Non-dyadic Mobile Multi-Robot System Interaction: A Systematic Literature Review

JACK RYAN, University of Bath, UK
ADWAIT SHARMA, University of Bath, UK
MICHAEL WRIGHT, University of Bath, UK
JASON ALEXANDER, University of Bath, UK

The increasing deployment of low-cost mobile robotic platforms offers transformative potential for addressing dull, dirty, and dangerous tasks across sectors such as defence, nuclear, logistics, and agriculture. Forecasts project the global multi-robot systems (MRS) market to exceed $5.9 billion by 2028, driven by demand for scalable, autonomous solutions. As Mobile Multi-Robot Systems (MMRS) transition from research prototypes to operational assets, the design of effective human–robot interaction (HRI) becomes a critical enabler of safe, efficient, and mission-relevant deployment. Non-dyadic deployment, where the number of robots exceeds the number of human operators, is the natural extension to maximise benefit by delivering systems at increased scales. Thus, this paper is the first systematic review to focus on the application of HRI paradigms exclusively to non-dyadic MMRS.

The systematic literature review employed the PRISMA methodology, analysing 2,314 peer-reviewed papers sourced from ACM and IEEE repositories. This paper identifies three persistent gaps in the MMRS interaction literature: (1) the limited treatment of Manned-Unmanned Teaming (MUM-T) as an operational scenario, and (2) the under-utilisation of Consumer-Off-The-Shelf (COTS) wearable devices as viable interaction paradigms. Additionally, (3) the paper introduces a structured reporting framework to enhance consistency, replicability, and contextual richness in MMRS interaction studies, elevating the comparative and practical value of future work.

Identified gaps, and the proposed reporting framework, are particularly salient in light of expanding global investment and the trajectory toward increasingly complex, cross-sector MMRS deployments. By supporting structured future research directions and enhancing reporting practices, this paper contributes to advancing the state-of-the-art in MMRS interaction, fostering more impactful and practically relevant research outcomes. For researchers and practitioners alike, this review provides a foundation for designing scalable multi-robot interaction that meets the demands of tomorrow's autonomous systems, where human oversight must remain effective, as robot numbers grow.

CCS Concepts: • **Human-centered computing** → *Interaction design theory, concepts and paradigms.*

Additional Key Words and Phrases: Human-robot interaction, multi-robot systems, mobile robots, systematic, literature review

1 Introduction

1.1 Background

Mobile multi-robot systems (MMRS) are increasingly utilised across various sectors, including defence, emergency services, security, and industrial logistics [151]. The ability for a human to control and coordinate multiple robots simultaneously has become a key capability across diverse applications, especially as tasks grow in

Authors' addresses: Jack Ryan, jr2247@bath.ac.uk, University of Bath, Bath, UK; Adwait Sharma, as5339@bath.ac.uk, University of Bath, Bath, UK; Michael Wright, maew20@bath.ac.uk, University of Bath, Bath, UK; Jason Alexander, jma73@bath.ac.uk, University of Bath, Bath, UK.

ACM 2573-9522/2026/5-ART
https://doi.org/10.1145/3819075

complexity and operational environments become more unpredictable. Traditionally, research has focused on dyadic human-robot interactions (HRI), where one operator controls one robot. However, robotic platforms are being deployed in increasingly large numbers [2], forming growing MMRS which utilise system scalability to increase capability. In these modern non-dyadic contexts, the number of robots routinely exceeds the number of human operators, presenting unique interaction challenges. In this review, non-dyadic interaction is defined solely by a robot–operator ratio greater than one, i.e., scenarios in which the number of robotic platforms exceeds the number of human operators. This definition is agnostic to the number of operators involved and does not impose assumptions about whether operator roles or responsibilities are static or dynamic, which are treated as contextual variables within the reviewed literature.

To support human interaction, there is a simultaneous and ongoing development in advanced interaction paradigms such as gesture-based control [7, 18, 82], brain-computer interfaces (BCIs) [9, 58, 67], virtual reality (VR) [85, 107, 116], and wearables [16, 39, 136] which are enabling operators to interact with multiple robots in novel ways. Despite these advances, there remains a significant gap in understanding how these interaction techniques perform in non-dyadic settings with MMRS, where a single operator must manage several robots across various operational domains. Huang et al. [56] identified the limiting factor in human-multi-robot teams as "user interface design and the cognitive load of the human supervisor". This key challenge highlights the need for further research into HRI with MMRS.

Overcoming this knowledge gap, and realising non-dyadic MMRS deployments as an everyday tool, will revolutionise operations across a range of mission-critical applications. The ability for a single operator to manage and collaborate with multiple robots enables significant gains in scalability, efficiency, and responsiveness. In disaster zones, swarms of ground or aerial robots can rapidly survey damage, locate survivors, and deliver aid with minimal human risk. In agriculture, MMRS can autonomously monitor crops, deploy treatments, and optimise yields at scale. Defence, infrastructure inspection, and warehouse logistics likewise stand to benefit from MMRS' capacity to perform complex, distributed tasks simultaneously, with reduced manpower and greater precision. These benefits are only realisable, however, with interaction paradigms that allow humans to remain situationally aware, in control, and confident in their decision-making, despite the increasing autonomy and number of platforms involved.

As MMRS continue to grow in scale and relevance, so too does the need to understand how interaction paradigms must evolve to support their effective deployment. This review seeks to bridge that gap by systematically analysing existing research on human interaction with MMRS in non-dyadic contexts. By exploring current trends, identifying technological enablers, and exposing new research challenges, this paper aims to provide a foundation of knowledge for the next generation of HRI paradigms in MMRS.

1.2 Problem Statement

In non-dyadic contexts, operators face a wide range of challenges, including increased cognitive load, reduced situational awareness, and heightened interaction complexity. These challenges emerge as the number of robots under an operator's control increases, particularly in dynamic and unpredictable environments. Unlike dyadic interaction, where attention and control can be devoted to a single system, MMRS demand that an operator divide their cognitive resources across multiple uncrewed systems, each potentially carrying out different tasks in different locations. The need to monitor system health, interpret status updates, make mission-critical decisions, and intervene across multiple platforms simultaneously places significant strain on human cognitive and perceptual capacities [75]. These demands are compounded in mobile systems, where spatial reasoning, coordination, and navigation introduce further layers of complexity compared to static robotic installations [151].

Crucially, these challenges manifest through the interaction itself, which becomes a limiting factor in effective MMRS supervision and control. The requirement to provide timely identification and response to individual robots within a larger group, and to oversee MMRS while engaged in alternative primary tasks, increases the risk of cognitive overload and delayed or suboptimal interventions. As MMRS scale, the design of effective interaction modalities becomes critical in enabling operators to manage these systems safely and efficiently.

Recent survey work has begun to consolidate understanding of specific interaction technologies relevant to HRI at scale. For example, Wang et al. provide a comprehensive systematic review of XR-enabled remote HRI systems, synthesising interaction modalities, interface designs, and evaluation practices across a broad set of predominantly dyadic and small-scale scenarios [145]. While such surveys offer valuable insights into the role of XR in remote HRI, they do not explicitly frame interaction challenges around non-dyadic control or analyse how interaction paradigms must adapt as the robot-to-operator ratio exceeds one, particularly for MMRS.

Thus, this paper seeks to review the state-of-the-art in interaction paradigms applied to MMRS, in order to identify opportunities and challenges in the academic literature. This work is intended to guide future research directions in optimising HRI in non-dyadic scenarios with MMRS.

To the best of our knowledge, this is the first systematic literature review to specifically examine *non-dyadic* interaction with *MMRS*. This work builds upon prior foundational studies, such as the survey by Dahiya et al. [28], which used a systems and interaction graph approach to classify multi-agent systems and explore computational characteristics of robot control. However, previous studies have not focused specifically on the interaction challenges that arise when the number of robots per human operator exceeds one (robot-operator ratio > 1), a framing particularly critical in scalable MMRS deployments. Moreover, while prior surveys encompass a broad range of robotic systems, our review specifically focuses on *mobile* platforms, where motion, spatial coordination, and task variability are themselves unique HRI challenges distinct from those of stationary systems such as robot arms.

In this paper, we take a technology-driven approach to analysis, surveying how advanced interaction paradigms (such as gesture-based input, Extended Reality (XR), and BCIs) are enabling new forms of control and supervision in non-dyadic MMRS contexts. We provide a consolidated understanding of the current state of research and establish a foundation for innovation in future interaction design for scalable robotic systems.

The primary objective of this paper is to understand the current state-of-the-art interaction paradigms used in non-dyadic MMRS. The specific objectives are:

- To identify the opportunities and challenges in non-dyadic MMRS HRI in the academic literature.
- To provide recommendations for future research directions aimed at improving HRI in large-scale, non-dyadic MMRS across a diverse set of applications.

1.3 Overview of the Paper

Based on a systematic literature review of major academic repositories, this paper applies both a contextual (top-down) analysis and a technological (bottom-up) analysis to the complex field of MMRS HRI. These two perspectives provide: (1) a thorough understanding of the current research landscape; (2) identification of key research challenges and gaps; and (3) the proposal of a new reporting framework for future MMRS HRI research. The paper is structured as follows: Section 2 outlines the review methodology. Section 3 presents the contextual analysis, examining operator roles, platform characteristics, and interaction dynamics. Section 4 delivers the technological analysis of interaction techniques. Section 5 summarises key findings and introduces the proposed reporting framework. Section 6 discusses parallel challenges to MMRS HRI technique development, including platform autonomy levels, team dynamics, and multi-modal interaction approaches. Finally, Section 7 offers concluding remarks, limitations, and recommendations for future research.

2 Methodology

2.1 Approach

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework [108]. The PRISMA flow diagram presented in Figure 1 illustrates the process from the initial search to the final selection of papers. This process consisted of four main stages:

(1) **Identification:** A comprehensive search of the ACM Digital Library and IEEE Xplore databases yielded 2,314 records, with 2,303 unique results. The search strategy is outlined in Section 2.2.
(2) **Filtering:** Titles and abstracts were screened against inclusion and exclusion criteria, presented in Section 2.3, removing studies not focused on MRS or interaction paradigms. This stage retained 239 papers.
(3) **Screening:** Full-text reviews, described in Section 2.4, ensured alignment with inclusion criteria, particularly on mobility and publication type. This reduced the set to 139 eligible studies.
(4) **Inclusion:** These 139 papers were selected for in-depth analysis, reflecting a 6.00% acceptance rate from the initial 2,314. The data extraction from the included papers is described in Section 2.5.

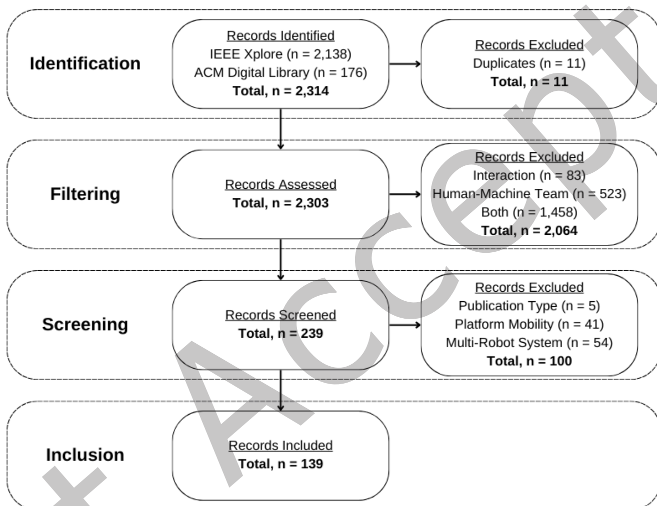

Fig. 1. PRISMA flow diagram showing how 2,314 records from ACM and IEEE were reduced to 139 included papers.

2.2 Search Strategy

Searches were conducted within the two most relevant academic research repositories, the ACM Digital Library and the IEEE Xplore database. These repositories were chosen due to their extensive collections of peer-reviewed papers in HRI and MMRS, by virtue of hosting the largest conferences and journals in these fields.

The search strategy was constructed to comprehensively capture literature at the intersection of six core dimensions: (1) robotics and autonomous systems, (2) multi-robot systems, (3) human-robot interaction, (4) interaction paradigms, (5) timeliness of research, and (6) publication quality. The set of keywords under each dimension was developed iteratively through collaborative workshops within the academic team, drawing on domain expertise and prior literature. As the search progressed, additional relevant terms were identified from paper abstracts and keyword metadata, and the search string was refined accordingly to improve recall without sacrificing relevance.

The final search strategy comprised the following six components:

- **Component 1:** A broad reference to robotic or uncrewed platforms, including terms such as "robot*", "unmanned", "uncrewed", "UAV", "UGV", "USV", "UUV", "autonomous", "platform*", "vehicle*", and "drone*".
- **Component 2:** A focus on studies concerning multiple platforms, using terms like "multi?robot*", "swarm*", "collective*", "team*", "distributed", "cooperative", and "collaborative".
- **Component 3:** Emphasis on HRI, encompassing terms such as "interface*", "interaction*", "command and control", "teleoperat*", "remote operation", and "user experience".
- **Component 4:** A specific focus on interaction paradigms, incorporating terms such as "touch*", "gesture*", "haptic", "augmented reality", "virtual reality", "brain-computer interface", and "natural language processing".
- **Component 5**: Relevance in timeliness of publication through a time filter as only articles published since the turn of the millennia are included, between January 1, 2000 and April 1, 2024.
- **Component 6:** Publication type was limited to full conference and journal articles, excluding short papers, posters, and technical reports. This ensures only high-quality, and peer-reviewed, papers were included.

The final search strings and formatting details for both repositories are provided in Appendix A. The search returned 2,314 papers: 176 from ACM, and 2,138 from IEEE. The disproportionate distribution between the two repositories may reflect broader topical coverage in IEEE for robotic systems and control engineering. After removing 11 duplicates, 2,303 papers were subject to further filtering.

2.3 Inclusion and Exclusion Criteria

The following inclusion and exclusion criteria were applied to the papers returned by the search engines.

Inclusion Criteria

- **Human-Robot Team and/or Multi-Robot System:** Papers involving two or more autonomous or uncrewed robots, in non-dyadic settings (robot-operator ratio > 1).
- **Interaction Paradigms as the Primary Focus:** Only papers in which the development and/or evaluation of interaction paradigms was the core contribution are included.
- **Platform Mobility:** Only mobile robots (UAV, UGV, USV, UUV) are included.

Exclusion Criteria

- **Single-Robot Systems:** Papers that involved only one robot were excluded.
- **1-to-1 Operator-Robot Mappings:** Papers in which each robot was controlled by a dedicated human operator (1-to-1 interaction), even if multiple robots were involved, were excluded.
- **Incidental Interaction:** Papers mentioning interaction without substantive focus were excluded.

Applying these criteria, 2,064 papers were excluded (1,458 for failing both criteria, 83 for lacking interaction focus, and 523 for not being MRS-related), resulting in 239 papers proceeding to the screening stage (10.76% acceptance).

2.4 Screening Process

The screening process involved a full review of each paper that passed through the initial filtering stage. We used this to ensure the correct application of the inclusion and exclusion criteria and to verify that the actual content of the research aligned with the claims made in the abstract.

A common issue encountered during this stage was the discrepancy between the abstract and the full content of the paper. In 54 cases, the plurality of robots referred to in the abstract did not reflect the actual implementation described in the research, where only a single robot was involved. These papers did not fulfil the MRS requirement and so were removed. Additionally, particular attention was given to the application of the publication type filter.

Despite including this filter in the initial search, the databases returned extended abstracts, contents pages, and other documents that did not meet the requisite standards of academic rigour. It was only possible to ascertain this on a full read of the papers, thus they were removed during the screening process.

100 records were excluded throughout the full-text screening. The reasons for exclusion comprised: 5 were not full academic papers; 41 lacked mobile platforms; and 54 were not considered MMRS. By the end of the screening process, 139 papers remained from the initial 2,314 search results, representing a 6.00% acceptance rate overall. These 139 papers form the core set for the literature review.

2.5 Data Extraction and Analysis

From the final set of 139 papers, data extraction was performed using a structured coding framework. Each paper was read in full, and relevant information was systematically recorded in a bespoke spreadsheet using predefined data categories including: empiricism, interaction paradigm, autonomy level, and robot-operator ratio. Free-text notes captured additional methodological nuances and context-specific insights. This approach enabled both quantitative aggregation and qualitative thematic analysis, ensuring a comprehensive and reproducible synthesis of the literature. The selection and importance of these categories is discussed further in Section 5.3, where they form the basis of our proposed reporting framework. Specifically, the following data categories were extracted:

- **Empiricism:** Yes/No classification of whether the study was grounded in empirical data (e.g., through experimental testing with robots and human operators) or whether it was more theoretical in nature.
- **Type of Testing:** Classification of whether testing was conducted in physical environments using actual robotic platforms; only in simulated environments; or if no testing took place at all.
- **Interaction Paradigm:** Documenting the diversity of interaction paradigms, including traditional control (such as joystick or GUI-based interfaces) as well as novel paradigms (such as BCIs, gesture control, voice commands, haptic feedback, and XR interfaces).
- **Application Context:** The application domains for the interaction paradigms. This encompassed the specific operational context in which the robots were being used, providing insights into the performance requirements and constraints of the interaction paradigms in different settings.
- **Operational Domain:** Referring to the type of robotic systems being studied, categorised into aerial (UAVs), ground (UGVs), surface (USVs), and underwater (UUVs) domains. The domain of operation was important to distinguish the different challenges and needs for HRI across these diverse environments, which have varying levels of complexity.
- **Number and Ratio of Operators and Robots:** Non-dyadic interaction contexts necessitate that the number of robots exceeds the number of human operators. Therefore, the number of robots and the number of operators in each study were extracted. From these figures, the robot-operator ratio is derived.
- **Operational Positionality:** Identified whether the operator was within line-of-sight (LOS) of the platforms being controlled, or were operating beyond-visual-line-of-sight (BVLOS).
- **Operational Skill:** Consideration given to the skill-set of the operator, by classification as trained or untrained, irrespective of how minimal the training requirement might be.
- **Autonomy Level:** Platforms exhibit different levels of autonomy, with the classification system defined in Section 6.1, resulting in a numeric score between 0 and 5.
- **Description of MMRS:** The specific terminology used by authors to describe the collective robotic system, such as *swarm*, *fleet*, *squad*, or *team*, was also recorded.
- **Operator Challenges:** Explicit identification of challenges faced by operators in controlling or overseeing the robotics platforms, for example operator cognitive overload.

