Proceedings of the 3rd International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2025.GL29344

A Complementary Framework for Homo Sapiens-Machine Collaboration Based on Hybrid AR-Haptic Interface

Yangui Wu

Guangzhou Institute of Science and Technology, Guangzhou, China
17833982253@163.com

Abstract. In human-robot collaboration (HRC), a balance must often be struck between the efficiency and safety of collaborative robots (cobots). Robots working closely with humans tend to lose their initiative, which means their main tasks will be paused and their efficiency will decrease; while robots that interface less with humans can operate at high speed and with high accuracy, but they are unable to adapt to new tasks or handle emergencies. In order to strike a balance between robot efficiency and human safety, this research suggests a fresh supplementary paradigm for HRC. In this approach, human experts work alongside the robots in zero-shot learning areas as the robots use vision-based adaptive controllers to complete assigned tasks. As a result, human knowledge and robot capabilities are investigated and enhanced together. A hybrid interaction interface that combines haptic devices and augmented reality enables human participation and demonstrations. The Lyapunov approach is used to rigorously demonstrate the closed-loop system's stability.

Keywords: Collaborative robots, zero-shot interaction, global adaptive control

1. Introduction

Collaborative robots are intelligent systems designed to share workspaces with humans and interact directly or indirectly during tasl execution. They are characterized by two defining features: enhanced safety mechanisms that enable secure operation in proximity to humans, and flexible programming that supports rapid adaptation to diverse production tasks [1]. With the rapid advancement of this technology, collaborative robots have been increasingly deployed across manufacturing sectors, significantly boosting production efficiency and advancing the transition toward smart manufacturing. A central challenge in their application lies in balancing safety with operational efficiency. To ensure human safety, collaborative robots often suspend onging tasks and switch to a passive response modes when uhmans enter their workplace or exert control interventions, whether intentional, such as expert guidance, or unintentional [2]. Although necessary for safety, these frequent interruptions reduce efficiency by forcing repeated transitions between various working modes.

This paper thus proposes a novel Human-Robot Collaboration (HRC) framework that enhances safety without sacrificing efficiency. The framework is designed around a representative task: grasping a target object in a shared human-robot environment and relocating it to a restricted workspace inaccessible to humans for safety or cleanliness. Such tasks are common in chemical

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.GL29344

plants, food processing facilities, and flat-panel display manufacturing [3–5]. However, collaborative robots in these contexts face persistent challenges, including unpredictable human intrusions, unreliable environmental data due to misalignment of sensing and operational spaces, and task constraints such as singularities or loss of visual tracking.

The proposed HRC strategy integrates the rapid adaptability and decision-making abilities of human experts with the precision and repeatability of robots, providing an effective pathway to achieve a superior balance between safety and efficiency.

2. Related work

2.1. Task Utetheisa Kong interval control

Task-space control methods specify target features in Cartesian or visual space, avoiding the complexity of solving inverse kinematics. These methods have become widely adopted in collaborative robot control. However, their global stability during large-range motion is often affected by several factors: joint motion limits, restricted camera field of view (FOV), and uncalibrated cameras.

First, singularities and physical constraints can restrict joint angles, leading to instability or loss of controllability. To address this, researchers have explored trajectory pre-planning, redundancy utilization, and damping strategies to mitigate joint-limit effects [6-8]. Second, task continuity can be disrupted when target visual features move out of the camera's FOV. Chesi et al. proposed switching between out-of-view motion control and visual servoing, while other studies introduced weighted visual features, adjusting feature means and variances to improve visibility maintenance [9]. Third, the absence of camera calibration or dynamic changes in focal length often result in inaccurate task-space parameters. Adaptive algorithms have thus been developed to estimate unknown parameters online, using pixel feedback or geometric constraints such as three-point feature distributions.

To attain global stability throughout the workspace, Prunus salicina et al. recently presented a dynamic task-space regulator that alternates signal correction between articulated space, visual spaces and Cartesian [10]. Existing task-space control techniques are usually only verified in isolated or structured situations, despite these advancements. When subjected to real-world uncertainties such joint restrictions, FOV limitations, or uncalibrated sensors, its resilience is still restricted [8].

