作。完整架构如图 2 所示。机器人摄像头捕捉机器人周围环境的图像，并将其显示在智能手机屏幕上，从而向操作员提供视觉反馈。这使操作员能够观察机器人的行进方向，并据此调整其运动。由于用户输入指令受到限制，机器人最初被视为以 $\mathbb{R}^2$ 恒速运动的刚体。设 a_θ 和 a_ψ 分别为头部俯仰和头部偏航指令，u_θ 和 u_ψ 分别为摄像头倾斜指令和机器人基座的偏航指令。设 v_o 和 v_r 分别为操作员和机器人在 $[m/s]$ 中的线性前进速度。此外，我们认为，操作员和遥控水下航行器的深度可分别通过 $[m]$ 中的动作变量 a_z 和 u_z 直接设定。我们将操作员的动作空间定义为

$$\mathcal{A} = (a_{\theta\,\min}, a_{\theta\,\max}) \times (a_{\psi\,\min}, a_{\psi\,\max}) \times (v_{o\,\min}, v_{o\,\max}) \times (a_{z\,\min}, a_{z\,\max}) \quad (8)$$

以及机器人的动作空间为

$$\mathcal{U} = (u_{\theta\,\min}, u_{\theta\,\max}) \times (u_{\psi\,\min}, u_{\psi\,\max}) \times (v_{r\,\min}, v_{r\,\max}) \times (u_{z\,\min}, u_{z\,\max}) \quad (9)$$

人机动作空间转换的示例如图 1 所示。

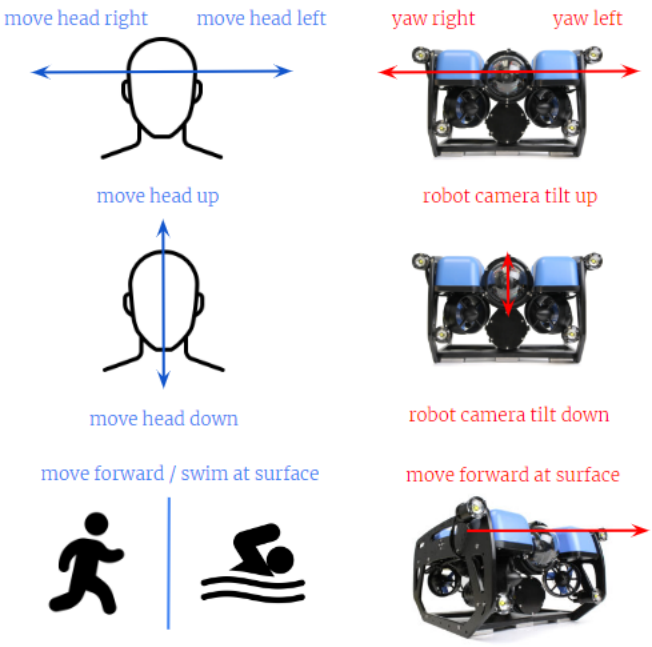

图 1. 人机动作空间转换示例

为评估这些映射，我们修改了 [13] 中定义的泛函，以反映与水下远程操作相关的具体因素。当 ψ 可用时，可使用公式 (2) 或 (3) 对所提出的问题进行表述。第一种公式对应的是最短距离问题，其中

$$\begin{aligned} \min_{g,a} \int_0^T L(x,g,Dg)dt &= \int_0^T ||\dot{x}(t)||dt \\ &= \int_0^T ||f(x(t), g(x(t), a(t)))||dt \\ \text{s.t } \dot{x}(t) &= f(x(t), g(x(t), a(t))) \\ x(0) &= x_{initial},\ x(T) = x_{final}. \end{aligned} \quad (10)$$

