Multi-fingered Hand Grasps with Visuo-Tactile Fusion via Multi-Agent Deep Reinforcement Learning

Peida Jia, Xuanheng Li*, Tianqiang Zhu, Rina Wu, Xiangbo Lin, Yi Sun

School of Information and Communication Engineering, Dalian University of Technology, China
{jiapd201883113, zhutq, hswrn}@mail.dlut.edu.cn, {xhli, linxbo, lslwf}@dlut.edu.cn

Abstract

Humans achieve contact-rich dexterous grasping through the synergy of visual and tactile information. However, the high-dimensional action space of high DoF multi-fingered hands poses significant challenges to this operation. In this study, we address this complexity by controlling the robotic hand at the reduced dimensional level of individual fingers instead of the entire hand, and develop a finger-based multi-agent deep reinforcement learning strategy by regarding the wrist, arm, and each finger of the hand as intelligent agents. We commence by applying a single-agent reinforcement learning algorithm to guide the whole hand to reach the feasible approaching direction and distance to the object. Then, we develop neuroscience-inspired visuo-tactile fusion networks to train multiple agents to control their assigned fingers by effectively leveraging visual and tactile feedback. This enables dynamic and collaborative adjustments of finger-object interactions, ultimately achieving precise contact with specific areas of the objects. The grasping results on 8 objects show that our approach can achieve stable and compliant grasps. To the best of our knowledge, this is the first work that employs a finger-based multi-agent reinforcement learning approach to control the dexterous grasping process under the guidance of both visual and tactile feedback.

Introduction

Humans naturally combine visual and tactile feedback to effortlessly grasp objects. Visual cues facilitate the identification and precise positioning of objects, while tactile feedback enables rapid grip adjustments to maintain stability. However, replicating the seamless coordination of sensory information for dynamic manipulation in dexterous robotic hands, especially those with over 20 joints, is significantly challenging. On one hand, the complexity of hand-object interactions escalates exponentially with the increase in joints and contact points, complicating motion planning. On the other hand, integrating different senses presents considerable obstacles because tactile feedback from finger sensors is sparse, whereas visual information regarding object size, shape, and hand-object distance is dense.

Numerous model-based methods have been developed to optimize effective grasping trajectories for dexterous grasping (Chen et al. 2021; Wei et al. 2023). However, these methods exhibit limited adaptability to environmental changes or novel tasks, especially in scenarios requiring the integration of visual and tactile data. Deep reinforcement learning (DRL) shows promise by employing trial and error to eliminate reliance on precise dynamic models, demonstrating strong adaptability across various objects. Given the huge dimensionality introduced by dexterous hands, some recent work has incorporated human demonstrations into DRL to facilitate stable grasping. However, acquiring these demonstrations is laborious, as they are collected either in Virtual Reality (VR) using the expensive CyberGlove III (Rajeswaran et al. 2018) or by a complex video tracking system (Mandikal and Grauman 2022). Additionally, due to the control complexity arising from the combination of visual and tactile, current DRL-based methods are mostly limited to two-fingered hands (Calandra et al. 2018; Gao et al. 2023), rather than five-fingered dexterous hands. Thus, leveraging DRL to guide dexterous grasping with visuo-tactile fusion remains a challenging issue.

In this paper, we introduce the Visuo-Tactile based Multi-Agent Grasp with Intra-Collaboration (VT-MAGIC) scheme, conceptualizing the grasping process of dexterous hands as a Multi-Agent DRL (MADRL) task guided by both visual and tactile information. Instead of relying on sequential demonstration data, we employ only a single static reference grasp, which is easier to obtain. The VT-MAGIC scheme is designed as a hierarchical framework to facilitate the exploration of grasping trajectories, segmented into an *approaching* phase and a *grasping* phase. The *approaching* phase aims to guide the hand to a pre-grasp position without object collision. Since this phase does not require close collaboration among different fingers, we treat the dexterous hand as a whole and guide it by single-agent reinforcement learning. While, for the *grasping* phase, inspired by the anatomical mechanisms of the human hand (Liu et al. 2016), we regard the five fingers, wrist, and arm as seven distinct agents, each of which undertakes a sub-task of the original hand. We employ MADRL methods to enhance inter-agent cooperation and reduce exploration in the high-dimensional space. By decomposing the grasping task into finger-based sub-tasks, each finger agent can learn more compliant behaviors towards various objects, achieving more natural grasps.

*Corresponding Author.

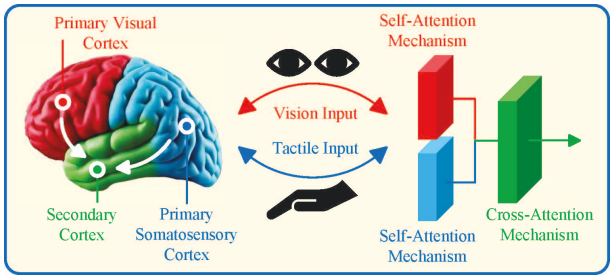

Figure 1: The visuo-tactile fusion network inspired by the dual-layer processing mechanism of the human brain.