3 Contextual Analysis

This section examines the overarching contextual factors that shape non-dyadic HRI, from a top-down perspective. These factors were systematically explored through four key research questions:

- **What applications are MMRS being applied to? (Section 3.1)** - Which application domains (e.g., disaster response, agriculture, security) drive the requirements for interaction paradigms, and how do these domains impose functional or environmental constraints?
- **Who interacts with MMRS? (Section 3.2)** - What are the roles, team structures, and expertise levels of operators in non-dyadic scenarios? How do collaborative operator setups influence control strategies?
- **Where are the operators located in relation to the robots? (Section 3.3)** - How does spatial separation (e.g., co-located, line-of-sight, remote, VR-mediated) affect interaction modality choice and system responsiveness?
- **How do operators interact with MMRS? (Section 3.4)** - What techniques are used to distribute cognitive load, assign control authority, and enable scalability in controlling many robots simultaneously?

These four questions provide a structured analytical lens that emphasises the human-side of interaction and the operational demands placed on MMRS in real-world deployments. Each of the following subsections addresses one of these questions in turn, drawing on findings from the reviewed literature to unpack the implications for MMRS interaction design.

3.1 What applications are MMRS being applied to?

We identified Search and Rescue (SAR), and Military and Defence use cases as the most commonly studied. Figure 2 shows the distribution of the most common application contexts identified in our paper set, presenting 45 of the 139 (32%) papers analysed. For clarity, only application contexts with an occurrence of three or more papers are shown. All other papers either specified more unique application contexts (39%) or did not specify an application context at all (18%). Notably, beyond SAR or Military use, all other categories were similarly represented at low levels (with four or fewer papers), indicating relatively little depth of research in applied HRI for MMRS across most application contexts.

Application context is a key driver of requirements for interaction system design, providing functional and environmental constraints. Where SAR and military use are prominent, this infers HRI designed for deployment of dispersed robotic systems over wide areas for tasks akin to wide area surveillance.

However, the contextual links between the research presented in papers and indicated application contexts, can be tenuous. As an example, Jeston-Fenton *et al.* present an Android-based platform for visualising swarm metrics such as grouping, alignment, fragmentation, and coverage in real time during human-swarm interaction [63]. Their use of boid-based swarming and interactive “breadcrumb” controls offers valuable insights into how users might influence and interpret swarm behaviours. However the work lacks grounding in a specific operational context. The experiments are framed generically in terms of “coverage missions” without reference to real-world scenarios (such as search-and-rescue operations, environmental monitoring, or military patrols) where such metrics would directly inform human decision-making. This exemplifies a broader issue identified in the literature. While HRI research often introduces compelling interaction paradigms or interfaces, it frequently abstracts away from the mission-specific requirements, constraints, and utility that would validate the relevance of such systems in applied MMRS contexts. In our dataset, for instance, 25 papers (18%) provided vague or no application context, categorised as “None” in our analysis, while a further 54 papers (39%) were grouped as “Other” reflecting niche or unclear operational framings. Combined, this indicates that over half of the papers reviewed lacked clear anchoring to specific, mission-focused application domains. Without this contextual anchoring, the practical value of the interaction method, however well designed, remains speculative.

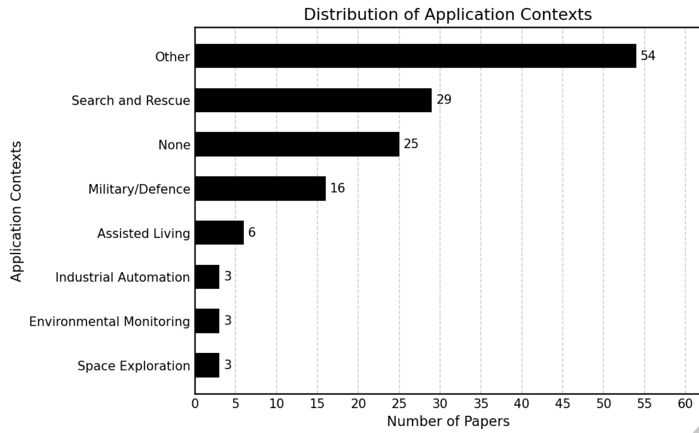

Fig. 2. Distribution of application contexts - showing the wide distribution of applications, with Search and Rescue being most frequently occurring.

3.2 Who interacts with MMRS?

The operator's themselves, their background, training and expertise, are key variables shaping MMRS interaction. This subsection focuses specifically on operator expertise, particularly the distinction between trained and untrained operators, as a contextual factor in shaping interaction outcomes. The majority of MMRS control scenarios analysed in this review involved individual operators, with only 7% of studies examining collaborative team-based control. Broader reflections on operator team structures are presented in Section 6.2.

To assess expertise levels, operators were classified as either "trained" or "untrained". A trained operator is defined as having prior experience in robotic control or having received explicit training as part of the study protocol. If no such experience or training was described, or if the study explicitly stated otherwise, the operator was considered untrained.

Only 32% of studies in the review involved trained operators, with the remaining 68% using untrained participants. While the impact of this split is difficult to quantify, prior literature in human-autonomy teaming highlights that operator experience significantly affects performance, interface preferences, and error rates [19, 96].

Two examples illustrate this disparity. In Dudek and Schulte's study on UAV task delegation, trained fighter-jet pilots conducted twelve simulated missions following a dedicated three-hour training programme [37]. By contrast, Abdi and Paley's work on gesture and haptic control of an aerial swarm used untrained operators, specifically the authors themselves, with no structured training or performance baseline [1]. While both papers contribute meaningful insights into HRI methods, the latter typifies a broader issue in the field: interaction paradigms are often evaluated in isolation from operational context and user expertise, limiting their practical applicability.

3.3 Where are the operators located in relation to the robots?

Operator positionality, or where the operator is located relative to the robotic platforms during interaction, was assessed across three categories: **Onboard** (the operator is co-located on a platform within the MMRS), **Line of Sight (LOS)** (the operator has a direct visual of the robots), and **Beyond Line of Sight (BLOS)** (the robots are outside the operator's direct view, requiring mediated feedback through screens, headsets, or interfaces).

Overall, most papers examined interaction with MMRS in LOS or BLOS contexts, with 47% featuring BLOS control, 29% LOS, and only 2% considering onboard operation (see Figure 3). A notable 22% of papers did

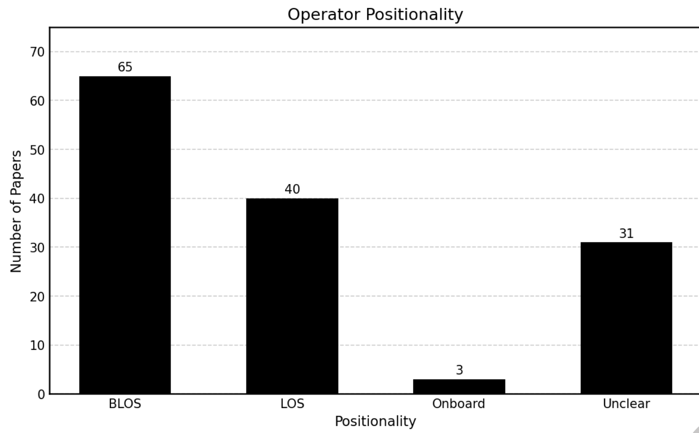

Fig. 3. Distribution of operator positionality in reviewed papers. BLOS: Beyond Line of Sight, LOS: Line of Sight.

not articulate operator positionality, particularly when interaction was evaluated only in simulation without real-world operational framing.

BLOS scenarios were the most frequently studied (47%), reflecting the need to supervise and control MMRS remotely using abstracted data representations such as 2D maps or XR overlays. For instance, VR headsets are sometimes used even when operators are near the platforms, effectively shifting an LOS scenario to BLOS by obscuring direct vision of the robots during operation [54].

LOS studies explored interactions where the operator directly observes the robots, such as UAV swarms operating in a confined environment or collaborative UGV tasks in an industrial or field setting. These papers leveraged gesture-based or tangible interfaces, benefitting from the direct perception of robot actions for effective command and supervision [105].

Onboard operation was considered in only 3 papers (2%), framed under "Manned-Unmanned Teaming" (MUM-T). For example, one paper explored pilots delegating tasks to UAVs via GUIs, voice commands, and gaze tracking, concluding that voice commands provided the best performance and resilience [37]. Another paper examined a pilot coordinating six UAVs using a tactical monitor while supported by a ground control station [70]. In a third case, operators inside an Autonomous Ground Vehicle supervised a UAV and UGV, with findings indicating that combining tactile and visual alerts supported multitasking without increasing perceived workload [114]. While these studies offer valuable insights into MUM-T interaction, onboard contexts remain underexplored within MMRS research.

A substantial portion (22%) of papers were categorised as **unclear** regarding operator positionality. This typically occurred in studies trialling new interaction paradigms exclusively in simulation, without specifying or constraining operator positioning in a plausible real-world scenario. For example, third-person, birds-eye 3D views of UAVs in collaborative manipulation simulations are feasible for evaluating interaction designs but are impractical in real operations, unlike GPS-based map overviews used for BLOS supervision [97]. This lack of contextual clarity limits the interpretability of findings and the practical transfer of proposed interaction paradigms to applied MMRS settings.

3.4 How do operators interact with MMRS?

Understanding how operators interact with MMRS, particularly roles and control strategies, is central to advancing scalable HRI. Operators typically engage with MMRS in either "active" or "supervisory" roles, a distinction shaped

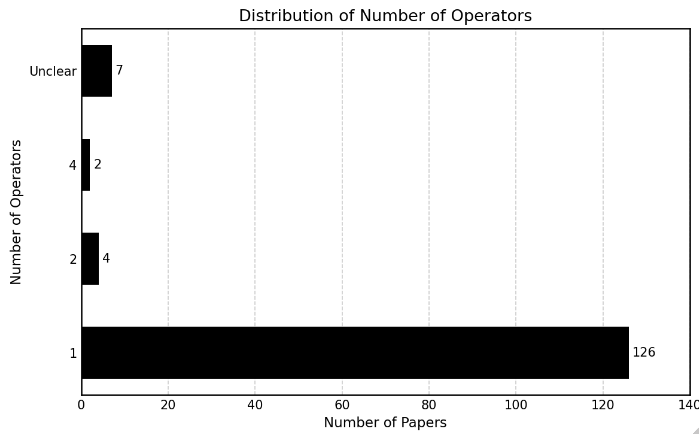

Fig. 4. Distribution of number of operators - showing that 1 operator is used in 126/139 papers surveyed.

by system autonomy levels [54]. Active roles involve direct control of robot actions (human-in-the-loop), while supervisory roles focus on higher-level tasking and oversight (human-on-the-loop). Achieving scalable MMRS control necessitates moving from dyadic, direct control to supervisory oversight, supported by increased autonomy of the robot team.

Figure 4 shows that in the vast majority of reviewed papers (91%), a single operator was responsible for controlling the MMRS, with only 3% involving two operators, 2% using four operators, and 5% unclear. This distribution strongly suggests that most studies assume supervisory rather than active control, as it is practically impossible for a single person to actively control multiple complex mobile robots at the same time (just as one cannot drive several cars simultaneously). Consequently, while supervisory MMRS control has seen substantial exploration, there is limited research addressing the structuring of high-performance human teams for active MMRS control, where distributing cognitive and control load across multiple operators is essential.

For example, the ARROCH system by Chandan *et al.* [21] enables a single user to supervise a robot team through augmented reality, providing high-level guidance while robots execute tasks autonomously. Conversely, Walker *et al.* [141] demonstrate a rare instance of multi-operator active control, using three operators assigned to piloting, communications, and mixed reality supervision during an outdoor mission.

These examples highlight that while supervisory control for MMRS is progressing, comparatively little research addresses sharing cognitive load and control authority across operator teams in active control scenarios. This challenge is further explored in Section 6.2, which discusses robot-operator ratios, role delineation, and human-robot team structuring for scalable MMRS deployment.

4 Technological Analysis

A technological analysis of MMRS interaction adopts a bottom-up approach, examining individual interaction paradigms and enabling technologies in detail. This approach focuses on the specific tools and modalities used to facilitate effective HRI, highlighting their respective strengths, limitations, and suitability for non-dyadic control scenarios. It further identifies the relative research intensity associated with each interaction paradigm within the MMRS literature.

This analysis is guided by two key research questions:

(1) How do the **techniques**, **capabilities**, and **limitations** of each interaction paradigm influence their ability to address core challenges in MMRS control, including operator workload, situational awareness, and scalable command of heterogeneous multi-robot teams?
(2) What is the **relative suitability** of different interaction paradigms for non-dyadic MMRS control, particularly in terms of their **integration potential**, their ability to support **scalability** (i.e., high robot-to-operator ratios), and the **operational challenges** they introduce in practical deployment?

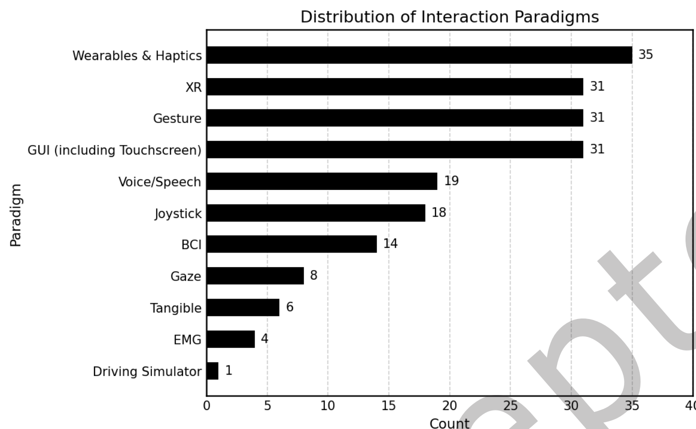

Fig. 5. Distribution of interaction paradigms - showing the distribution from popular to fringe interaction in the MMRS context.

Interaction paradigms vary widely in their design and application, ranging from tactile hardware such as joysticks to immersive software environments including XR. Each paradigm contributes uniquely to addressing operator workload, situational awareness, and scalability challenges, as identified in the Section 3 contextual analysis. The most frequently employed paradigms, as shown in Figure 5, include “wearables and haptics,” observed in 25% of the reviewed studies, followed by XR (including Augmented Reality (AR), Mixed Reality (MR), and VR), gesture control, and graphical user interfaces (GUIs), each appearing in approximately 22% of the papers. Notably, the total number of occurrences exceeds the 139 papers analysed, as many studies employed multi-modal approaches and reported the use of multiple paradigms. The concentration of studies around a small set of interaction paradigms suggests that MMRS HRI research remains focused on a limited design space, with relatively little exploration of how alternative or hybrid paradigms might better support scalability and non-dyadic supervision.

A comparative analysis of these paradigms was developed through a synthesis of all 139 reviewed papers, focusing on studies that implemented, evaluated, or discussed each respective interaction paradigm in MMRS contexts. The comparative assessment presented in Table 1 is derived from recurring patterns reported across the literature, rather than from an external or subjective evaluation. In this context, *suitability* reflects how well a paradigm’s interaction characteristics align with MMRS control demands such as precision, abstraction level, and operator workload; *integration* captures the extent to which a paradigm is reported to interface with existing robotic architectures, middleware, and multi-modal control frameworks; *scalability* reflects the degree to which a paradigm supports increasing robot-to-operator ratios, either directly or through abstraction and autonomy; and *challenges* summarise recurring limitations, trade-offs, and operational constraints identified across the papers reviewed.

Noting, whilst other interaction modalities such as gaze tracking [26, 90], mobile and web-based control [63, 126], tangible interfaces [32, 48, 128], deformable interfaces [34], and driving simulators [83] are also present in the literature, these appeared too infrequently to warrant dedicated subsections. The analysis therefore focuses on the most prevalent paradigms to provide meaningful insight into the current state-of-the-art in MMRS interaction.

The subsequent subsections explore each interaction paradigm in detail: joysticks and controllers (Section 4.1); wearables and haptics (Section 4.2); gesture control (Section 4.3); extended reality (Section 4.4); voice interfaces (Section 4.5); and brain-computer interfaces (Section 4.6). Each subsection outlines the technique's operational characteristics, capabilities, and limitations. Taken together, the comparison reveals a trade-off between interaction richness and scalability, where paradigms offering fine-grained control tend to increase operator workload at scale, while more abstracted paradigms shift cognitive and trust-related demands onto the operator.

4.1 Joysticks and Controllers

4.1.1 Technique. Joysticks and controllers represent a **direct manipulation interaction paradigm** characterised by hand-held devices that translate physical movements into continuous control inputs. This paradigm is commonly employed in traditional robotic teleoperation, leveraging analogue sticks, buttons, and triggers to enable real-time control over robotic systems. In MMRS contexts, joysticks are predominantly utilised in low-autonomy configurations or direct teleoperation modes, particularly where high-fidelity control of individual robots is required, or when direct supervisory override is necessary. Examples of joystick-based controllers are shown in Figure 6.

