Virtual Reality

https://doi.org/10.1007/s10055-026-01455-7

Article in Press

Grasping techniques: innovations in robotics and virtual reality

Received: 3 October 2025

Accepted: 21 July 2026

Published online: 10 August 2026

Cite this article as: Zhou M. & Aburumman N. Grasping techniques: innovations in robotics and virtual reality. *Virtual Reality* (2026). https://doi.org/10.1007/s10055-026-01455-7

Mingzhao Zhou & Nadine Aburumman

Grasping Techniques: Innovations in Robotics and Virtual Reality

Mingzhao Zhou[1] and Nadine Aburummanb[2*]

[1]Department of Computer Science, Brunel University of London, Kingston Ln, Uxbridge, UB8 3PH, London, UK.
[2*]Department of Computer Science, Brunel University of London, Kingston Ln, Uxbridge, UB8 3PH, London, UK.

*Corresponding author(s). E-mail(s): Nadine.Aburumman@brunel.ac.uk ;
Contributing authors: Mingzhao.Zhou@brunel.ac.uk ;

Abstract

The act of grasping is a fundamental mode of interaction when manipulating objects, in both physical and virtual environments. Robotic grasping has been studied for more than three decades, leading to the development of sophisticated frameworks. In contrast, achieving believable grasping in virtual reality (VR) requires a complex interplay of graphics, physics, and perception, where realistic haptic feedback and high-quality rendering are crucial for user immersion. This paper reviews grasping techniques in both robotics and VR, analysing VR grasping from visual and haptic perspectives and identifying its current challenges. We then compare robotic and VR grasping, showing how the relatively mature evaluation metrics of robotics can inspire potential directions for improving VR grasping and support the development of solutions that are more stable, natural, and efficient. Beyond surveying existing methods, this paper provides an outlook on upcoming challenges and opportunities in grasping research across robotics and VR. We set out both a high-level vision towards natural, and reliable grasping in real and virtual domains and priorities, including benchmark creation, robust evaluation metrics, and improvements in haptic fidelity and latency.

Keywords: Virtual Reality (VR), Robotics, Grasping, Haptic, Human-computer interaction (HCI), Interaction techniques.

1 Introduction

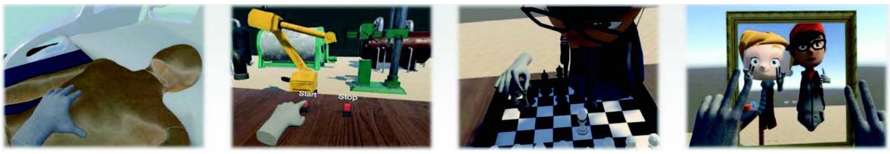

Fig. 1: Realistic interaction with virtual objects within an immersed environment is crucial for many applications, such as industrial and medical training, entertainment and virtual social interaction [1].

Understanding human hand movements is crucial for robotics, virtual simulation, teleoperated manipulation and rehabilitation [2]. When a human interacts with an object, grasping is a fundamental and intuitive action in many circumstances. This natural capability allows the hand to grasp a wide variety of objects with different shapes, weights, and frictional properties.

Grasping is furthermore a fundamental operation in robotics [3]. Robotic grasping has found widespread applications in various domains, including assembly line processing, machinery manufacturing, and agriculture [4]. In robotic grasping, factors such as the shape, state, and material of the object being grasped, as well as the external environment, must be considered to improve work efficiency [5, 6]. Over several decades of development, robotic grasping methods have matured, in addition to traditional robotic grasping methods, recent advances have introduced data-driven approaches. These methods can now effectively accomplish a wide range of grasping tasks.

An ideal grasping action must take into account the geometry and dynamic characteristics of the object. These considerations, which are essential in both physical and robotic grasping, are equally relevant in virtual environments [7]. In recent decades, with the advancement in Virtual Reality (VR) devices and Human-Computer Interaction (HCI) studies, the need for simulation of human behaviour with realistic interaction has become vital for many applications (see Fig. 1), such as industrial training, medical and surgical simulation, and rehabilitation [1, 8, 9]. High-fidelity graphics in such virtual environments have been used to provide users with a relatively immersive virtual experience [10]. Obtaining a fully immersive experience requires realistic interaction with objects within the environment, where also contact forces and weight of the objects should be felt during the interaction [11]. Haptic devices for either fingertips or the whole hand enable users to feel and manipulate the 3D objects and explore the virtual environment through kinesthetic and cutaneous perception [12].

Haptic techniques have evolved considerably, with glove-based interfaces [13] and controller-based systems [8, 14] becoming the dominant solutions. These device enhance the haptic capabilities of virtual hand interactions, and when combined with Head-Mounted Displays (HMD), allow users to grasp virtual objects more naturally. However, the human hand possesses more than 20 Degrees of Freedom (DoFs), far exceeding the capabilities of current haptic devices. In addition, to achieve clear object recognition, virtual grasping often requires multiple contact points, which introduces complex modelling and considerable computational overhead, particularly in non-penetration grasping scenarios, where the virtual fingers are constrained to remain on the object surface without interpenetrating it [15]. Therefore, further technical solutions are needed to ensure stable and controllable grasping in virtual reality.

Moreover, grasping techniques in VR differ broadly in complexity according to the virtual object's properties, such as mass, size, and materials. The object's stiffness also plays a significant role in the complexity of simulating the interaction with the object [16]. Most existing techniques focus on reducing the complexity of the grasping targets by simulating rigid bodies or objects with relatively simple and similar properties [17]. However, unlike rigid bodies, deformable bodies havehighly dynamic force attributes when the fingertips make contact with them (see Fig. 2). Therefore, achieving stable and realistic grasping simulations of deformable bodies with high-quality visuals and haptic feedback in virtual reality remains a challenging problem.

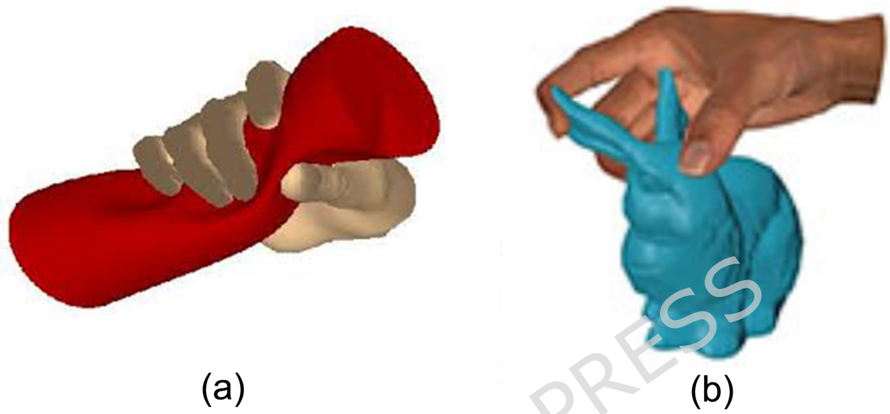

Fig. 2: Virtual hand grasping objects with different stiffness properties: (a) Grasping a deformable body, where penetration often occurs, while the object lacks stability [18]. (b) Grasping a rigid body, which remains stable during grasping [19].

In this paper, we review existing methods in both robotic and virtual grasping, with the aim of identifying key design principles. We introduce well-established robotic grasping techniques as a comparison framework for VR grasping methods. By consolidating knowledge from both robotics and VR, we propose a set of requirements for improving VR grasp quality and outline potential directions for future research. Section 2 describes the review methodology and literature selection process used to identify and analyse the relevant studies.

2 Methodology

This study adopts a structured literature review approach to analyse existing grasping techniques in robotics and virtual reality (see Fig. 3). The review follows principles similar to PRISMA-style selection to ensure a transparent and reproducible study identification process [20]. The objective of the review is to identify key developments in robotic grasping methods and to examine how these techniques relate to or inform grasping interaction within virtual reality environments.

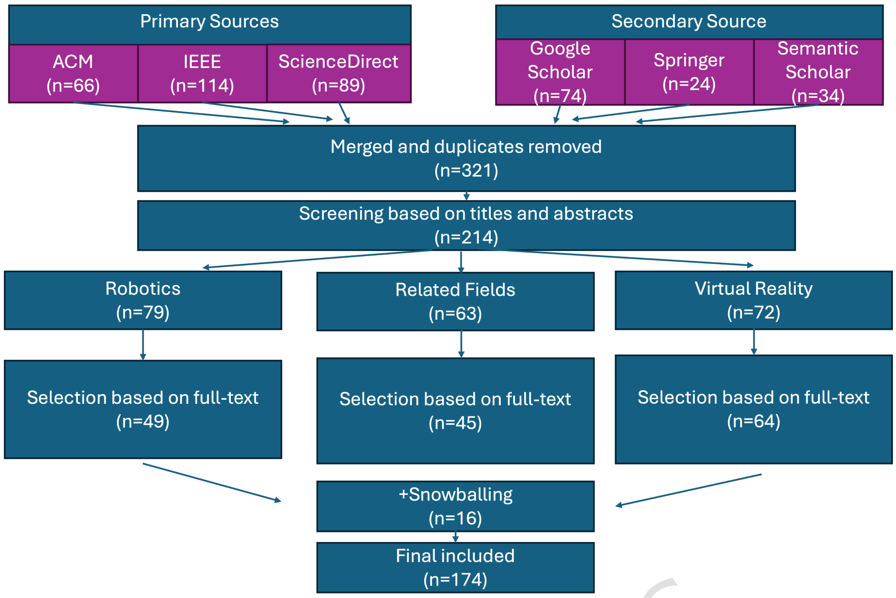

Fig. 3: PRISMA-guided for the selection process. Records were identified from primary databases (ACM, IEEE, ScienceDirect) and secondary sources (Google Scholar, Springer, Semantic Scholar). After merging and removing duplicates (n=321), papers were screened by titles and abstracts (n=214) and grouped into Robotics (n=79), Related Fields (n=63), and Virtual Reality (n=72). Following full-text review and snowballing (n=16), a total of 174 studies were included in the final analysis.

Relevant publications were identified through searches in major academic databases including IEEE Xplore, ACM Digital Library, and ScienceDirect, which provide extensive coverage of research in robotics, computer graphics, and human–computer interaction. Google Scholar and Semantic Scholar were additionally used as complementary search tools to broaden the search coverage. The search strategy combined keywords related to grasping, robotics, and virtual interaction techniques to ensure broad coverage of both domains.

After retrieving candidate papers, titles and abstracts were screened to determine relevance. Studies were included if they presented methods, algorithms, models, or frameworks related to robotic grasping, virtual grasping, or hand–object interaction techniques. The selected literature was subsequently analysed and organised into conceptual categories reflecting the major methodological directions identified in the field. As reflected in this review, robotic grasping techniques are primarily classified into analytic-based, geometric-based, and data-driven approaches, while virtual reality grasping research is discussed in terms of challenges, including visual modelling, visual artefacts, etc., and grasping methods are classified similarly to robotic classification: physics-based methods, animation-based methods, and data-driven methods. This categorisation provides a structured basis for comparing developments in robotics with emerging approaches to realistic grasping in virtual environments, as discussed throughout the remainder of this paper.

3 Robotic Grasping Overview

In many practical applications, robotic grasping is an action requiring precision, aimed at accomplishing specific tasks while requiring a high degree of adaptability to object shapes and stability in the grasping process [21, 22]. The aims of different robotics studies vary, leading to various categories of robotic grasping techniques. In this review, we divide robotic grasping methods into three classifications: analytic methods, geometric methods and data-driven methods. Under this classification, grasping can be formulated as a problem of identifying optimal contact points and hand configurations, based on the object's and hand's geometric and kinematic information, subject to force closure and equilibrium criteria, as well as reachability constraints [23] (see Fig. 4). In this section, we offer a concise overview of these methods, which serves as a comparison for the following chapter on VR grasping.

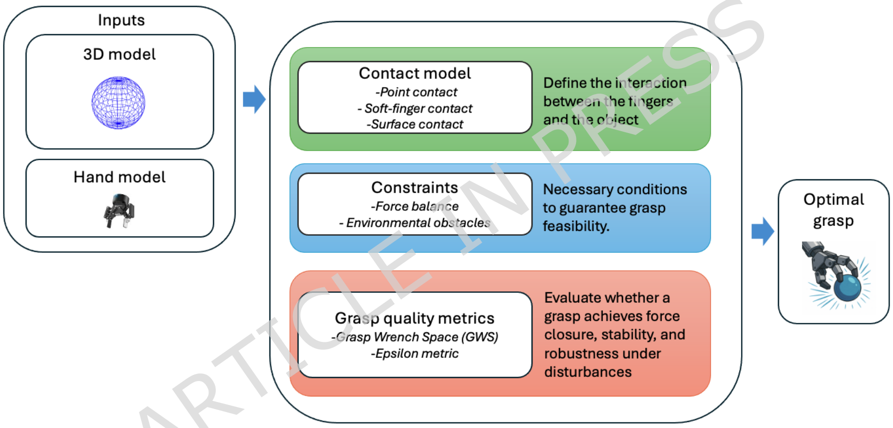

Fig. 4: Pipeline illustrating the robotic grasping formulation. The inputs consist of the object model and the hand (gripper) model. The contact model defines the interaction between the fingers and the object, while constraints specify the necessary conditions to guarantee grasp feasibility. Grasp quality metrics are applied to evaluate stability and robustness, ultimately leading to an optimal grasp.

3.1 Analytic-Based Grasping

Analytic methods are applicable to situations in which the information of the target shape has been defined, with constraints to find the suitable grasping pose [21]. In the determination process of the grasping movement, analytic methods need to be considered with the kinematics modelling (e.g., the joint configuration of the hand and the geometric relation of fingers to the object surface), dynamics modelling (e.g., the distribution of contact forces, torques). Specifically speaking, based on the positions of the hand and object, the grasping system identifies candidate contact points on the object model and synthesises the grasping configuration. This synthesis is further constrained by criteria such as closure properties and stability measures [24]. Under these criteria, stability depends on conditions defined by the number and distribution of contact points, the friction cones associated with these contacts, and the balance of forces and torques acting on the object. If these conditions satisfy the requirements for closure properties, the object can remain stable under grasping even in the presence of disturbances. Conversely, if these requirements are not met, the grasp will not be secure.

3.1.1 Contact Model

In analytic-based methods, it is essential to understand the contact models that describe the interaction between fingers and objects. Basically, different models lead to different implications for grasp stability and the conditions required for achieving force closure [25]. We focus on a widely adopted classification into two categories: Frictionless Point Contact (FPC) and Point Contact with Friction (PCWF).