To fully leverage visual and tactile inputs for precise robot manipulation in MADRL, we develop a neuroscience-inspired visuo-tactile fusion network, drawing on the dual-layer processing mechanism of the human brain (Stein, Stanford, and Rowland 2009). As illustrated in figure 1, we employ self-attention mechanisms to process raw data from visual and tactile sensors, similar to primary cortical processing in humans. Subsequently, the cross-attention mechanism effectively integrates these multisensory signals, reflecting the process of sensory input fusion in the secondary cortex.

To the best of our knowledge, this is the first work that employs visuo-tactile based MADRL to enhance the stability and improve finger cooperation in dexterous grasping tasks. The contributions of this work are summarized as follows:

- To address the challenge of controlling the high-dimensional space of dexterous hands through visuo-tactile fusion, we treat each finger, arm, and wrist as individual agents and develop the VT-MAGIC scheme. This scheme allows a dexterous hand to achieve a stable and compliant grasping pose with only a single-frame reference grasp. Compared to directly controlling the whole hand, this MADRL approach mitigates exploration complexity in the vast action space and facilitates successful grasp trajectories through enhanced internal cooperation.
- Inspired by the human brain's dual-layer processing mechanism, we develop a neuroscience-inspired visuo-tactile fusion network to effectively integrate visual and tactile feedback, leveraging their complementary characteristics to produce more stable grasping postures.
- We perform experiments on 8 representative objects, demonstrating that the proposed MADRL-based grasping method can generate a stable approaching and grasping trajectory, and achieve a compliant grasping pose from a single-frame reference grasp.

Related Work

With the development of DRL, many researchers have considered employing it to facilitate grasp control. VPG (Zeng et al. 2018) utilized a Q-learning framework to control the cooperation between push and grasp. A novel DRL-based moving object grasping framework (Chen and Lu 2021) achieved grasp control using the Soft Actor-Critic (SAC) algorithm. However, these studies were limited to two-fingered hands, where the training is significantly less challenging than that of the dexterous hand.

To achieve dexterous grasping, more adaptive design and processing are needed. Researchers focused on combining imitation learning and DRL to help reduce exploration in huge action spaces by collecting human demonstration data. DAPG (Rajeswaran et al. 2018) combined Natural Policy Gradient (NPG) with a small amount of demonstration data, successfully achieving complex manipulation tasks with dexterous hands. Reach-and-grasp tasks have also been accomplished by incorporating imperfect demonstrations into a replay buffer (She et al. 2022). Although vision-based grasping systems have achieved remarkable success, they often fail to incorporate tactile feedback, which is crucial for overcoming obstacles such as occlusions and environmental noise. Furthermore, integrating tactile input can significantly enhance grasping stability.

An end-to-end action-conditional model (Calandra et al. 2018) learned grasping strategies directly from raw visuo-tactile data, but it was limited to two-fingered hands. The tactile contact points presented in the form of point clouds (Dikhale et al. 2022) effectively integrate visuo-tactile input. Nonetheless, these methodologies primarily cater to grippers with limited fingers and struggle to realize the grasping operation of five-finger dexterous hand.

In recent years, MADRL has attracted research interests in robot control and has been leveraged to accomplish various tasks. Multi-Agent Deep Deterministic Policy Gradient (MADDPG) was applied to a collaborative shaft slot assembly task in a dual-manipulator system (Yao et al. 2022). However, their model designated the robotic arm as the agent rather than the robotic fingers. MAGCLA (Tao et al. 2023) applied MADRL to dexterous hands to realize the rotation of objects like blocks and eggs. Their approach focused on in-hand manipulations, thus assuming a fixed arm position. Additionally, both studies neglected the role of tactile feedback. Unlike these existing works, we develop the VT-MAGIC scheme, which treats the dexterous grasping process as a finger-based MADRL task, guided by visuo-tactile input.

Methodology

The objective of this work is to effectively integrate visual and tactile information using MADRL methods to achieve stable dexterous grasps with only a single-frame reference grasp. This reference can be obtained from recent static grasp synthesis methods (Zhu et al. 2021) or a grasp dataset (Wang et al. 2023). To this end, we propose the VT-MAGIC scheme, as illustrated in figure 2. This scheme is divided into two main phases. During the *approaching* phase, we employ a standard DRL method to train the dexterous hand to achieve a pre-grasp pose. The *grasping* phase is the core of our scheme, where the key challenge is to achieve stable and dexterous grasping in a vast exploration space under multi-sensory guidance. To tackle this, we employ the MADRL method, treating the five fingers, arm, and wrist of the dexterous hand as seven agents, allowing for dynamic adjustments at the finger level.