已有多项研究指出，在远程操作任务中需要为用户保持舒适的环境 [2], [30]。因此，我们希望减少操作者的动作次数，以提高操作机器人时的舒适度。例如，[30] 通过尽量减少头部运动来提升用户的舒适度。这可以视为简化曲线的曲率。直观而言，我们考虑转弯次数较少的路径。这可表示为$\kappa(t) = ||\ddot{x}(t)||$，平均曲率则为

$$\begin{aligned} \min_{g,a} \int_0^T L(x,g,Dg)dx &= \frac{1}{T}\int_0^T ||\ddot{x}(t)||dt \\ &= \frac{1}{T}\int_0^T ||f(f_x + f_u g_x) + f_u g_a \dot{a}||dt \\ \text{s.t. } \dot{x}(t) &= f(x(t), g(x(t), a(t))) \\ x(0) &= x_{initial},\ x(T) = x_{final}. \end{aligned} \quad (11)$$

为提高可读性，省略了函数参数；下标表示偏导数。但需要注意的是，这些泛函的计算开销可能较大。此外，范数函数在原点处不可导。尽管存在这一局限，我们仍希望将这些函数用作评估器，以在不同的人机动作空间之间转换动作。

图1展示了操作员动作与机器人动作之间的映射。该映射具有连续性和一致性等关键特性，这意味着操作员的小幅动作会引起机器人相应的小幅动作，并且这些动作可逆且一致。因此，即使机器人的可达性有所限制（因为并非操作员的所有动作都能转换为机器人动作），机器人仍然非常直观易控。

B. 问题*2*：转换观测空间

为解决转换观测空间这一第二个问题，我们认为机器人的观测空间$\mathcal{Y}_r$ 仅由RGB摄像头传输的视频帧组成，可定义如下

$$\mathcal{Y}_r = \mathbb{R}^{M\times N\times 3} \times [0, T] \quad (12)$$

其中，M 和 N 分别表示视频帧的宽度和高度，T 表示时间戳。由于通信约束，来自机器人（therobot）的帧会经过降采样和缩放，以减小其大小，同时

保留足够的信息供操作员处理。从数据传输效率来看，黑白（二值）图像通常比灰度图像或语义分割更高效，因为表示图像时，每个像素只需一位。主要的权衡在于图像可呈现的细节程度与表示图像所需的数据量之间。彩色图像通常每个像素使用24位，以表示每个通道上的256个强度级别。灰度图像只使用一个通道，因此每个像素需要8位；黑白图像则每个像素仅使用1位来表示黑色或白色。根据通信约束，我们可以选择一种映射，或根据式（7）中的I（y）和 c（t）在不同映射间切换。同样，语义分割算法通常会输出多个数据通道，从而进一步增加需要传输的数据量。图4展示了水下环境的模拟图像，以及相应的灰度和黑白版本。

通过这种方式，操作员的观测空间 $\mathcal{Y}_o$ 由处理后的视频帧组成，可定义如下：

$$\mathcal{Y}_r = \mathbb{R}^{M' \times N'} \times [0, T] \tag{13}$$

其中，$M' \leq M$和 $N' \leq N$。

评估所提出方法性能的一种自然候选方案，是通过计算这些映射所损失的信息量。衡量图像之间相似度的一个常用概念是互信息，它可以理解为通过观测一个随机变量而从另一个随机变量中获得的信息量。在我们的例子中，考虑两个图像 X 和 Y；它们定义了概率分布 $p_X(x)$、 $p_Y(y)$ 及其联合分布 $p_{(X,Y)}(x, y)$。因此，通过观测 X 从 Y 获得的互信息定义为

$$\mathcal{I}(X,Y) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} p_{(X,Y)}(x,y) \log \left(\frac{p_{(X,Y)}(x,y)}{p_X(x) p_Y(y)} \right) dxdy. \tag{14}$$

具体而言，我们可以考虑最大化与变换所保留的平均信息相关的泛函（7）

$$K(h) = \mathbb{E}_{y \sim p(y)}[\mathcal{I}(h(y), y)] = \int_{\mathcal{Y}_r} p(y) \mathcal{I}(h(y), y) dy \tag{15}$$