FPC assumes that no friction is present at the interface between the fingertip and the object. Under this condition, the contact can only transmit a force along the surface normal, without any tangential component. The associated wrench (i.e., the generalised force consisting of a force vector and a torque vector) can be expressed as

$$F_{ci} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} f_{ci}, \qquad f_{ci} \geq 0 \tag{1}$$

where F_{ci} represents the unit force in the normal direction, and f_{ci} is the contact force. In this case, the force can only be compressive.

Generally, the FPC model is rarely considered in practice. Instead, the PCWF model is commonly applied when friction exists between the finger and objects. In this case, the contact point can exert forces within a cone of directions around the surface normal, which is referred to as the friction cone. This cone is derived from the Coulomb friction model [25].

$$F_{ci} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} f_{ci}, \quad f_{ci} \in FC_{ci}, \tag{2}$$

where

$$f_{ci} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \in \mathbb{R}^3$$

denotes the contact force vector at contact c_i. $FC_{ci} \in \mathbb{R}^3$ is the corresponding generalised wrench (force and torque) on the object. The 6×3 contact mapping matrix expands the contact force into the wrench space, specifically, the upper three rows represent the force components, while the lower three rows means the torque contribution of the contact point is neglected in this simplified formulation.

In practical frictional contacts, the torque contribution needs to be considered. In this case, the generalised contact wrench at contact c_i is

$$F_{ci} = \begin{bmatrix} I_3 \\ [r_{ci}]_\times \end{bmatrix} f_{ci},$$

where I_3 is the 3×3 identity matrix mapping the contact force to the translational component, and $r_{ci} \in \mathbb{R}^3$ is the position vector of the contact point relative to the object's centre of mass. If the contact point is not located at the centre of mass, then $r_{ci} \neq 0$, with $[r_{ci}]_\times$ is the skew-symmetric matrix of r_{ci}.

The admissible set of contact forces is given by the Coulomb friction cone:

$$FC_{ci} = \left\{ f \in \mathbb{R}^3 : \sqrt{f_1^2 + f_2^2} \leq \mu f_3,\ f_3 \geq 0 \right\}, \tag{3}$$

where μ is the coefficient of friction at the contact. Here, f_3 denotes the normal force, with the condition $f_3 \geq 0$ enforcing that the contact can only sustain compression. The tangential force magnitude $\sqrt{f_1^2 + f_2^2}$ cannot exceed the maximum static friction limit μf_3.

Compared with the FPC model, the PCWF model can resist external disturbances in multiple directions, thereby enhancing grasp stability.

3.1.2 Closure Properties

Closure is an important property of a grasp [26]. The grasp closure property can be divided into form closure and force closure. The earliest formal definition was given by Antonio et al. (1995) [27]. Subsequently, Elon et al. (1996) introduced the precise notions of both form closure and force closure [28].

Form closure is recognised as the ability to completely immobilise an object using only geometry, without relying on friction. We define a grasp with N contact points as (u^N, q^N), where $u^N = (u_1, \ldots, u_N)$ denotes the set of contact normal directions, and

$q^N = (q_1, \ldots, q_N)$ denotes the corresponding set of contact positions on the object. For each contact i, let $\phi_i(u_i, q_i)$ denote the *gap function*, which measures the signed distance between the object surface and the contact along direction u_i. The vector $\phi = (\phi_1, \ldots, \phi_N)$ then stacks the gap functions for all contacts. A grasp (u^N, q^N) achieves *form closure* if and only if the following implication holds:

$$\phi(u^N + \delta u, q^N) \geq 0 \quad \Rightarrow \quad \delta u = 0, \tag{4}$$

meaning that the only feasible infinitesimal displacement δu consistent with all contact constraints is the trivial one.

Specifically, the gap function $\phi_i(u, q)$ describes the state of contact between the object and the hand:

$$\begin{aligned} \phi_i &= 0 \quad \text{(contact)} \\ \phi_i &> 0 \quad \text{(separation)} \\ \phi_i &< 0 \quad \text{(penetration)} \end{aligned}$$

The interpretation is that the object cannot undergo any infinitesimal motion without causing at least one gap function to be violated (e.g., resulting in penetration), meaning that it is completely constrained by the geometric arrangement of the contacts [29].

On the other hand, force closure means that a grasp can resist any arbitrary external wrench applied to the object, provided that friction is available at the contact points [30]. Unlike form closure, which immobilises the object purely through the geometric arrangement of the contacts, force closure exploits friction cones at the contact points. This makes force closure a weaker requirement in terms of geometry, but stronger in terms of grasping capability: it guarantees stability against all possible external disturbances. For instance, a 3D object with six DoFs usually requires at least seven contact points to achieve form closure [31]. Under the assumption of hard finger closure, only three non-collinear contact points are sufficient to achieve force closure [29]. According to Bicchi et al. (2000), the stability refers to the ability of the grasp to maintain equilibrium under perturbations, which is affected not only by the locations of the contacts and their normals, but also by the local geometry and the distribution of contact forces [26].

Thus, a grasp is said to have force closure if and only if the following condition is satisfied:

$$\begin{cases} G\lambda = -g \\ \lambda \in \mathcal{F} \end{cases} \qquad \forall g \in \mathbb{R}^{n_v}, \tag{5}$$

where n_v is the dimension of the generalised wrench space (e.g., $n_v = 6$ for a rigid body in 3D). $G \in \mathbb{R}^{n_v \times m}$ is the grasp matrix mapping the contact force vector λ to the object's generalized wrench (force and torque), where m is the number of contact points multiplied by the number of basis vectors used to approximate each friction cone. Considering computational tractability, the nonlinear friction cone is often approximated by a polyhedral linearised cone spanned by a finite set of basis vectors $\{d_j\}$. In the linearised formulation, $\lambda \in \mathbb{R}^m$ denotes the vector of non-negative weighting coefficients associated with the basis vectors. $g \in \mathbb{R}^{n_v}$ denotes an arbitrary external generalised wrench applied to the object. In other words, force closure implies that for any arbitrary external generalised wrench g, there exists a feasible contact force distribution λ within the friction cone constraints such that the grasp matrix satisfies $G\lambda = -g$.

$\mathcal{F}$ is the composite friction cone defined as

$$\mathcal{F} = F_1 \times F_2 \times \cdots \times F_{n_c} = \{\lambda \in \mathbb{R}^m \mid \lambda_i \in F_i,\ i = 1, \ldots, n_c\}. \tag{6}$$

where n_c is the number of contact points, and each F_i denotes the friction cone associated with contact i.

It is worth mentioning that force closure only ensures that any external wrench can be resisted in a static sense, it does not guarantee dynamic stability of the grasp, meaning that a grasp satisfying force closure may still fail to remain stable under dynamic conditions such as impacts or sliding [30].

3.1.3 Grasp Quality Metrics

Grasp quality metrics are used to evaluate the performance of a grasp, once the grasp satisfies the conditions of form closure or force closure. These metrics allow a further assessment and can be used to optimise the grasp's stability [32]. We introduce the Grasp Wrench Space (GWS) as a fundamental framework that characterises grasping capabilities, and present the ϵ-metric as a widely used quality measure derived from it.

The GWS is defined as the convex hull of all possible wrenches that could be applied through the contact points $\{p_i\}_{i=1}^N$ of grasp g_i [33]. It represents the range of external forces and torques that the grasp can resist. The minimum distance between the origin and the boundary of the GWS, called the Largest Minimum Resisted Wrench (LRW), also referred to as the ϵ-metric, represents the minimum external wrench that could destabilise the object.

$$\epsilon = \min_{w \in \partial\,\mathrm{GWS}} \|w\| \tag{7}$$

The ϵ-metric measures the minimum external wrench that a grasp can resist in its most vulnerable direction. A larger ϵ value indicates a more stable grasp.

3.2 Geometric-based Grasping

Analytic methods are strictly based on force definitions, while geometric methods rely on geometric heuristics to efficiently generate candidate grasps. Among them, caging-based techniques represent a classical geometric approach that has been widely employed to address the problem of grasping arbitrary objects. Unlike analytic methods, caging does not depend on force modelling but on geometric constraints [34].

Caging refers to positioning the grippers so as to restrict the object within a region from which it cannot escape, thereby defining a feasible grasping space through geometric constraints. Wan et al. (2012) present a caging-based technique in which robotic fingers progressively shrink around an object until immobilisation is achieved [35]. Kim et al. (2019) formulate the required geometric caging conditions using a 2D planar abstraction of the object and hand configuration, enabling tractable mathematical analysis and algorithmic implementation [36]. Su et al. (2017) investigate caging grasping of 3D objects and propose the concept of a form-closure grasping cage. They construct a matrix representation using an attractive function in the configuration space to detect cage boundaries and analyse the transition from caging to form closure [34] (see Fig. 5).

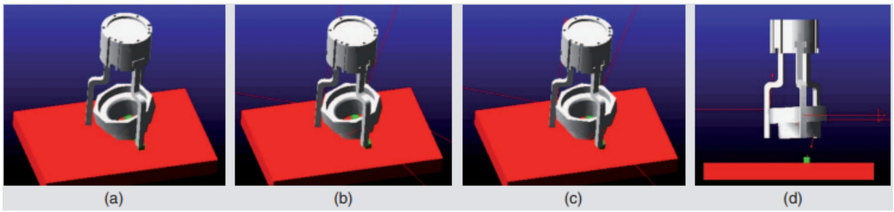

Fig. 5: After determining an initial caging configuration, the gripper subsequently executes a form-closure grasp, successfully securing the target object [34].

However, the gripper cage essentially provides a coarse configuration, which is not suitable for precise grasping tasks. In practice, heuristic search is often used in conjunction with geometric candidate generation, including caging-based approaches: starting from initial contacts generated by geometric heuristics and terminating once a force-closure grasp is identified [24]. Heuristic search has also been used to find optimal grasps, for example, by exploring combinations of candidate contact points when objects are modelled as a set of vertices [37].

3.3 Data-driven Grasping

Data-driven methods eliminate the need for explicit analytic modelling of objects. Instead, they rely on training with large-scale datasets. Compared to analytic methods, data-driven approaches can avoid complex computations [23]. In general, optimal grasping requires accurate perception and detailed analysis of geometric information. However, when facing unknown grasping targets, traditional analytic approaches often struggle [38]. With the development of machine learning and deep learning, prior knowledge extracted from data enables robotic grasping to move beyond a strict reliance on object models. We broadly categorise data-driven methods into two classes based on their learning mechanisms, which will be introduced in the following subsections.

3.3.1 Feature-based Methods

Feature-based methods aim to learn the association between object characteristics and grasp configurations, in order to select a set of grasp candidates for the task object [24]. This approach can generalise to previously unseen objects.

Saxena et al. (2008) presented an algorithm that identifies characteristic points in each image to obtain a 3D grasping location; however, it does not perform well on elongated objects [39]. Lenz et al. (2015) employed a sparse autoencoder (SAE) to automatically learn features from data, construct a two-stage cascaded detection system, and propose group regularisation to encourage each hidden unit to focus on a subset of input modalities, thereby avoiding the learning of spurious cross-modal correlations [40]. Redmon et al. (2015) directly feed RGB-D images into an AlexNet-based network, where convolutional layers automatically extract image features, and fully connected layers then output the grasp rectangle coordinates in a single forward pass, eliminating the need for sliding window scanning [41]. Mousavian et al. (2019) propose 6-DOF GraspNet, which employs a VAE to learn the conditional distribution of grasp poses given the object point cloud, generating a diverse set of grasp candidates; the sampled grasps are further optimised through a grasp evaluator network and gradient-based iterative refinement. This method significantly outperforms the baseline method [42]. Yang et al. (2025) proposed an attribute-based grasping framework that integrates visual and textual command inputs into a gated-attention encoder, decoding pixel-wise grasp affordances in an end-to-end architecture [43].

These feature-based methods do not rely on precise object models, support cross-category and cross-task generalisation, and reduce the need for manual teaching. Nevertheless, they rely heavily on visual and semantic features while largely ignoring physical constraints, which may lead to grasp failures when perception errors occur. They are also highly sensitive to the scale and quality of the training data, as the model tends to overfit without sufficient pretraining data and noisy reconstructions or biased web images can significantly degrade performance.

3.3.2 Demonstration-based Methods

Another category is demonstration-based methods, which transfers grasping knowledge from humans to robots. Operators provide demonstrations through devices such as data gloves or cameras, recording data including joint angles, hand configurations, and wrench information. The robot then learns grasping strategies from these demonstrations, enabling it to perform grasps in new scenarios [24].

Wang et al. (2024) present a representative Learning from Demonstration approach, in which human-demonstrated contact regions and approach directions are incorporated as priors into an ISF-based optimisation pipeline, offering a practical engineering implementation [44]. Ye et al. (2022) proposed the Continuous Grasping Function (CGF), an implicit function that takes continuous time, a latent code, and object point cloud features as inputs to predict robot joint positions at arbitrary timesteps, thereby generating smooth and continuous grasping trajectories. The method is built upon a Conditional Variational Autoencoder framework and trained on large-scale 3D human demonstration data, achieving significant improvements in both trajectory smoothness and real-robot deployment success rate [45]. Both addressing few-shot robotic grasping, DemoGrasp [46] employs explicit geometric based approach to resolve grasping in parallel gripper scenarios, while AdaDexGrasp [47] targets the more demanding setting of dexterous manipulation. By combining trajectory-guided reinforcement learning with curriculum learning, AdaDexGrasp efficiently acquires grasping skills from a single human demonstration per skill, and further leverages a vision-language model to adaptively select the appropriate grasping strategy based on user instructions.

Demonstration-based methods directly benefit robots by inheriting human dexterity and task knowledge. However, their main limitation lies in the difficulty of mapping human hand configurations to robotic hands with different degrees of freedom and mechanical constraints, which introduces additional complexity in real applications.

Kinematic mismatch arises primarily from discrepancies in the structure of the finger base joints and the location of the thumb [48]. Such a mismatch makes direct replication of human hand motion on a robot hand infeasible, necessitating retargeting, during which additional constraints further complicate the problem: joint limits restrict the reachable configuration space, meaning that some demonstrated gestures are infeasible even after optimisation [49]. In addition, collision and self-collision constraints require retargeted motions to avoid penetration, altering the approach direction and contact sequence, which can lead to contact timing differences [50], contact point offsets, and changes in contact normals [51]. Even when geometric alignment is achieved, contact feasibility must be ensured: the selected contacts should admit physically realisable forces and torques to provide stable grasps and successful downstream manipulation [49].