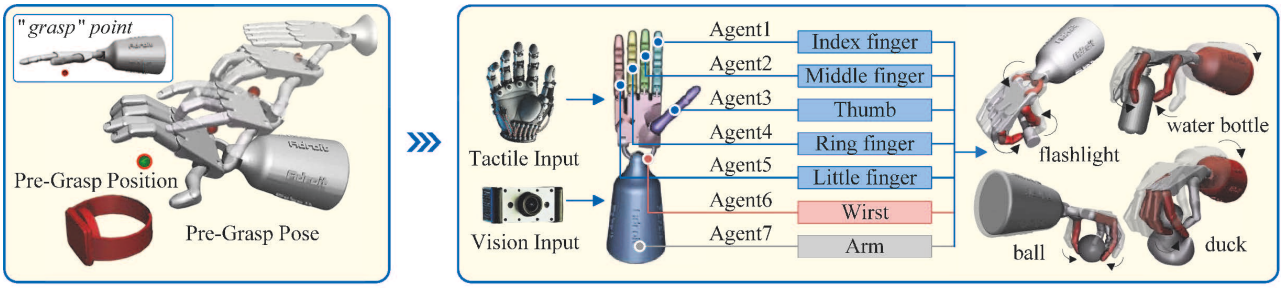

Figure 2: An overview of our VT-MAGIC structure, including the *approaching* phase, which utilizes single-agent deep reinforcement learning, and the *grasping* phase, which utilizes multi-agent deep reinforcement learning.

Our proposed MADRL method replicates the kinematics and movement patterns of human fingers by mimicking the coordinated movements of the human hand (Chen, Xiong, and Yue 2014; Li et al. 2013). This method reduces the exploration space and enhances finger cooperation, resulting in more stable grasps. For example, in figure 2, the cooperation mode of the fingers varies for different grasping tasks: the thumb and index finger need to firmly pinch a slender flashlight, while grasping an irregularly shaped duck requires the fingers to wrap around the object.

Additionally, recognizing the limitations of relying solely on visual information for grasping, which can lead to instability and knocking over objects, we incorporate a visuo-tactile fusion mechanism in MADRL: visual guidance enables the hand to make contact with the object with an optimal pose, while tactile feedback ensures that excessive force on any joint does not cause the object to be knocked over or dropped. To effectively merge visual and tactile information, we developed a neuroscience-inspired visuo-tactile fusion network that employs both self-attention and cross-attention mechanisms.

Next, we detail the design of the VT-MAGIC scheme. We first introduce the *approaching* and *grasping* phase, and then describe our network training and key design elements.

The Approaching Phase

The *approaching* phase aims to navigate the robotic hand to a pre-grasp position approximately 5 cm above the static reference grasp and maintain the pre-grasp pose. This is critical to avoid potential collisions and facilitates learning in the subsequent *grasping* phase. The pre-grasp pose, with five fingers extended, is illustrated in figure 2 and is determined by the NPG algorithm, a single-agent reinforcement learning algorithm. The state, action, reward, and termination conditions during the *approaching* phase are detailed below.

State space. The state of the *approaching* phase includes joint angles and the distance between the hand and the pre-grasp. Specifically, it encompasses angles for 6 arm joints and 24 hand joints, along with the distance from a designated "grasp" point, located 4 cm below the center of the palm, to the pre-grasp position as shown in figure 2. The distance between the five fingertips and their respective positions in the pre-grasp pose are also integral to this state space.

Action space. The action space consists of joint torques for the arm's 6 joints and the dexterous hand's 24 joints. To reduce control complexity, the robotic arm's movements are limited to three-dimensional translation without three-dimensional rotation.

Reward function. We design two reward functions, $R_{position}$ and R_{pose}, to guide the robotic hand's movements. $R_{position}$ is designed to direct the robotic hand to reach and stay at the pre-grasp position:

$$R_{position} = -\alpha_1 d_{palm} + \alpha_2 T_t, \tag{1}$$

where α_1 and α_2 are positive weighting coefficients, d_{palm} is the distance between the palm's "grasp" point and the pre-grasp position. T_t, a time-related variable, rewards the hand for staying stably at the pre-grasp position for an extended period, as shown in (2):

$$T_t = \begin{cases} 0 & t = 0 \\ T_{t-1} + 1 & t > 0 \text{ and } d_{palm} \leq 0.02m \\ T_{t-1} & \text{otherwise} \end{cases}, \tag{2}$$

the integer t is the index of the step in an epoch. At each step, T_t increases if the distance from the "grasp" point to the pre-grasp position is less than 0.02m, otherwise it remains unchanged.

The reward function R_{pose} adjusts the hand's posture to match the intended pre-grasp pose, as shown in (3):

$$R_{pose} = \sum_{i=1}^{5} -\beta_i d_i, \tag{3}$$

where β_i is a positive weighting coefficient, and d_i is the distance between the fingertip of the ith finger and its target position in the pre-grasp pose. The closer they are, the greater the reward obtained.

The total reward is the sum of $R_{position}$ and R_{pose}, guiding the hand toward a collision-free pre-grasp:

$$R_{all} = R_{position} + R_{pose}. \tag{4}$$

Termination condition. The *approaching* phase terminates when either the number of epochs reaches a preset threshold or T_t exceeds 30, indicating that the dexterous hand has successfully learned the pre-grasp pose and can remain stably in the pre-grasp position.