$p(y)$ 是定义在 $\mathcal{Y}_r$上的概率分布。

例如，使用图4（上）中的图像，计算原始图像与其灰度版本和黑白版本之间的互信息。灰度图像与原始图像之间的互信息为 2.008，黑白图像与实际图像之间为 0.16，原始图像与其自身之间为 4.9。此外，还可以对图像进行其他变换，例如语义分割 [15]，或受便于人类理解的导航表示方式启发的变换 [21]，。

V. 原型的实现

本研究采用了两种不同的方法，即软件在环（SIL）和硬件在环（HIL）。SIL配置使用 BlueSim 硬件仿真器 [9] ，而非实际硬件。SIL配置在软件组件开发和配置过程中使用 BlueSim 硬件仿真器[9] ，而非真实硬件。BlueSim 模拟 BlueROV2 硬件 [3]，并提供一个虚拟摄像头单元，用于测试和改进系统。图2展示了初始原型的架构以及 SIL 和 HIL 的连接。

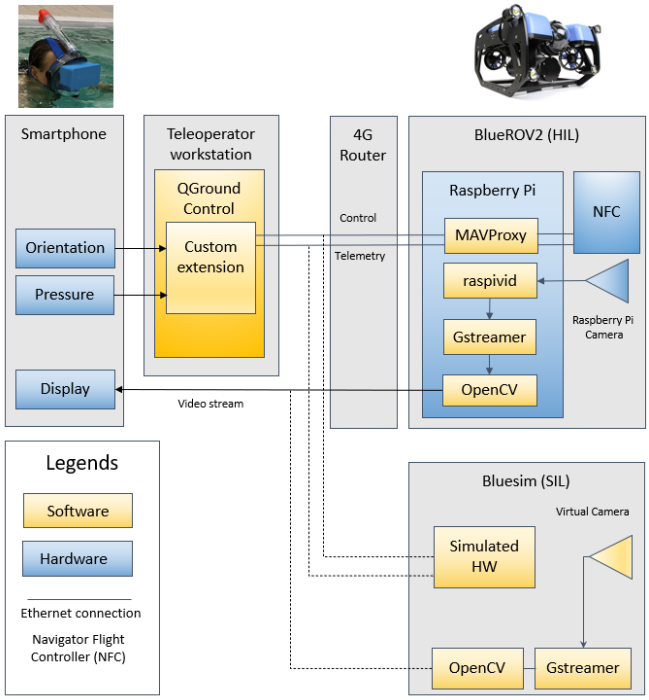

图2. 初始原型架构

与SIL配置相比，HIL配置使用实际硬件来控制遥控水下航行器。位于FIU实验室的遥控水下航行器通过有缆连接与水面上的4G路由器相连，从而能够传输黑白视频并控制遥控水下航行器的功能。为传输水下视频，BlueROV2使用GStreamer [10]，并开发了一款基于OpenCV [24]的应用程序，用于处理视频流并仅向远程操作员传输黑白图像，以节省带宽。BlueROV2通过4G路由器调制解调器连接互联网，使远程操作员能够通过公共静态IP地址，不受位置限制地遥操作遥控水下航行器。因此，安装在远程操作员工作站上的修改版QGroundControl（QGC） [8] 可用于控制遥控水下航行器。最后，智能手机连接到QGC，使远程操作员能够通过身体运动控制遥控水下航行器。

远程操作控制设备由潜水面罩、外壳和智能手机组成。该潜水面罩是标准的全脸式潜水面罩。智能手机外壳是定制的——

设计部件，可将潜水面罩和智能手机组合成头戴式显示器。设计外壳前，使用三维激光扫描仪提取了面罩的点云，并在三维 CAD 软件中进行设计，以确保面罩与外壳准确贴合。该外壳原型由芬兰 Centria大学生产自动化实验室使用三维打印机以聚乳酸（PLA）材料打印而成。

为了访问智能手机的惯性测量单元（IMU）姿态和压力数据，我们使用了由 Lorenz [20] 开发的 Sensorstream IMU+GPS 应用程序。该应用程序读取 IMU 和压力数据，并将其流式传输到远程操作员工作站的 UDP端口。我们为 QGC 开发了自定义扩展，以接收传感器数据流，并将姿态和压力数据转换为遥控水下航行器的方向指令，以及遥控水下航行器摄像机的上升和下降指令。