Principle-driven retargeting seeks to reduce mapping-induced deviations by encoding task-relevant geometric priors while avoiding expensive online optimisation. Yin et al. (2025) propose a formulation that learns a mapping from human fingertip keypoints to robot joint targets, guided by five geometric criteria: (I) Motion preservation: The retargeted fingertips preserve local motion directions, so small human fingertip displacements induce robot fingertip motions in consistent directions; (II) C-space coverage: The retargeted fingertip trajectories are encouraged to span the robot's reachable fingertip space, promoting an approximately surjective mapping from human to robot workspace; (III) High flatness: The mapping is regularized to have low curvature / low gain variation, yielding more uniform sensitivity across the workspace; (IV) Pinch correspondence: Pinch gestures are mapped to robot fingertip proximity to preserve pinch semantics crucial for grasping; and (V) Collision-free: Self-collisions are penalized to encourage physically feasible postures during teleoperation [52]. The design is well-suited for high-rate teleoperation and scalable data collection, but it does not explicitly reason about grasping force feasibility or physically realisable contact wrenches, so downstream stability may still benefit from contact-aware refinement when physically realisable wrench generation is critical.

In practice, the retargeting deviations manifest primarily as point offsets, normal misalignment, and altered contact sequences, which degrade grasp stability by reducing force-closure margins and increasing frictional requirements, often resulting in slip, push-away, or missed closure. Specifically, contact point offsets change the effective lever arms and required frictional support. Under the same normal squeezing force, the tangential demand can more easily exceed the friction limit, leading to slippage [53]. Normal misalignment alters the direction of the applied contact forces. Instead of producing stabilising opposition, the grasp may generate unintended pushing motions [54]. Finally, contact-sequence deviations can perturb the object early in the interaction, so that subsequent finger closure cannot recover the intended geometry, often resulting in failed grasp closure [55].

Therefore, while [52]'s principle-driven retargeting improves kinematic consistency and controllability at high rates, mapping errors can still be amplified during physical interaction through contact point offsets, normal misalignment, and contact-sequence deviations. To keep such error propagation within acceptable bounds, practical downstream, contact-aware refinement is often required. Liu et al. (2022) propose a differentiable and efficient surrogate grasp-quality metric, which reformulates force-closure conditions into a continuous, soft-constrained objective via a contact-normal–based residual. This enables direct synthesis of diverse grasps within a gradient-guided sampling/optimisation framework. The approach generalises to different hand and gripper morphologies as long as the hand's forward kinematics and geometry are differentiable [56]. Garcia et al. (2009) note that, for tasks involving contact with the environment, it is common to adopt task-space hybrid control and impedance control. In particular, along the direction perpendicular to the contact surface, the interaction force can be regulated via an impedance relationship, allowing the end-effector trajectory to adapt locally so as to suppress excessive contact-force peaks and improve contact stability [57]. Zhao et al. (2025) proposed a slip-detection model aimed at generalising across five-finger dexterous hands, diverse grasp types, and objects with varying materials and shapes, while remaining fast and robust enough for stable real-time deployment [58].

3.4 Damage-Aware Grasping

As the existing techniques can achieve grasping in a stable way in many circumstances, the complexity of grasping increases as additional constraints come into play. Some specific grasping tasks focus on highly deformable and even fragile bodies, which must be taken into account in addition to the stability of object-hand interaction. Most flexible deformable objects, basically, will go through three kinds of states: elastic deformation, plastic deformation, and compressive failure [59]. In such cases, it is often important to avoid damage to the grasping targets. Yu et al. (2023) propose a method that jointly predicts the deformation properties and grasp pose of the objects, with the aim of enabling safe grasping and avoiding damage to deformable items [60]. Their method controls the deformation change in the range below the elastic limit, in order to avoid plastic deformation or damage of the grasping targets, such as an orange or a paper cup. Zaidi et al. (2017) introduce a non-linear model, which is able to handle the interaction between the deformable bodies and robotic hands [61]. Xu et al. (2020) also try to solve the problem of damage and liquid spillage in the process of grasping hollow objects, like plastic bottles and paper cups [62]. However, these methods are only applicable to some specific objects, and they do not have much versatility. For some special fragile objects, like tofu, grasping strategies could be more challenging. there is still no perfect way to solve this problem, One solution is to use fluid fingertips [63, 64]. Another solution is to use micro-patterned pads, an interesting method which is motivated by the microstructure of the footpalm of some organisms in nature, like geckos. Nguyen et al. (2020) refers to the foot palm of the tree frog to design the cushion of the gripper [65], which simplifies the complex deformation analysis of fragile deformable objects during the grasping.

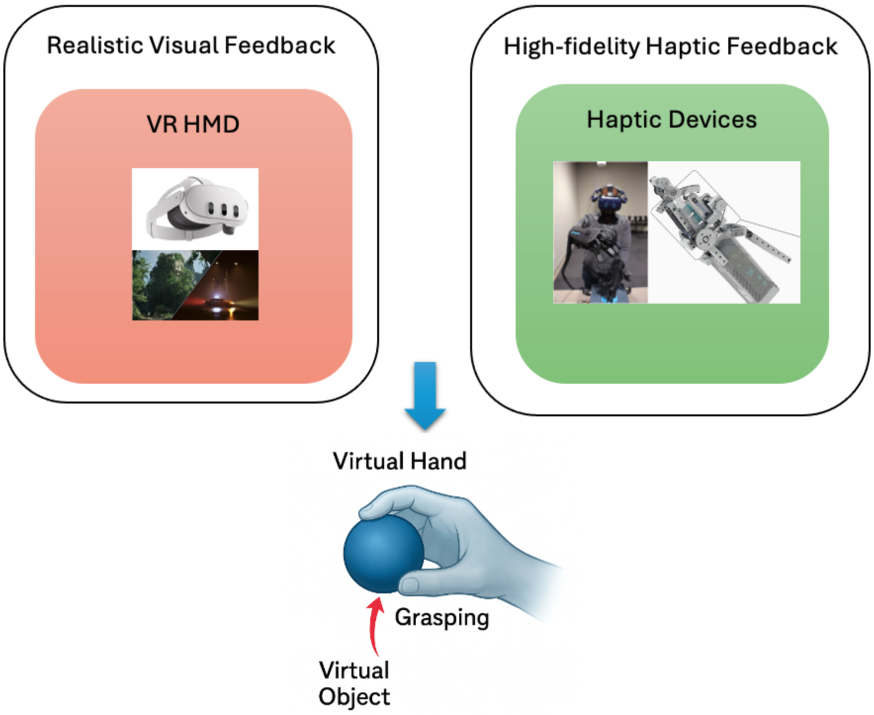

Fig. 6: Overview of the standard VR grasping setup. The user achieves realistic grasping through the integration of photorealistic visual feedback and high-fidelity haptic feedback [66–68].

4 VR Grasping Techniques

Realistic grasping is inherently governed by the laws of physics [69]. As a result, human hands are capable of easily grasping objects with different shapes, masses, spatial positions, and friction properties, often without requiring too much effort [70]. In contrast, when grasping is considered in virtual environments, this changes drastically. According to Dalia et al. (2021), users tend to rely predominantly on power grasps when interacting with virtual objects, mainly due to the lack of physical constraints, whereas in the real-world, grasping relies more heavily on precise movements [71].

When realistically simulating grasping in VR, the virtual environment cannot constrain human hand motion according to the laws of physics, as in the real environment [69], which may lead to visually implausible results. To address these issues, significant progress has been made in recent decades in terms of visual feedback, enabling the rendering of photorealistic objects [66]. Meanwhile, various technological devices have been introduced to enhance haptic feedback and improve the plausibility of interactions, such as haptic gloves [67], handheld controllers [68], and hand exoskeletons [72]. As illustrated in Fig. 6, VR grasping requires the combined contribution of both visual and haptic modalities. Nevertheless, a thorough understanding of grasping mechanisms is particularly important for safety-critical applications, including teleoperation [73] and surgical operations [74]. In this section, we review virtual grasping techniques with a focus on their key challenges and methodological approaches.

4.1 Challenge

Despite substantial progress in visual feedback and haptic feedback, several challenges remain in grasping simulation, particularly in the reproduction of precise grasping behaviours and the handling of deformable bodies [17]. For large deformable bodies and heterogeneous deformation bodies–defined here as objects with spatially varying material properties such as stiffness or density–grasping becomes increasingly complex [75]. In some approaches, deformable objects are simplified and approximated as rigid bodies [76], although such approximations remain limited in their ability to capture realistic interactions. Consequently, much of the existing literature has focused on rigid bodies, since they exhibit relatively simple physical properties [77]. Furthermore, with the continuous improvement of virtual environments, VR applications place growing demands on the ability to grasp arbitrary targets. This underscores the significance of clearly defining the challenges of grasping simulation and developing effective solutions to address them.

4.1.1 High-Fidelity Modeling

High-fidelity hand modeling is a crucial input for virtual grasping, and the virtual hand should reproduce, as closely as possible, the degrees of freedom and dexterity of the human hand [78]. The whole hand has 27 DoFs, and each finger except the thumb has three bones: the distal, middle, and proximal phalanges. Also, these fingers have three joints, which are the MCP (metacarpophalangeal), PIP (proximal interphalangeal), and DIP (distal interphalangeal) joints [79, 80] (see Fig. 7). Such properties induce a large configuration space for the hand, requiring virtual grasping algorithms to handle various groups of DoFs. According to Mulatto et al. (2013), the human skill of manipulating objects depends on the thumb's kinematics and on its capability of opposing the other fingers [81]. The thumb has 5 DoFs, which allows humans to perform complex spatial movements [82]. Physiologically, tendon and ligament constraints prevent finger joints from moving completely independently [83]. Hence, additional constraints are necessary to reproduce natural hand poses. Altogether, real-time rendering of high-fidelity virtual hands does incurs substantial computational costs.

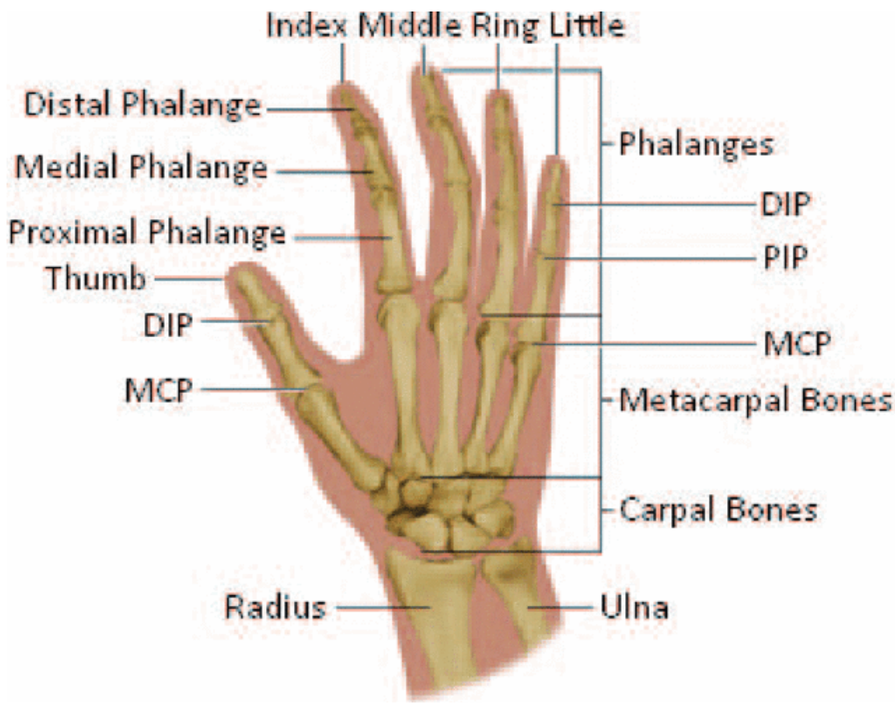

Fig. 7: The structure of the human hand skeleton. Each finger except the thumb has three phalanges and joints [79].

To address these issues, Kuch et al. (1995) presented a universal hand model in which the surfaces of the palm and fingers are represented using cubic B-splines, thereby reducing the number of control points and improving rendering efficiency [84]. Burton et al. (2011) introduce a kinematic model of the human hand, focusing on the precise mapping of thumb movements [79]. Oprea et al. (2019) introduced a visually realistic grasping system that allows users to interact with or manipulate non-predefined objects based on collision detection [76]. Unfortunately, this method suffers from penetration issues, which reduce the quality of the interaction. Although relatively realistic models of human hands have been developed [78, 85], reproducing surface stress remains difficult due to the limitations of current connection models between the hand and objects. Modelling a precise and stable contact mechanism is still a challenging task [17].

Notably, most researchers rely on virtual rigid fingers to simulate hand–object interactions. According to Verschoor et al. (2018), adopting soft finger models can provide more realistic feedback (see Fig. 8). However, high-resolution soft finger models are limited by the trade-off between computational cost and physical realism. Other approaches approximate deformable skin by modelling a soft pad attached to a rigid skeleton [86–88]. Jacobs et al. (2011) proposed a novel soft body model based on Lattice Shape Matching (LSM) [86]. While this approach increases the robustness of object manipulations, the soft pad may collapse in some cases. Moreover, compliance between the bones and the tracker, as well as inter-penetration artefacts, remain important factors to consider. These challenges suggest that existing soft finger modelling remains incompletely resolved, particularly with respect to numerical stability, contact fidelity, and real-time performance. It worth mentioning that the robotics literature on soft-finger contact provides a more physically grounded and calibratable contact semantics: reliable contact is not merely the absence of penetration, but the ability to stably realize a contact wrench space at the fingertip. A soft fingertip forms a finite contact patch, which enables not only a tangential friction force f_t but also a frictional moment about the contact normal, denoted τ_n (torsional friction moment) [89]. Ciocarlie et al. (2007) model the coupled feasibility of tangential force and torsional moment using a friction ellipsoid,

$$f_t^2 + \frac{\tau_n^2}{e_n^2} \leq (\mu P)^2, \tag{8}$$

where μ is the friction coefficient and P is the normal load (larger P increases the available friction budget μP). The parameter

$$e_n = \frac{\max(\tau_n)}{\max(f_t)} \tag{9}$$

acts as a torsional-capacity scale, Using the Winkler foundation model, the maximum torsional friction moment about the contact normal satisfies

$$\frac{\tau_n^{\max}}{P} = \frac{8}{15}\,\mu\,\sqrt{ab}, \tag{10}$$

and therefore

$$\tau_n^{\max} = \frac{8}{15}\,\mu\,P\,\sqrt{ab}, \tag{11}$$

where a and b are the semi-axes of the elliptical contact patch, μ is the friction coefficient, and P is the normal load. Since the maximum tangential friction force is $f_t^{\max} = \mu P$, the torsional-capacity scale (eccentricity parameter) is

$$e_n = \frac{\tau_n^{\max}}{f_t^{\max}} = \frac{\frac{8}{15}\mu P\sqrt{ab}}{\mu P} = \frac{8}{15}\sqrt{ab}, \tag{12}$$

i.e., $e_n \propto \sqrt{ab}$, scaling with the effective radius of the contact patch. Hence, e_n is closely tied to the effective size of the contact patch. This formulation makes explicit that, for a given P, f_t and τ_n draw from the same friction budget: allocating more capacity to resist tangential slip leaves less margin to resist torsional slip, and vice versa. Consequently, by explicitly targeting (i) contact patch geometry (e.g., a, b or $\sqrt{ab}$, which determines e_n), (ii) the tangential friction threshold (μP), and (iii) the torsional friction threshold $\tau_n^{\max} \approx e_n \mu P$, we can turn soft-finger realism into calibratable objectives that provide a principled basis for fingertip calibration in terms of contact patch size, shear resistance, and torsional resistance.