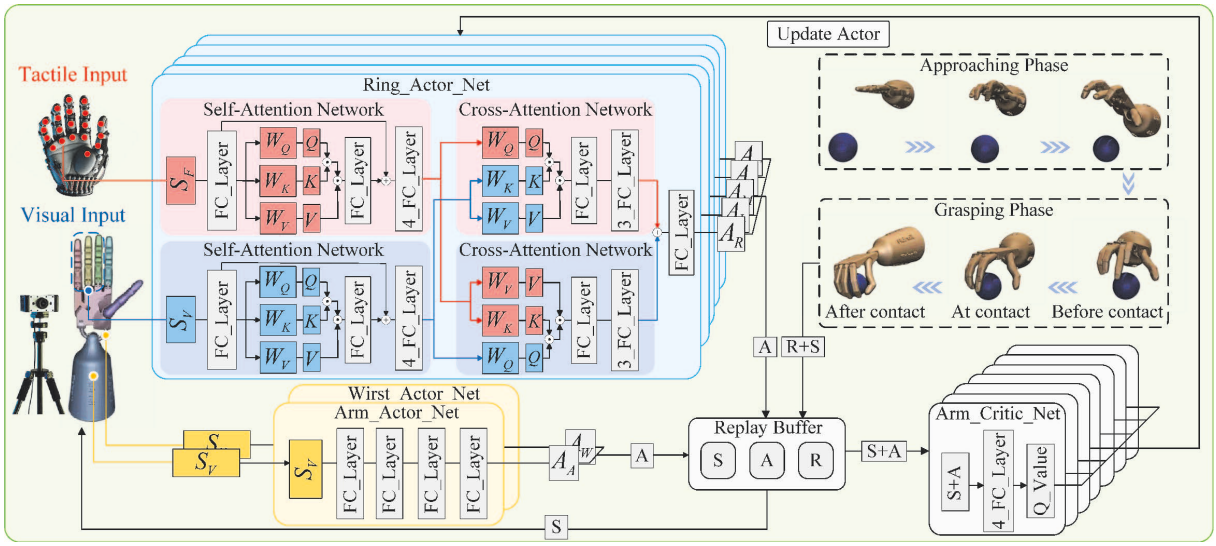

Figure 3: The network of the grasping phase in our VT-MAGIC scheme: The agents' Actor networks produce action A based on the state S. The Critic network then inputs the global state and action and outputs a value Q to update the Actor network.

The Grasping Phase

In the *grasping* phase, our model conceptualizes each finger, wrist, and arm of the dexterous hand as individual agents. The objective is to find a physically plausible grasping pose during hand-object interactions using visuo-tactile based MADRL. Initially, before the fingers make contact with the object, each agent's movements are guided by visual information. Once contact occurs, tactile feedback is activated, and the hand's current configuration is recorded as the initial state for the neuroscience-inspired visuo-tactile fusion network. The network then trains each agent to integrate visual and tactile inputs to guide the grasping process.

To achieve this functionality, each agent is equipped with dual Actor networks, i.e., a Main-Actor network and a Target-Actor network (which are the previously mentioned visuo-tactile fusion networks), and dual Critic networks, i.e., a Main-Critic network and a Target-Critic network, as illustrated in figure 3. For simplicity, only the Main-Actor and Main-Critic networks are depicted in the figure; the target networks, which have the same structure as their respective main networks, are used solely for parameter updates. 1) The Actor networks determine the agents' actions based on the input state. Specifically, each Actor network includes a tactile module and a visual module. Each module consists of two parts: the self-attention component, which extracts raw features from the sensory inputs, and the cross-attention component, which integrates the sensory features. The fused outputs processed by the tactile and visual modules are concatenated and then mapped through a fully connected layer to produce the agent's action. Note that since the arm and wrist agents do not process tactile information, their Actor networks are composed of fully connected layers. 2) The Critic networks assess these actions by evaluating the potential rewards for each agent based on the state and actions.

The state, action, reward, and termination conditions during the *grasping* phase are defined as follows:

State space. For the five fingers, the state includes both visual and tactile states, whereas for the arm and wrist, it includes only visual states. Specifically, for each finger, the state includes its own joint angles, adjacent fingers' joint angles, the magnitude of contact force feedback from sensors on the finger, and the distance from each fingertip to its corresponding point in the static reference grasp. For the arm and wrist, the state is defined by each joint angle and the distance from the "grasp" point to the static reference grasp.

Action space. Actions for each agent are defined by the joint torques of their respective joints, with specific counts for each part of the hand: 6 for the arm, 2 for the wrist, and 5, 4, 4, 4, and 5 for the thumb, index, middle, ring, and little finger, respectively. This segmentation of the action space facilitates precise control over each part of the hand, essential for dexterous manipulations.

Reward function. We design individual reward functions for each agent to ensure they complete their unique tasks.

For a finger agent, the objective is to closely align with the target fingertip position as defined in the static reference grasp. To this end, we introduce finger-guided rewards as detailed in (5):

$$\begin{aligned} R_{thumb} = R_{index} = R_{middle} = R_{ring} \\ = R_{little} = f_i(d_i)(i = 1, 2, 3, 4, 5), \end{aligned} \tag{5}$$

where f_i is a quadratic function of the distance d_i from the ith fingertip to the target point; the smaller the d_i, the greater the reward.