Fig. 6. Joystick controller examples - to include consumer grade (left); custom (middle); and commercial grade (right).

Some of the earliest included papers [73] used a joystick controller, and this has continued in various forms throughout the study period [5, 124, 150, 154]. Joysticks were observed as a primary interaction paradigm in only 13% of the papers analysed, but took a number of different forms, including:

- **Game Controllers** [3, 67]
- **XR-based Hand Controllers** [98, 111]
- **Multi-DOF Controllers** [73, 154]

4.1.2 Capabilities. The continued prevalence of joystick-based paradigms in MMRS control is tied to their longstanding use in teleoperation scenarios, particularly in 1:1 robot-to-operator contexts [3]. One key advantage is their low-cost accessibility, as operators can leverage widely available commercial gaming hardware, such as Xbox or PlayStation controllers. This reduces overall development costs and simplifies hardware procurement, while also benefiting from robust and well-established supply chains. Another contributing factor is the familiarity of these devices; many operators have prior exposure to gaming controllers, which reduces cognitive barriers

Technique	References	Suitability	Integration	Scalability	Challenges
Joysticks & Controllers	[3, 5, 9, 10, 41, 50, 67, 73, 86, 98, 107, 111, 116, 120, 124, 150, 154, 155]	Highly suitable for low to moderate robot counts where direct control is needed. Best suited for applications requiring precise control and low latency.	Easily integrated with existing control architectures using standard protocols. Often requires minimal modification to robot firmware and interfaces.	Limited scalability due to the direct control nature. Effective for small teams but becomes impractical for large swarms due to limited control channels.	Operator cognitive load increases with robot count.
Wearables & Haptics	[1, 3, 4, 15–17, 24, 36, 39, 42, 49, 51, 52, 60, 68, 69, 86, 90, 92, 93, 97, 114, 115, 117–120, 124, 131, 136, 137, 142, 143, 146, 150]	Suitable for intuitive control and feedback in moderate to complex MMRS environments. Best for applications involving close operator-robot interaction, such as telepresence and hazard scenarios.	Integration complexity varies with device sophistication. Common standards (e.g., Bluetooth, Zigbee) help, but real-time feedback synchronization can be challenging.	Limited by the number of sensory channels and operator feedback resolution. Effective for moderate robot groups but less suited for large-scale operations without abstraction layers.	Balancing sensory feedback without overwhelming the operator.
Gestures	[7, 11, 15, 17, 18, 26, 31, 36, 42, 46, 47, 49, 64, 65, 68, 77, 82, 100, 104, 105, 107, 113, 122, 123, 130, 131, 133, 140, 144, 146, 149]	Suitable for intuitive, high-level control, especially in small to moderate MMRS. Best for tactical tasks where operators need a hands-free method of interaction.	Requires integration of computer vision and gesture recognition software. Effective for line-of-sight operations but challenging for complex scenarios with occlusions.	Limited by the accuracy of gesture interpretation and control mapping complexity. Works better with smaller robot groups where symbolic gestures can convey group commands.	Consistency in gesture recognition under varying environmental conditions.
Extended Reality (XR)	[21, 22, 25, 30, 31, 36, 38–40, 46, 54, 56, 61, 66, 85, 98, 99, 107, 110, 111, 113, 116, 127, 129, 132, 140–142, 146, 149, 152]	Highly suitable for supervisory control, mission planning, and immersive control interfaces. Best for complex MMRS scenarios requiring spatial awareness and enhanced situational understanding.	Requires high-performance graphical processing units (GPUs) and low-latency networks. Modern XR platforms support standard robotics middleware like ROS, improving integration ease.	Moderate scalability due to graphical limitations and real-time rendering constraints. Best when combined with autonomous behaviours and abstracted control layers for large robot teams.	Hardware limitations, especially for AR in field conditions.
Voice	[16–18, 31–33, 35, 37, 43, 46, 47, 53, 64, 98, 101, 106, 107, 109, 112, 113]	Best suited for high-level task delegation rather than fine-grained control. Effective in structured environments with limited background noise.	Relatively straightforward using voice command software and NLP engines (e.g., Alexa Skills, Google Assistant APIs). Accuracy can be affected by environmental noise, requiring noise-cancelling systems.	Limited for fine control but scalable for issuing broad task-level instructions to MMRS. Works best when paired with a hierarchical command system.	Speech ambiguity, multilingual support, and noise interference.
Brain-Computer Interface (BCI)	[9, 20, 23, 29, 55, 58, 62, 67, 71, 72, 79, 80, 88, 148]	Suitable for specialised use cases requiring hands-free control, such as in medical or defence scenarios. Often experimental and requires significant calibration for real-world deployments.	Complex integration requiring specialised hardware and signal processing pipelines. EEG-based systems need highly controlled environments for reliable data interpretation.	Currently limited to controlling single to small groups of robots due to signal complexity. Scalability is hindered by signal accuracy and cognitive strain on the user.	Signal accuracy, real-time responsiveness, and user training.

Table 1. Comparative Analysis of Interaction Techniques.

during initial training and promotes faster operator proficiency. Additionally, joysticks enable high-precision, low-latency control, making them particularly effective for MMRS applications that require fine-grained manipulation or navigation control of individual robots.

Beyond standard game controllers, specialised implementations of the joystick paradigm have been explored in high-fidelity use cases. A notable example is the European Space Agency's "haptic telepresence" system, which employed a bespoke three-degree-of-freedom joystick with integrated haptic feedback to overcome the sensory and control challenges associated with microgravity on the International Space Station [120]. However, such highly tailored solutions remain relatively uncommon in MMRS deployments, where simpler, off-the-shelf controllers are typical.

4.1.3 Limitations. Despite these advantages, joystick-based interaction paradigms present several significant limitations when applied to MMRS, especially in non-dyadic contexts with high robot-to-operator ratios. A primary limitation is their poor scalability. Since joysticks are designed for direct control of individual platforms, operator cognitive workload increases rapidly as the number of controllable robots grows, rendering the paradigm unsuitable for large-scale deployments [3]. Additionally, joystick paradigms offer limited situational feedback. They do not inherently provide operators with environmental or swarm status information, which necessitates reliance on separate visual or auditory interfaces and can fragment operator attention across multiple display systems. Finally, the joystick paradigm is predominantly associated with low-level control inputs, such as velocity and heading adjustments. This characteristic makes them ill-suited to higher-autonomy MMRS scenarios where abstracted, task-based, or supervisory-level interactions are preferable, limiting their applicability in modern, scalable multi-robot systems.

4.2 Wearables and Haptics

4.2.1 Technique. Wearables and haptics constitute an **embodied interaction paradigm** where human operators engage with robotic systems through body-worn devices and tactile feedback mechanisms. Wearables are defined as body-mounted devices that capture human movement or physiological signals through sensors such as inertial measurement units (IMUs), electromyography (EMG), or capacitive touch inputs. Haptics refers to interaction modalities that provide physical feedback to the user, typically via vibration or force cues, enhancing situational awareness by engaging the tactile sensory channel.

These two modalities are grouped within a single interaction paradigm for MMRS control because they are most commonly co-implemented within the same systems. In the majority of the MMRS literature reviewed, haptic feedback was delivered via wearable devices such as gloves, armbands, or torso suits. While a smaller number of studies integrated haptic feedback through alternative interfaces, such as haptic-enabled joysticks, the overall trend was for haptics to be embedded within wearable devices. Given this observed overlap in implementation, it was not appropriate to categorise haptics as a fully standalone interaction paradigm within the MMRS context.

From the 139 papers analysed, examples of wearable-based paradigms include:

- **Armbands**, incorporating EMG, IMU, and vibrotactile feedback modalities [1, 3, 47, 49, 114, 119, 131].
- **Haptic torso suits**, providing distributed vibrotactile feedback across the body [1].
- **Gloves**, which enable hand gesture recognition with haptic feedback on the fingertips/ palms [15, 136, 137].

Although consumer-grade wearables such as smartwatches and smart rings are increasingly prevalent in adjacent sectors [121], no MMRS studies analysed employed such devices. Instead, research predominantly used bespoke or laboratory-grade wearables with extended sensing and feedback capabilities, reflecting a current gap between consumer/commercial off-the-shelf (COTS) wearable markets and MMRS control applications.

4.2.2 Capabilities. Wearable and haptic interaction paradigms offer distinct advantages for MMRS control, particularly in scenarios where operators must remain mobile, multitask, or operate away from fixed stations.

Compared to joystick-based paradigms, wearables reduce physical encumbrance and allow operators to engage with robots while moving freely through operational environments [119].

A significant strength of this paradigm is the redistribution of sensory load. By delivering information via haptics, wearable systems reduce the reliance on already overloaded visual and auditory channels, which is especially beneficial in complex MMRS deployments. For example, Scheggi et al. [119] demonstrated that haptic feedback provided via armbands improved performance by reducing cognitive load in multitasking conditions.

Different wearable form factors serve varying purposes within MMRS control. **Armbands** are commonly employed for detecting arm movements and muscle activation patterns via IMU and EMG sensing, often supporting symbolic or gestural control of robot collectives [1, 3]. **Torso suits**, typically fitted with vibrotactile actuators, are used to deliver broad spatial feedback across the body, providing operators with distributed swarm status updates without relying on head-mounted or hand-held displays [1]. **Gloves** integrate fine-grained hand gesture recognition with localised fingertip haptics, allowing for nuanced input and high-resolution feedback, particularly in formation management and swarm cohesion tasks [15, 136]. Embodied advantages are reported across all three wearable categories, supporting operators with non-disruptive, body-synchronous control mechanisms in non-dyadic MMRS contexts.

4.2.3 Limitations. Despite these advantages, wearables and haptics present several notable limitations in MMRS control. One commonly reported issue relates to signal processing requirements. Many wearable systems, particularly those using continuous sensing modalities like EMG or IMU, require data processing pipelines and sometimes machine learning (ML) techniques to interpret raw sensor data into control commands [68]. This can introduce latency, increase system complexity, and lead to occasional misclassification of user intent. However, this is not universally the case. Simpler wearable systems, such as those employing discrete inputs via button presses or basic capacitive touch sensing, can operate without ML inference layers, providing direct and immediate command execution. The trade-off, however, is a corresponding reduction in input richness or adaptability to user movement.

A second limitation concerns the limited interaction bandwidth of wearable paradigms. The number of reliably distinguishable gestures or haptic cues is inherently constrained, particularly in mobile or operationally noisy environments. This limits their scalability for MMRS deployments involving large heterogeneous teams or high-frequency command inputs [49]. Many wearable systems, especially those focused on gesture-based input, are optimised for group-level or supervisory commands, rather than fine-grained, continuous teleoperation of individual units.

Finally, the literature remains mixed regarding cognitive load. Some studies report increased subjective workload due to the need to remember gesture vocabularies and respond to distributed haptic cues [1], while others find no significant negative effects, particularly when wearables are integrated into multimodal control frameworks [114]. This suggests that the cognitive burden is highly dependent on the wearable design, the MMRS task context, and the level of autonomy afforded to the robots.

This review focuses primarily on interaction paradigms and reporting practices; deeper empirical treatment of trust dynamics, ethical responsibility, and long-term human factors remains limited in the underlying literature and warrants dedicated investigation.

4.3 Gesture Control

4.3.1 Technique. Gesture control forms a symbolic interaction paradigm where human operators issue commands to MMRS through predefined physical movements. These movements are detected and translated into machine-readable commands via a range of sensing modalities, enabling operators to convey high-level intent to robot collectives without relying on physical contact or traditional hardware controllers. In MMRS contexts, gesture control is primarily employed for issuing symbolic commands to entire groups of robots or dynamically assigning behaviours such as formation changes, navigation goals, or mission phase transitions.

Within the reviewed literature, gesture-based interaction can be broadly distinguished between symbolic gestures and physical (continuous) gestures. Symbolic gestures consist of discrete, lexically defined hand postures or movements that map directly to predefined commands or system states (e.g., start, stop, disperse, select). These gestures dominate MMRS research and are well suited to non-dyadic interaction, enabling efficient group-level command of multiple robots with low ambiguity. In contrast, physical gestures encode intent through continuous motion, such as hand or arm trajectories, where control parameters (e.g., direction or magnitude) are inferred from the gesture dynamics. Such physical gestures are less frequently observed and are limited to coarse supervisory influence over collective motion, reflecting scalability and robustness constraints in MMRS settings.

The dominant sensing modalities for gesture recognition in MMRS research include **vision-based systems** [133], **wearable sensor systems** (e.g. instrumented gloves) using IMU or EMG data [15, 42, 49], **extended reality (XR) interfaces** incorporating optical hand tracking [36, 122, 146, 149], and **pure EMG-based muscle sensing** [65, 130, 131].

Figure 7 shows examples of physical gestures used in MMRS studies, highlighting the variety of symbolic and movement-based input methods deployed in this interaction paradigm.

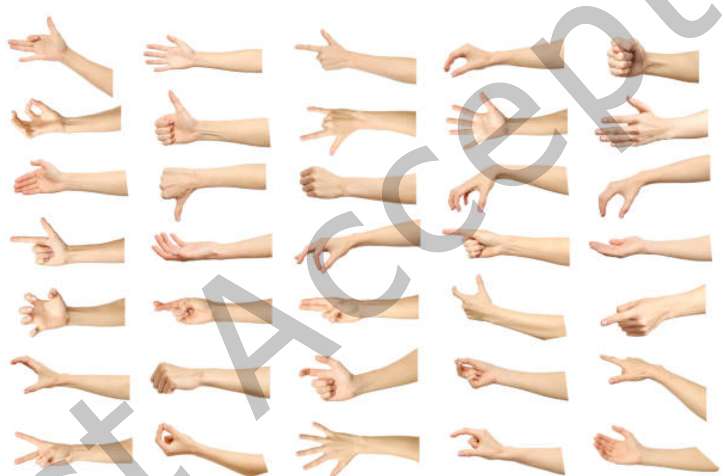

Fig. 7. Example symbolic (discrete) and physical (continuous) hand gestures used in MMRS control. Image © Filistimlyanin at iStock, reproduced under licence.

Across the reviewed literature, gesture control in MMRS is most frequently applied in supervisory modes where operators command multiple robots synchronously. Common applications include controlling swarm formations [15], directing navigation goals [42], and issuing task-level commands such as initiating surveillance or payload deployment routines [133]. In a minority of cases, asynchronous control was observed, with gestures used to select individual robots and issue targeted commands [65]. However, the primary design pattern remains symbolic, group-level control optimised for non-dyadic operator-robot interactions.

4.3.2 Capabilities. Gesture control offers benefits for MMRS interaction, particularly in environments where hands-free, mobility-compatible control is desired. Foremost is natural alignment between gestures and human motor skills, allowing operators to issue commands in an intuitive, body-synchronous manner without reliance on intermediary devices [15, 119]. This reduces interface friction and lowers cognitive load compared to indirect control paradigms.

In MMRS deployments, gesture paradigms are especially well-suited to simultaneous multi-robot control. For example, [15] demonstrated how symbolic gestures could dynamically adjust swarm formations and movement directions without requiring operators to address each robot individually. This collective control capability enhances scalability in MMRS applications. Hand-worn gesture interfaces may also incorporate explicit disengagement and re-engagement modes, allowing operators to suspend and resume gesture input as needed, thereby reducing inadvertent commands and supporting task switching in non-dyadic MMRS control.

Another advantage is the potential for increased spatial situational awareness, particularly when gesture control is embedded within XR environments (see Section 4.4). In these cases, gestures are combined with immersive spatial visualisation, enabling operators to manipulate virtual representations of the robot swarm in 3D space [36, 149]. This combination of interaction modalities improves operator engagement with multi-robot environments.

Finally, non-contact operation is a significant benefit in MMRS deployments within hazardous environments. Vision-based gesture systems allow operators to issue commands without physical interfaces, which is especially advantageous in sterile, contaminated, or safety-critical domains such as nuclear decommissioning or disaster response [42, 133].

4.3.3 Limitations. Despite these advantages, gesture-based interaction paradigms present several limitations in MMRS control.

First, gesture recognition accuracy remains a critical issue. Vision-based systems are vulnerable to misclassification under poor lighting, occlusions, and dynamic backgrounds, limiting their reliability in outdoor or unstructured MMRS environments [65, 133]. Wearable-based systems reduce some environmental dependencies but introduce their own challenges related to sensor drift and calibration [15].

Second, user fatigue and ergonomic strain are well-documented limitations of gesture paradigms. Studies such as [15, 36] note that prolonged use of arm or hand gestures leads to physical fatigue, limiting session durations and negatively impacting operator performance. While wearable and XR-based gestures can mitigate some ergonomic issues, sustained gesture input remains physically demanding in high-tempo MMRS operations.

A third limitation is the lack of standardisation and associated training burden. Unlike button- or joystick-based interfaces, gesture paradigms frequently lack universal symbolic mappings. As observed in [42], operators are required to learn system-specific gesture vocabularies, which increases training time and risks errors during mission execution, especially in time-critical MMRS deployments.

Finally, latency and computational load are practical concerns, particularly in XR-based gesture interfaces (see Section 4.4). The need for real-time gesture recognition and fusion with other sensory inputs increases processing demands, which can introduce latency and decrease control responsiveness in complex MMRS environments [36, 149].

An additional ergonomic consideration is the hidden cost of hands-free control. Although gestures are categorised as hands-free, they still monopolise the operator's upper limbs, reducing the ability to perform other manual tasks concurrently. This limitation is highlighted by [125], who propose alternatives such as micro-gestures (e.g., SoloFinger) that allow MMRS control using subtle finger movements while leaving the hands free for primary tasks. Adaptation of such approaches could offer promising solutions for multitasking operators in non-dyadic MMRS deployments.