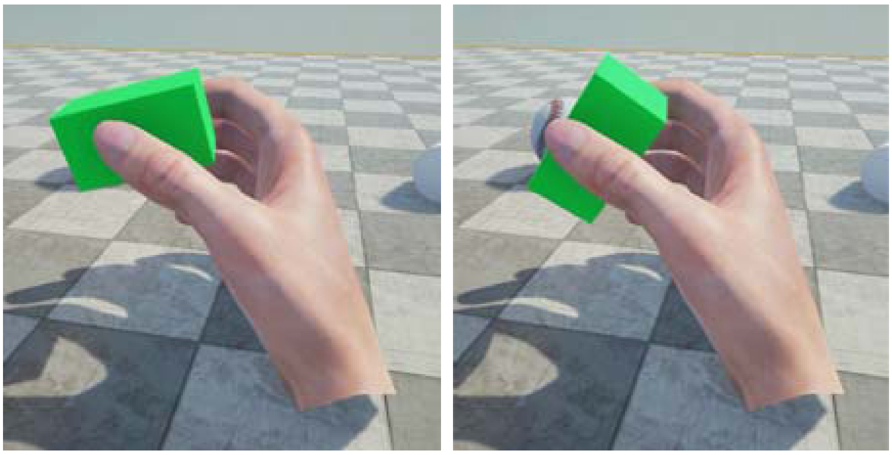

Fig. 8: Soft hand simulations while interacting with virtual objects within virtual reality. The proposed method does not support self-collisions of the hand nor support contact between deformable objects [7].

4.1.2 Visual Artefacts and Hand-Object Overlapping

Current VR techniques do not allow the virtual environment to constrain real hand motion as in the physical world. This often leads to unrealistic grasping behaviours and visually distracting artefacts [69], such as penetration between the virtual hand models and objects (see Fig. 9). These artefacts degrade the quality of visual feedback and diminish the immersive experience. Moreover, hand-object overlapping is a persistent problem in computer animation. It occurs not only in 3D object grasping but also in 2D interactions [90]. when the grasping targets have complex geometric features, such as vases, stuffed toys, or handicrafts, penetration is likely to occur [91, 92], since bounding boxes or spheres are often used as coarse approximations of such objects. In some applications, discrete collision detection is employed, which tests for intersections only at fixed time intervals and can therefore miss collisions occurring between frames, resulting in visual artefacts such as hand-object interpenetration. This phenomenon is known as tunneling [93]. Despite our desire to adopt low-cost solutions to prevent penetration, in dexterous manipulation tasks such cost-driven approximations in collision and contact handling often exacerbate other equally critical failure modes. These include contact jitter—high-frequency frame-to-frame oscillations of the contact point [94], or missed time-of-impact (TOI) events, where collisions occurring between discrete time steps go undetected [95], and erroneous stick–slip mode switching, which introduces spurious stickiness and abrupt force/velocity transitions that undermine stable, controlled manipulation [96].

Pelletier et al. (2025) cast triangle–SDF collision detection as an explicit search for the earliest time-of-impact (TOI) over a continuous time interval [93]. Their procedure jointly minimises the signed distance over the continuous time variable and the barycentric coordinates of the triangle, locating the earliest contact across the full time interval. In practice, robustness remains coupled to the quality of the SDF as well as numerical precision and the available iteration budget, which can reintroduce conservative steps or residual failure cases in cost-constrained settings.

In contrast, constraint-based contact handling treats collision and contact directly as feasibility constraints. For example, Müller et al. (2007) describe a projection-based strategy that resolves penetrations by projecting points back onto the feasible set, enforcing non-interpenetration at the position level and bypassing TOI detection entirely [97]. A different line of work accelerates collision queries by organising geometry into hierarchical bounding structures: Klosowski et al. (1998) propose bounding volume hierarchies of k-DOPs that prune non-overlapping subtree pairs early in traversal, reserving exact primitive-level intersection tests for the leaves [98]. The performance gain is achieved through hierarchical culling, though the tightness of the bounding volumes and the granularity of the hierarchy still govern sensitivity to thin structures and near-contact configurations.

More recent robust formulations integrate contact into the global time-stepping solve. Li et al. (2020) propose Incremental Potential Contact (IPC), which combines unsigned distance with a smooth log-barrier and enforces intersection-free iterates by bounding each Newton step via CCD-based line search. IPC further stabilises frictional behaviour by smoothing static friction and using a lagged dissipative potential, mitigating spurious stick–slip mode switching while maintaining non-penetration throughout the optimisation [99].

Collision detection and contact handling are not only responsible for preventing interpenetration or generating reaction forces; more importantly, they delineate the feasible interaction set induced by object geometry, kinematic reachability, and frictional constraints. By bounding which contact modes are physically admissible at each state, this structured characterisation provides the discrete scaffolding on which planners and controllers can organise and reason over long-horizon contact sequences. For instance, in the task-driven hybrid model reduction framework of Jin et al. (2024), contact feasibility is encoded implicitly through the complementarity constraints of a linear complementarity system (LCS), where the active or inactive inequalities partition the state-input space into distinct hybrid modes. Rather than reasoning over all combinatorially many modes, the framework identifies a small set of task-relevant modes through on-policy learning, and a trust-region LCS-MPC replans in receding-horizon fashion to stably realise contact transitions. Structurally, this explicitly elevates contact mode selection and contact evolution to first-class decision variables in both the model representation and the controller [100].

This reliance on contact variables as decision inputs implies that the quality of the underlying contact detection directly conditions the integrity of both the learned model and the MPC rollout, as misclassified contact events corrupt the training data from which task-relevant modes are identified and perturb the complementarity constraints during online replanning.

From this planning-variable viewpoint, low-cost collision and contact approximations carry consequences that extend well beyond minor geometric artefacts such as small interpenetrations. Rather, such approximations can significantly degrade the quality of contact variables, undermining the assumptions on which the decision-making stack depends. Specifically, contact jitter corresponds to high-frequency noise or discontinuities in estimated contact points and normals, making the contact state difficult to track consistently over time. Missed time-of-impact (TOI) events perturb hybrid event timing and lead to phase errors in contact transitions. Erroneous stick–slip switching is effectively a misclassification of a discrete contact-mode variable, which propagates downstream and degrades long-horizon reasoning over contact sequences. Since these local errors compound and propagate across contact transitions, they can systematically undermine geometry-aware long-horizon manipulation methods that depend on stable contact sequences.

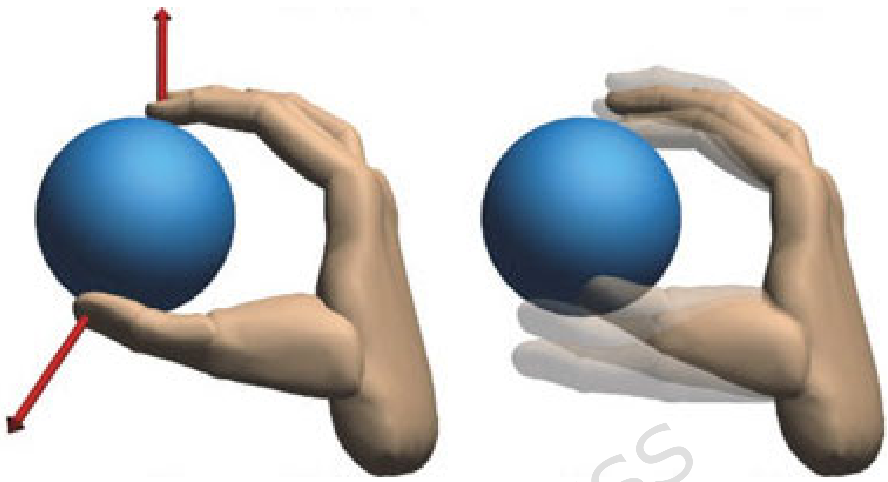

Fig. 9: Illustration of unrealistic interpenetration that may occur during the interaction between the virtual hand and the object [92].

4.1.3 Device Limitations

In virtual environments, visual and haptic feedback are critical factors influencing grasping performance. Modern head-mounted displays (HMDs) can already provide photo-realistic rendering of visual scenes, distinguishing them from early VR systems [101]. Such high-fidelity visual scenes can assist users in reproducing realistic grasping actions [102]. However, the development of haptic devices continues to encounter substantial limitations. Haptic devices are essential for simulating grasping interactions and have been widely investigated to enhance user immersion [103]. Through haptic interaction, users are able to perceive force, material properties, and object shape during virtual manipulation, which in turn provides essential cues for realistic grasping [104]

Early systems, such as the force-feedback arm exoskeleton introduced by Frisoli et al (2005), achieved a high control update rate of 2 kHz but was limited by a low number of DoFs, constraining realistic operation [105]. Standard VR controllers typically provide only vibrotactile cues triggered by collisions or discrete events, which lack the ability to convey continuous and directional force information [106, 107].

Work	**mitigation**	**Main sensitivity**
Pelletier et al. [93]	Triangle–SDF collision detection formulated as a spatio-temporal local optimisation; jointly minimises signed distance over the continuous time variable and triangle barycentric coordinates to locate the earliest TOI without discrete threshold checks.	Robustness depends on SDF quality (gradient regularity/accuracy, thin features) and on numerical precision / iteration budget
Müller et al. [97]	Contact handled as feasibility constraints; penetration removed by projecting points/states back onto the feasible set (state-level non-interpenetration rather than TOI-only handling)	Quality depends on projection accuracy and iteration budget; under-resolution can yield non-physical artifacts or jitter
Klosowski et al. [98]	Organises geometry into k-DOP bounding volume hierarchies to prune non-overlapping subtree pairs early in traversal; exact primitive-level intersection tests are reserved for the leaves.	Sensitivity governed by BVH tightness and hierarchy granularity; thin structures and near-contact configurations may escape early culling.
Li et al. [99]	IPC: unsigned distance + smooth log-barrier; CCD-bounded Newton line search ensures intersection-free iterates; friction forces resolved via lagged potential within the same solver, reducing spurious stick–slip transitions arising from force–velocity coupling.	Higher per-step cost (CCD + nonlinear solve); performance depends on contact-pair count and solver settings

Table 1: Representative methods for mitigating contact-variable failure modes.

Haptic glove technologies have demonstrated the ability to render force feedback with higher accuracy. For example, the HaptX Glove G1 uses a steady flow of compressed air to provide resistive force feedback, delivering high-resolution, high-displacement tactile sensation [108]. While this allows for enhanced feedback, the weight of the device makes it challenging for users to wear for extended periods. In contrast, grounded devices like Phantom Premium and Touch X are able to transfer realistic elastic forces, as well as simulate the textures of object surfaces [106]. Nevertheless, they are limited to a single input modality, which only supports touching or poking behaviour.

While visual rendering in VR has reached photo-realistic quality, haptic devices are still hindered by difficulties in reproducing fine tactile cues, high development and maintenance costs, and poor portability or wearability. Addressing these issues to develop lightweight and stable haptic solutions with high-fidelity force feedback remains central to enabling realistic grasping in VR.

4.1.4 Incomplete Theoretical Framework

A complete theoretical system for assessing the quality VR grasping has not yet been established. In particular, there is currently no universally accepted metric [109]. Grasping performance in VR is often evaluated qualitatively, for instance through user studies and questionnaires, which makes it difficult to assess grasping quality in a consistent and quantitative manner. In contrast, in robotic grasping, Ferrari et al. (1992) introduced grasp quality metrics based on GWS, which have since been widely adopted and extended to provide guarantees of grasp stability under external disturbances [33]. Furthermore, in robotic assembly research, the National Institute of Standards and Technology (NIST) has presented a set of performance metrics and benchmarking tools that contribute to the technical specification of robotic systems [110]. The NIST benchmarks can also be applied to the manipulation and grasping of deformable bodies in assembly tasks. However, to the best of our knowledge, although simulation-based benchmark such as DaXBench [111] exist for evaluating virtual grasping of deformable objects, no benchmark comparable to those for industrial assembly has yet been established to systematically define performance metrics and evaluation procedures for VR grasping. Hence, the evaluation framework for virtual grasping still requires significant improvement.

4.2 Grasping Methods

With increasing user demands, symbolic substitution–where grasping is represented by simplified attach/detach events rather than realistic physical interaction [112]–can no longer be regarded as a sufficient means of grasping. To achieve grasping with stable and realistic interaction, a variety of methods have been proposed. This subsection outlines the main approaches to virtual grasping in detail.

4.2.1 Physics-Based Methods

In most VR grasping applications, form closure is sufficient to visually and geometrically constrain objects, as it provides the appearance of a stable grasp without the computational cost of simulating detailed contact forces. However, for applications requiring physically accurate interaction and haptic realism — such as surgical training or robotic teleoperation — force closure becomes essential, since the stability of the grasp depends not only on geometry but also on the distribution of contact forces and friction. Such methods rely on physical input devices to achieve grasping behaviours, which are constrained by the laws of physics and thus referred to as physics-based methods [113–115]. Specifically, given a hand model, the VR system couples the tracked hand motion to a physics engine that enforces dynamics, non-penetration constraints, and friction. Stable grasps naturally emerge when the simulated contact forces and torques drive the hand–object system to static equilibrium, rather than from precomputed pose generation [116].