The arm and wrist agents aim to assist fingers in attaining a stable grasp pose and lifting the object. We define the arm and wrist reward as:

$$R_{arm} = R_{wrist} = -\alpha_3 d_6 + R_l, \tag{6}$$

where α_3 is a positive weighting coefficient, d_6 measures the distance between the palm's "grasp" point and the static reference grasp pose. R_l is used to encourage the dexterous hand to pick up the object, as represented by (7):

$$R_l = -\alpha_5 h^2 + \alpha_6 h + \alpha_7, \tag{7}$$

where α_5, α_6 and α_7 are positive weighting coefficients, and h represents the height of the object. This reward is a quadratic function; the closer the object is lifted to the set height, the greater the reward.

Termination condition. The *grasping* phase terminates either when the maximum number of training epochs is reached or when the object is grasped with a stable pose.

Visuo-Tactile Based MADRL

We adopt the MADDPG algorithm within MADRL to train each agent and modify the Actor network to our designed neuroscience-inspired visuo-tactile fusion network to seamlessly integrate visual and tactile inputs.

Starting from the initial state, the agent i generates joint torques as actions a_i based on the current state s_i received from its Main-Actor network. After interacting with the environment, the actions a_i lead to a new state s_i' and a reward R_i. The s_i, a_i, R_i, and s_i' are stored in a replay buffer.

To train the networks, a batch of experience data is randomly sampled from the replay buffer. For agent i, the Target-Critic network calculates the target Q-value Q_T, which is an estimate of the expected future rewards:

$$Q_T = R_i + \gamma Q_T(s'_{all}, a'_{all}|\mu_T), \tag{8}$$

where γ is the discount factor, s'_{all} is the next state observed by the entire dexterous hand, and a'_{all} is the action taken by all agents in the next state, as predicted by the Target-Actor networks. μ_T is the parameter of the Target-Critic network.

The Main-Critic network then updates its parameters by minimizing the mean squared error between its current Q-value predictions and the target Q-values:

$$L(\mu) = E[(Q(s_{all}, a_{all}|\mu) - Q_T)^2], \tag{9}$$

where μ is the parameter of the Main-Critic network, s_{all} is the state observed by the entire dexterous hand, and a_{all} is the action of all agents. This loss function helps the network to better approximate the future rewards Q_T.

The Main-Actor network is updated using (10):

$$L(\theta) = -E[Q(s_{all}, a_{all}|\mu)], \tag{10}$$

where θ is the Main-Actor network's parameter. This loss encourages the Main-Actor network to select actions that increase the Q-values predicted by the Main-Critic network.

Lastly, the two target networks are updated employing the soft update method as depicted in (11) and (12):

$$\theta_T = (1-\tau)\theta_T + \tau\theta, \tag{11}$$

and

$$\mu_T = (1-\tau)\mu_T + \tau\mu. \tag{12}$$

These equations gradually integrate the latest network parameters, θ and μ, into the target networks, reducing the risk of abrupt changes that could destabilize the learning process.

In summary, to apply MADRL for promoting intra-hand cooperation and achieving stable grasps with a dexterous hand, our design incorporates three key features: 1) we assign individual rewards to each agent, enabling them to specialize in tasks that best suit their capabilities during grasping, thereby enhancing learning efficiency; 2) each finger agent's Actor networks have access to the state information of adjacent fingers and react accordingly, improving behavioral synchronization; 3) the Critic network receives comprehensive state and action data from the entire dexterous hand, allowing each agent to optimize its actions based on the overall system state and the anticipated actions of other agents, promoting a stable grasping pose.

Simulations and Experiments

We verify our VT-MAGIC scheme using the MuJoCo physics simulator. Our setup includes a dexterous hand with 24 joints attached to a 6-DoF robotic arm. We apply VT-MAGIC to 8 different objects, including a ball, watch, rubber duck, elephant toy, light bulb, water bottle, flashlight, and hammer. The hardware required for the experiments includes an Intel® Core^{TM} i9-10920 CPU @3.50GHz $\times$ 24, and an NVIDIA® Geforce RTX 3090 graphics card.

Compared Methods

Current dexterous grasping methods based on deep reinforcement learning mostly rely on single-agent DRL. Therefore, we compare VT-MAGIC with several single-agent DRL methods. First, we compare it with DAPG (Rajeswaran et al. 2018). To ensure a fair comparison, we modify the DAPG method by removing the sequential hand-object demonstration data it typically uses and instead train it using our static reference grasps and custom rewards designed to promote effective grasp pose generation. We refer to this modified version as SAPG (Single-Agent Policy Gradient).

We also compare VT-MAGIC with PPO, a commonly used single-agent DRL algorithm in robotic grasping, which we refer to as SAPPO (Single-Agent PPO). Considering the impact of tactile information on the final grasping results, we design two additional baselines that integrate both visual and tactile guidance, namely VT-SAPG (Visuo-Tactile based SAPG) and VT-SAPPO. All methods are trained under the same experimental settings, including reward functions, iteration counts, and static reference grasps.

Evaluation Metrics

We evaluate the quality of grasps using two metrics: 1) Success Rate. This metric measures the ability of the robotic hand to lift an object up to 10 cm and maintain it within grasp for at least 5 seconds without dropping. We evaluate this across 50 trials for each object and calculate the average success rate for all objects combined in table 1.