VI. 实验结果

A. 最优界面

图 1 展示了一个人机接口示例，但潜在的界面设计可能有很多种。我们的目标是说明，这种“自然”的界面在某种意义上是最优的。为此，我们明确界定操作员和机器人的动作空间。由于摄像机角度不会影响机器人的运动，我们考虑一位能够执行两种动作的操作员：移动头部和移动身体（例如，通过游泳或行走）。这些动作分别表示为 $a_{head}(t)$ 和 $a_{body}(t)$，我们用 $a(t) = [a_{head}(t)\ a_{body}(t)]^\top$ 表示操作员的动作空间，其中 $\mathcal{A} \subseteq \mathbb{R}^2$ 和 $\mathcal{C}_0 \subseteq \mathbb{R}^2$。机器人的运动学模型为

$$\begin{aligned} \dot{x}_{pos}(t) &= v(t)\cos(\theta(t)) \\ \dot{y}_{pos}(t) &= v(t)\sin(\theta(t)) \\ \dot{\theta}(t) &= w(t) \end{aligned} \tag{16}$$

其中，$v(t)$为前进速度，$\omega(t)$为角速度。我们用 $=$ $[\dot{x}_{pos}\ t\ \ \dot{y}_{pos}\ t\ \ \dot{\theta}\ t\ ]^\top$ $(t)\ =\ [v(t)\ (t)]^\top$ 和 $\dot{x}(t)\ (\ )\ (\ )\ (\ )$分别表示机器人的动作空间和状态空间，其中 $\mathcal{U} \subseteq \mathbb{R}^2$ 和 $\mathcal{C}_r \subseteq \mathbb{R}^2 \times S^1$。

如问题1所述，我们的目标是找到映射g。在这种情况下，假设 g 是由 $u(t)\ =\ Ga(t)$给出的线性变换，并且在某些条件下具有若干性质，包括连续性、线性和一致性。我们考虑将机器人从初始点 $x_{initial}$ 移动到最终点 x_{final} 的任务，因为在进行环境探索时，这是最常执行的任务之一。我们的目标是找到最优接口 g 和操作员应采用的控制策略 $a(t)$。为此，我们参考式(10)和式(11)，定义以下优化问题

$$\begin{aligned} \min_{g,a}\quad & \alpha||x_{final} - x(T)||^2 + \beta \int_0^T a(t)^\top M a(t) dt \\ & + \gamma \int_0^T ||\dot{x}(t)|| dt + \delta\, dist(G, O(2)). \\ \text{S.t.}\quad & \dot{x}(t) = \begin{bmatrix} \cos(\theta(t)) & 0 \\ \sin(\theta(t)) & 0 \\ 0 & 1 \end{bmatrix} u(t), \quad x(0) = x_{initial}. \end{aligned} \tag{17}$$

其中，M 是一个正定矩阵，$O(2)$ 是大小为 2×2 的正交矩阵集合，而系数 α，β，γ 和 δ 是非负正则化系数，用于确定各项的相对重要性。第一项确保在操作员的控制策略经界面变换后，能够到达期望点。第二项衡量用户付出的努力；为使用户使用的界面舒适，对头部运动赋予的代价高于身体运动。第三项考虑机器人行进的距离，并促使其选择最优路径。第四项鼓励线性变换保持角度。因此，为更好地满足一致性准则，G 应为正交矩阵。该项表示 G与正交矩阵集合之间的弗罗贝尼乌斯范数距离，即 $||UV^\top - G||_F$，其中 $G = U\Sigma V^\top$ 是矩阵 G的奇异值分解，$||\cdot||_F$是弗罗贝尼乌斯范数。