4.4 Extended Reality (XR)

4.4.1 Technique. Extended Reality (XR) is an umbrella term encompassing Augmented Reality (AR), Mixed Reality (MR), and Virtual Reality (VR), representing a spectrum of immersive digital experiences that blend real and virtual environments to varying degrees. In the context of MMRS control, XR enables spatially aware, intuitive, and immersive interactions, enhancing operator situational awareness and cognitive engagement. The technology required to implement these interaction techniques is now available off-the-shelf, with VR from the likes of Meta Quest, AR from Google Glass, and MR from Microsoft Hololens [152].

XR incorporates all instances of AR, MR, and VR observed. As a breakdown, for rates of observation in papers analysed, this is weighted in favour of VR (52%), then AR (32%), and lastly MR (16%). XR was not observed as a stand-alone interaction technique, relying on multi-modal interaction through incorporation of one or more other interaction techniques - such as:

- **Hand controllers** [98, 111, 116]
- **Gestures** [46, 113, 140]
- **Gaze/head tracking** [31, 146]
- **Voice control** [31, 46, 98]

4.4.2 Capabilities. XR technologies, particularly VR and MR, significantly enhance situational awareness and mission efficiency in MMRS control. Popović et al. [111] highlighted that immersive XR environments allow operators to perceive threats and respond more quickly, especially in high-pressure scenarios. The study confirms that shared control in XR reduces the number of operator instructions, optimising human-swarm teaming efficiency in time-sensitive environments. Similarly, Huang et al. [56], demonstrated that integrating robot sensor data with VR visualisations improves operator spatial awareness, reducing cognitive workload and improving mission success rates. A key strength of XR systems is their ability to allow an operator to take control of any robot within a team through an immersive interface, a capability that significantly enhances flexibility and responsiveness, particularly in search-and-rescue applications. Furthermore, in the use of MR, [152] emphasises that MR can be used to provide virtual training and development environments, reducing the need for costly physical prototypes and accelerating robotic system development.

4.4.3 Limitations. Despite these benefits, XR-based interaction for MMRS control presents several technical and operational limitations. One major drawback is the high cognitive load imposed on operators, particularly in time-sensitive scenarios where multiple visual streams must be managed simultaneously [111]. While XR enhances awareness, stress and information overload in immersive settings can reduce human performance under pressure. Another concern is motion sickness, which remains a challenge for sustained XR use, especially in long-duration or high movement scenarios [56]. Additionally, computational demands pose a limitation; XR systems require high processing power, making them difficult to deploy on low-power or mobile platforms [56]. Furthermore, Yu et al. [152] notes that occlusion processing and depth perception remain significant challenges in MR, leading to inconsistencies when blending real and virtual elements. Finally, hardware constraints such as limited field-of-view, low resolution, and rendering latency can reduce usability and the precision of XR-based robotic control systems, necessitating further advancements in headset technology and interaction design.

4.5 Voice Control

4.5.1 Technique. Voice control is a **natural language interaction paradigm**, where operators interact with MMRS through spoken language processed by automatic speech recognition (ASR) systems. This paradigm enables command-and-control functionality through speech-based interfaces, allowing operators to issue commands without physical input devices. In the MMRS context, voice control is primarily used to facilitate hands-free interaction, often in combination with other modalities such as gesture or graphical user interfaces, forming part of multimodal interaction schemes.

Voice control within MMRS systems typically follows two distinct approaches. First, command-based models rely on pre-defined keywords or phrases that trigger discrete actions within the robotic system. Second, flexible language models allow for more natural sentence structures, using natural language understanding techniques to extract intent from operator speech. The majority of MMRS studies favour simplified command-based approaches to reduce computational complexity and mitigate recognition errors in noisy environments. For example, [18] employed a flexible parsing method, where the system recognised key action words within sentences without requiring strict grammar or word order, supporting more adaptable speech interactions during robot tasking.

4.5.2 Capabilities. Voice interaction paradigms offer several practical advantages for MMRS control, especially in field environments where operators must remain mobile or have their hands occupied with other tasks. A key advantage is the hands-free nature of voice commands, which removes the need for physical interfaces and allows operators to maintain mobility during mission execution. This is particularly beneficial in high-pressure environments such as search and rescue (SAR) or hazardous industrial inspection tasks [16].

Another strength of voice control in MMRS is its low-cognitive load for simple commands. Spoken instructions align naturally with human communication patterns, allowing for rapid, intuitive issuance of high-level commands without needing complex manual inputs. This facilitates quick transitions from operator tasks to tasking of robot teams.

Voice control also offers potential for scalable group command issuance. By framing instructions at the task or behaviour level (e.g., "all drones proceed to waypoint two"), operators can control multiple robots simultaneously without needing to sequentially address each unit. Studies also highlight that multimodal integration - combining voice with gestures or visual interfaces - further enhances system robustness by allowing operators to disambiguate ambiguous speech commands through complementary modalities [18].

Empirical results indicate that even in field environments, well-designed MMRS voice systems can achieve high recognition accuracy. For instance, [18] demonstrated a speech command success rate exceeding 96% in outdoor operational scenarios.

4.5.3 Limitations. Despite its benefits, voice control presents several limitations for MMRS control. The most significant is sensitivity to environmental noise. Studies consistently report degraded ASR performance in outdoor or operational environments characterised by background noise, crowds, or machinery, which are common conditions in MMRS deployments [16, 18]. For example, [18] noted that their ASR system became unreliable in the presence of public address announcements and other ambient sounds.

Another limitation is latency and error recovery. While voice input is fast to deliver, real-time system responsiveness is frequently compromised by recognition lag and the need for error confirmation or correction, especially in high-noise environments. This can introduce undesirable delays in fast-paced MMRS operations.

Additionally, generalisation across speakers remains a challenge. Systems often require adaptation or training to handle speaker-specific variations such as accent, pitch, or speaking style [18]. This impacts the ease of deployment in dynamic MMRS contexts where operator teams may rotate or vary between missions.

Finally, interaction bandwidth is inherently limited by the nature of voice input. While effective for high-level commands, voice interfaces are less suited for precise or continuous control of individual robots. They work best when paired with higher-autonomy MMRS configurations or integrated within multimodal frameworks, supporting supervisory-level rather than direct teleoperation.

4.6 Brain-Computer Interfaces (BCIs)

4.6.1 Technique. BCI techniques typically rely on measurement of electroencephalographic signals - referred to as an **electroencephalogram (EEG)**. An EEG is a measurement of brain activity taken by small sensors attached on, or close to, the surface of the skin on a person's head. In the papers assessed, headwear is used to position the sensors on the operator's head, typically on or very close to the surface of the skin.

BCI use in MMRS control remains limited, with EEG-based non-invasive methods explored for tasks such as issuing high-level commands, selecting formations, or providing directional input to swarms. Invasive BCI methods have not been studied in this context, but from a technical perspective could enable faster and more precise control.

4.6.2 Capabilities. Throughout, it was widely recognised that "Brain Machine Interfaces (BMI) can offer intuitive control in a plethora of applications where other interfaces alone (e.g. joysticks) are inadequate or impractical" [67]. The ability of BCIs to facilitate a transpose between visual imagery and robot control is a powerful concept, since "using the visual imagery paradigm that imagines the formation change of the swarm drone itself is more intuitive than the motor imagery paradigm that imagines the movements of hands, feet, and tongue" [71].

Looking at the use of BCI in swarm control, [20] presents a BCI system that can efficiently control a high-complexity robot swarm with simple mental commands. The approach "constructs a scalable dictionary of robotic behaviours that can be searched simply and efficiently by a BCI user". Unlike traditional continuous control BCIs, which struggle with more than three degrees of freedom, this method can "control upwards of 6 separate degrees of freedom". This is facilitated by an interaction algorithm that is based on posterior matching, which refines user inputs iteratively and accounts for errors, making it "robust to system noise in measurement processing". To this end, the system was tested with "non-stationary input errors" and still achieved a 75.7% overall configuration selection accuracy in swarm control.

4.6.3 Limitations. BCI use in MMRS is constrained by three primary limitations:

(1) Cost – Several studies [62, 72, 79] employed high-end equipment (e.g., BrainAmp), while open-source alternatives (e.g. OpenBCI) remain expensive. This high entry cost limits access for many research teams, likely contributing to the low representation of BCIs in MMRS HRI studies (10% of papers in this review).
(2) Accuracy and Latency – Signal quality and classification accuracy are critical challenges. Several studies [20, 58, 72] report accuracies ranging from near-chance levels [62] to 60–70% even with advanced models [72]. The highest performance (82.4%) was achieved using ensemble deep reinforcement learning [55]. However, delays remain: models often operate on 2-second windows, introducing systematic lag [55]. Real-time performance is thus constrained by a trade-off between processing time and control accuracy [58].
(3) Cognitive Load and User Variability – Effective BCI use demands significant training. Experienced users outperform novices [58], but even trained users show substantial variability [62]. This is likely due to inter-user differences and limited generalisability of EEG-based ML models [71, 79].

Additionally, BCI-based swarm control introduces cumulative input errors with increased command frequency [20], attributed to fatigue and signal degradation. Dictionary-based systems, while intuitive, suffer from performance bottlenecks as vocabulary size grows and requires users to learn complex configuration sorting methods. Finally, non-invasive EEG suffers from inherently low signal-to-noise ratios, limiting accuracy compared to invasive techniques.

5 Results, Key Metrics & Reporting Framework

5.1 Results

Through a systematic literature review and subsequent contextual and technological analyses of non-dyadic MMRS interaction, a number of key findings have emerged:

- **Inconsistent Reporting of Key Variables** - Many studies fail to report essential variables, hindering comparative analysis and replication. Future research should explicitly state application context, operator training level, operator positionality, number of robots, number of operators, autonomy levels, and use consistent terminology. To this end, a *Reporting Framework* is presented in Section 5.3.

- **Limited Research on Onboard Control** - Research into MUM-T scenarios, where operators control MMRS from onboard moving platforms, remains scarce, presenting an opportunity for novel contributions in designing and evaluating interaction paradigms under these conditions.
- **Underexplored Multi-Operator Team Models** - Research predominantly focuses on single-operator MMRS control, with virtually no studies exploring teams larger than four operators. Investigating multi-operator control models could address scalability and workload challenges in larger MMRS deployments.
- **Expanding Hands-Free Control** - Joysticks, while common in teleoperation, are inadequate for large-scale MMRS due to limited feedback and scalability. Alternative methods, including wearables, gestures, voice, and BCIs, enable hands-free, multi-task-capable control suitable for MUM-T and non-dyadic scenarios. Linked to MUM-T scenarios, enabling operators to multi-task, where oversight of a highly autonomous MMRS is of secondary importance, is a particularly interesting research direction.
- **Utilising Off-the-Shelf Technologies** - Advanced consumer devices such as smart watches, rings, and VR/MR headsets offer haptics, gesture control, and multimodal inputs but remain underutilised in MMRS research. Incorporating these technologies could improve replicability, scalability, and real-world readiness of MMRS interaction techniques.
- **Need for Context-Driven Research** - Robotics research must be grounded in clear application contexts to guide effective interaction design and establish meaningful performance metrics, avoiding retrofitting interaction paradigms, and corresponding methods of assessment, to loosely defined tasks.
- **Importance of Using Trained Operators** - Although robots are being deployed with greater ubiquity in professional industries, most studies still use untrained operators. Future MMRS research should prioritise testing with suitably qualified and experienced personnel (SQEP) to align with real-world deployment conditions.
- **Managing Cognitive Load and Workload** - Managing operator cognitive load remains a recurrent challenge in MMRS interaction. Potential solutions include using multi-operator teams, increasing platform autonomy (explored in the context of the robot-operator ratio in Section 5.2), and adopting multi-modal interaction techniques (discussed in Section 6.2) to distribute workload and improve effectiveness.

5.2 The Key Metric: Robot-Operator Ratio

A critical consideration in MMRS interaction is the robot-operator ratio, which directly influences cognitive workload, scalability, and mission effectiveness. While Figure 8 shows the overall distribution of robot-operator ratios across the dataset, analysis confirms that the modal ratio is 3:1, reflecting a prevailing focus on single-operator control of small-scale multi-robot teams. This aligns with our underlying data, where three robots is the most commonly studied, indicating that much of the MMRS research to date focuses on lower-complexity, lesser in number, and thus more manageable configurations. The dominance of low robot–operator ratios highlights a persistent gap between experimental MMRS research and envisioned large-scale deployments, limiting empirical insight into interaction breakdowns, workload saturation, and trust calibration at higher levels of scale. It is important to note that this observed modal ratio may also be influenced by pragmatic research constraints, including the affordability, availability, and maintainability of robotic platforms within academic and laboratory settings, which can limit experimentation at larger scales irrespective of intended application contexts.

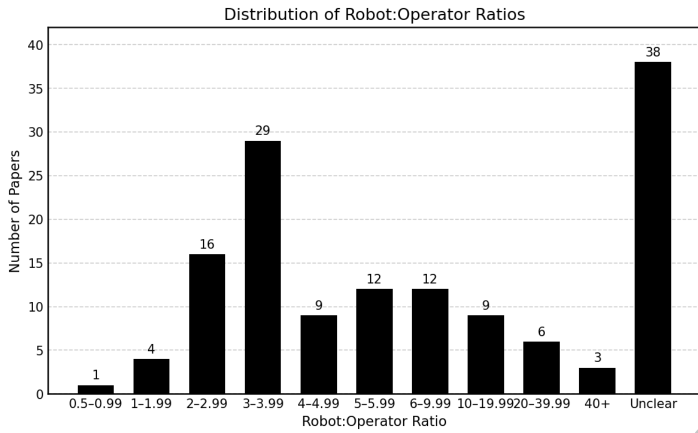

Fig. 8. Distribution of Robot:Operator Ratios - highlighting that the most common case was "Unclear" from the papers surveyed.

However, in 38 of 139 papers (27%), it was not possible to determine the robot-operator ratio due to absent or inconsistent reporting of the number of robots and operators. This lack of clarity hinders comparative analysis across MMRS studies and limits the ability to assess how different interaction strategies scale across varying levels of complexity.

Establishing the robot-operator ratio as a standard reporting metric would enable meaningful comparisons across a diverse range of MMRS systems, supporting clearer evaluation of HRI techniques across application domains, platform types, and autonomy levels. For example, a gesture-based interface that is effective when controlling three drones may not scale when applied to a heterogeneous team of 20 aerial and ground platforms, but without explicit reporting, these limitations remain hidden.

Future research should adopt the robot-operator ratio as a guiding metric when designing, evaluating, and reporting MMRS studies. As the field progresses towards higher autonomy and larger-scale deployments, clear articulation of how many robots an operator can effectively manage, and under what conditions, will be essential for advancing scalable and usable MMRS frameworks. This focus will also support the shift from single-operator paradigms to multi-operator approaches that distribute cognitive load and enhance mission resilience in complex, non-dyadic MMRS environments, as discussed further in Sections 6.1 and 6.2.

5.3 Reporting Framework

During this systematic literature review, a number of reporting variables were found to be ambiguously defined or inconsistently documented across studies, significantly diminishing their analytical utility and hindering direct comparison. These shortcomings were encountered repeatedly during screening, data extraction, and synthesis, thereby directly affecting this review's ability to interpret and compare findings across non-dyadic MMRS studies. As such, a key contribution of this paper is the proposal of a reporting framework intended to increase the impact of future work in this area by improving transparency, comparability, and interpretability in MMRS HRI research.

Based on the challenges encountered during analysis, we recommend that researchers publishing in HRI — particularly in non-dyadic contexts — explicitly and consistently report the following core variables: **(1)** number of robotic platforms; **(2)** autonomy level; **(3)** number of operators; **(4)** extent of operator training; **(5)** operator positionality; **(6)** definitions of system descriptor terminology.

These variables were selected because their absence or ambiguity repeatedly constrained the interpretation of study outcomes. In particular, the review revealed that many studies failed to define the autonomy level of the robots employed, making it difficult to compare findings across experiments. Without explicit autonomy definitions, it remains unclear whether operators were interacting with teleoperated, semi-autonomous, or fully autonomous systems, limiting reproducibility, generalisability, and the downstream impact of otherwise valuable empirical work.

Inconsistencies in terminology further complicate comparative analysis. Figure 9 illustrates the distribution of phraseology used to describe MMRS within the reviewed literature. The term "team" was used in 84 cases (60.43%), over four times more frequently than the next most common descriptor, "swarm." In many cases, such terms were not formally defined, leading to ambiguity between studies that may in practice be examining different coordination models, autonomy structures, or interaction paradigms. This variability in terminology reflects a lack of shared conceptual framing within the field, complicating cross-study comparison and reinforcing the need for clearer reporting and definitional consistency in MMRS HRI research.

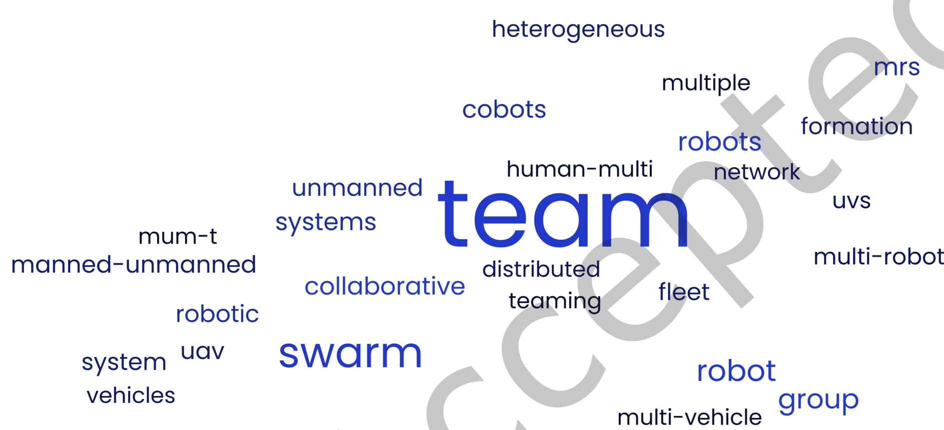

Fig. 9. Distribution of phraseology used to describe MMRS - with no consensus on terms to describe different MMRS team set-ups.