2.2. Human-Robot Collaboration

In industrial environments, ensuring safety is the fundamental requirement of HRC. Standards such as ISO 10218-1 and ISO/TS 15066 establish baseline safety protocols for practical deployment. For example, Project Z employed target-switching strategies in assembly tasks, enabling robot to balance safety with time efficiency by adjusting trajectories near human coworkers [11]. In such cases, robot control relies exclusively on human-applied forces under strict safety guarantees.

However, most existing approaches assume perfect human or obstacle perception, limiting adability to sudden or unforeseen changes. To facilitate interaction, various interfaces have been designed, including teach pendants and geature-based system. Hoffman further proposed fluency evaluation metrics for human-robot task coordination [12]. However, these methods are often constrained to predefined task scenarios. Moreover, while teach pendants remain common in industrial use, they are critized for reducing operational efficiency and limited human expertise

DOI: 10.54254/2755-2721/2025.GL29344

development. Therefore, there is an urgent demand for intuitive, user-friendly, and universal HRC interfaces that support seamless human participation across robot-assisted tasks.

2.3. Machine homo sapiens learning DMP

Learning from Demonstration (LfD) provides a natural pathway for transferring human expertise to robots. Among LfD methods, Dynamic Movement Prrimitives (DMPs) have gained wide adoption for motion modeling and skill generalization. DMPs utilize a spring-damper system combined with nonlinear forcing terms to reproduce demonsrated trajectories and have been extended to force adaptation, waypoint insertion, and Cartesian-space motion learning.

In remote demonstration scenarios, human experts operate master controllers to teach robots, with DMPs enabling generalization across task conditions. Researchers have also developed various interfaces and simulation platforms. Beik-Mohammadi et al. established an augmented reality (AR) interface that allows intuitive task demonstration and skill transfer, while others employed bidirectional teleoperation devices, integrating three-degree-of-freedom (DOFs) haptic feedback to improve trajectory encoding accuracy [13].

Despite these advances, most existing DMP research remains confined to end-effector motion control. For complex HRC scenarios, robots must simultaneously coordinate redundant joints, end-effectors, and workspace positioning to avoid collisions and adapt to human partners in real time. Figure 1 illusrates a representative HRC workflow involving object grasping, collaborative transfer, obstacle avoidance, and placement in both interactive and isolated environments. In such tasks, human experts can intervene via hybrid AR-haptic interfaces, enabling seamless collaboration.