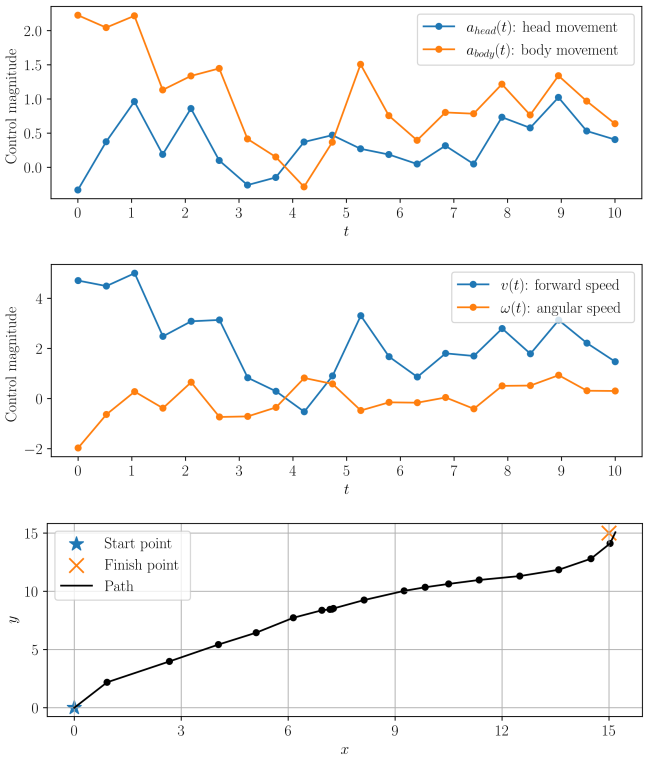

图 3. 求解问题 (17) 的实验结果。 $[0, 0, \pi/2]^\top$ $x_{final}\ =\ [15, 15, -\pi/2]^\top$ $M\ =$ $x_{initial}\ =$ $10\ 1/2$,,diag(,) $\alpha\ =$ 250 $\beta = 10$ $\gamma = 5$ $\delta = 10$ $a\ t$,,以及 。操作员执行的控制 ( )（上）；映射到机器人的控制 $u(t)$（中）；以及模拟路径任务（下）。

进行了仿真实验，以确定操作员所采用的最优动作集合（a（t），图 3 上图）；我们注意到，操作员无需过度依赖头部运动，仅通过身体运动即可完成指定任务。此外，我们还求得最优界面 g、施加于机器人的控制 u（t）$= Ga$（t）（u（t），图 3 中图），以及机器人所走的路径（x（t），图 3 下图）。更重要的是，映射 g，结果表明它

$$G = \begin{bmatrix} 0.24 & 2.15 \\ 1.73 & -0.62 \end{bmatrix} \approx \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}. \tag{18}$$

假设采用公式（18）中所表示的粗略近似，可以观察到，人类控制 $a(t)$与机器人控制 $u(t)$之间存在如下关系：

$$\begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} = u(t) = Ga(t) \approx \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} a_{head}(t) \\ a_{body}(t) \end{bmatrix}$$
$$\begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \approx 2 \begin{bmatrix} a_{body}(t) \\ a_{head}(t) \end{bmatrix}. \tag{19}$$

根据公式（17）可以推断出，“自然”的映射方式——操作员通过头部朝向控制机器人的移动，并通过自身的前向运动控制机器人的前向运动——不仅直观、便于用户使用，而且是最优的。

为了收集有关所提出方案有效性的可量化数据，我们根据 [30]的指南，由一组志愿者对该框架进行测试。实验流程包含三个任务。第一项任务中，为每位用户提供一个空泳池场景，并给予其3分钟时间熟悉头戴设备和模拟器（图4，上）。第二项任务中，为用户提供来自机器人前置摄像头的RGB视频流，并要求其通过将机器人摄像头指向相应形状，在泳池中识别一个立方体形状。泳池的一个角落中仅有一个立方体形状（图4，中）。第三项任务中，为用户提供泳池的黑白视频流，并要求其通过将机器人摄像头指向相应形状，识别一个椭圆形状（图4，下）。在此步骤中，泳池的一个角落中有一个椭圆形状，另一个角落中有一个立方体形状。关于用户指令策略的实施，我们使用了一个定长缓冲区，在仿真开始时收集足够的样本，以避免模拟遥控水下航行器进行不必要的旋转。为了移动遥控水下航行器，我们计算缓冲区的移动平均值，其中每个样本表示连续两次头部朝向测量值之间的差异。随后，将每个时间步的移动平均值与预设阈值进行比较，以决定遥控水下航行器应向哪个方向（左或右）转动。在操控过程中记录完成任务2和任务3所需的时间。