Beyond system-level descriptors, the review also identified limited reporting of operator-related variables. As discussed in Section 3.2, 68% of papers did not specify whether operators were trained professionals, experienced users, or novices, nor did they consistently state whether operators were remote, co-located, or onboard the system (Section 3.3). These factors are critical for interpreting workload, cognitive demand, and control efficiency, particularly in non-dyadic MMRS contexts where interaction demands scale with system size.

To address these issues, Figure 10 proposes a broader reporting framework that supplements the core variables with additional contextual information spanning: System & Task Design, Operator-Centric Information, and Experimental Design & Analysis. The framework is intentionally lightweight and descriptive, with the aim of maximising impact by enabling future studies to be more easily situated, interpreted, and synthesised within the growing MMRS HRI literature, rather than prescribing experimental methods or evaluation protocols. By explicitly highlighting commonly omitted contextual variables, the framework aims to increase the interpretability and downstream impact of future MMRS HRI studies rather than to prescribe experimental design or evaluation methodology.

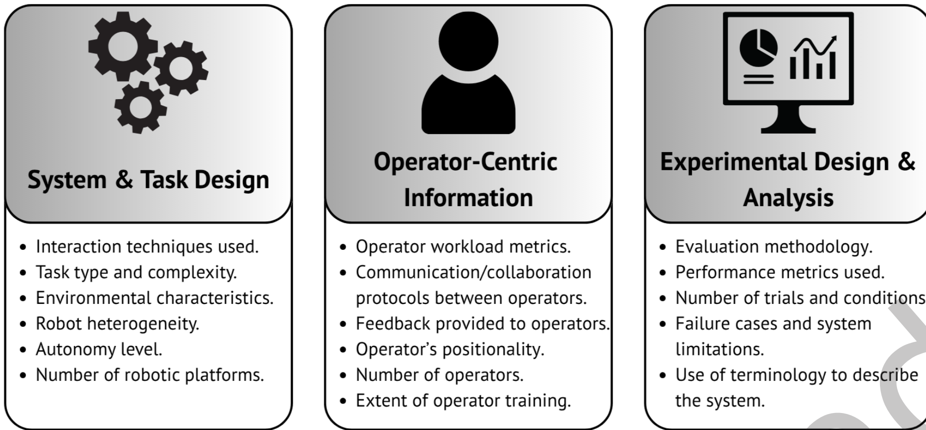

Fig. 10. Proposed Reporting Framework For MMRS HRI Research, organising essential reporting variables across three categories: *System & Task Design*, *Operator-Centric Information*, and *Experimental Design & Analysis*.

The proposed framework has not yet been empirically assessed for consistency of application across independent researchers, nor has it been evaluated through large-scale adoption across multiple studies. It should therefore be understood as an impact-oriented reporting aid, grounded in the practical challenges encountered during this systematic review, rather than as a definitive or formally validated standard.

Several limitations of the framework should be acknowledged. First, it is derived from the scope of the reviewed literature and may not capture all reporting needs in domains beyond mobile MMRS, such as tightly coupled dyadic systems or purely stationary robotic installations. Second, the framework does not enforce a specific autonomy taxonomy, terminology ontology, or operator role model, reflecting the current fragmentation of the field and avoiding premature standardisation. Finally, the framework does not explicitly address long-term deployment factors, such as learning effects, organisational integration, or sustained autonomy, which remain underrepresented in the existing literature.

Despite these limitations, clearer and more consistent reporting of these variables has the potential to substantially increase the impact of individual MMRS HRI studies and of the present review itself, by improving study comparability, supporting reproducibility, and enabling more robust synthesis across future work. The framework is therefore intended as a foundation for community discussion and iterative refinement, with future research required to assess its uptake, utility, and longer-term influence on MMRS HRI research practice.

6 Discussion

Synthesising the contextual (Section 3) and technological analyses (Section 4), this discussion examines how autonomy level mediates the relationship between interaction design, operator workload, and trust in non-dyadic MMRS. Across the reviewed literature, increasing robot–operator ratios and autonomy levels consistently shift interaction from direct control toward supervisory oversight, amplifying challenges in situational awareness, prioritisation, and cognitive workload. These findings highlight that effective MMRS deployment depends not only on advances in interaction technologies, but on their alignment with autonomy assumptions, operational context, and human team structures.

So, the discussion first explores the interplay between autonomy levels, risk appetite, and accuracy in MMRS HRI research (Section 6.1). It then examines how structured multi-operator teams and multi-modal interaction approaches can support scalable, effective MMRS control in real-world contexts (Section 6.2).

6.1 Autonomy Levels, Risk, and Accuracy

Across the interaction paradigms identified in the technological analysis (Section 4) limitations in raw data processing and input classification accuracy interact directly with contextual challenges identified in Section 3, including timely identification and response to robot states, maintenance of situational awareness in dynamic environments, and management of operator cognitive workload.

In high-stakes MMRS domains such as SAR, input classification errors directly undermine mission effectiveness, constraining reductions in operator numbers and limiting adoption of lower-TRL interaction techniques. Overcoming this risk–accuracy trade-off requires improved methods and stronger empirical evidence in realistic deployments.

A comparable trade-off exists in adopting higher autonomy levels within MMRS. Autonomy levels, ranging from Level 0 (no automation) to Level 5 (full automation) [59], determine the independence of robotic platforms and directly affect operator roles, shifting from direct control to supervisory oversight as autonomy increases. For reference, the Autonomy Levels Framework [59] is presented in Figure 11, defining each level from zero to five in accordance with the Society of Automotive Engineers' standard SAE J3016 and International Standards Organisation ISO-PAS 22736. Noting, that whilst SAE J3016 provides a widely adopted and useful reference point for discussing autonomy progression, it was originally developed for road vehicle automation and does not fully capture MMRS characteristics such as distributed autonomy, heterogeneous platforms, or team-level decision-making. Its use here is pragmatic, providing a common vocabulary for autonomy discussion, while recognising that richer autonomy models are required to describe MMRS.

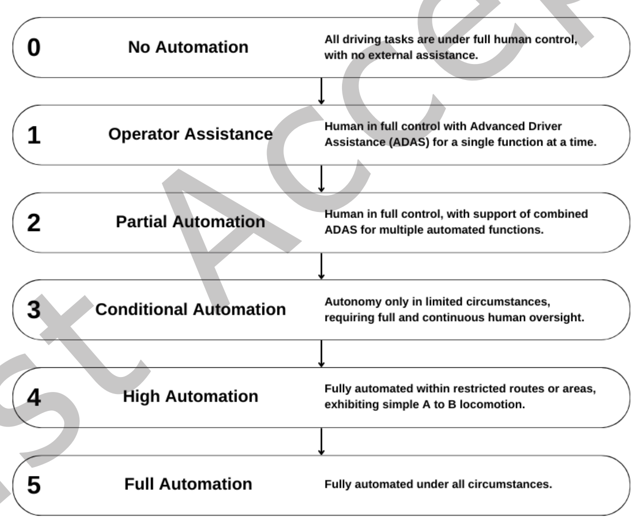

Fig. 11. Autonomy Levels Framework based on SAE J3016 and ISO-PAS 22736.

Higher autonomy is essential for scalable, non-dyadic MMRS, enabling single operators to manage larger fleets while maintaining situational awareness and coordination. At lower levels (0-2), operators directly control each robot, limiting scalability and increasing cognitive load, with typical caps of two to three robots per operator [54]. At higher levels (3-5), robots execute tasks based on high-level commands, reducing intervention needs and enabling higher robot-operator ratios that are critical for large-scale MMRS deployments [132]. Coupling higher autonomy with multi-modal interfaces (discussed in Section 6.2) can improve performance, robustness, and ease of use in MMRS control [131].

However, the pace of adopting higher autonomy is constrained by regulatory limitations, trust, and safety case requirements. For example, FAA regulations often mandate a single pilot per UAV, limiting autonomy applications [12]. Autonomy is also not a universal solution, as human adaptability and oversight remain critical in complex, dynamic environments [56]. This was evident during the DARPA Subterranean Challenge, which demonstrated that despite the deployment of advanced autonomy, human supervisors were frequently the limiting factor in mission performance, underscoring the continuing need for effective human-robot teaming even at higher autonomy levels [56].

In summary, advancing MMRS HRI requires balancing the benefits of higher autonomy with reliability and accuracy in high-risk applications. Building a robust evidence base to support the safe and effective use of advanced autonomy will be key to increasing operational confidence and enabling scalable, efficient, and practical non-dyadic MMRS control.

6.2 Team Control & Multi-Modal Interaction

Sections 3 and 4 show that transitioning from dyadic to non-dyadic MMRS control shifts operator responsibility from direct control to supervisory oversight, increasing workload, decision-making complexity, and coordination demands. However, the current MMRS research landscape remains heavily focused on single-operator scenarios (91% from Section 3.4), with very limited exploration of multi-operator team formations - only 3% consider two operators, and just 2% a team of four.

As illustrated in Figure 12, MMRS control exists along a continuum, ranging from teams of operators managing a single robot to hierarchical teams managing large multi-robot systems where the number of robots significantly exceeds the number of operators. As the robot-operator ratio increases, control must shift from direct interaction toward scalable supervisory frameworks.

A core challenge in non-dyadic MMRS control is identifying which platforms require intervention, as operators must monitor multiple robots simultaneously and determine where engagement is most critical. Unlike dyadic control, non-dyadic contexts necessitate selective intervention, often relying on operator intuition due to the lack of structured prioritisation frameworks. As robot-operator ratios increase, robots may simultaneously request guidance, confirmation, or control, leading to cognitive strain, particularly when robots are heterogeneous (e.g., aerial surveillance alongside ground-based manipulation). The absence of standardised prioritisation mechanisms in the literature highlights the need for future research into automated algorithms that can rank control demands by urgency, risk, and autonomy state to support structured decision-making and mitigate cognitive overload.

Utilising shared experiences, MR, and tabletop interfaces [32] can alleviate cognitive demands in non-dyadic MMRS control. MR, digital twins, and tabletop simulations provide intuitive, spatially organised representations of robot formations, improving situational awareness and command execution efficiency. These methods reduce the cognitive load of mental modelling and abstract interpretation, allowing operators to visualise and manipulate robot behaviour within a controlled virtual environment.

Some papers propose promising multi-modal solutions to these challenges. For example, Huang et al. [56] present a VR-based system facilitating multi-user collaboration, which they highlight as essential for sharing task load when the "human supervisor is overwhelmed." Comparative studies demonstrated that single-operator task completion times were 70% longer than multi-operator teams, despite successful task completion in both cases. The authors further note that "increasing robot autonomy is critical to enhance the performance of heterogeneous robot teams". This observation links directly to the preceding discussion on autonomy levels, reinforcing that effective multi-modal and team-based approaches are most impactful when paired with increased platform autonomy to support scalable, non-dyadic MMRS.

While over 70% of papers surveyed used a multi-modal approach, none explicitly considered how combinations of interaction methods should be tailored to specific robotic platforms, operator roles, or mission contexts. This

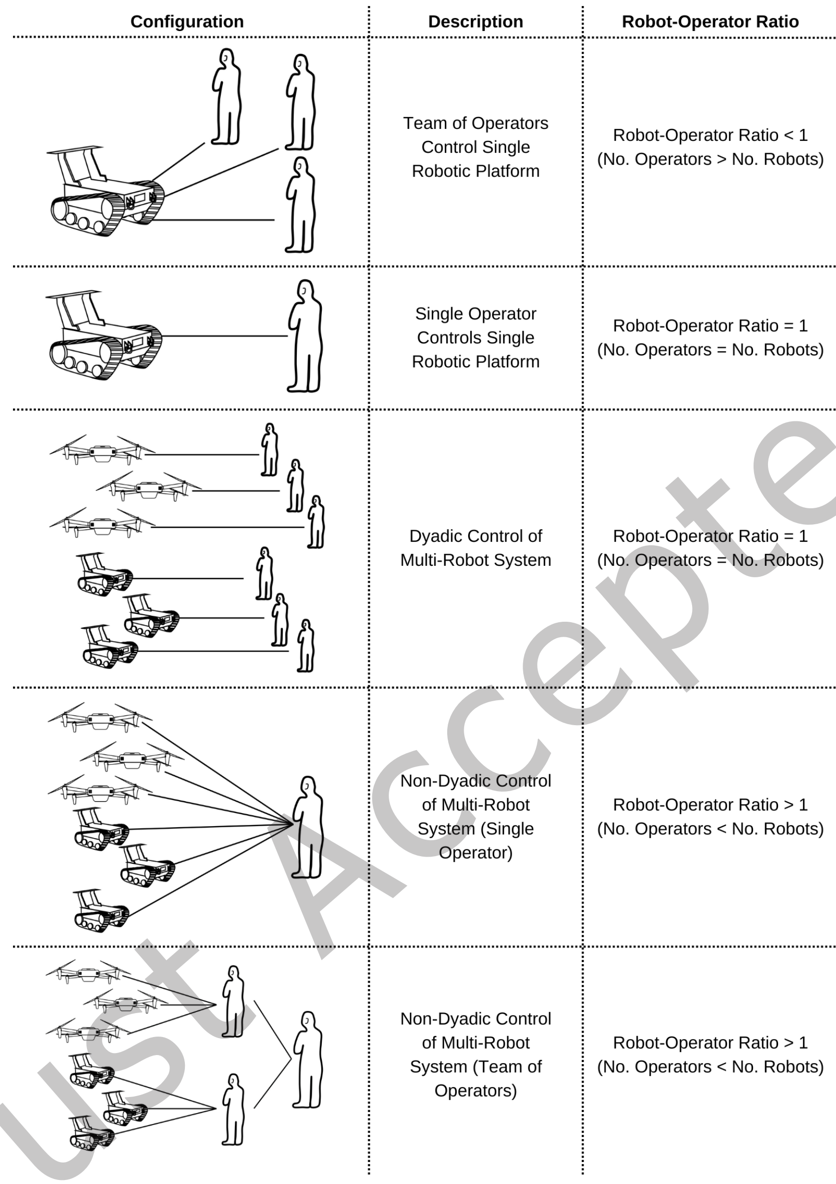

Fig. 12. Continuum of Control Configurations by Robot-Operator Ratio: Permutations of Ground/Aerial Robots with Operators.

gap presents an opportunity for future research to systematically investigate the optimal selection and integration of multi-modal interaction techniques to support effective human-MMRS teaming in diverse operational environments.

Further work is needed to align technological advancements in interaction techniques with context-driven design of high-performance MMRS operator teams. An interesting avenue for exploration is the use of wearable devices (bespoke or COTS) that synthesise multiple interaction modalities (including touchscreens, gesture recognition, voice input, and physiological sensing) into a single unit. These wearable systems enable seamless, hands-free interaction while multitasking, offering a practical, scalable interface for high robot-operator ratio deployments.

Building on the insights gained from this review, we propose that structuring MMRS operator teams around three core roles could address the challenges of scalable, non-dyadic control:

- **Supervisory control** - Overseeing high-level mission execution and delegating control actions.
- **Tactical operators** - Providing intermediate guidance to groups of semi-autonomous robots as needed.
- **Direct controllers** - Engaging in real-time, low-level control adjustments for individual robots.

While this structure is not explicitly defined in the papers reviewed, it synthesises patterns observed across the literature and aligns with team structuring practices in other complex, high-stakes domains, such as air traffic control and surgical teams. By integrating multi-modal interaction approaches with structured team control, future MMRS research can address the scalability, cognitive load, and operational challenges critical to effective non-dyadic deployments in complex real-world scenarios.

7 Conclusion

This paper presented a systematic literature review of non-dyadic MMRS interaction, complemented by contextual and technological analyses, to map the current research landscape and identify opportunities for advancing the field.

The review highlights a recurrent lack of comparative studies and standardised reporting/assessment methods, making it difficult to determine which interaction paradigms best support effective, scalable MMRS control. In response, this paper proposed a clear reporting framework for MMRS HRI research, outlining essential variables such as the number of robots, autonomy levels, number of operators, operator training, positionality, and consistent system terminology. Adoption of this framework will support more transparent, reproducible, and comparable studies across the field, laying the groundwork for systematic evaluation of interaction strategies.

Hands-free and multi-modal interaction techniques, including the use of COTS wearable devices, emerged as underexplored yet promising pathways for enabling scalable, intuitive MMRS control, particularly in dynamic, multitasking scenarios such as MUM-T and onboard control. These approaches can contribute to reducing cognitive load, enhance situational awareness, and improve operator effectiveness when managing high robot-operator ratio systems.

Additionally, current research heavily favours single-operator MMRS control, with minimal exploration of large and diverse operator teams. Investigating hierarchical and distributed team structures, including the delineation of supervisory and active roles, will be critical for future deployments of large-scale MMRS.

Finally, the paper emphasises the importance of grounding MMRS HRI research in real-world applications, using clear task requirements and realistic contexts to guide the development and evaluation of interaction techniques. Collaborations with industry can further support this goal, providing the scale, complexity, and operational constraints necessary to test and refine innovative approaches.

By addressing these challenges, and by adopting clear reporting frameworks, future research can move beyond isolated technical demonstrations toward the development of practical, scalable MMRS interaction solutions capable of supporting the complex missions and operational contexts where these systems are increasingly deployed.