In physics-based methods, force is the most significant physical constraint. Simulating haptic feedback while grasping a 3D virtual object requires three-dimensional forces and torques, with sufficient magnitude to reproduce the contact forces involved in dexterous manipulation, thereby enabling the perception of power grasps of virtual objects [117]. To capture subtle variations in contact between fingers and objects, both high force resolution and rapid dynamic response are essential. In tasks demanding accurate force feedback, such as rotating a virtual tumour in a surgical simulation, the error in force rendering should remain below the human discrimination threshold [118]. Under such conditions, users can infer tumour stiffness through the interaction of resistance forces and movement [119].

Ideally, the force rendering error e should remain below the human force discrimination threshold $F_{\text{threshold}}$, i.e.,

$$e < F_{\text{threshold}}. \tag{13}$$

In practice, sampling and zero-order hold (ZOH), computation/communication delay, and limited actuator DoF introduce distortions that prevent perfect transparency [120–122].

Therefore, we formulate a perceptual requirement that the force-rendering error remains below the human discrimination threshold (JND):

$$e_{\text{total}} < \text{JND}. \tag{14}$$

The JND is the minimum change in stimulus required for reliable perception. This measure corresponds to the maximum level of augmentation that an individual can experience without becoming aware that augmentation has occurred [123]. Prior studies suggest that JND depends on multiple factors, including force direction, increment versus decrement, interaction context, and hand velocity [124, 125].

We thus represent the scenario-specific JND as

$$\text{JND} = f(F_0,\ v,\ \mathbf{d},\ \sigma,\ \ell,\ k)\,, \tag{15}$$

where F_0 is the reference force, v is hand velocity, $\mathbf{d}$ is force direction, $\sigma \in \{+1, -1\}$ indicates increment/decrement, ℓ is contact location, and k indexes the subject. For a given task, we define the worst-case (most sensitive) threshold as

$$\mathrm{JND}^* = \min_{\substack{F_0 \in \mathcal{F}_{\text{task}},\, v \in \mathcal{V}_{\text{task}} \\ \mathbf{d} \in \mathcal{D}_{\text{task}},\, \sigma,\, \ell,\, k}} f(F_0,\ v,\ \mathbf{d},\ \sigma,\ \ell,\ k)\,. \tag{16}$$

The design requirement $e_{\text{total}} < \mathrm{JND}^*$ then provides a conservative, task-specific bound that is guaranteed to hold across all operating conditions within the task.

Let $F_{\text{sim}}(t)$ denote the task-relevant scalar contact force produced by the simulator. The haptic device cannot reproduce the full simulated wrench [126]; instead it generates a finite-dimensional actuator stimulus vector $\mathbf{u}(t) \in \mathbb{R}^{n_a}$ (e.g., vibration amplitude/force level per actuator), which is transformed into an equivalent rendered force through a calibration map:

$$F_{\text{rend}}(t) = \mathbf{c}^\top \mathbf{u}(t), \tag{17}$$

where $\mathbf{c} \in \mathbb{R}^{n_a}$ is identified from device calibration (linearised around the operating range).

It is worth noting that human perception further filters the rendered force:

$$F_{\text{perc}}(s) = H_{\text{human}}(s)\, \mathbf{c}^\top \mathbf{u}_{\text{act}}(s), \tag{18}$$

where the human perception filter is modelled as a first-order low-pass system:

$$H_{\text{human}}(s) = \frac{1}{1 + s/\omega_h}. \tag{19}$$

Because the actuator space is limited, we introduce a mapping function that converts the simulated force to a desired actuator stimulus:

$$\mathbf{u}(t) = g(F_{\text{sim}}(t))\,. \tag{20}$$

The mismatch between the simulator force and rendered force defines the mapping error:

$$e_{\text{map}}(t) = |F_{\text{rend}}(t) - F_{\text{sim}}(t)|\,. \tag{21}$$

This term captures information loss due to finite actuator DoF and any approximation within $g(\cdot)$.

The desired stimulus $\mathbf{u}(t)$ is delivered through a sampled-data haptic pipeline. Let T_s be the sampling period and T the one-way communication latency. We model sampling and ZOH by:

$$H_{\text{zoh}}(s) = \frac{1 - e^{-sT_s}}{sT_s}, \tag{22}$$

and communication delay by:

$$H_{\text{comm}}(s) = e^{-sT}. \tag{23}$$

We separate *transmission* effects (sampling/ZOH, delay, and command quantisation) from the *device/closed-loop* dynamics:

$$\mathbf{u}_1(s) = Q(H_{\mathrm{tf}}(s)\,\mathbf{u}(s))\,, \qquad \mathbf{u}_{\mathrm{act}}(s) = H_{\mathrm{dyn}}(s)\,\mathbf{u}_1(s), \tag{24}$$

where

$$H_{\mathrm{tf}}(s) = H_{\mathrm{zoh}}(s)\,H_{\mathrm{comm}}(s), \tag{25}$$

$Q(\cdot)$ denotes command quantisation (e.g., a uniform quantiser with step size Δ_u), and $H_{\mathrm{dyn}}(s)$ denotes the closed-loop device dynamics (actuators + embedded controller), which is a stable causal transfer matrix. $H_{\mathrm{dyn}}(s)$ is identified from the closed-loop input–output response $\mathbf{u}_1 \to \mathbf{u}_{\mathrm{act}}$. The key point is that H_{tf} captures sampled-data and delay, while H_{dyn} captures the device bandwidth and closed-loop tracking limitations.

Using the intermediate signal $\mathbf{u}_1$, we define two additional error components in the equivalent-force domain:

$$e_{\mathrm{tf}}(t) = \left|\mathbf{c}^\top\mathbf{u}_1(t) - \mathbf{c}^\top\mathbf{u}(t)\right| = \left|\mathbf{c}^\top(\mathbf{u}_1(t) - \mathbf{u}(t))\right|, \tag{26}$$

$$e_{\mathrm{dyn}}(t) = \left|\mathbf{c}^\top\mathbf{u}_{\mathrm{act}}(t) - \mathbf{c}^\top\mathbf{u}_1(t)\right| = \left|\mathbf{c}^\top(\mathbf{u}_{\mathrm{act}}(t) - \mathbf{u}_1(t))\right|. \tag{27}$$

We define the total perceived error as the difference between the perceived force and the ideal perceived force:

$$e_{\mathrm{total}}(t) = |F_{\mathrm{perc}}(t) - F_{\mathrm{ideal}}(t)|\,, \qquad F_{\mathrm{ideal}}(s) = H_{\mathrm{human}}(s)\,F_{\mathrm{sim}}(s). \tag{28}$$

Noting that

$$\mathbf{c}^\top\mathbf{u}_{\mathrm{act}} - F_{\mathrm{sim}} = (F_{\mathrm{rend}} - F_{\mathrm{sim}}) + \mathbf{c}^\top(\mathbf{u}_1 - \mathbf{u}) + \mathbf{c}^\top(\mathbf{u}_{\mathrm{act}} - \mathbf{u}_1), \tag{29}$$

and applying the triangle inequality to $e_{\mathrm{total}}(t) = |H_{\mathrm{human}} * (\mathbf{c}^\top\mathbf{u}_{\mathrm{act}} - F_{\mathrm{sim}})|$ yields:

$$e_{\mathrm{total}}(t) \le \underbrace{|H_{\mathrm{human}}*(F_{\mathrm{rend}} - F_{\mathrm{sim}})|}_{\triangleq\,\tilde{e}_{\mathrm{map}}(t)} + \underbrace{\left|H_{\mathrm{human}}*\left(\mathbf{c}^\top\mathbf{u}_1 - \mathbf{c}^\top\mathbf{u}\right)\right|}_{\triangleq\,\tilde{e}_{\mathrm{tf}}(t)} + \underbrace{\left|H_{\mathrm{human}}*\left(\mathbf{c}^\top\mathbf{u}_{\mathrm{act}} - \mathbf{c}^\top\mathbf{u}_1\right)\right|}_{\triangleq\,\tilde{e}_{\mathrm{dyn}}(t)}. \tag{30}$$

Taking the supremum over t on both sides of Eq. (30) and applying Young's convolution inequality with $\|H_{\mathrm{human}}\|_{L_1} = 1$, we obtain

$$\sup_t e_{\mathrm{total}}(t) \;\le\; \sup_t e_{\mathrm{map}}(t) + \sup_t e_{\mathrm{tf}}(t) + \sup_t e_{\mathrm{dyn}}(t). \tag{31}$$

Therefore, we suggest a conservative design requirement for accurate haptic feedback is

$$\sup_t e_{\mathrm{total}}(t) \;<\; \mathrm{JND}^* \quad \Leftarrow \quad \sup_t \left(e_{\mathrm{map}}(t) + e_{\mathrm{tf}}(t) + e_{\mathrm{dyn}}(t)\right) \;<\; \mathrm{JND}^*. \tag{32}$$

where $\sup_t$ denotes the supremum of the error over the time interval of interest. To make Eq. (32) actionable, we allocate an error budget

$$\sup_t e_{\mathrm{map}}(t) < \alpha\,\mathrm{JND}^*, \qquad \sup_t e_{\mathrm{tf}}(t) < \beta\,\mathrm{JND}^*, \qquad \sup_t e_{\mathrm{dyn}}(t) < \gamma\,\mathrm{JND}^*, \tag{33}$$

with $\alpha + \beta + \gamma \leq 1$.

In force simulation, most researchers incorporate friction [103], with gravity also being considered, though typically to a lesser extent. Reproducing gravity is vital for enhancing the haptic experience, since perceiving the weight of virtual objects significantly enriches immersion. Conventional haptic devices, however, cannot directly reproduce gravitational forces; instead, users are typically informed of successful grasping through vibration or simple force cues [127]. One possible approach for simulating gravitational effects is to dynamically vary the device mass, for example by circulating liquid to emulate different weight levels [128].

By modulating the device's effective mass, they induced a perceived change in weight for the user. Their subjective evidence assessing perceived realism, comfort, and preference—suggested that this approach can support the simulation of gravity-related sensations.

Nevertheless, such techniques require bulky hardware and incur high costs. Skin-stretch feedback has been explored as a lightweight alternative for representing texture, friction, slip, and tangential forces. By deforming the fingertip skin, skin-stretch devices can effectively simulate tangential components, including gravitational cues. Based on this principle, Choi et al. (2017) designed a wearable haptic device capable of simulating gravity to enhance grasping sensations. The device employs voice-coil actuators driven with an asymmetric waveform, producing a cycle-wise asymmetric displacement/acceleration profile. This temporal asymmetry generates a directionally biased cutaneous cue (shear/skin-stretch) on the fingertip. Defining F_z as the gravitational (downward) component of the virtual contact force, the system maps F_z to the waveform amplitude. The system elicits an illusion of weight and gravity direction, rather than applying a sustained physical downward force [129]. Task-level validity was demonstrated through a VR weight-sorting task, in which users successfully rank-ordered virtual objects of different masses. However, qualitative feedback indicated that weight discrimination remained difficult in practice, with users reporting that vibration amplitude was more salient than perceived weight.

Additionally, fingertip placement and grip force can substantially modulate the perceived asymmetric skin-stretch cue, and it is difficult to keep these factors consistent across users or even across trials for the same user. In addition, inter-individual differences further affect perceptual gain [129]. Therefore, we recommend adopting the general psychophysical procedure of [130] to estimate each participant's JND. Specifically, a two-alternative forced-choice (2AFC) task with an adaptive staircase can be used, and the JND computed as the mean stimulus amplitude at the final N reversal points once a predefined total number of reversals is reached. The cue can then be discretised into a small set of amplitude levels separated by $m \times \mathrm{JND}$. This participant-specific quantisation preserves level-to-level discriminability in the presence of trial-to-trial variations in fingertip placement and grip force that effectively rescale cue transmission gain. Since effective cue transmission gain varies across grasp configurations due to differences in contact area and skin compliance, a grasp-specific correction factor should also be derived from a brief pilot calibration session.

Physics-based approaches can be extended to model the interaction between elastic surfaces and virtual fingertips (see Fig. 10). Luo et al. (2007) applied Bernoulli-Euler bending beam theory to simulate elastic shape deformation [131]. This non-linear model is capable of handling compliant motions with friction and complex variations in contact areas, but it remains limited to contiguous areas. To overcome this limitation, Tong et al. (2008) introduced a non-linear contact force model combined with a beam-skeleton representation for global shape deformation, allowing the modelling of multi-contact regions between the virtual hand and deformable objects [18]. Nevertheless, these physics-based approaches are strongly reliant on the precision of the devices and typically entail high computational requirements. In 2007, position-based approaches have been proposed. In contrast to the explicit force constraints described above, implicit methods can significantly reduce computational overhead. Position-Based Dynamics (PBD) [97] improves simulation stability through direct positional constraint enforcement, allowing large time steps and avoiding force integration. Macklin et al. [132] introduced XPBD, a compliant constraint formulation that enhances stability while enabling physically plausible deformation under interactive workloads. GPU-accelerated implementations [133] have further extended the scalability of position-based methods, and their integration into game engines such as Unity3D has been demonstrated [134].

In a real-world system, selecting among the four solver families requires considering four update loops: rendering, tracking, physics, and haptics.

The rendering loop is typically maintained above 60 FPS to ensure basic visual smoothness. The tracking loop rate is determined by the tracking hardware. For haptics, since human perception is highly sensitive to delay and discretisation, the control loop is commonly designed to run at around 1 kHz [135].

Regarding the physics loop, different solvers impose different trade-offs:

(1) Explicit force-based [136] integration produces physically interpretable forces, but in the presence of high stiffness or hard contact, it often requires small time steps to remain stable, increasing the total computational cost.

(2) Implicit force-based [136, 137] integration can use larger time steps with improved stability, but each step is more expensive because it typically requires solving linear and/or nonlinear systems.

(3) Constraint-based methods [138] can directly yield contact impulses/forces and can run at high rates in practice; however, their stability and performance depend on regularisation and iterative solver settings.

(4) PBD/XPBD [97, 132] methods are designed to remain stable under relatively large time steps and are computationally efficient for real-time applications. Equivalent constraint forces can be estimated from Lagrange multipliers or position corrections, but their physical fidelity depends on model assumptions and calibration.

Table 2 compares four solver classes in terms of their effects on soft contact, frictional slip, and penetration control.