用户群体的性别分布均衡，包括三名女性和三名男性，平均年龄为24.5岁（图5）。关于头戴式显示器，实验表明，增大眼睛与智能手机之间的距离可以提升远程操控舒适度。此外，我们注意到，延迟较高的通信会显著增加任务完成时间，因为用户必须等待手机屏幕上的图像更新。正如预期，向用户呈现RGB视频流时，他们能更快地检测出形状；而黑白视频流的检测时间相对更长。任务2和3的平均完成时间及标准差分别为 $\approx 14.77 \pm 5.18$ 秒和$\approx 32.44 \pm 5.62$ 秒。

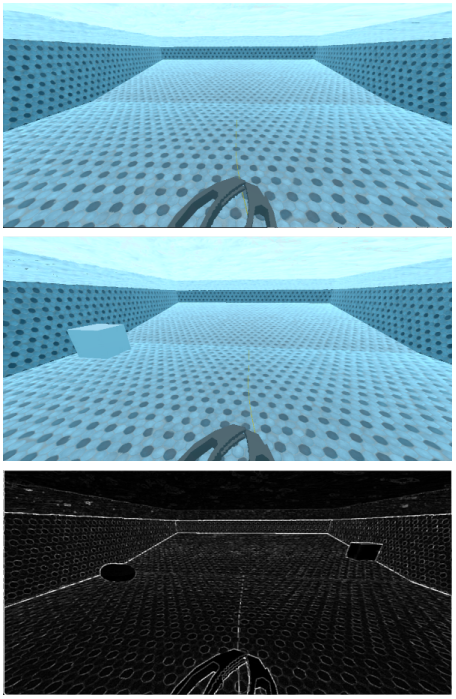

图4. 水下遥操作仿真的实验流程。任务1：熟悉空泳池（上）；任务2：通过RGB视频流寻找立方体形状（中）；任务3：通过边缘检测后的黑白视频流寻找椭圆形（下）。

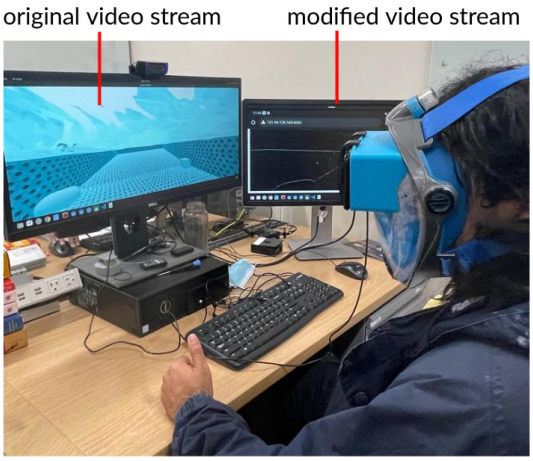

图5. 实验装置。任务1和任务2中，向用户的头戴式显示器显示原始视频流；任务3中则提供修改后的视频流。

VII. 结论与未来工作

在本文中，我们基于关键概念研究了机器人遥操作的人机界面设计。

例如线性、一致性、连续性和用户舒适度。作为一项概念验证，所提出的解决方案用于从安全且符合人体工学的位置（例如实验室环境）执行水下检查任务。我们构建了一个原型，利用人体作为控制设备，实现对遥控水下航行器的不受位置限制的遥操作。我们对一组志愿者进行了实验，以收集有关所提出解决方案有效性的可量化数据。所提出架构中的两个弱点需要重新设计。首先，缺少用户身份验证以及控制数据和视频数据加密等网络安全策略，使该架构容易受到数据窃取和遥控水下航行器未授权控制的威胁。其次，控制数据的通信方式不适用于存在延迟和断连的远距离通信。尽管如此，该原型仍为认识到这一需求奠定了基础：采用开放通信标准、促进遥控水下航行器安全遥操作的网络安全应用。未来的另一个研究方向是扩展所提出的动作空间和配置空间映射，以涵盖载具深度控制、平移和旋转运动，以及机器人夹爪操作。最后，未来工作的第三个方向是继续我们先前关于数字孪生[16]的研究，并将扩展现实用作水下机器人技术的教育培训工具。

致谢

本研究由欧盟地平线2020研究与创新计划下的欧洲研究理事会（ERC）资助（资助协议编号：825196），并获得美国国家科学基金会（NSF）资助项目 IIS-2034123、IIS-2024733 以及美国国土安全部 2017-ST-062000002 资助。