Acknowledgments

We would like to thank the reviewers for their constructive feedback and insights, which have helped strengthen this work. The authors also acknowledge the support of the University of Bath and the wider Human-Robot

Interaction research community, whose discussions and perspectives informed this review. Finally, we recognise the many researchers and practitioners working across robotics, autonomy, and human-robot interaction whose contributions underpin the body of literature analysed in this paper.

References

[1] Sydrak S. Abdi and Derek A. Paley. 2023. Safe Operations of an Aerial Swarm via a Cobot Human Swarm Interface. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. 1701–1707. https://doi.org/10.1109/ICRA48891.2023.10161343

[2] Julie A. Adams, Joshua Hamell, and Phillip Walker. 2024. Can A Single Human Supervise A Swarm of 100 Heterogeneous Robots? *IEEE Transactions on Field Robotics* (2024), 1–1. https://doi.org/10.1109/TFR.2024.3502316

[3] Marco Aggravi, Claudio Pacchierotti, and Paolo Robuffo Giordano. 2021. Connectivity-Maintenance Teleoperation of a UAV Fleet With Wearable Haptic Feedback. *IEEE Transactions on Automation Science and Engineering* 18, 3 (2021), 1243–1262. https://doi.org/10.1109/TASE.2020.3000060

[4] Marco Aggravi, Giuseppe Sirignano, Paolo Robuffo Giordano, and Claudio Pacchierotti. 2022. Decentralized Control of a Heterogeneous Human–Robot Team for Exploration and Patrolling. *IEEE Transactions on Automation Science and Engineering* 19, 4 (2022), 3109–3125. https://doi.org/10.1109/TASE.2021.3106386

[5] M. Alencastre-Miranda, L. Munoz-Gomez, and I. Rudomin. 2003. Teleoperating robots in multiuser virtual environments. In *Proceedings of the Fourth Mexican International Conference on Computer Science, 2003. ENC 2003*. 314–321. https://doi.org/10.1109/ENC.2003.1232910

[6] George Allison. 2021. *Britain to get air combat drones to fly alongside fighters.* https://ukdefencejournal.org.uk/britain-to-get-air-combat-drones-to-fly-alongside-fighters/ Accessed: 2024-12-02.

[7] J. Alonso-Mora, S. Haegeli Lohaus, P. Leemann, R. Siegwart, and P. Beardsley. 2015. Gesture based human - Multi-robot swarm interaction and its application to an interactive display. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 5948–5953. https://doi.org/10.1109/ICRA.2015.7140033

[8] Nora Ayanian, Andrew Spielberg, Matthew Arbesfeld, Jason Strauss, and Daniela Rus. 2014. Controlling a team of robots with a single input. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. 1755–1762. https://doi.org/10.1109/ICRA.2014.6907088

[9] Gregory Bales and Zhaodan Kong. 2017. Neurophysiological and behavioral studies of human-swarm interaction tasks. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 671–676. https://doi.org/10.1109/SMC.2017.8122684

[10] Daniel Barber, Larry Davis, Don Kemper, Peter Smith, and Denise Nicholson. 2007. Collaborative human robot interactions in combined arms operations. In *2007 International Symposium on Collaborative Technologies and Systems*. 88–92. https://doi.org/10.1109/CTS.2007.4621742

[11] Ahmed Baza, Ayush Gupta, Ekaterina Dorzhieva, Aleksey Fedoseev, and Dzmitry Tsetserukou. 2022. SwarMan: Anthropomorphic Swarm of Drones Avatar with Body Tracking and Deep Learning-Based Gesture Recognition. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1284–1289. https://doi.org/10.1109/SMC53654.2022.9945537

[12] Aabhash Bhandari, Tuan Dinh Luong, and Jordan Larson. 2023. Cooperative Guidance and Navigation of UAS for Common Midpoint Radar Remote Sensing. In *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. 280–285. https://doi.org/10.1109/PLANS53410.2023.10139947

[13] Juan A. Bonache-Seco, José A. Lopez-Orozco, Eva Besada-Portas, and José L. Risco-Martín. 2019. ART-GCS: An Adaptive Real-Time Multi-Agent Ground Control Station. In *2019 Spring Simulation Conference (SpringSim)*. 1–11. https://doi.org/10.23919/SpringSim.2019.8732874

[14] J. A. Bonache-Seco, J. A. Lopez-Orozco, Eva Besada Portas, and José L. Risco Martín. 2018. Adaptive event driven framework for real time multi-agent missions. In *Proceedings of the 22nd International Symposium on Distributed Simulation and Real Time Applications* (Madrid, Spain) *(DS-RT '18)*. IEEE Press, 255–262. https://doi.org/10.1109/DISTRA.2018.8600998

[15] Nuttapon Boonpinon and Attawith Sudsang. 2008. Formation Control for Multi-Robot Teams Using A Data Glove. In *2008 IEEE Conference on Robotics, Automation and Mechatronics*. 525–531. https://doi.org/10.1109/RAMECH.2008.4681481

[16] Jonathan Cacace, Riccardo Caccavale, Alberto Finzi, and Vincenzo Lippiello. 2016. Attentional multimodal interface for multidrone search in the Alps. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 001178–001183. https://doi.org/10.1109/SMC.2016.7844401

[17] Jonathan Cacace, Alberto Finzi, and Vincenzo Lippiello. 2016. Implicit robot selection for human multi-robot interaction in Search and Rescue missions. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 803–808. https://doi.org/10.1109/ROMAN.2016.7745211

[18] Jonathan Cacace, Alberto Finzi, and Vincenzo Lippiello. 2017. A robust multimodal fusion framework for command interpretation in human-robot cooperation. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 372–377. https://doi.org/10.1109/ROMAN.2017.8172329

[19] G. L. Calhoun, D. M. Draper, A. Ruff, and B. A. Knott. 2016. Comparing Mouse, Touch, and Voice Input for UAV Tasking. In *2016 IEEE Aerospace Conference*. https://doi.org/10.1109/AERO.2016.7500630

[20] Gregory Canal, Yancy Diaz-Mercado, Magnus Egerstedt, and Christopher Rozell. 2023. A Low-Complexity Brain–Computer Interface for High-Complexity Robot Swarm Control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 31 (2023), 1816–1825. https://doi.org/10.1109/TNSRE.2023.3257261

[21] Kishan Chandan, Vidisha Kudalkar, Xiang Li, and Shiqi Zhang. 2021. ARROCH: Augmented Reality for Robots Collaborating with a Human. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 3787–3793. https://doi.org/10.1109/ICRA48506.2021.9561144

[22] Mingxuan Chen, Ping Zhang, Xiaodan Chen, Yuliang Zhou, Fang Li, and Guanglong Du. 2018. A Human-swarm Interaction Method Based on Augmented Reality. In *2018 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*. 108–114. https://doi.org/10.1109/WRC-SARA.2018.8584252

[23] Xiao Chen. 2023. Application and prospect of brain computer interface technology in the future war. In *Proceedings of the 2022 4th International Conference on Robotics, Intelligent Control and Artificial Intelligence* (Dongguan, China) *(RICAI '22)*. Association for Computing Machinery, New York, NY, USA, 669–673. https://doi.org/10.1145/3584376.3584494

[24] Yushing Cheung, Jae H. Chung, and Norman P. Coleman. 2009. Semi-autonomous formation control of a single-master multi-slave teleoperation system. In *2009 IEEE Symposium on Computational Intelligence in Control and Automation*. 117–124. https://doi.org/10.1109/CICA.2009.4982792

[25] S.K. Choi, S.A. Menor, and J. Yuh. 2000. Distributed virtual environment collaborative simulator for underwater robots. In *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000) (Cat. No.00CH37113)*, Vol. 2. 861–866 vol.2. https://doi.org/10.1109/IROS.2000.893127

[26] Alex Couture-Beil, Richard T. Vaughan, and Greg Mori. 2010. Selecting and Commanding Individual Robots in a Multi-Robot System. In *2010 Canadian Conference on Computer and Robot Vision*. 159–166. https://doi.org/10.1109/CRV.2010.28

[27] Jacob Crossman, Robert Marinier, and Edwin B. Olson. 2012. A hands-off, multi-robot display for communicating situation awareness to operators. In *2012 International Conference on Collaboration Technologies and Systems (CTS)*. 109–116. https://doi.org/10.1109/CTS.2012.6261036

[28] Abhinav Dahiya, Alexander M. Aroyo, Kerstin Dautenhahn, and Stephen L. Smith. 2023. A survey of multi-agent Human–Robot Interaction systems. *Robotics and Autonomous Systems* 161 (2023), 104335. https://doi.org/10.1016/j.robot.2022.104335

[29] Wei Dai, Yaru Liu, Huimin Lu, Zhiqiang Zheng, and Zongtan Zhou. 2021. A Shared Control Framework for Human-Multirobot Foraging With Brain-Computer Interface. *IEEE Robotics and Automation Letters* 6, 4 (2021), 6305–6312. https://doi.org/10.1109/LRA.2021.3092290

[30] M. Daily, Youngkwan Cho, K. Martin, and D. Payton. 2003. World embedded interfaces for human-robot interaction. In *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the*. 6 pp.–. https://doi.org/10.1109/HICSS.2003.1174285

[31] Aditya Das, Patrik Kol, Cody Lundberg, Kris Doelling, Hakki Erhan Sevil, and Frank Lewis. 2018. A Rapid Situational Awareness Development Framework for Heterogeneous Manned-Unmanned Teams. In *NAECON 2018 - IEEE National Aerospace and Electronics Conference*. 417–424. https://doi.org/10.1109/NAECON.2018.8556769

[32] Sarah Grace Dellana, Sune Johansson, Simon Boel Poulsen, Anke van Oosterhout, Mikael Skov, and Timothy Merritt. 2020. Collaboration Around an Interactive Tabletop Map: Comparing Voice Interactions and a Tangible Shape-changing Controller. In *Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia* (Essen, Germany) *(MUM '20)*. Association for Computing Machinery, New York, NY, USA, 132–142. https://doi.org/10.1145/3428361.3428395

[33] Alessandro Di Nuovo, Frank Broz, Tony Belpaeme, Angelo Cangelosi, Filippo Cavallo, Raffaele Esposito, and Paolo Dario. 2014. A web based Multi-Modal Interface for elderly users of the Robot-Era multi-robot services. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2186–2191. https://doi.org/10.1109/SMC.2014.6974248

[34] Matteo Diana, Jean-Pierre de la Croix, and Magnus Egerstedt. 2013. Deformable-medium affordances for interacting with multi-robot systems. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 5252–5257. https://doi.org/10.1109/IROS.2013.6697116

[35] Patrick Doherty and Fredrik Heintz. 2011. A delegation-based cooperative robotic framework. In *2011 IEEE International Conference on Robotics and Biomimetics*. 2955–2962. https://doi.org/10.1109/ROBIO.2011.6181755

[36] Ekaterina Dorzhieva, Ahmed Baza, Ayush Gupta, Aleksey Fedoseev, Miguel Altamirano Cabrera, Ekaterina Karmanova, and Dzmitry Tsetserukou. 2022. DroneARchery: Human-Drone Interaction through Augmented Reality with Haptic Feedback and Multi-UAV Collision Avoidance Driven by Deep Reinforcement Learning. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 270–277. https://doi.org/10.1109/ISMAR55827.2022.00042

[37] Marius Dudek and Axel Schulte. 2024. Experimental Evaluation of UAV Task Delegation Methods. In *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*. 1–8. https://doi.org/10.1109/ICHMS59971.2024.10555792

[38] Ingy Fam, Hassan Soubra, and Noura Gamal. 2023. Human-Swarm Interaction Methods' Effect on Human Psychophysiology. In *2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS)*. 184–189. https://doi.org/10.1109/ICICIS58388.2023.10391199

[39] Aleksey Fedoseev, Ahmed Baza, Ayush Gupta, Ekaterina Dorzhieva, Riya Neelesh Gujarathi, and Dzmitry Tsetserukou. 2022. DandelionTouch: High Fidelity Haptic Rendering of Soft Objects in VR by a Swarm of Drones. In *2022 IEEE International Conference on*

Systems, Man, and Cybernetics (SMC). 1078–1083. https://doi.org/10.1109/SMC53654.2022.9945390

[40] Luigi Feola, Andreagiovanni Reina, Mohamed S. Talamali, and Vito Trianni. 2023. Multi-Swarm Interaction Through Augmented Reality for Kilobots. *IEEE Robotics and Automation Letters* 8, 11 (2023), 6907–6914. https://doi.org/10.1109/LRA.2023.3312031

[41] J. Fredslund and M.J. Mataric. 2002. Huey, Dewey, Louie, and GUI-commanding robot formations. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, Vol. 1. 175–180 vol.1. https://doi.org/10.1109/ROBOT.2002.1013357

[42] Alessandro Giusti, Jawad Nagi, Luca M. Gambardella, and Gianni A. Di Caro. 2012. Distributed consensus for interaction between humans and mobile robot swarms (demonstration). In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3* (Valencia, Spain) *(AAMAS '12)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1503–1504.

[43] Dylan F. Glas, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2012. Teleoperation of Multiple Social Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 42, 3 (2012), 530–544. https://doi.org/10.1109/TSMCA.2011.2164243

[44] Joseph J. Glavan, Ellen J. Bass, Julie A. Adams, and Christopher A. Sanchez. 2022. Measures of Attention in Autonomous and Semi-Autonomous Multi-Vehicle Supervision. In *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. 1–6. https://doi.org/10.1109/ICHMS56717.2022.9980754

[45] KS Gokulraj and J Manikandan. 2021. Design and development of simulator software for formation flight of drones. In *2021 Zooming Innovation in Consumer Technologies Conference (ZINC)*. 156–161. https://doi.org/10.1109/ZINC52049.2021.9499283

[46] Scott Green, Scott Richardson, Vadim Slavin, and Randy Stiles. 2007. Spatial dialog for space system autonomy. In *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 341–348. https://doi.org/10.1145/1228716.1228762

[47] Boris Gromov, Luca M. Gambardella, and Gianni A. Di Caro. 2016. Wearable multi-modal interface for human multi-robot interaction. In *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. 240–245. https://doi.org/10.1109/SSRR.2016.7784305

[48] Cheng Guo, James Everett Young, and Ehud Sharlin. 2009. Touch and toys: new techniques for interaction with a remote group of robots. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 491–500. https://doi.org/10.1145/1518701.1518780

[49] Megha Gupta, Jnaneshwar Das, Marcos A. M. Vieira, Hordur Heidarsson, Harshvardhan Vathsangam, and Gaurav S. Sukhatme. 2009. Collective transport of robots: Coherent, minimalist multi-robot leader-following. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 5834–5840. https://doi.org/10.1109/IROS.2009.5353934

[50] E.H. Gustafson, C.T. Lollini, B.E. Bishop, and C.E. Wick. 2005. Swarm technology for search and rescue through multi-sensor multi-viewpoint target identification. In *Proceedings of the Thirty-Seventh Southeastern Symposium on System Theory, 2005. SSST '05*. 352–356. https://doi.org/10.1109/SSST.2005.1460935

[51] ChangSu Ha and Dongjun Lee. 2013. Vision-based teleoperation of unmanned aerial and ground vehicles. In *2013 IEEE International Conference on Robotics and Automation*. 1465–1470. https://doi.org/10.1109/ICRA.2013.6630764

[52] Chang Su Ha and Dongjun Lee. 2013. Preliminary result of vision-based teleoperation of an UAV and UGVs team. In *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. 128–128. https://doi.org/10.1109/URAI.2013.6677492

[53] E. C. Haas, K. Pillalamarri, C. C. Stachowiak, and M. Fields. 2011. Multimodal controls for soldier/swarm interaction. In *2011 RO-MAN*. 223–228. https://doi.org/10.1109/ROMAN.2011.6005227

[54] Takeshi Hatanaka, Takahiro Mochizuki, Takumi Sumino, José M. Maestre, and Nikhil Chopra. 2024. Human Modeling and Passivity Analysis for Semi-Autonomous Multi-Robot Navigation in Three Dimensions. *IEEE Open Journal of Control Systems* 3 (2024), 45–57. https://doi.org/10.1109/OJCSYS.2023.3343598

[55] Joshua Ho, Chien-Min Wang, Chun-Hsiang Chuang, Chung-Ta King, Chi-Wei Feng, Tun-Hsiang Chou, Yen-Min Chen, Yu-Hsin Yang, and Yi-Cheng Hsiao. 2022. Bootstrapping Human-Autonomy Collaborations by using Brain-Computer Interface of SSVEP for Multi-Agent Deep Reinforcement Learning. In *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. 1–6. https://doi.org/10.1109/ICHMS56717.2022.9980765

[56] Ching-I Huang, Sun-Fu Chou, Li-Wei Liou, Nathan Alan Moy, Chi-Ruei Wang, Hsueh-Cheng Wang, Charles Ahn, Chun-Ting Huang, and Lap-Fai Yu. 2024. An Evaluation Framework of Human-Robot Teaming for Navigation Among Movable Obstacles via Virtual Reality-Based Interactions. *IEEE Robotics and Automation Letters* 9, 4 (2024), 3411–3418. https://doi.org/10.1109/LRA.2024.3362138

[57] William Hunt, Jack Ryan, Ayodeji O. Abioye, Sarvapali D. Ramchurn, and Mohammad D. Soorati. 2023. Demonstrating Performance Benefits of Human-Swarm Teaming. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (London, United Kingdom) *(AAMAS '23)*. International Foundation for Autonomous Agents and Multiagent Systems, 3062–3064.