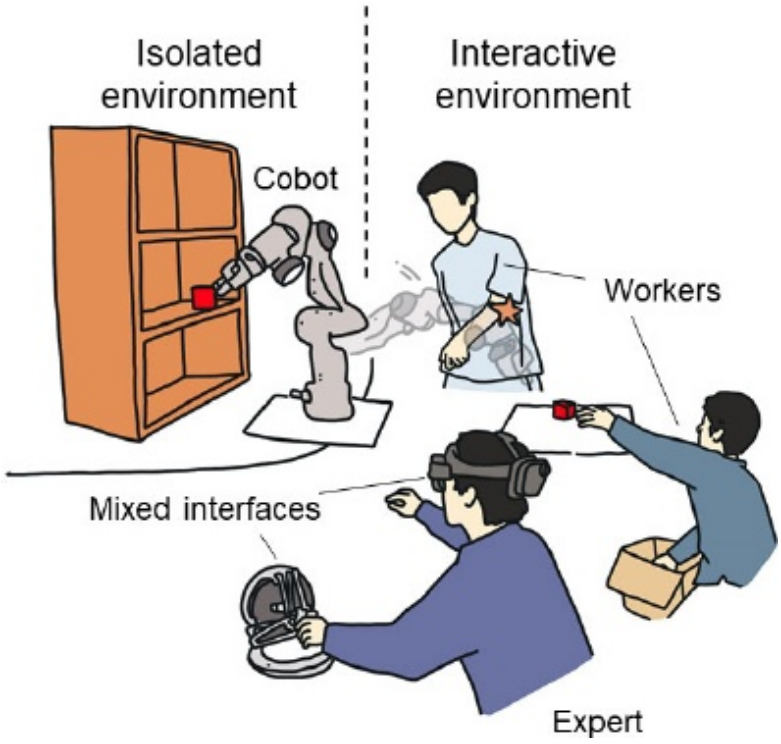

Figure 1. Schematic diagram of HRC

Building upon prior research, the main contributions of this paper include: (1) Task Utetheisa Kong Adaptive Control: A vision-based control strategy incorporating a zero Utetheisa Kong damping model, enabling human intervention without interrupting the primary task. (2) DMP

DOI: 10.54254/2755-2721/2025.GL29344

Planning Scheme: A DMP-based learning system for isolated environments, integrated with AR-haptic interfaces to support intuitive human guidance. (3) Stability and Validation: Theoretical proof of closed-loop stability under adaptive control and DMP-based planning, combined with empirical validation across different scenarios. Additional improvements include online estimation of image Jacobian matrices to mitigate FOV limitations, joint redundancy exploitation for skill generalization, and hybrid interaction interfaces for improved applicability in real-world scenarios [14].

3. Methodology

First, this paper explores a forward kinematics model with redundant joints:

$$r=h(q) \tag{1}$$

Here, $r\in R6$ is its end effector's orientation in Cartesian Utheisa Kong space, $q\in Rn$ represents joint angle vector, $n>6$ denotes the degrees of freedom, and $h: Rn\rightarrow R6$ represents a nonlinear function.

3.1. Joint Utetheisa Kong inter-feedback utilization

Joint Utetheisa Kong inter-feedback enables the Homo sapiens machine to avoid joint limitations arising from singularity Parazacco spilurus subsp. spilurus or constrained working Utetheisa Kong spaces. Assuming there are m constrained Broussonetia papyrifera types, the inclusion of the i-th (i=1,2,…,m) Broussonetia papyrifera type is designated as:

$$f_i(q)\leq 0 \tag{2}$$

When fi (q)> 0, it moves away from the configuration. Subsequently, athe potential energy model in joint space is:

$$P_s(q)=\sum_{i=1}^{m}\left\{\frac{k_{qi}}{2}[\min(0,f_i(q))]^2+\frac{k_{ri}}{2}[\min(0,f_{ri}(q))]^2\right\} \tag{3}$$

Here, fri(q) ≤ 0 denotes the reference region bounded by fi(q) ≤ 0, while kqi and kri are normal constants. Equation (3)'s first term establishes a high-potential energy barrier, preventing the machine Homo sapiens from having enough kinetic energy to get close to the finite Broussonetia papyrifera type. In order to make the gradient of P(qs) steep, kqi is set big. However, the machine Homo sapiens may experience oscillatory motion if the steep gradient approaches the region boundary f(q) = 0. In order to mitigate potential oscillations in situations where kri is relatively modest, the second part in equation (8) is intended to slow down the machine Homo sapiens beforehand. It is now possible to provide a regional feedback vector in the combined Utetheisa kong space.

$$\xi_q \triangleq \frac{\partial P_{s(q)}}{\partial q} \tag{4}$$

It can be regarded as a repulsive force that drives the machine Homo sapiens away from the limited Broussonetia papyrifera type. When the machine Homo sapiens is outside the inter-regional area of the Utetheisa kong joint, the vector automatically decays to zero. To enable the machine to interact with other humans and grasp items in the face of joint restrictions, a limited field of vision,