参考文献

[1] A. Barbie、N. Pech、W. Hasselbring、S. Flogel、F. Wenzhoefer、M. Wal- ter、E. Shchekinova、M. Busse、M. Turk、M. Hofbauer 和 S. Sommer。利用数字孪生原型开发水下海洋观测系统网络——波罗的海现场报告。*IEEE Internet Computing*，PP:1–1，2021年3月。[2] I. Becerra、M. Suomalainen、E. Lozano、K. J. Mimnaugh、R. Murrieta-Cid 和 S. M. LaValle。面向舒适基于虚拟现实的远程临场的人类感知优化规划。*IEEE Robotics and Automation Letters*，5(4):6489–6496，2020年。[3] Bluerobotics Inc.。BlueROV2。https://bluerobotics.com/store/rov/bluerov2/，2023年。访问日期：2023年2月2日。[4] J. Y. C. Chen、E. C. Haas 和 M. J. Barnes。远程操作机器人中的人类操作表现问题与用户界面设计。*IEEETransactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*，37(6):1231–1245，2007年。[5] R. Codd-Downey 和 M. Jenkin。使用 Li-Fi 对水下机器人进行无线远程操作。载于 *2018 IEEE International Conference on Information and Automation (ICIA)*，第859–864页，2018年。[6] J. Cohen、C. Small、A. Mellinger、J. Gallup、J. Sachs、P. Vitousek 和 H. Mooney。沿海人口估计。*Science*，278:1209–1213，1997年11月。[7] M. De la Cruz、G. Casa˜n、P. Sanz 和 R. Mar´ın Prades。用于水下作业任务的新型虚拟现实界面。*IFAC- PapersOnLine*，53:14600–14607，2020年1月。[8] Dronecode Project, Inc.。QGroundControl。http:// qgroundcontrol.com/，2023年。访问日期：2023年2月2日。[9] W. Galvani 和 P. Pereira。BlueSim。https://github.com/ bluerobotics/bluesim，2023年。访问日期：2023年2月23日。[10] GStreamer 团队。GStreamer 开源多媒体框架。https://gstreamer.freedesktop.org/，2023年。访问日期：2023年2月8日。[11] J. Halkola、M. Suomalainen、B. Sakcak、K. J. Mimnaugh、J. Kalliokoski、A. P. Chambers、T. Ojala 和 S. M. LaValle。基于倾斜控制沉浸式远程临场机器人。载于 2022 IEEE国际混合与增强现实研讨会（ISMAR），第576–583页。电气电子工程师学会，2022年。[12] K. Hauser。自由形式任务辅助遥操作中的识别、预测与规划。《自主机器人》，35，2013年11月。[13] K. Hauser。最优机器人用户界面设计。载于运动规划进展与未解问题研讨会、2014 IEEE国际机器人与自动化会议（ICRA），2014年。[14] I. Havoutis 和 S. Calinon。结合在线学习与最优控制的监督式遥操作。载于2017 IEEE国际机器人与自动化会议（ICRA），第1534–1540页，2017年。[15] M. J. Islam、C. Edge、Y. Xiao、P. Luo、M. Mehtaz、C. Morse、S. S. Enan 和 J. Sattar。水下图像语义分割：数据集与基准。载于2020 IEEE/RSJ智能机器人与系统国际会议（IROS），第1769–1776页。