Method	VT-MAGIC	VT-SAPG	VT-SAPPO	SAPG	SAPPO
Success	**86%**	75%	74%	70%	72%

Table 1: The comparison of average success rates

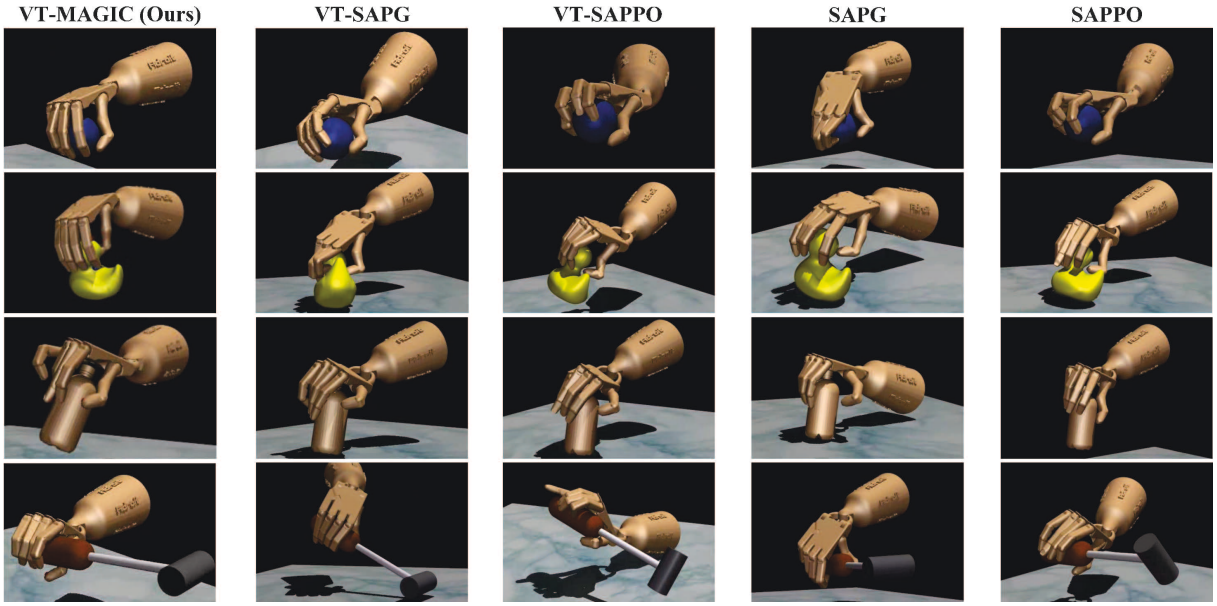

Figure 4: The final grasp poses of VT-MAGIC and other baseline methods on four different objects.

Method	Displacement ↓	Ball ↓	Hammer ↓	Elephant ↓	Light ↓	Watch ↓	Flashlight ↓	Duck ↓	Water bottle ↑
SAPG	x (cm)	0.548	Fall	Fall	**0.628**	Fall	Fall	Fall	0 Times
	y (cm)	Fall	**1.933**	Fall	Fall	Fall	Fall	Fall	0 Times
	z (cm)	0.340	Fall	Fall	0.858	Fall	**0.608**	Fall	0 Times
SAPPO	x (cm)	Fall	Fall	Fall	1.205	6.707	Fall	Fall	**1 Times**
	y (cm)	1.591	6.657	Fall	Fall	**1.435**	2.670	Fall	2 Times
	z (cm)	0.415	1.495	Fall	1.118	1.390	7.723	Fall	2 Times
VT-SAPG	x (cm)	0.495	Fall	Fall	1.216	1.632	Fall	Fall	0 Times
	y (cm)	**0.698**	2.619	4.840	1.795	1.615	Fall	Fall	0 Times
	z (cm)	0.265	0.867	Fall	1.497	1.756	Fall	Fall	0 Times
VT-SAPPO	x (cm)	Fall	Fall	Fall	1.974	4.201	10.907	Fall	0 Times
	y (cm)	Fall	4.742	**2.684**	2.395	1.882	**2.516**	Fall	0 Times
	z (cm)	Fall	Fall	Fall	Fall	1.824	0.751	1.433	0 Times
VT-MAGIC	x (cm)	**0.458**	**1.862**	**3.006**	0.639	**1.557**	**2.666**	**1.349**	**1 Times**
	y (cm)	0.826	2.228	2.986	**1.551**	1.468	2.618	**1.598**	**3 Times**
	z (cm)	**0.175**	**0.584**	Fall	**0.479**	**0.609**	0.894	**0.605**	**6 Times**

Table 2: The maximum relative displacement between the object and the hand

2) Stability Detection. This metric assesses the stability of the grasp pose. After achieving a successful grasp, we apply alternating forces to the arm in positive and negative directions along the x, y, and z axes to induce object shaking. We then measure the maximum relative displacement between the object and the hand over a set time period in table 2. A smaller displacement indicates a more stable grasp, reflecting the stability of the generated hand postures. Notably, due to the significant mass and volume of the water bottle, all methods result in the object's dropping under shaking conditions. Thus, we count the number of shakes before the water bottle drops, using this as a stability indicator.