Table 2: Implications of solver selection

Solver	Soft contact	Frictional slip	Penetration control
Force-based (Explicit integration)	Achieving hard contact typically requires high contact stiffness, which increases numerical stiffness and forces smaller time steps or sub-stepping to maintain stability. In contrast, when soft contact is acceptable, this approach is often efficient and robust.	Prone to stick–slip chatter and tangential oscillations unless the time step is small.	Reducing it typically means increasing contact stiffness, which hurts stability.
Force-based (Implicit integration)	Support stiffer contacts than explicit schemes at the same step size, at the cost of heavier per-step solves and potentially increased numerical damping.	Usually less chatter than explicit methods.	Better than explicit with the same model, but still not true non-penetration.
Constraint-based	Hard contact by default; non-penetration enforced as a constraint. Soft contact is achieved via regularisation/compliance.	Friction handled as constraints enables more consistent stick vs. slip, depending on iteration quality.	Near non-penetration is achievable.
PBD/XPBD	Stable and visually clean contact.	Often stable but can be over-sticky or over-damped.	Very strong visually: penetrations are removed by position projection; geometric non-penetration is straightforward to enforce.

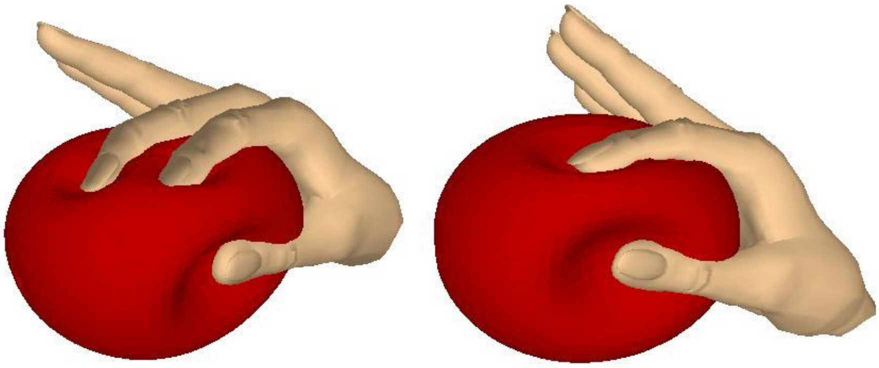

Fig. 10: Illustrates contact force and deformation modelling for haptic simulation of grasping a deformable object with a realistic virtual human hand [18], which runs at an interactive rate and needs efficient methods to be able to run in real-time.

Recent designs have shifted attention toward synchronising the virtual and tracked hands and developing haptic devices to optimise the user experience, since earlier devices often failed to ensure proper coordination, leading to visual disharmony [139]. To address this, Delrieu et al. (2020) proposed an enhanced approach that couples a virtual kinematic hand with the visual hand tracking system, thereby enabling precision grasping, particularly for small objects [17]. More recently, Zhang et al. (2025) improved embodiment by enforcing a 1:1 alignment between the hand, body, and physical environment across users in location-based VR, which significantly enhances immersion and social presence [140].

Beyond tracking, hardware-based haptic feedback has been widely recognised as a promising strategy to optimise user experience [13]. For example, Jain et al. (2016) developed a VR system capable of rendering fluid-related cues such as buoyancy, drag, and temperature changes [141]. Similarly, Liu et al. (2019) employed vibration motors to deliver finger-specific tactile sensations, while Choi et al. (2016) introduced a wearable brake-based device that provides binary force feedback [142]. Although these devices successfully deliver stable haptic cues, they inevitably suffer from limitations—including restricted fidelity, high cost, and limited portability—that continue to hinder their widespread adoption.

4.2.2 Animation-based Methods

In VR grasping, animation-based methods are widely employed. The main categories of such approaches are summarised in 3. Example-based animation can produce compelling visual results during interaction, thereby significantly improving the quality of the experience. However, such approaches are limited to objects with predefined animations [143]. This means that prior knowledge about the grasped objects is required. Nasim et al. (2016) introduced a method for handling complex tasks, by rigidly binding the object to the hand and avoiding jitter or instability during manipulation. In this approach, the virtual hand is deactivated once it collides with the object, a concept similar to "freezing" [144]. Similarly, a strategy involves snapping virtual objects directly to the virtual hand. By predefining grasp anchor points on objects, the system attaches the object to the virtual hand upon grasp detection, without requiring complex collision handling or shape fitting. While this can achieve relatively stable grasping, the resulting interactions tend to be unnatural, limiting dexterous operations and reducing user experience [76].

Most of the aforementioned methods are primarily suited to rigid bodies. For objects undergoing large deformations, however, efficient strategies are still lacking, since reproducing physically plausible and visually convincing deformations in real time remains highly challenging. Some relatively mature deformation techniques have been proposed in the field of computer animation, which may serve as potential guidance for future research on grasping deformable objects. One representative approach is deformation transfer, which employs mesh sequences to represent 3D models. Typically, deformation gradients over triangle meshes are used to transfer deformation information from a source model to a target model [145–147], making this method suitable for complex shapes such as toys and handicrafts. Another widely adopted technique is cage-based deformation. For instance, Chen et al. (2010) proposed a method requiring manual cage construction for each target object [146]. To reduce this burden, Le et al. (2017) presented an optimised approach for automatic cage generation [148], while Sacht et al. (2015) introduced a nested cage-based method with hierarchical levels, making it particularly effective for complex geometry and topology [149]. Cage-based approaches have been demonstrated to optimise natural interactions in VR, and for deformable bodies, they offer a structured way to represent deformations.

Although deformation transfer and cage-based deformation are well established in computer animation and modelling, their application to VR grasping remains largely unexplored. Nevertheless, these methods could potentially inspire future approaches for representing complex deformations of grasped objects at lower computational cost compared to physics-based methods.

4.2.3 Data-Driven Methods

Physics-based simulations can preserve plausibility by simulating interaction forces. However, the main challenge lies in designing effective controllers that can both replicate natural human-like motions and maintain stable physical interactions [143].

Early approaches also relied on rule-based methods, where grasping rules were pre-defined for basic geometric primitives and extended to other shapes. Although such methods can be easily applied to large collections of objects, they often lack generalisation to complex geometries and tend to produce unnatural hand poses.

To overcome the limitations of rule-based methods, researchers have shifted toward data-driven approaches, with a representative strategy being the construction of grasp databases. The most direct way to construct such a database was to collect demonstrations from human volunteers, typically recorded using digital gloves or optical motion capture systems to capture finger joint angles, contact points, and hand configurations with precision. For example, Liu et al. (2024) introduce a reconfigurable data glove that can automatically record finger joint angles and contact points in both physical and virtual environments, providing data that can subsequently be used to drive grasp synthesis and analysis [150]. Subsequently, various techniques were explored to

Table 3: Animation-based methods

Method Category	Representative Studies	Application Objects	Advantages	Limitations
Animation-based grasping				
Example-based animation	Pollard et al. (2005)	Predefined objects	Realistic visual effects, natural operations	Limited to captured examples, lacks explicit force control, low generalisation capability
Snap-to-hand	Blomgren et al. (2021)	Rigid bodies	Simple implementation, low computational overhead, minimal user effort	Lacks physical realism, restricts immersive manipulation, constrained to predefined grasp points
Freeze-based grasping	Nasim et al. (2016)	Complex tasks	Avoids penetration, jitter, or instability	Unnatural interaction, reduced immersion, limited finger readjustment
Animation-driven deformation techniques				
Deformation transfer	Müller et al. (2005), Chen et al. (2010), Casti et al. (2019)	Deformable bodies (e.g., toys, handicrafts)	Ability to transfer complex geometric deformations	Requires mesh correspondence, limited to similar topology
Cage-based deformation	Chen (2010), Le (2017), Sacht (2015)	Complex geometry and topology	Low computational cost, intuitive control	Manual or semi-automatic cage construction, less flexible for dynamic interaction

generalise downstream grasp synthesis. Li et al. (2005) introduce a shape-matching algorithm that adapts hand configurations to novel objects by identifying collections of features based on similar relative placements and surface normals [151, 152] (see Fig. 11). Tian et al. (2019) propose a real-time virtual grasping algorithm that leverages machine learning and optimisation techniques to compute a learned grasp space. This grasp space is precomputed using support vector machines and represents a set of stable grasp configurations, which can be used at runtime to generate plausible grasping motions [19]. More recently, Zhang et al. (2024) propose a diffusion-based framework for whole-body grasping sequence generation, trained on the GRAB [153] and ARCTIC datasets [154], which is capable of synthesising natural, stable, and non-penetrating bimanual grasping motions [155]. The aforementioned representative approaches are summarised in Table 4.

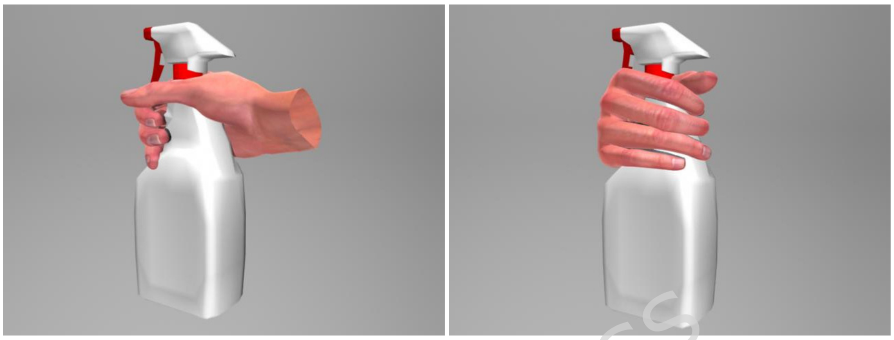

Fig. 11: Shows grasping of a complex virtual object, where the virtual hand matches the object geometry based on pre-defined poses [151].

Data-driven approaches for virtual grasping have so far received relatively limited attention in the scientific literature. While such methods can generate visually natural grasping motions, their construction often incurs higher costs compared to physics-based or purely animation-based techniques. Consequently, current industrial applications still predominantly rely on physics- and animation-driven methods [116]. Nevertheless, data-driven approaches demonstrate strong potential and are likely to become increasingly competitive as more efficient data acquisition and learning techniques are developed.

4.2.4 Collision Detection

Regardless of the method employed, collision detection is indispensable in virtual grasping. Collision detection is the essential step toward achieving realistic grasping, referring to the determination of intersections between two or more spatial targets. From a visual perspective, once a collision between the virtual hand and object is detected, a grasping event is triggered [69]. From a haptic perspective, collision detection streams real-time data such as position, orientation, and joint angles to the haptic rendering module, which utilises them to generate appropriate grasping events and

Table 4: Representative Downstream Grasp Synthesis Techniques

Approach	**Representative Studies**	**Advantages**	**Limitations**
Shape matching	Li et al. (2005, 2007)	Targets grasp pose synthesis; adapts learned grasps to previously unseen objects	May fail on complex geometries
Learning-based grasp space	Tian et al. (2019)	Oriented toward grasp planning with the objective of identifying optimal grasp configurations; achieves real-time performance while producing physically plausible grasps	Limited to trained configurations
Diffusion-based synthesis	Zhang et al. (2024)	Extends diffusion-based generation to whole-body grasp sequences, synthesising stable, natural, and non-penetrating motions	Requires large-scale training datasets

feedback. This process enables users to perceive forces, material properties, and shapes during virtual object manipulation [156]. Generally, the collision detection algorithm can be described as a two-stage process: (1) Broad phase: rapidly identifying potentially colliding object pairs using simplified representations (e.g., bounding boxes, bounding spheres, or spatial partitioning). (2) Narrow phase: performing precise intersection tests between the actual geometric primitives (e.g., meshes) of the candidate pairs. Once a collision is detected, the algorithm computes the contact information, including the intersection point, normal vector, and penetration depth, which are subsequently used to trigger grasping or physical simulation [157].

Specifically, the procedure that describes how collision detection is responded to, in the context of grasping, can be formalised as presented in Algorithm 1.

Algorithm 1 Collision Detection Response Algorithm

Inputs: Virtual hand model $\mathbf{H}_i$ with colliders, VR objects $\mathbf{O}_i$ with velocities $\mathbf{v}_i$
Initialize: Collision set $\mathcal{K} \leftarrow \emptyset$
while simulation running **do**
  for each object $\mathbf{O}_i$ in scene **do**
    for virtual hand model $\mathbf{H}_i$ **do**
      if collision detected$(\mathbf{H}, \mathbf{O}_i)$ **then**
        Compute contact point $\mathbf{c}$, normal $\mathbf{n}$, time-of-impact $\mathbf{t}$ and penetration depth pd
        $\mathcal{K} \leftarrow \mathcal{K} \cup \{(c, n, \mathrm{pd}, t)\}$
      end if
    end for
  end for
  Output $\mathcal{K}$ {used to trigger grasp events, haptic cues, or physics responses}
end while

Although many existing game engines provide built-in collision detection algorithms, the complexity of virtual grasping scenarios demands more stable and efficient solutions tailored to VR (see Fig. 12).

Wang et al. (2023) proposed an improved collision detection approach based on mesh simplification and particle swarm optimisation, which reduces the search space and improves efficiency [158]. Macklin et al. (2014) introduced a particle-based method that handles both contacts and collisions in a unified framework [159]. In contrast, Höll et al. (2018) applied a simple ray-based technique derived from the classical Coulomb contact model, although it was limited to unconstrained hand–object interactions [116].

Although detecting rigid-body contacts is well-established, interactions with deformable objects are more complex due to continuous changes in geometry during manipulation [134]. Zou et al. (2016) proposed replacing the traditional virtual plane with volumetric data as the basis for collision detection, recording the initial collision point and tracking proxy point displacement in real time to calculate soft tissue deformation and force feedback, thereby avoiding the erroneous feedback output that occurs when surgical instruments penetrate soft tissue in conventional methods [160]. Fazioli et al. (2016) proposed to mitigate the object penetration issue inherent in the default soft-rigid collision detection of Bullet Physics by integrating a ray casting-based hybrid collision detection algorithm into the engine, extending it from rectangular rigid bodies to arbitrary convex geometries. In addition, a Hunt-Crossley nonlinear viscoelastic model was adopted to compute the haptic force feedback [161].

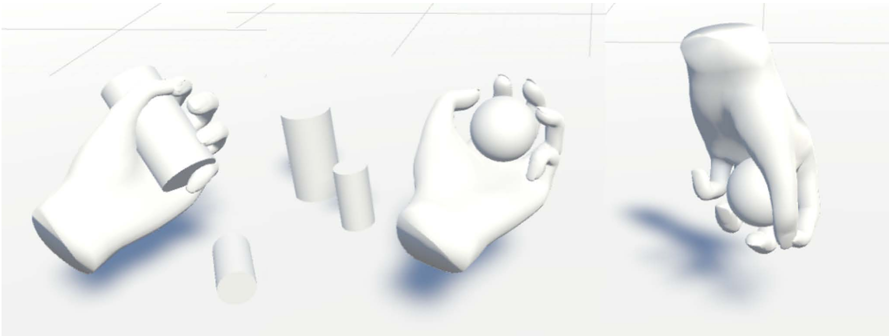

Fig. 12: Shows grasping of handheld objects using geometric colliders, which employs GPU PhysX to detect collisions [162].