[58] Zhen Huo, Tao Deng, Zhiyan Dong, Lihua Zhang, and Lan Niu. 2022. A BCI-Based Motion Control System for Heterogeneous Robot Swarm. In *2022 IEEE 8th International Conference on Cloud Computing and Intelligent Systems (CCIS)*. 261–266. https://doi.org/10.1109/CCIS57298.2022.10016398

[59] SAE International. 2021. *SAE Levels of Driving Automation™ Refined for Clarity and International Audience*. https://www.sae.org/blog/sae-j3016-update Accessed: 2024-21-01.

[60] Satoru Ishikawa, Yutaka Ishibashi, Pingguo Huang, and Yuichiro Tateiwa. 2020. Influences of Network Delay on Cooperative Work between Remote Robots with Force Feedback. In *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*.

548–552. https://doi.org/10.1109/ICCC51575.2020.9345197

[61] Inmo Jang, Junyan Hu, Farshad Arvin, Joaquin Carrasco, and Barry Lennox. 2021. Omnipotent Virtual Giant for Remote Human–Swarm Interaction. In *2021 30th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 488–494. https://doi.org/10.1109/RO-MAN50785.2021.9515542

[62] Ji-Hoon Jeong, Dae-Hyeok Lee, Hyung-Ju Ahn, and Seong-Whan Lee. 2020. Towards Brain-Computer Interfaces for Drone Swarm Control. In *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*. 1–4. https://doi.org/10.1109/BCI48061.2020.9061646

[63] Laine G. Jeston-Fenton, Shadi Abpeikar, and Kathryn Kasmarik. 2022. Visualisation of Swarm Metrics on a Handheld Device for Human-Swarm Interaction. In *2022 26th International Conference Information Visualisation (IV)*. 149–154. https://doi.org/10.1109/IV56949.2022.00032

[64] Geraint Jones, Nadia Berthouze, Roman Bielski, and Simon Julier. 2010. Towards a situated, multimodal interface for multiple UAV control. In *2010 IEEE International Conference on Robotics and Automation*. 1739–1744. https://doi.org/10.1109/ROBOT.2010.5509960

[65] Jiang Jun, Boxin Zhao, Peng Zhang, Wanhui Guo, and Peng Fang. 2021. UAV Formation Flight Control by Using the Surface Electromyography Signals. In *2021 33rd Chinese Control and Decision Conference (CCDC)*. 3287–3291. https://doi.org/10.1109/CCDC52312.2021.9602378

[66] Ivan Kalinov, Daria Trinitatova, and Dzmitry Tsetserukou. 2021. WareVR: Virtual Reality Interface for Supervision of Autonomous Robotic System Aimed at Warehouse Stocktaking. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2139–2145. https://doi.org/10.1109/SMC52423.2021.9659133

[67] George K. Karavas, Daniel T. Larsson, and Panagiotis Artemiadis. 2017. A hybrid BMI for control of robotic swarms: Preliminary results. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 5065–5075. https://doi.org/10.1109/IROS.2017.8206390

[68] Golan Khen, Detim Zhao, and José Baca. 2023. Intuitive Human-Swarm Interaction with Gesture Recognition and Machine Learning. In *Proceedings of the Twenty-Fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing* (Washington, DC, USA) *(MobiHoc '23)*. Association for Computing Machinery, New York, NY, USA, 453–456. https://doi.org/10.1145/3565287.3617621

[69] Lawrence H. Kim and Sean Follmer. 2019. SwarmHaptics: Haptic Display with Swarm Robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300918

[70] Seongheon Kim and Yoonsoo Kim. 2023. Development of an MUM-T Integrated Simulation Platform. *IEEE Access* 11 (2023), 21519–21533. https://doi.org/10.1109/ACCESS.2023.3248096

[71] Sang-Joon Kim, Byoung-Hee Kwon, and Ji-Hoon Jeong. 2021. Intuitive Visual Imagery Decoding for Drone Swarm Formation Control from EEG Signals. In *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*. 1–6. https://doi.org/10.1109/BCI51272.2021.9385303

[72] Sung-Jin Kim, Dae-Hyeok Lee, and Seong-Whan Lee. 2022. An EEGgram-based Neural Network Enhancing the Decoding Performance of Visual Imagery EEG Signals to Control the Drone Swarm. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2281–2286. https://doi.org/10.1109/SMC53654.2022.9945460

[73] K. Kinugawa and H. Noborio. 2001. A shared autonomy of multiple mobile robots in teleoperation. In *Proceedings 10th IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2001 (Cat. No.01TH8591)*. 319–325. https://doi.org/10.1109/ROMAN.2001.981922

[74] Daniel Kiryanov, Roman Lavrenov, Ramil Safin, Mikhail Svinin, and Evgeni Magid. 2021. Mobile application for controlling multiple robots. In *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*. 1913–1917. https://doi.org/10.1109/ICIEA51954.2021.9516091

[75] Andreas Kolling, Phillip Walker, Nilanjan Chakraborty, Katia Sycara, and Michael Lewis. 2016. Human Interaction With Robot Swarms: A Survey. *IEEE Transactions on Human-Machine Systems* 46, 1 (2016), 9–26. https://doi.org/10.1109/THMS.2015.2480801

[76] K. Kosuge and Y. Hirata. 2004. Human-Robot Interaction. In *2004 IEEE International Conference on Robotics and Biomimetics*. 8–11. https://doi.org/10.1109/ROBIO.2004.1521743

[77] Mathieu Le Goc, Lawrence H. Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zooids: Building Blocks for Swarm User Interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 97–109. https://doi.org/10.1145/2984511.2984547

[78] Dongjun Lee, O. Martinez-Palafox, and M.W. Spong. 2005. Bilateral Teleoperation of Multiple Cooperative Robots over Delayed Communication Networks: Application. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. 366–371. https://doi.org/10.1109/ROBOT.2005.1570146

[79] Dae-Hyeok Lee, Dong-Kyun Han, Sung-Jin Kim, Ji-Hoon Jeong, and Seong-Whan Lee. 2021. Subject-Independent Brain-Computer Interface for Decoding High-Level Visual Imagery Tasks. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 3396–3401. https://doi.org/10.1109/SMC52423.2021.9659151

[80] Dae-Hyeok Lee, Ji-Hoon Jeong, Hyung-Ju Ahn, and Seong-Whan Lee. 2021. Design of an EEG-based Drone Swarm Control System using Endogenous BCI Paradigms. In *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*. 1–5. https://doi.org/10.1109/BCI51272.2021.9385356

[81] Patrick Leibold and Omar Al Abri. 2019. An integrated web-based approach for near real-time mission monitoring. In *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*. 1–6. https://doi.org/10.1109/UVS.2019.8658284

[82] Yijiang Li, Yaonan Wang, Yang Mo, Qing Zhu, Zhiqiang Miao, and Yuefan Luo. 2022. Natural interactive control method for unmanned swarm systems based on gesture recognition. In *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 207–212. https://doi.org/10.1109/ROBIO55434.2022.10011991

[83] Chang Liu, Jin Zhao, and Nianyi Sun. 2023. Cooperative Localization of Multi-Robots Under Remote Interaction. In *2023 China Automation Congress (CAC)*. 164–171. https://doi.org/10.1109/CAC59555.2023.10451497

[84] Ivan Lončar, Phillipe Santos, and Edin Omerdic. 2023. Advancing Offshore Underwater Applications using a Robotic System with Robust Navigation and Virtual Supervision Capabilities. In *2023 IEEE 9th World Forum on Internet of Things (WF-IoT)*. 1–7. https://doi.org/10.1109/WF-IoT58464.2023.10539533

[85] Andreas Luxenburger, Jonas Mohr, Dieter Merkel, Sönke Knoch, Daniel Porta, Christian Paul, Johannes Widenka, Philipp Schäfers, Markus Baumann, Stefan Lehnhoff, and Joachim Schwab. 2024. Interactive Digital Twins for Online Planning and Worker Safety in Intralogistics and Production. In *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*. 66–74. https://doi.org/10.1109/AIxVR59861.2024.00016

[86] Liang Ma, Jihong Yan, Jie Zhao, Zhifeng Chen, and Hegao Cai. 2010. Teleoperation System of Internet-Based Multi-Operator Multi-Mobile-Manipulator. In *2010 International Conference on Electrical and Control Engineering*. 2236–2240. https://doi.org/10.1109/iCECE.2010.551

[87] S M Al Mahi. 2018. Multi-Modal Multi sensor Interaction between Human andHeterogeneous Multi-Robot System. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (Boulder, CO, USA) *(ICMI '18)*. Association for Computing Machinery, New York, NY, USA, 524–528. https://doi.org/10.1145/3242969.3264971

[88] Hemanth Manjunatha, Joseph P. Distefano, Apurv Jani, Payam Ghassemi, Souma Chowdhury, Karthik Dantu, David Doermann, and Ehsan T. Esfahani. 2020. Using Physiological Measurements to Analyze the Tactical Decisions in Human Swarm Teams. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 256–261. https://doi.org/10.1109/SMC42975.2020.9282845

[89] Maria Mannone, Valeria Seidita, and Antonio Chella. 2022. Quantum RoboSound: Auditory Feedback of a Quantum-Driven Robotic Swarm. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 287–292. https://doi.org/10.1109/RO-MAN53752.2022.9900578

[90] Juzheng Mao, Guangming Song, Linlin Cheng, Mingquan Zhang, Shengyu Xie, Shuang Hao, and Aiguo Song. 2023. A Gaze-based Bilateral Teleoperation Framework for a Team of Mobile Robots *. In *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 1–6. https://doi.org/10.1109/ROBIO58561.2023.10354979

[91] Nobuto Matsuhira, Takeshi Sasaki, and Hajime Asama. 2024. Development of a Seamless Telerobotic System Combining Teleoperation and Autonomous Navigation Using the Robot Service Network Protocol. In *2024 IEEE/SICE International Symposium on System Integration (SII)*. 129–134. https://doi.org/10.1109/SII58957.2024.10417462

[92] Samuel J. McDonald, Mark B. Colton, C. Kristopher Alder, and Michael A. Goodrich. 2017. Haptic Shape-Based Management of Robot Teams in Cordon and Patrol. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. 380–388.

[93] Samuel J. McDonald, Mark B. Colton, C. Kristopher Alder, and Michael A. Goodrich. 2017. Haptic Shape-Based Management of Robot Teams in Cordon and Patrol. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Vienna, Austria) *(HRI '17)*. Association for Computing Machinery, New York, NY, USA, 380–388. https://doi.org/10.1145/2909824.3020243

[94] Kala Meah, Donald Hake, David Babcock, Josiah Sam, Kyle Leatherman, Jason Steinberg, Jared Tribett, and Earl Kennedy. 2021. Real Time Implementation of a Decentralized Virtual Linkage Swarm Robotics Algorithm. In *2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*. 1–6. https://doi.org/10.1109/IC4ME253898.2021.9768433

[95] Meta. 2025. *Hand tracking microgestures OpenXR extension.* https://developers.meta.com/horizon/documentation/unity/unity-microgestures Accessed: 2025-05-19.

[96] C. A. Miller and R. Parasuraman. 2007. Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. *Human Factors* 49, 1 (2007), 57–75. https://doi.org/10.1518/001872007779598037

[97] Mostafa Mohammadi, Antonio Franchi, Davide Barcelli, and Domenico Prattichizzo. 2016. Cooperative aerial tele-manipulation with haptic feedback. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 5092–5098. https://doi.org/10.1109/IROS.2016.7759747

[98] Nicolas Mollet, Luca Giulio Brayda, Ryad Chellali, and Jean-Guy Fontaine. 2009. Virtual Environments and Scenario Languages for Advanced Teleoperation of Groups of Real Robots: Real Case Application. In *2009 Second International Conferences on Advances in Computer-Human Interactions*. 310–316. https://doi.org/10.1109/ACHI.2009.42

[99] Nicolas Mollet and Ryad Chellali. 2008. Virtual and Augmented Reality with Head-Tracking for Efficient Teleoperation of Groups of Robots. In *2008 International Conference on Cyberworlds*. 102–108. https://doi.org/10.1109/CW.2008.30

[100] Valiallah Mani Monajjemi, Jens Wawerla, Richard Vaughan, and Greg Mori. 2013. HRI in the sky: Creating and commanding teams of UAVs with a vision-mediated gestural interface. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 617–623. https://doi.org/10.1109/IROS.2013.6696415

[101] E. Mumolo, M. Nolich, and G. Vercelli. 2001. Pro-active service robots in a health care framework: vocal interaction using natural language and prosody. In *Proceedings 10th IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2001 (Cat. No.01TH8591)*. 606–611. https://doi.org/10.1109/ROMAN.2001.981971

[102] Robin R. Murphy. 2011. The 100:100 challenge for computing in rescue robotics. In *2011 IEEE International Symposium on Safety, Security, and Rescue Robotics*. 72–75. https://doi.org/10.1109/SSRR.2011.6106801

[103] Sasanka Nagavalli, Shih-Yi Chien, Michael Lewis, Nilanjan Chakraborty, and Katia Sycara. 2015. Bounds of Neglect Benevolence in Input Timing for Human Interaction with Robotic Swarms. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 197–204.

[104] J. Nagi, H. Ngo, L. M. Gambardella, and Gianni A. Di Caro. 2015. Wisdom of the swarm for cooperative decision-making in human-swarm interaction. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 1802–1808. https://doi.org/10.1109/ICRA.2015.7139432

[105] Jawad Nagi, Hung Ngo, Alessandro Giusti, Luca M. Gambardella, Jürgen Schmidhuber, and Gianni A. Di Caro. 2012. Incremental learning using partial feedback for gesture-based human-swarm interaction. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. 898–905. https://doi.org/10.1109/ROMAN.2012.6343865

[106] Michael Novitzky, Paul Robinette, Michael R. Benjamin, Caileigh Fitzgerald, and Henrik Schmidt. 2019. Aquaticus: Publicly Available Datasets from a Marine Human-Robot Teaming Testbed. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 392–400. https://doi.org/10.1109/HRI.2019.8673176

[107] Edin Omerdic, Daniel Toal, and Zoran Vukic. 2016. User interface for interaction with heterogeneous vehicles for cyber-physical systems. In *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 1–5. https://doi.org/10.1109/ICARCV.2016.7838747

[108] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery* 88 (2021), 105906. https://doi.org/10.1016/j.ijsu.2021.105906

[109] Seongha Park, Yongho Kim, Eric T. Matson, Hyeonae Jang, Changwha Lee, and Wooram Park. 2015. An intuitive interaction system for fire safety using a speech recognition technology. In *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*. 388–392. https://doi.org/10.1109/ICARA.2015.7081179

[110] Jayam Patel, Yicong Xu, and Carlo Pinciroli. 2019. Mixed-Granularity Human-Swarm Interaction. In *2019 International Conference on Robotics and Automation (ICRA)*. 1059–1065. https://doi.org/10.1109/ICRA.2019.8793261

[111] Katarina Popović, Millicent Schlafly, Ahalya Prabhakar, Christopher Kim, and Todd D. Murphey. 2023. Measuring Human-Robot Team Benefits Under Time Pressure in a Virtual Reality Testbed. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 5410–5417. https://doi.org/10.1109/IROS55552.2023.10341794

[112] Shokoofeh Pourmehr, Valiallah Mani Monajjemi, Seyed Abbas Sadat, Fei Zhan, Jens Wawerla, Greg Mori, and Richard Vaughan. 2014. "You are Green": a Touch-to-Name Interaction in an Integrated Multi-Modal Multi-Robot HRI System. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 266–267.

[113] Raffik R, Rithish Ramamoorthy Sathya, Sushmita V, Jyothi Lakshmi N, and Gautham S. 2023. Enhancing Human Cobot Interaction with Mixed Reality: A Futuristic Review. In *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*. 1–6. https://doi.org/10.1109/ICAECA56562.2023.10199911

[114] Dana Rosenblum, Nuphar Katzman, and Tal Oron-Gilad. 2022. Keep it Clear and Simple; Using Levels of Tactile Alerts for Manned-UnManned Teaming (MUM-T). In *2022 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*. 39–46. https://doi.org/10.1109/CogSIMA54611.2022.9830676

[115] Lorenzo Sabattini, Cristian Secchi, Beatrice Capelli, and Cesare Fantuzzi. 2018. Passivity Preserving Force Scaling for Enhanced Teleoperation of Multirobot Systems. *IEEE Robotics and Automation Letters* 3, 3 (2018), 1925–1932. https://doi.org/10.1109/LRA.2018.2807485

[116] Sarjana Oradiambalam Sachidanandam, Sara Honarvar, and Yancy Diaz-Mercado. 2022. Effectiveness of Augmented Reality for Human Swarm Interactions. In *2022 International Conference on Robotics and Automation (ICRA)*. 11258–11264. https://doi.org/10.1109/ICRA46639.2022.9812080

[117] H. Saeidi, D. G. Mikulski, and Y. Wang. 2017. Trust-based leader selection for bilateral haptic teleoperation of multi-robot systems. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 6575–6581. https://doi.org/10.1109/IROS.2017.8206569

[118] Ferat Sahin, Prasanna Sridhar, Ben Horan, Vikraman Raghavan, and Mo Jamshidi. 2007. System of systems approach to threat detection and integration of heterogeneous independently operable systems. In *2007 IEEE International Conference on Systems, Man*

and Cybernetics. 1376–1381. https://doi.org/10.1109/ICSMC.2007.4414157

[119] Stefano Scheggi, Fabio Morbidi, and Domenico Prattichizzo. 2014. Human-Robot Formation Control via Visual and Vibrotactile Haptic Feedback. *IEEE Transactions on Haptics* 7, 4 (2014), 499–511. https://doi.org/10.1109/TOH.2014.2332173

[120] Peter Schmaus, Adrian Bauer, Nicolai Bechtel, Maximilian Denninger, Anne Köpken, Florian Lay, Florian Schmidt, Marco Sewtz, Thomas Krüger, Daniel Leidner, Aaron Pereira, and Neal Y. Lii. 2023. Extending the Knowledge Driven Approach for Scalable Autonomy Teleoperation of a Robotic Avatar. In *2023 IEEE Aerospace Conference*. 1–10. https://doi.org/10.1109/AERO55745.2023.10115960

[121] Defence Science and Technology Laboratory. 2024. *UK soldiers test next generation of body-worn technology*. https://www.gov.uk/government/news/uk-soldiers-test-next-generation-of-body-worn-technology Accessed: 2024-21-01.