Robot Grasp Evaluation

Qualitative results. We first visualize generated grasps for four different objects. The final grasp poses generated by VT-MAGIC are shown in the first column of figure 4. As can be seen, our method produces stable grasps with natural hand postures for all objects. In contrast, except for VT-SAPPO, all baseline methods struggle to successfully grasp complex-shaped objects (such as the duck) and have difficulty lifting larger, heavier objects (such as the water bottle, with SAPPO consistently failing to lift the elephant toy). Additionally, compared to our method, where the fingers tightly grip the object, the baselines tend to only have the fingertips in contact with the object, resulting in unnatural hand poses.

Quantitative results. The average success rates for all objects across the five methods are shown in table 1. The results demonstrate that the average success rate of the VT-MAGIC scheme significantly outperforms that of the baseline methods, highlighting its stability and superior performance. The

Method	Displacement ↓	Ball ↓	Hammer ↓	Elephant ↓	Light ↓	Watch ↓	Flashlight ↓	Duck ↓	Water bottle ↑
VT-MAGIC-NF	x (cm)	2.380	3.018	Fall	1.549	1.568	3.629	**0.903**	**1 Times**
	y (cm)	1.125	Fall	3.612	Fall	**1.256**	**1.658**	Fall	1 Times
	z (cm)	1.311	0.852	**0.550**	0.601	0.618	Fall	0.621	1 Times
VT-MAGIC-MLP	x (cm)	1.194	Fall	Fall	1.275	1.604	4.079	Fall	0 Times
	y (cm)	1.029	**1.885**	Fall	Fall	1.596	Fall	Fall	0 Times
	z (cm)	0.835	0.874	1.094	0.523	0.957	Fall	Fall	0 Times
VT-MAGIC	x (cm)	**0.458**	**1.862**	**3.006**	**0.639**	**1.557**	**2.666**	1.349	**1 Times**
	y (cm)	**0.826**	2.228	**2.986**	**1.551**	1.468	2.618	**1.598**	**3 Times**
	z (cm)	**0.175**	**0.584**	Fall	**0.479**	**0.609**	**0.894**	**0.605**	**6 Times**

Table 3: The maximum relative displacement between the object and the hand in ablation experiments

Method	VT-MAGIC	VT-MAGIC-NF	VT-MAGIC-MLP
Success	**86%**	82%	77%

Table 4: The ablation experiments' average success rates

baselines often had lower success rates due to their difficulty in grasping specific objects.

Table 2 details the Stability Detection metrics. The entries in this table include the relative displacement for 8 objects of our VT-MAGIC scheme, the SAPG method, the SAPPO method, and their respective variations that utilize both visual and tactile feedback (denoted as VT-SAPG and VT-SAPPO). In the table, smaller relative displacement indicates a more stable grasp posture. The occurrence of "Fall" signifies instances where the object slipped or the grasp was lost. The results demonstrate that with our VT-MAGIC scheme, relative displacements for almost all objects are consistently lower than those in other baselines, and there are almost no drops during shaking.

The superior performance of VT-MAGIC can be attributed to its closer emulation of human hand kinematics and movement patterns. Specifically, in the MADRL method, the independence of each finger minimizes the number of joint angles that need to be controlled, thereby reducing the complexity of the control problem. Additionally, this method promotes cooperation among all agents, resulting in more compliant postures. The integration of tactile feedback further enhances grasp stability.

Ablation Studies

In this section, we will verify the contribution of tactile feedback to grasping control and the effectiveness of our novel neuroscience-inspired visuo-tactile fusion network compared to a conventional MLP. To do these, we record the stability detection metrics of the ablation studies VT-MAGIC-NF (VT-MAGIC without force) and VT-MAGIC-MLP (VT-MAGIC with a conventional MLP) in table 3, and test the grasping success rates as shown in table 4.

The contribution of tactile feedback. Table 3 reflects the stability of the generated grasp poses: without tactile feedback, VT-MAGIC-NF struggles to apply optimal joint forces during training, leading to consistently high relative displacement and even object slippage during shaking tests. VT-MAGIC, on the other hand, demonstrates almost no falls across all objects (except for the water bottle and the z-axis of the elephant) and shows lower relative displacement values for all objects. This emphasizes VT-MAGIC's superior grasp stability under varied and dynamic environments. Table 4 clearly shows that tactile feedback increases grasping success rates. This is because the inclusion of tactile feedback allows the robotic hand to adjust its grasping strategy in real-time to prevent excessive force that could knock over objects, thereby improving the success rate. The inclusion of tactile information enhances the robotic hand's response to the physical properties of objects, ensuring they remain consistently under control.

The neuroscience-inspired visuo-tactile fusion network. We further examine the effectiveness of our neuroscience-inspired visuo-tactile fusion network by comparing it with a conventional MLP model (VT-MAGIC-MLP). The results in table 3 indicate that our VT-MAGIC method results in fewer drops and lower relative displacements, demonstrating that only a refined integration of visual and tactile information can fully leverage the complementarity of multi-sensory data. A simple mix of multi-sensory, like VT-MAGIC-MLP, can disrupt the robot's judgment, leading to more failed grasps. Table 4 shows that the conventional MLP model negatively impacts grasping success rates.