5 Discussion

The purpose of virtual grasping is to simulate real-world interactions and provide users with a realistic virtual experience, which also includes scenarios that do not exist in physical reality, or are difficult to realise in daily life. In virtual reality, visual feedback is primarily used to ensure stable grasping. Even haptic feedback is often designed to enhance the plausibility of interactions within visual scenes, as no lightweight device is yet capable of reproducing the full richness of human haptic sensations [163]. Importantly, techniques from computer vision (CV), particularly tracking, are indispensable to virtual grasping, providing the essential link that aligns real-world hand motions with their virtual representations. Tracking refers to the continuous estimation of the position and orientation of an object [164], and recent years have witnessed significant breakthroughs in hand tracking and hybrid tracking [165–167]. However, due to the high number of degrees of freedom of the human hand and frequent occlusions, robust hand tracking remains challenging [168]. Beyond this, in terms of deployment, cloud-based tracking enhances multi-device coordination and supports the efficient processing of complex models [169]; in terms of tracking targets, the introduction of eye tracking technology further expands the interactive possibilities of VR [170]. Specifically, gaze-based methods can effectively reduce the user's operational effort in virtual grasping by eliminating the need for conventional scaling techniques, which are often associated with hand fatigue and unnatural control [171].

Table 5: Evaluation dimensions and typical standards for robotic grasping

Evaluation Dimension	**Typical Standards in Robotic Grasping**
(i) Stability	Force-closure / Form-closure; GWS ϵ-metric; Drop rate / Slip rate
(ii) Plausibility	Similarity of hand posture to human grasps; Task-appropriate grasp type; Comparison against grasping databases
(iii) Efficiency	Grasp planning time; Execution completion time; Task success rate

Although both robotic grasping and VR grasping aim to enable dexterous object manipulation, they differ fundamentally in their interaction medium and sensing–actuation loop. Compared to mechanical grippers, which physically interact with real objects and directly transmit reaction forces through rigid or compliant mechanisms [172], the haptic device relies on collision detection in a virtual environment, coupled with force, vibrotactile, or thermal actuators to approximate the grasping sensation [173]. Such devices involve a trade-off, with some designs offering high force fidelity and bandwidth, while others prioritise immersion and spatial freedom [174].

Further, mechanical grippers measure and regulate contact force in real time via embedded sensors, achieving stable force control even under dynamic loading [26]. In contrast, the fidelity of VR grasping interactions depends heavily on tracking precision and modelling fidelity to produce plausible visual results. Meanwhile, for haptics, the physics engine is required to compute computation of contact forces and render them through the haptic device's actuators, which have a limited dynamic range. This inevitably introduces reduced peak force capability and a mismatch between visual deformation and tactile sensation [175].

In addition, no efficient and robust strategy for VR grasping has been established, and all existing approaches whether physics-based, animation-driven, or data-driven suffer from inherent limitations. Moreover, from the perspective of grasping outcomes, we suggest that VR grasping should adopt evaluation criteria analogous to those used in robotic grasping, namely:

1. is the grasp stable?
2. is the grasp plausible?
3. can the grasp be completed efficiently?

As summarised in Table 5, these criteria are particularly important in haptic-enabled teleoperation and domain-specific training scenarios such as surgical simulation, where stability, plausibility, and efficiency determine task success.

6 Conclusion

In this paper, we reviewed existing grasping techniques in both robotics and VR. Through comparison, we suggest that VR grasping should draw on relatively mature practices from robotics. In particular, VR grasping also requires a mathematical framework–similar to the GWS used in robotics–that abstracts grasping behaviour into contact points and force/torque representations, thereby enabling precise evaluation of grasp stability. Furthermore, findings from contact modelling in robotics, such as contact patch geometry and torsional friction, can provide a theoretical foundation for soft hand simulation in VR, enabling more realistic grasping interactions. While VR grasping has so far emphasised immersion, with simple symbolic grasping often sufficient for most users, the introduction of rigorous evaluation metrics can enhance physical plausibility and help avoid unrealistic aspects, which is especially important in domains such as medical training and engineering simulation, where accurate physical behaviour is critical. Current studies rely largely on users' subjective impressions, including aspects such as realism and immersion. A promising future direction is to develop a hybrid evaluation framework that combines widely accepted mathematical evaluation metrics with qualitative measures. Importantly, such a framework should evolve beyond rigid-body assumptions and be extended to deformable bodies: this remains one of the main challenges to address for researchers, since local deformation, energy and material properties need to be taken into account, especially as the distribution of contact points is dynamic. In addition, for certain application domains, such as gaming, even fluid-like objects may need to be considered.

Currently, the VR grasping in virtual environments continues to face fundamental challenges in both visuals and haptics. Immersive grasping requires haptic devices with high tracking accuracy, low latency and lightweight design, avoiding penetration grasping simulation. The future of VR grasping should not be limited to visual outcomes but instead focus on unifying visual and haptic rendering into a coherent multi-sensory framework, while ensuring scalability as environments and interaction tasks become increasingly complex. Consequently, addressing these challenges will demand cross-disciplinary advances across materials science, robotics, kinematics, computer vision, and the biology of human behaviour.

Declarations

The authors declare no financial interests/personal relationships which may be considered as potential competing interests. This research project did not receive funding from any agency.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors consent for the publication of this manuscript.

Funding

Not applicable.

Authors contributions

Mingzhao Zhou prepared the first draft of this paper. The authors contributed equally to all remaining aspect of this study.

Acknowledgements

Not applicable.

References

[1] Qi, J., Gao, F., Sun, G., Yeo, J.C., Lim, C.T.: Haptglove—untethered pneumatic glove for multimode haptic feedback in reality–virtuality continuum. Advanced Science **10**(25), 2301044 (2023) https://doi.org/10.1002/advs.202301044

[2] Liu, Y., Jiang, L., Liu, H., Ming, D.: A systematic analysis of hand movement functionality: Qualitative classification and quantitative investigation of hand grasp behavior. Frontiers in Neurorobotics **15**, 658075 (2021) https://doi.org/10.3389/fnbot.2021.658075

[3] Thrun, S.: Probabilistic robotics. Commun. ACM **45**(3), 52–57 (2002) https://doi.org/10.1145/504729.504754

[4] Zheng, W., Guo, N., Zhang, B., Zhou, J., Tian, G., Xiong, Y.: Human grasp mechanism understanding, human-inspired grasp control and robotic grasping planning for agricultural robots. Sensors **22**(14), 5240 (2022) https://doi.org/10.3390/s22145240

[5] Deng, Z., Jonetzko, Y., Zhang, L., Zhang, J.: Grasping force control of multi-fingered robotic hands through tactile sensing for object stabilization. Sensors **20**(4) (2020) https://doi.org/10.3390/s20041050

[6] Li, T., Yan, Y., Yu, C., An, J., Wang, Y., Chen, G.: A comprehensive review of robot intelligent grasping based on tactile perception. Robotics and Computer-Integrated Manufacturing **90**, 102792 (2024) https://doi.org/10.1016/j.rcim.2024.102792

[7] Verschoor, M., Lobo, D., Otaduy, M.A.: Soft hand simulation for smooth and robust natural interaction. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 183–190 (2018). https://doi.org/10.1109/VR.2018.8447555

[8] Shi, Y., Zhao, L., Lu, X., Hoang, T., Wang, M.: Grasping 3d objects with virtual hand in vr environment. In: Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. VRCAI '22, pp. 8–188. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3574131.3574428 . https://doi.org/10.1145/3574131.3574428

[9] Chessa, M., Maiello, G., Klein, L.K., Paulun, V.C., Solari, F.: Grasping objects in immersive virtual reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 1749–1754 (2019). https://doi.org/10.1109/VR.2019.8798155

[10] Mizuho, T., Narumi, T., Kuzuoka, H.: Effects of the visual fidelity of virtual environments on presence, context-dependent forgetting, and source-monitoring

error. IEEE Transactions on Visualization and Computer Graphics **29**(5), 2607–2614 (2023) https://doi.org/10.1109/TVCG.2023.3247063

[11] Lim, J., Choi, Y.: Force-feedback haptic device for representation of tugs in virtual reality. Electronics **11**(11) (2022) https://doi.org/10.3390/electronics11111730

[12] WANG, D., GUO, Y., LIU, S., ZHANG, Y., XU, W., XIAO, J.: Haptic display for virtual reality: progress and challenges. Virtual Reality & Intelligent Hardware **1**(2), 136–162 (2019) https://doi.org/10.3724/SP.J.2096-5796.2019.0008 . Haptic Interaction

[13] Liu, H., Zhang, Z., Xie, X., Zhu, Y., Liu, Y., Wang, Y., Zhu, S.-C.: High-fidelity grasping in virtual reality using a glove-based system. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 5180–5186 (2019). https://doi.org/10.1109/ICRA.2019.8794230

[14] Choi, I., Ofek, E., Benko, H., Sinclair, M., Holz, C.: Claw: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, pp. 1–13. Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3174228 . https://doi.org/10.1145/3173574.3174228

[15] Tong, Q., Wei, W., Zhang, Y., Xiao, J., Wang, D.: Survey on hand-based haptic interaction for virtual reality. IEEE Transactions on Haptics **16**(2), 154–170 (2023) https://doi.org/10.1109/TOH.2023.3266199

[16] Najdovski, Z., Nahavandi, S.: Extending haptic device capability for 3d virtual grasping. In: Ferre, M. (ed.) Haptics: Perception, Devices and Scenarios, pp. 494–503. Springer, Berlin, Heidelberg (2008)

[17] DELRIEU, T., Weistroffer, V., Gazeau, J.P.: Precise and realistic grasping and manipulation in virtual reality without force feedback. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 266–274 (2020). https://doi.org/10.1109/VR46266.2020.00046

[18] Cui, T., Xiao, J., Song, A.: Simulation of grasping deformable objects with a virtual human hand. In: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3965–3970 (2008). https://doi.org/10.1109/IROS.2008.4651080

[19] Tian, H., Wang, C., Manocha, D., Zhang, X.: Realtime hand-object interaction using learned grasp space for virtual environments. IEEE Transactions on Visualization and Computer Graphics **25**(8), 2623–2635 (2019) https://doi.org/10.1109/TVCG.2018.2849381

[20] Moher, D., Liberati, A., Tetzlaff, J., Altman, D.: Group, tp, & the prisma group.(2009). preferred reporting items for systematic reviews and meta-analyses: The prisma statement. PLoS medicine **6**(7), 1000097

[21] Kleeberger, K., Bormann, R., Kraus, W., Huber, M.F.: A survey on learning-based robotic grasping. Current Robotics Reports **1**(4), 239–249 (2020) https://doi.org/10.1007/s43154-020-00021-6

[22] Du, G., Wang, K., Lian, S., Zhao, K.: Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. Artificial Intelligence Review **54**(3), 1677–1734 (2021) https://doi.org/10.1007/s10462-020-09888-5

[23] Bohg, J., Morales, A., Asfour, T., Kragic, D.: Data-driven grasp synthesis—a survey. IEEE Transactions on Robotics **30**(2), 289–309 (2014) https://doi.org/10.1109/TRO.2013.2289018

[24] Sahbani, A., El-Khoury, S., Bidaud, P.: An overview of 3d object grasp synthesis algorithms. Robotics and Autonomous Systems **60**(3), 326–336 (2012) https://doi.org/10.1016/j.robot.2011.07.016 . Autonomous Grasping

[25] Murray, R.M., Sastry, S., Li, Z.: A mathematical introduction to robotic manipulation. (1994). https://api.semanticscholar.org/CorpusID:108605633

[26] Bicchi, A., Kumar, V.: Robotic grasping and contact: a review. In: Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065), vol. 1, pp. 348–3531 (2000). https://doi.org/10.1109/ROBOT.2000.844081

[27] Bicchi, A.: On the closure properties of robotic grasping. The International Journal of Robotics Research **14**(4), 319–334 (1995) https://doi.org/10.1177/027836499501400402

[28] Rimon, E., Burdick, J.W.: On force and form closure for multiple finger grasps. In: Proceedings of the 1996 IEEE International Conference on Robotics and Automation (ICRA), Minneapolis, Minnesota, pp. 1795–1800 (1996). https://doi.org/10.1109/ROBOT.1996.506972

[29] Prattichizzo, D., Trinkle, J.C.: Grasping. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, 2nd edn., pp. 955–988. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32552-1_38

[30] Nguyen, V.-D.: Constructing force-closure grasps. In: Proceedings. 1986 IEEE International Conference on Robotics and Automation, vol. 3, pp. 1368–1373 (1986). https://doi.org/10.1109/ROBOT.1986.1087483

[31] Dizioğlu, B., Lakshiminarayana, K.: Mechanics of form closure. Acta Mechanica

52(1), 107–118 (1984) https://doi.org/10.1007/BF01175968

[32] Roa, M.A., Suárez, R.: Grasp quality measures: review and performance. Autonomous Robots **38**(1), 65–88 (2015) https://doi.org/10.1007/s10514-014-9402-3

[33] Ferrari, C., Canny, J.: Planning optimal grasps. In: Proceedings 1992 IEEE International Conference on Robotics and Automation, pp. 2290–22953 (1992). https://doi.org/10.1109/ROBOT.1992.219918

[34] Su, J., Qiao, H., Liu, C., Song, Y., Yang, A.: Grasping objects: The relationship between the cage and the form-closure grasp. IEEE Robotics & Automation Magazine **24**(3), 84–96 (2017) https://doi.org/10.1109/MRA.2016.2615332

[35] Wan, W., Fukui, R., Shimosaka, M., Sato, T., Kuniyoshi, Y.: Grasping by caging: A promising tool to deal with uncertainty. In: 2012 IEEE International Conference on Robotics and Automation, pp. 5142–5149 (2012). https://doi.org/10.1109/ICRA.2012.6224676

[36] Kim, D., Maeda, Y., Komiyama, S.: Caging-based grasping of deformable objects for geometry-based robotic manipulation. ROBOMECH Journal **6**(1), 3 (2019) https://doi.org/10.1186/s40648-019-0131-4

[37] Laguna, M., Martí, R.: Heuristics. In: Gass, S.I., Fu, M.C. (eds.) Encyclopedia of Operations Research and Management Science, pp. 695–703. Springer, Boston, MA (2013). https://doi.org/10.1007/978-1-4419-1153-7_1184 . https://doi.org/10.1007/978-1-4419-1153-7_1184