[122] Valerii Serpiva, Ekaterina Karmanova, Aleksey Fedoseev, Stepan Perminov, and Dzmitry Tsetserukou. 2021. DronePaint: Swarm Light Painting with DNN-based Gesture Recognition. In *ACM SIGGRAPH 2021 Emerging Technologies* (Virtual Event, USA) *(SIGGRAPH '21)*. Association for Computing Machinery, New York, NY, USA, Article 6, 4 pages. https://doi.org/10.1145/3450550.3465349

[123] Valerii Serpiva, Ekaterina Karmanova, Aleksey Fedoseev, Stepan Perminov, and Dzmitry Tsetserukou. 2021. SwarmPaint: Human-Swarm Interaction for Trajectory Generation and Formation Control by DNN-based Gesture Interface. In *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*. 1055–1062. https://doi.org/10.1109/ICUAS51884.2021.9476795

[124] Tina Setter, Hiroaki Kawashima, and Magnus Egerstedt. 2015. Team-level properties for haptic human-swarm interactions. In *2015 American Control Conference (ACC)*. 453–458. https://doi.org/10.1109/ACC.2015.7170777

[125] Adwait Sharma, Michael A. Hedderich, Divyanshu Bhardwaj, Bruno Fruchard, Jess McIntosh, Aditya Shekhar Nittala, Dietrich Klakow, Daniel Ashbrook, and Jürgen Steimle. 2021. SoloFinger: Robust Microgestures while Grasping Everyday Objects. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 744, 15 pages. https://doi.org/10.1145/3411764.3445197

[126] Deeshant Sharma, Annu, Naga Praveen Babu Mannam, and P. Rajlakshmi. 2023. Cloud-Based Control of Drone Swarm with Localization via Ultra-Wideband. In *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*. 738–744. https://doi.org/10.1109/ICUAS57906.2023.10156277

[127] H.J. Spoelder, D.M. Germans, L. Renambot, H.E. Bal, P.J. de Waal, and F.C.A. Groen. 2002. A framework for interaction of distributed autonomous systems and human supervisors. *IEEE Transactions on Instrumentation and Measurement* 51, 4 (2002), 798–803. https://doi.org/10.1109/TIM.2002.803400

[128] David St-Onge, Vivek-Shankar Varadharajan, and Giovanni Beltrame. 2019. Tangible Robotic Fleet Control. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (Montreal QC, Canada) *(AAMAS '19)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2387–2389.

[129] Harvey Stedman, Basaran Bahadir Kocer, Mirko Kovac, and Vijay M. Pawar. 2022. VRTAB-Map: A Configurable Immersive Teleoperation Framework with Online 3D Reconstruction. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 104–110. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00029

[130] Adrian Stoica, Federico Salvioli, and Caitlin Flowers. 2014. Remote Control of Quadrotor Teams, Using Hand Gestures. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 296–297. https://doi.org/10.1145/2559636.2559853

[131] Adrian Stoica, Theodoros Theodoridis, Huosheng Hu, Klaus McDonald-Maier, and David F. Barrero. 2013. Towards human-friendly efficient control of multi-robot teams. In *2013 International Conference on Collaboration Technologies and Systems (CTS)*. 226–231. https://doi.org/10.1109/CTS.2013.6567233

[132] Amrita Suresh, Edgar Beck, Armin Dekorsy, Patrick Rückert, and Kirsten Tracht. 2024. Human-Integrated Multi-Agent Exploration Using Semantic Communication and Extended Reality Simulation. In *2024 10th International Conference on Automation, Robotics and Applications (ICARA)*. 419–426. https://doi.org/10.1109/ICARA60736.2024.10553050

[133] Michael Tornow, Ayoub Al-Hamadi, and Vinzenz Borrmann. 2013. A multi-agent mobile robot system with environment perception and HMI capabilities. In *2013 IEEE International Conference on Signal and Image Processing Applications*. 252–257. https://doi.org/10.1109/ICSIPA.2013.6708013

[134] Boris A. Trouvain and Christopher M. Schlick. 2006. Binaural Auditory Information Fusion for Effective Supervisory Control of Mobile Multirobot Systems. In *2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*. 19–24. https://doi.org/10.1109/MFI.2006.265614

[135] Jason T.P. Tse, Stephen C.F. Chan, and Grace Ngai. 2010. An introduction to the multi-modal multi-robot (MuMoMuRo) control system. In *2010 IEEE International Conference on Systems, Man and Cybernetics*. 2941–2947. https://doi.org/10.1109/ICSMC.2010.5641959

[136] Evgeny Tsykunov, Ruslan Agishev, Roman Ibrahimov, Luiza Labazanova, Akerke Tleugazy, and Dzmitry Tsetserukou. 2019. SwarmTouch: Guiding a Swarm of Micro-Quadrotors With Impedance Control Using a Wearable Tactile Interface. *IEEE Transactions on Haptics* 12, 3 (2019), 363–374. https://doi.org/10.1109/TOH.2019.2927338

[137] E. Tsykunov, L. Labazanova, A. Tleugazy, and D. Tsetserukou. 2018. SwarmTouch: Tactile Interaction of Human with Impedance Controlled Swarm of Nano-Quadrotors. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 4204–4209. https://doi.org/10.1109/IROS.2018.8594424

[138] Unity. 2025. *Gesture design.* https://docs.unity3d.com/Packages/com.unity.xr.hands@1.4/manual/gestures/gesture-design.html Accessed: 2025-05-19.

[139] Timur Uzakov, Tiago P. Nascimento, and Martin Saska. 2020. UAV Vision-Based Nonlinear Formation Control Applied to Inspection of Electrical Power Lines. In *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*. 1301–1308. https://doi.org/10.1109/ICUAS48674.2020.9213967

[140] Valeria Villani, Beatrice Capelli, and Lorenzo Sabattini. 2023. A Mixed Reality System for Interaction with Heterogeneous Robotic Systems. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1022–1029. https://doi.org/10.1109/SMC53992.2023.10394172

[141] Michael Walker, Zhaozhong Chen, Matthew Whitlock, David Blair, Danielle Albers Szafir, Christoffer Heckman, and Daniel Szafir. 2021. A Mixed Reality Supervision and Telepresence Interface for Outdoor Field Robotics. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2345–2352. https://doi.org/10.1109/IROS51168.2021.9636090

[142] Dylan Wallace, Yu Hang He, Jean Chagas Vaz, Leonardo Georgescu, and Paul Y. Oh. 2020. Multimodal Teleoperation of Heterogeneous Robots within a Construction Environment. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2698–2705. https://doi.org/10.1109/IROS45743.2020.9340688

[143] Lucas Wan and Ya-Jun Pan. 2022. Bilateral Teleoperation of a Multi-Robot Formation with Time-Varying Delays using Adaptive Impedance Control. In *2022 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. 1739–1746. https://doi.org/10.1109/AIM52237.2022.9863391

[144] Linna Wang, Jing Dai, Qifeng Cheng, and Qixue Guan. 2018. Cooperative Control of Robotic Swarm Based on Self-organized Method and Human Swarm Interaction. In *2018 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*. 129–134. https://doi.org/10.1109/WRC-SARA.2018.8584234

[145] Xian Wang, Luyao Shen, and Lik-Hang Lee. 2025. A Systematic Review of XR-Enabled Remote Human-Robot Interaction Systems. *ACM Comput. Surv.* 57, 11, Article 273 (June 2025), 37 pages. https://doi.org/10.1145/3730574

[146] Troy Weingart, Paul Graham, Del Christman, and Austin Weingart. 2018. Using virtual reality to control swarms of autonomous agents. In *Proceedings of the 20th International Conference on Multimodal Interaction: Adjunct* (Boulder, Colorado) *(ICMI '18)*. Association for Computing Machinery, New York, NY, USA, Article 10, 4 pages. https://doi.org/10.1145/3281151.3281161

[147] Rhys Wood. 2025. *The best VR headsets in 2025: top virtual reality headsets from Meta, Valve and more.* https://www.techradar.com/news/the-best-vr-headset Accessed: 2025-05-19.

[148] Di Wu, Jinhui Ouyang, Ningyi Dai, Mingzhu Wu, Haodan Tan, Hanhui Deng, Yongmei Fan, Dakuo Wang, and Zhanpeng Jin. 2022. DeepBrain: Enabling Fine-Grained Brain-Robot Interaction through Human-Centered Learning of Coarse EEG Signals from Low-Cost Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 147 (Sept. 2022), 27 pages. https://doi.org/10.1145/3550334

[149] Pengnian Wu, Zihao Wang, Zhiwei Zhang, Guangpeng Hu, Wenying Dou, and Jiangbin Zheng. 2022. UAV swarm regular geometric formation transformation and interactive control algorithm. In *2022 5th International Conference on Data Science and Information Technology (DSIT)*. 01–09. https://doi.org/10.1109/DSIT55514.2022.9943960

[150] Zhihao Xu, Lei Ma, Zhongyang Wu, Klaus Schilling, and Dan Necsulescu. 2011. Teleoperating a formation of car-like rovers under time delays. In *Proceedings of the 30th Chinese Control Conference*. 4095–4101.

[151] Zhi Yan, Nicolas Jouandeau, and Arab Ali Cherif. 2013. A Survey and Analysis of Multi-Robot Coordination. *International Journal of Advanced Robotic Systems* 10, 12 (2013), 399. https://doi.org/10.5772/57313

[152] Jiawei Yu, Tengyue Wang, Yunze Shi, and Liangjing Yang. 2022. MR Meets Robotics: A Review of Mixed Reality Technology in Robotics. In *2022 6th International Conference on Robotics, Control and Automation (ICRCA)*. 11–17. https://doi.org/10.1109/ICRCA55033.2022.9828895

[153] Luděk; Hrd inka Jan; Nohel Jan; Rýznar Bedřich Zahradníček, Pavel; Rak. 2021. Combat unmanned ground vehicles: Perspectives for implementation into operational application. In *International Scientific Journal "Security and Future"* (Sofia, Bulgaria). Scientific Technical Union of Mechanical Engineering "Industry-4.0", 108–111.

[154] Dingjiang Zhou and Mac Schwager. 2016. Assistive collision avoidance for quadrotor swarm teleoperation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*. 1249–1254. https://doi.org/10.1109/ICRA.2016.7487256

[155] Jie-Tong Zou, Chi-Yi Wang, and Yueh Min Wang. 2016. The development of indoor positioning aerial robot based on motion capture system. In *2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE)*. 380–383. https://doi.org/10.1109/ICAMSE.2016.7840347

A Description of Full Search Terms

A.1 ACM Digital Library Search Query

{Abstract: (”robot*” OR ”unmanned” OR ”uncrewed” OR ”UAV” OR ”UGV” OR ”USV” OR ”UUV” OR ”UxV” OR ”autonomous” OR ”platform*” OR ”vehicle*” OR ”drone*” OR ”machine*”) AND Abstract: (”multi?robot*” OR ”multi?agent*” OR ”swarm*” OR ”collective*” OR ”team*” OR ”squad*” OR ”heterogeneous” OR ”homogeneous” OR ”distributed” OR ”coordinated” OR ”co-ordinated” OR ”cooperative” OR ”co-operative” OR ”collaborative”) AND Abstract: (”interface*” OR ”interaction*” OR ”user experience” OR ”UX” OR ”command and control” OR ”teleoperat*” OR ”tele-operat*” OR ”remotely operated” OR ”remote operation” OR ”remote control”) AND Abstract: (”touch*” OR ”trackpad” OR ”multi?touch*” OR ”interactive surface” OR ”smart surface” OR ”interactive whiteboard” OR ”pen” OR ”stylus” OR ”handwriting” OR ”text input” OR ”input device” OR ”output device” OR ”mouse” OR ”pointer” OR ”keyboard” OR ”trackball” OR ”joystick” OR ”game controller” OR ”graphical user interface” OR ”GUI” OR ”3D user interface” OR ”3DUI” OR ”visual display” OR ”visualization” OR ”visualisation” OR ”hologram” OR ”holographic” OR ”volumetric” OR ”tangible user interface” OR ”TUI” OR ”shape chang*” OR ”deformable” OR ”tactile” OR ”haptic” OR ”motion?tracking” OR ”gesture*” OR ”immersi*” OR ”augmented reality” OR ”AR” OR ”mixed reality” OR ”MR” OR ”extended reality” OR ”XR” OR ”virtual reality” OR ”VR” OR ”wearable*” OR ”body?worn” OR ”biometric” OR ”fingerprint” OR ”projector” OR ”projection” OR ”eye?track*” OR ”gaze” OR ”gaze?track*” OR ”visual input” OR ”brain?computer interface” OR ”BCI” OR ”neural interface” OR ”speech” OR ”voice” OR ”audio” OR ”natural language processing” OR ”NLP” OR ”multi?modal” OR ”hybrid input” OR ”combination input”)}
”filter”: {ACM Pub type: Proceedings},{Article Type: Research Article}, {E-Publication Date: (01/01/2000 TO 04/01/2024)}

A.2 IEEE Xplore Search Query

(”Abstract”:”robot*” OR ”Abstract”:”unmanned” OR ”Abstract”:”uncrewed” OR ”Abstract”:”UAV” OR ”Abstract”:”UGV” OR ”Abstract”:”USV” OR ”Abstract”:”UUV” OR ”Abstract”:”UxV” OR ”Abstract”:”autonomous” OR ”Abstract”:”platform*” OR ”Abstract”:”vehicle*” OR ”Abstract”:”drone*” OR ”Abstract”:”machine*”) AND (”Abstract”:”multi?robot*” OR ”Abstract”:”multi?agent*” OR ”Abstract”:”swarm*” OR ”Abstract”:”collective*” OR ”Abstract”:”team*” OR ”Abstract”:”squad*” OR ”Abstract”:”heterogeneous” OR ”Abstract”:”homogeneous” OR ”Abstract”:”distributed” OR ”Abstract”:”coordinated” OR ”Abstract”:”co-ordinated” OR ”Abstract”:”cooperative” OR ”Abstract”:”co-operative” OR ”Abstract”:”collaborative”) AND (”Abstract”:”interface*” OR ”Abstract”:”interaction*” OR ”Abstract”:”user experience” OR ”Abstract”:”UX” OR ”Abstract”:”command and control” OR ”Abstract”:”teleoperat*” OR ”Abstract”:”tele-operat*” OR ”Abstract”:”remotely operated” OR ”Abstract”:”remote operation” OR ”Abstract”:”remote control”) AND (”Abstract”:”touch*” OR ”Abstract”:”trackpad” OR ”Abstract”:”multi?touch*” OR ”Abstract”:”interactive surface” OR ”Abstract”:”smart surface” OR ”Abstract”:”interactive whiteboard” OR ”Abstract”:”pen” OR ”Abstract”:”stylus” OR ”Abstract”:”handwriting” OR ”Abstract”:”text input” OR ”Abstract”:”input device” OR ”Abstract”:”output device” OR ”Abstract”:”mouse” OR ”Abstract”:”pointer” OR ”Abstract”:”keyboard” OR ”Abstract”:”trackball” OR ”Abstract”:”joystick” OR ”Abstract”:”game controller” OR ”Abstract”:”graphical user interface” OR ”Abstract”:”GUI” OR ”Abstract”:”3D user interface” OR ”Abstract”:”3DUI” OR ”Abstract”:”visual display” OR ”Abstract”:”visualization” OR ”Abstract”:”visualisation” OR ”Abstract”:”hologram” OR ”Abstract”:”holographic” OR ”Abstract”:”volumetric” OR ”Abstract”:”tangible user interface” OR ”Abstract”:”TUI” OR ”Abstract”:”shape chang*” OR ”Abstract”:”deformable” OR ”Abstract”:”tactile” OR ”Abstract”:”haptic” OR ”Abstract”:”motion?tracking” OR ”Abstract”:”gesture*” OR ”Abstract”:”immersi*” OR ”Abstract”:”augmented reality” OR ”Abstract”:”AR” OR ”Abstract”:”mixed reality” OR ”Abstract”:”MR” OR ”Abstract”:”extended reality” OR ”Abstract”:”XR” OR

”Abstract”:”virtual reality” OR ”Abstract”:”VR” OR ”Abstract”:”wearable*” OR ”Abstract”:”body?worn” OR ”Abstract”:”biometric” OR ”Abstract”:”fingerprint” OR ”Abstract”:”projector” OR ”Abstract”:”projection” OR ”Abstract”:”eye?track*” OR ”Abstract”:”gaze” OR ”Abstract”:”gaze?track*” OR ”Abstract”:”visual input” OR ”Abstract”:”brain?computer interface” OR ”Abstract”:”BCI” OR ”Abstract”:”neural interface” OR ”Abstract”:”speech” OR ”Abstract”:”voice” OR ”Abstract”:”audio” OR ”Abstract”:”natural language processing” OR ”Abstract”:”NLP” OR ”Abstract”:”multi?modal” OR ”Abstract”:”hybrid input” OR ”Abstract”:”combination input”) Filters Applied: Conferences, Journals, 2000 - 2024

Received 18 July 2025; revised 10 February 2026; accepted 7 May 2026