The advantage of the neuroscience-inspired fusion network lies in its ability to mimic the human brain's process of integrating visual and tactile data. By processing different sensory inputs through the self-attention mechanism, it retains effective data while filtering out extraneous data. The cross-attention mechanism then captures the most relevant features from both visual and tactile inputs, efficiently integrating them. Additionally, by allowing each finger agent to observe the state information of adjacent fingers, we enhance the cooperation between neighboring fingers, ultimately improving the grasping performance.

Conclusions

To enable dexterous hands to integrate visual and tactile feedback and generate stable grasps based on a single static reference grasp, we propose the VT-MAGIC scheme. This scheme effectively reduces the need for extensive exploration in the high-dimensional action space, while excelling at capturing complex inter-finger relationships. Next, we will investigate how to perform sim-to-real policy transfer.

Acknowledgments

This work was supported by the National Natural Science Fundation of China (62373075), the Science and Technology Program of Liaoning Province under Grant 2023JH2/101700366, the Fundamental Research Funds for the Central Universities under Grant DUT24ZD127, the Xiaomi Young Talents Program.

References

Calandra, R.; Owens, A.; Jayaraman, D.; Lin, J.; Yuan, W.; Malik, J.; Adelson, E. H.; and Levine, S. 2018. More than a feeling: Learning to grasp and regrasp using vision and touch. *IEEE Robotics and Automation Letters*, 3(4): 3300–3307.

Chen, C.; Culbertson, P.; Lepert, M.; Schwager, M.; and Bohg, J. 2021. Trajectotree: Trajectory optimization meets tree search for planning multi-contact dexterous manipulation. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 8262–8268. IEEE.

Chen, P.; and Lu, W. 2021. Deep reinforcement learning based moving object grasping. *Information Sciences*, 565: 62–76.

Chen, W.; Xiong, C.; and Yue, S. 2014. Mechanical implementation of kinematic synergy for continual grasping generation of anthropomorphic hand. *IEEE/ASME Transactions on mechatronics*, 20(3): 1249–1263.

Dikhale, S.; Patel, K.; Dhingra, D.; Naramura, I.; Hayashi, A.; Iba, S.; and Jamali, N. 2022. VisuoTactile 6D Pose Estimation of an In-Hand Object Using Vision and Tactile Sensor Data. *IEEE Robotics and Automation Letters*, 7(2): 2148–2155.

Gao, J.; Huang, Z.; Tang, Z.; Song, H.; and Liang, W. 2023. Visuo-Tactile-Based Slip Detection Using A Multi-Scale Temporal Convolution Network. arXiv:2302.13564.

Li, K.; Nataraj, R.; Marquardt, T. L.; and Li, Z.-M. 2013. Directional coordination of thumb and finger forces during precision pinch. *PloS one*, 8(11): e79400.

Liu, M.; Xiong, C.; Xiong, L.; and Huang, X. 2016. Biomechanical Characteristics of Hand Coordination in Grasping Activities of Daily Living. *PloS one*, 11(1): 1–16.

Mandikal, P.; and Grauman, K. 2022. Dexvip: Learning dexterous grasping with human hand pose priors from video. In *Conference on Robot Learning*, 651–661. PMLR.

Rajeswaran, A.; Kumar, V.; Gupta, A.; Vezzani, G.; Schulman, J.; Todorov, E.; and Levine, S. 2018. Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations. arXiv:1709.10087.

She, Q.; Hu, R.; Xu, J.; Liu, M.; Xu, K.; and Huang, H. 2022. Learning High-DOF Reaching-and-Grasping via Dynamic Representation of Gripper-Object Interaction. arXiv:2204.13998.

Stein, B. E.; Stanford, T. R.; and Rowland, B. A. 2009. The neural basis of multisensory integration in the midbrain: its organization and maturation. *Hearing research*, 258(1-2): 4–15.

Tao, L.; Zhang, J.; Bowman, M.; and Zhang, X. 2023. A Multi-Agent Approach for Adaptive Finger Cooperation in Learning-based In-Hand Manipulation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 3897–3903. IEEE.

Wang, R.; Zhang, J.; Chen, J.; Xu, Y.; Li, P.; Liu, T.; and Wang, H. 2023. Dexgraspnet: A large-scale robotic dexterous grasp dataset for general objects based on simulation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 11359–11366. IEEE.

Wei, M.; Huang, Y.; Xu, Z.; Liu, N.; Che, Z.; Zhang, X.; Shen, C.; Feng, F.; Shan, C.; and Tang, J. 2023. CMG-Net: An End-to-End Contact-based Multi-Finger Dexterous Grasping Network. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 9125–9131. IEEE.

Yao, J.; Wang, X.; Li, R.; Wang, W.; Ping, X.; and Liu, Y. 2022. Dual manipulator collaborative shaft slot assembly via MADDPG. In *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 1047–1052. IEEE.

Zeng, A.; Song, S.; Welker, S.; Lee, J.; Rodriguez, A.; and Funkhouser, T. 2018. Learning synergies between pushing and grasping with self-supervised deep reinforcement learning. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4238–4245. IEEE.

Zhu, T.; Wu, R.; Lin, X.; and Sun, Y. 2021. Toward human-like grasp: Dexterous grasping via semantic representation of object-hand. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 15741–15751.