电气电子工程师学会，2020年。[16] T. Kaarlela、P. Padrao、T. Pitk¨aaho、S. Piesk¨a 和 L. Bobadilla。利用XR技术的数字孪生作为机器人训练工具。《机器》，11(1)，2023年。[17] H. Kaushal 和 G. Kaddoum。水下无线光通信。《IEEE开放获取》，4:1518–1547，2016年。[18] J. Kofman、X. Wu、T. Luu 和 S. Verma。使用基于视觉的人机交互界面对机器人机械臂进行遥操作。《IEEE工业电子学汇刊》，52(5):1206–1219，2005年。[19] D. Lee 和 M. Spong。具有恒定时延的无源双边遥操作。《IEEE机器人学汇刊》，22(2):269–281，2006年。[20] A. Lorenz。Sensorstream imu+gps。https://play.google. com/store/apps/details?id=de.lorenz_fenster. sensorstreamgps，2023年。访问日期：2023年2月21日。[21] A. M. MacEachren。世界的线性视图：带状地图作为一种独特的制图表达形式。《美国制图师》，13(1):7–26，1986年。[22] K. J. Mimnaugh、M. Suomalainen、I. Becerra、E. Lozano、R. Murrieta- Cid 和 S. M. LaValle。沉浸式远程临场中用户对机器人运动偏好的分析。载于2021 IEEE/RSJ智能机器人与系统国际会议（IROS），第4252–4259页，2021年。[23] M. Moniruzzaman、A. Rassau、D. Chai 和 S. M. S. Islam。移动机器人遥操作方法与增强技术：综合综述。《机器人学与自主系统》，150 103973，2022年：。[24] OpenCV.org。OpenCV主页。https://opencv.org/，2023年。访问日期：2023年2月21日。[25] P. Padrao、J. Fuentes、L. Bobadilla 和 R. N. Smith。通过强化学习估计时空场。《机器人学与人工智能前沿》，9，2022年。[26] A. Quattrini Li、I. Rekleitis、S. Manjanna、N. Kakodkar、J. Hansen、G. Dudek、L. Bobadilla、J. Anderson 和 R. N. Smith。利用异构水下机器人团队对珊瑚礁上多传感器数据进行相关性分析与比较。载于实验机器人国际研讨会，第717–728页。施普林格，2016年。[27] V. Ranadive 和 T. Sheridan。视频帧率、分辨率和灰度在海底遥控操作器控制中的权衡。载于OCEANS 81，第1222–1222页，1981年。[28] A. Shukla 和 H. Karki。机器人学在海上油气行业中的应用——综述（第二部分）。《机器人学与自主系统》，75，2015年9月。[29] B. Siciliano 和 O. Khatib。《Springer机器人学手册》。Springer- Verlag，柏林、海德堡，2007年。[30] M. Suomalainen、A. Q. Nilles 和 S. M. LaValle。面向机器人的虚拟现实。载于2020 IEEE/RSJ智能机器人与系统国际会议（IROS），第11458–11465页，2020年。[31] D. S. Terracciano、L. Bazzarello、A. Caiti、R. Costanzi 和 V. Manzari。用于水下监视的海洋机器人。《机器人学当前报告》，1，2020年12月。[32] W. Wu、C. Yang、Z. Xu、X. Wu、Y. Zhu 和 Q. Wei。人形水下机器人的开发与控制。载于2020年第六届机电一体化与机器人工程国际会议（ICMRE），第6–11页，2020年。[33] S. Zhu、X. Chen、X. Liu、G. Zhang 和 P. Tian。水下无线光通信的最新进展与展望。《量子电子学进展》，73:100274，2020年。