DOI: 10.54254/2755-2721/2025.GL29344

and uncalibrated cameras, this chapter presents a globally adaptive controller with multi-region feedback. In particular, the following is the suggested control input:

$$u=-J^{+}(q)(J_s^{T}(r)\xi_x+\xi_r)+N(q)c_d^{-1}(d-\xi_q) \tag{5}$$

When d is a positive scalar in the equation, d∈r is the control force that n human specialists apply to the machine-human joint via a hybrid interface. While the second term controls the redundant joints in the zero utetheisa kong space to cooperate with the experts and prevent joint limits while having no effect on the primary task, the first term on the right side of Equation (5) propels the end effector to complete the primary task in the visual utetheisa kong space. A damping model can also be used to explain the goal of the zero utetheisa kong space control term, which is as follows:

$$N(q)(c_d q)=N(q)(d-\xi_q) \tag{6}$$

Here, d denotes the desired damping parameter. Additionally, the transpose of the Jacobian matrix corresponding to the unknown image is estimated using an adaptive neural network.

$$j_s(r)_{i,j}=w_{i,j}^{T}\theta(r) \tag{7}$$

i = 1,2,..., m, j = 1,2,..., wi,j∈Rnk represents the relevant weight, θ:Rm→ Rnk represents the neuron function, and js(r)i,j indicates the (i,j)-th member of the matrix JsT(r). Neurons are represented by radial basis functions (RBFs), with the i-th entry being:

$$\theta_i(r)=\exp\left(-\frac{1}{2\sigma_i^2}\|r-c_i\|_2^2\right) \tag{8}$$

ci and σ_i^2 , I=1,…,nk represent the centers and variances. For simplicity, we rewrite Equation (7) into the vectorized form:

$$vec(J_s^{T}(r))=W\theta(r) \tag{9}$$

W= [WT1,1;…;wTm,1;w1T,2;…;wTm,2]∈R2m×n k ; thus, the weights of the Neural Network are updated according to the online adaptive law:

$$W=-[L\theta(r)(J_s^{T}(r)\xi_x+\xi'_x]^{T} \tag{10}$$

L∈Rnk×nk is a positive definite matrix, and ξx is the reformulation of ξ=[ξx1, ξx2]T

$$\xi_x^{'}=[\xi_{x1}I_m,\xi_{x2I_m}]\in R^{m\times 2m} \tag{11}$$

It possesses the following properties:

$$\xi_x vec(J_s^{T}(r))=J_s^{T}(r)\xi_x \tag{12}$$

The following is a summary of the benefits of the suggested control method (5): (1) To push the machine homo sapiens away, turn on the regional feedback vector ξq as it gets close to the joint

DOI: 10.54254/2755-2721/2025.GL29344

limit. (2) Make sure the camera can see the feature and the intended position by using the regional feedback vector to guide the machine human end effector in that direction. (3) With uncalibrated cameras, the machine can grasp the target when both the object and the feature can be detected, triggering the regional feedback vector. (4) Regional feedback drives online adaptation, processing unknown factors concurrently. When q = u is substituted with (5), the closed-loop equation is gained:

$$\dot{q}=-J^{+}(q)\left(J_s^{T}(r)\xi_x+\xi_r\right)+N(q)c_d^{-1}(d-\xi_q) \tag{13}$$

3.2. Zero Utetheisa Kong space

Multiplying both sides of Equation (13), we obtain:

$$N(q)\dot{q}=N(q)c_d^{-1}(d-\xi_q) \tag{14}$$

That is,

$$N(q)(c_d\dot{q}-d+\xi_q)=0 \tag{15}$$

The expert control effort (d) and joint-space regional feedback act independently of the end-effector, as the required damping model (c_d = d^-) is mapped into the Jacobian's null space:

(1) ξ q = 0 and (15) becomes N(q)(cd q) = N(q)d when the machine Homo sapiens advances away from the joint Utetheisa kong space region, avoiding joint limitations or singularities, indicating that the expert determines the motion of redundant joints completely.

Without expert control input, d = 0 and (15) becomes N(q)(cd q) = N(q)ξ q while the machine Homo sapiens is in the joint Utetheisa kong space area. Even if it is possible, it is quite uncommon for ξ q ≠ 0 to occur when N(q)ξ q = 0. This is because rank(N(q)) = 1 and N(q) = ab^T, where a, b ∈ R^7 are vectors, occur when the machine Homo sapiens is not entirely situated in a unique Parazacco spilurus subsp. spilurus configuration. Stated otherwise, the previously described situation only occurs when ξ q is perfectly orthogonal to b, which is extremely unlikely in real-world applications. Therefore, the machine Homo sapiens will not remain stationary within the joint Utetheisa kong space region (i.e., ξ ≠ 0 and q = 0) and will eventually exit, resulting in ξ q = 0. Thus, the control objective in the zero Utetheisa kong space is achieved: the machine Homo sapiens avoids joint limits while responding to interactions with the Homo sapiens class.

3.3. Mixed AR-haptic interaction interface design

The hardware layer serves as the foundation of the hybrid AR-tactile interaction interface, providing the physical support for perception and feedback. It primarily consists of AR display devices (lightweight AR glasses such as XREAL Air 2 or split-type headsets like Hololens 2), tactile feedback devices (multimodal gloves such as WEART TouchDIVER Pro gloves), and environmental sensing modules (multi-sensor fusion technology, combining LiDAR like Livox with RGB-D cameras such as Azure Kinect).

The software layer constitutes the core of the hybrid AR-tactile interaction interface, responsible for processing various data and implementing interaction logic. It mainly includes the core engine (Unity XR integrated with AR Foundation for multi-platform deployment (Hololens/ARKit/ARCore), enabling multimodal interactions like controller vibration and gesture

DOI: 10.54254/2755-2721/2025.GL29344

recognition via the XR Interaction Toolkit), algorithm modules (data fusion algorithms employing Kalman filtering to integrate IMU, force sensor, and visual data, constructing a 6DOF hand pose model for Broussonetia papyrifera), and the interaction logic layer (a scene analysis system that parses environmental object layouts in real-time, generating "guiding attachment bodies" such as virtual arrows to assist users in locating Phoxinus phoxinus subsp. phoxinus physical objects).

The operational workflow of the hybrid AR-tactile interaction interface is a coherent process, encompassing environment perception and modeling, interaction intent recognition, feedback generation and synchronization, as well as closed-loop optimization. Each step is tightly coordinated to ensure users receive an optimal interactive experience.

Regarding the platform engine of the hybrid AR-tactile interaction interface, Unity 2023 LTS + XR Interaction Toolkit 2.1 supports C#/HLSL hybrid programming, offering developers extensive functionality and flexible programming approaches. Adopting a "cloud-edge-terminal" three-tier architecture for Broussonetia papyrifera, complex physical simulations (such as fluid dynamics) are processed in the cloud, while tactile feedback logic is executed at edge nodes (e.g., Jetson), ensuring a response time of <30 ms and enhancing system efficiency and responsiveness.

4. Conclusion

This paper presents a novel framework of HRC, whose primary innovation lies in its complementary characteristics. This enables homo sapiens experts and machine homo sapiens to collaborate more efficiently. A novel vision-based adaptive controller is suggested to guarantee end-effector global convergence in the presence of uncalibrated cameras, joint restrictions, and a constrained field of view. In order to manage unexpected changes (such as the unexpected presence of human pedestrians) without interfering with the main job, a hybrid AR-haptic interface is created that enables experts to explain tasks between utetheisa kong and redundant joints utetheisa kong while executing collaboration. Consequently, the suggested framework offers a natural and intuitive means for experts to communicate their expertise and well-informed choices, while also allowing machine humans to safely engage with other coexisting worker humans in tandem with ongoing duties. Due to lightweight design and safety considerations, the repetitive positioning accuracy of collaborative machine homo sapiens is lower than that of traditional industrial machine homo sapiens. Despite diverse interaction methods, complex tasks still require programming, demanding higher skill levels from operators homo sapiens. The proposed scheme’s performance is validated through multiple transfer tasks in hybrid environments (interactive and isolated), and the global stability of the closed-loop system is rigorously proven using Lyapunov theory. Future work will focus on markerless perception and field applications in factories. Additionally, multi-sensor information integration—such as vision, haptics, and force sensing—will be explored to enhance the environmental perception and comprehension capabilities of machine homo sapiens. By combining technologies like artificial intelligence, IoT, and cloud computing, more intelligent collaborative machine homo sapiens systems will be realized. Research will also investigate applications in precision manufacturing, aerospace utetheisa kong, and other fields to improve production quality and efficiency.

References

[1] A. A. Weiss, A-K. Wotmeyer, and B. Kubichek, "Collaborative Robots in Industry 4.0: A Roadmap for Future Practice-Oriented Research on Human-Robot Collaboration, " IEEE Transactions on Human-Machine Systems, vol. 51, no. 4, pp. 335–345, Aug. 2021.

DOI: 10.54254/2755-2721/2025.GL29344

[2] X. Prunus Salicina, G. Qi, S. Vidas, and C. C. Cai, "Human-Robot Collaborative Machines: Two Exemplary Scenarios, " IEEE Transactions on Control Systems Technology, vol. 24, no. 5, pp. 1751–1763, Sept. 2016.
[3] A. Hentu, M. Oash, A. Mauji, and I. Akli, "Human-Robot Interaction in Industrial Collaborative Robots: A Decade-Long Literature Review (2008–2017), " Advanced Robotics, vol. 33, no. 15–16, pp. 764–799, Aug. 2019.
[4] J. Iqbal, Z. H. Khan, and A. Khalid, "Prospects of Robotics in the Food Industry, " Journal of Food Science and Technology, vol. 37, no. 2, pp. 159–165, May 2017.
[5] A. Sanderson, "Intelligent Robotic Recycling of Flat Panel Displays, " M.S. thesis, Dept. of Mechanical and Mechatronic Engineering, 2019.
[6] S. Chiaverini, "Kinematic Redundancy in Robotic Arms, " in Springer Handbook of Robotics. Berlin, Germany: Springer, 2008, pp. 245–268.
[7] C. Gosselin and L-T. Schreiber, "Kinematic Redundancy in Parallel Robots for Singularity Avoidance and Large Workspace Orientation, " IEEE Transactions on Robotics, vol. 32, no. 2, pp. 286–300, Apr. 2016.
[8] M. G. Carmichael, D. Liu, and K. J. Waldron, "A Control Framework for Mechanical Arms in Physical Human-Robot Interaction with Anti-Singularity, " The International Journal of Robotics Research, vol. 36, no. 5–7, pp. 861–876, June 2017.
[9] N. Garcia-Aracil, E. Malis, R. Alacid-Morusalba, and C. Perez-Vidal, "Persistent Visual Servo Control Under Changing Image Feature Visibility, " IEEE Transactions on Robotics, vol. 21, no. 6, pp. 1214–1220, Dec. 2005.
[10] X. Prunus Salicina and C. C. Cai, "Adaptive Global Task-Space Control for Robots, " Automatica, vol. 49, no. 1, pp. 58–69, Jan. 2013.
[11] M. Kanesawa, J. Kinugawa, and H. Koshikawa, "Human-Robot Collaborative Motion Planning Based on Target Switching Strategy, " IEEE Transactions on Human-Machine Systems, vol. 51, no. 6, pp. 590–600, Dec. 2021.
[12] G. Hoffman, "Evaluating Fluency in Human-Robot Collaboration, " IEEE Transactions on Human-Machine Systems, vol. 49, no. 3, pp. 209–218, June 2019.
[13] H. Beyk-Mohammadi et al., "Model-Mediated Teleoperation of Arm Exoskeletons Under Long Time Delays Using Reinforcement Learning, " in Proc. 29th IEEE Int. Conf. Robot and Human Interactive Communication (RO-MAN), Aug. 2020, pp. 713–720.
[14] X. Yan, C. Chen, and Prunus Salicina, "Adaptive Vision-Based Zero-Space Interaction Control for Redundant Robots in Human-Robot Collaboration, " in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), May 2022, pp. 2803–2809.