[38] Miller, E., Durner, M., Humt, M., Quere, G., Boerdijk, W., Sundaram, A.M., Stulp, F., Vogel, J.: Unknown object grasping for assistive robotics. 2024 IEEE International Conference on Robotics and Automation (ICRA), 18157–18163 (2024)

[39] Saxena, A., Driemeyer, J., Ng, A.Y.: Robotic grasping of novel objects using vision. The International Journal of Robotics Research **27**(2), 157–173 (2008) https://doi.org/10.1177/0278364907087172

[40] Lenz, I., Lee, H., Saxena, A.: Deep learning for detecting robotic grasps. The International Journal of Robotics Research **34**(4-5), 705–724 (2015) https://doi.org/10.1177/0278364914549607

[41] Redmon, J., Angelova, A.: Real-time grasp detection using convolutional neural networks. In: 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 1316–1322 (2015). IEEE

[42] Mousavian, A., Eppner, C., Fox, D.: 6-dof graspnet: Variational grasp generation for object manipulation. 2019 IEEE/CVF International Conference on

Computer Vision (ICCV), 2901–2910 (2019)

[43] Yang, Y., Yu, H., Lou, X., Liu, Y., Choi, C.: Attribute-based robotic grasping with data-efficient adaptation. IEEE Transactions on Robotics **40**, 1566–1579 (2025)

[44] Wang, K., Fan, Y., Sakuma, I.: Robot grasp planning: A learning from demonstration-based approach. Sensors **24**(2) (2024) https://doi.org/10.3390/s24020618

[45] Ye, J., Wang, J., Huang, B., Qin, Y., Wang, X.: Learning continuous grasping function with a dexterous hand from human demonstrations. IEEE Robotics and Automation Letters **8**, 2882–2889 (2022)

[46] Wang, P., Manhardt, F., Minciullo, L., Garattoni, L., Meie, S., Navab, N., Busam, B.: Demograsp: Few-shot learning for robotic grasping with human demonstration. 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 5733–5740 (2021)

[47] Shi, L., Liu, Y., Zeng, L., Ai, B., Hong, Z., Su, H.: Learning adaptive dexterous grasping from single demonstrations. 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 9456–9463 (2025)

[48] Rohling, R.N., Hollerbach, J.M.: Optimized fingertip mapping for teleoperation of dextrous robot hands. In: [1993] Proceedings IEEE International Conference on Robotics and Automation, pp. 769–7753 (1993). https://doi.org/10.1109/ROBOT.1993.292238

[49] Qi, L., Popoola, O., Imran, M.A., Ahmad, W.: Genhand: generalised human grasp kinematic retargeting. npj Robotics **4**(1), 19 (2026) https://doi.org/10.1038/s44182-026-00076-1

[50] Zhao, H., Khadiv, M.: Trajectory optimization under contact timing uncertainties. In: 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids), pp. 1064–1071 (2024). https://doi.org/10.1109/Humanoids58906.2024.10769861

[51] Tang, M., Manocha, D., Otaduy, M.A., Tong, R.: Continuous penalty forces. ACM Trans. Graph. **31**(4) (2012) https://doi.org/10.1145/2185520.2185603

[52] Yin, Z.-H., Wang, C., Pineda, L., Bodduluri, K., Wu, T., Abbeel, P., Mukadam, M.: Geometric retargeting: A principled, ultrafast neural hand retargeting algorithm. In: 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 17376–17382 (2025). https://doi.org/10.1109/IROS60139.2025.11247700

[53] Zhu, F., Wang, L., Wen, Y., Yang, L., Pan, J., Wang, Z., Wang, W.: Failure

handling of robotic pick and place tasks with multimodal cues under partial object occlusion. Frontiers in Neurorobotics **15**, 570507 (2021) https://doi.org/10.3389/fnbot.2021.570507

[54] Garg, N.P., Hsu, D., Lee, W.S.: Learning to grasp under uncertainty using pomdps. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 2751–2757 (2019). https://doi.org/10.1109/ICRA.2019.8793818

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

[56] Liu, T., Liu, Z., Jiao, Z., Zhu, Y., Zhu, S.-C.: Synthesizing diverse and physically stable grasps with arbitrary hand structures using differentiable force closure estimator. IEEE Robotics and Automation Letters **7**(1), 470–477 (2022) https://doi.org/10.1109/LRA.2021.3129138

[57] Garcia, G.J., Corrales, J.A., Pomares, J., Torres, F.: Survey of visual and force/tactile control of robots for physical interaction in spain. Sensors **9**(12), 9689–9733 (2009) https://doi.org/10.3390/s91209689

[58] Zhao, C., Yu, Y., Ye, Z., Tian, Z., Zhang, Y., Zeng, L.L.: Universal slip detection of robotic hand with tactile sensing. Frontiers in Neurorobotics **19**, 1478758 (2025) https://doi.org/10.3389/fnbot.2025.1478758

[59] Zhang, B., Han, Q., Zhang, J., Han, Z., Niu, S., Ren, L.: Advanced bio-inspired structural materials: Local properties determine overall performance. Materials Today **41**, 177–199 (2020) https://doi.org/10.1016/j.mattod.2020.04.009

[60] Yu, X., Huang, R., Zhao, C., Zhou, L., Ou, L.: Def-grasp: A robot grasping detection method for deformable objects without force sensor. Neural Processing Letters **55**(8), 11739–11756 (2023) https://doi.org/10.1007/s11063-023-11398-8

[61] Zaidi, L., Corrales, J.A., Bouzgarrou, B.C., Mezouar, Y., Sabourin, L.: Model-based strategy for grasping 3d deformable objects using a multi-fingered robotic hand. Robotics and Autonomous Systems **95**, 196–206 (2017) https://doi.org/10.1016/j.robot.2017.06.011

[62] Xu, J., Danielczuk, M., Ichnowski, J., Mahler, J., Steinbach, E., Goldberg, K.: Minimal work: A grasp quality metric for deformable hollow objects. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 1546–1552 (2020). https://doi.org/10.1109/ICRA40945.2020.9197062

[63] Adachi, R., Fujihira, Y., Watanabe, T.: Identification of danger state for grasping delicate tofu with fingertips containing viscoelastic fluid. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 497–503

(2015). https://doi.org/10.1109/IROS.2015.7353418

[64] Nishimura, T., Fujihira, Y., Adachi, R., Watanabe, T.: New condition for tofu stable grasping with fluid fingertips. In: 2016 IEEE International Conference on Automation Science and Engineering (CASE), pp. 335–341 (2016). https://doi.org/10.1109/COASE.2016.7743425

[65] Van Nguyen, P., Luu, Q.K., Takamura, Y., Ho, V.A.: Wet adhesion of micro-patterned interfaces for stable grasping of deformable objects. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 9213–9219 (2020). https://doi.org/10.1109/IROS45743.2020.9341095

[66] Geng, Y., Wang, X., Nishizawa, W., Negreiro, N., Lin, Y.-J., Atlas, C., Liu-Dujardin, A., Schliemann, F., Gunderson, A., Crisp, R., Hu, D., Lanman, D., Abrash, M.: Hyperrealistic vr: A 90-ppd, 1400-nit, high-contrast headset. In: ACM SIGGRAPH 2025 Emerging Technologies. SIGGRAPH '25, pp. 8–182. Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3721257.3734037 . https://doi.org/10.1145/3721257.3734037

[67] Hornsby, J.H., Arnwine, A., Stowe, A., Lin, M., Mason, C., Eakins, M., Veller, D., Mendoza, C., Metcalf, D., Oriaifo, V., Dunar, A., Crockett, B.: Vr haptics for glovebox operations training. Frontiers in Virtual Reality **Volume 6 - 2025** (2025) https://doi.org/10.3389/frvir.2025.1560713

[68] Shi, K., Chen, T., Xiang, Y., Li, Y., Zhu, L., Song, A.: igripper: A semi-active handheld haptic vr controller based on variable stiffness mechanism. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. CHI '25, pp. 813–181315. Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3706598.3714291 . https://doi.org/10.1145/3706598.3714291

[69] Borst, C.W., Indugula, A.P.: Realistic virtual grasping. In: IEEE Proceedings. VR 2005. Virtual Reality, 2005., pp. 91–98 (2005). https://doi.org/10.1109/VR.2005.1492758

[70] Zhao, W., Zhang, J., Min, J., Chai, J.: Robust realtime physics-based motion control for human grasping. ACM Trans. Graph. **32**(6) (2013) https://doi.org/10.1145/2508363.2508412

[71] Dalia Blaga, A., Frutos-Pascual, M., Creed, C., Williams, I.: A grasp on reality: Understanding grasping patterns for object interaction in real and virtual environments. In: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 391–396 (2021). https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00090

[72] Bartalucci, L., Secciani, N., Brogi, C., Topini, A., Della Valle, A., Ridolfi, A., Allotta, B.: An original mechatronic design of a kinaesthetic hand exoskeleton

for virtual reality-based applications. Mechatronics **90**, 102947 (2023) https://doi.org/10.1016/j.mechatronics.2023.102947

[73] Milstein, A., Alyagon, L., Nisky, I.: Grip force control during virtual interaction with deformable and rigid objects via a haptic gripper. IEEE Transactions on Haptics **14**(3), 564–576 (2021) https://doi.org/10.1109/TOH.2021.3060507

[74] Wang, Z., Fratarcangeli, M., Ruimi, A., Srinivasa, A.R.: Real time simulation of inextensible surgical thread using a kirchhoff rod model with force output for haptic feedback applications. International Journal of Solids and Structures **113-114**, 192–208 (2017) https://doi.org/10.1016/j.ijsolstr.2017.02.017

[75] Berenson, D.: Manipulation of deformable objects without modeling and simulating deformation. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4525–4532 (2013). https://doi.org/10.1109/IROS.2013.6697007

[76] Oprea, S., Martinez-Gonzalez, P., Garcia-Garcia, A., Castro-Vargas, J.A., Orts-Escolano, S., Garcia-Rodriguez, J.: A visually realistic grasping system for object manipulation and interaction in virtual reality environments. Computers & Graphics **83**, 77–86 (2019) https://doi.org/10.1016/j.cag.2019.07.003

[77] Kyota, F., Saito, S.: Fast grasp synthesis for various shaped objects. Computer Graphics Forum **31**(2pt4), 765–774 (2012) https://doi.org/10.1111/j.1467-8659.2012.03035.x

[78] Mangalam, M., Oruganti, S., Buckingham, G., Borst, C.W.: Enhancing hand-object interactions in virtual reality for precision manual tasks. Virtual Reality **28**(4), 166 (2024) https://doi.org/10.1007/s10055-024-01055-3

[79] Burton, T.M.W., Vaidyanathan, R., Burgess, S.C., Turton, A.J., Melhuish, C.: Development of a parametric kinematic model of the human hand and a novel robotic exoskeleton. In: 2011 IEEE International Conference on Rehabilitation Robotics, pp. 1–7 (2011). https://doi.org/10.1109/ICORR.2011.5975344

[80] Birouaş, F.I., Ţarcă, R.C., Dzitac, S., Dzitac, I.: Preliminary results in testing of a novel asymmetric underactuated robotic hand exoskeleton for motor impairment rehabilitation. Symmetry **12**(9) (2020) https://doi.org/10.3390/sym12091470

[81] Mulatto, S., Formaglio, A., Malvezzi, M., Prattichizzo, D.: Using postural synergies to animate a low-dimensional hand avatar in haptic simulation. IEEE Transactions on Haptics **6**(1), 106–116 (2013) https://doi.org/10.1109/TOH.2012.13

[82] Gabardi, M., Solazzi, M., Leonardis, D., Frisoli, A.: Design and evaluation of a novel 5 dof underactuated thumb-exoskeleton. IEEE Robotics and Automation

Letters **3**(3), 2322–2329 (2018) https://doi.org/10.1109/LRA.2018.2807580

[83] Aristidou, A.: Hand tracking with physiological constraints. The Visual Computer **34**, 213–228 (2018) https://doi.org/10.1007/s00371-016-1327-8

[84] Kuch, J.J., Huang, T.S.: Vision based hand modeling and tracking for virtual teleconferencing and telecollaboration. In: Proceedings of IEEE International Conference on Computer Vision, pp. 666–671 (1995). https://doi.org/10.1109/ICCV.1995.466875

[85] Karunratanakul, K., Prokudin, S., Hilliges, O., Tang, S.: Harp: Personalized hand reconstruction from a monocular rgb video. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 12802–12813 (2022)

[86] Jacobs, J., Froehlich, B.: A soft hand model for physically-based manipulation of virtual objects. In: 2011 IEEE Virtual Reality Conference, pp. 11–18 (2011). https://doi.org/10.1109/VR.2011.5759430

[87] Jacobs, J., Stengel, M., Froehlich, B.: A generalized god-object method for plausible finger-based interactions in virtual environments. In: 2012 IEEE Symposium on 3D User Interfaces (3DUI), pp. 43–51 (2012). https://doi.org/10.1109/3DUI.2012.6184183

[88] Rumman, N.A., Schaerf, M., Bechmann, D.: Collision detection for articulated deformable characters. In: Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games. MIG '15, pp. 215–220. Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2822013.2822034 . https://doi.org/10.1145/2822013.2822034

[89] Ciocarlie, M., Lackner, C., Allen, P.: Soft finger model with adaptive contact geometry for grasping and manipulation tasks. In: Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07), pp. 219–224 (2007). https://doi.org/10.1109/WHC.2007.103

[90] Prachyabrued, M., Borst, C.W.: Virtual grasp release method and evaluation. Int. J. Hum.-Comput. Stud. **70**(11), 828–848 (2012) https://doi.org/10.1016/j.ijhcs.2012.06.002

[91] Boulic, R., Rezzonico, S., Thalmann, D.: Multi-finger manipulation of virtual objects. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology. VRST '96, pp. 67–74. Association for Computing Machinery, New York, NY, USA (1996). https://doi.org/10.1145/3304181.3304195 . https://doi.org/10.1145/3304181.3304195

[92] Prachyabrued, M., Borst, C.W.: Design and evaluation of visual interpenetration cues in virtual grasping. IEEE Transactions on Visualization and

Computer Graphics **22**(6), 1718–1731 (2016) https://doi.org/10.1109/TVCG.2015.2456917

[93] Pelletier-Guénette, J., Mercier-Aubin, A., Andrews, S.: Real-time triangle-sdf continuous collision detection. Proc. ACM Comput. Graph. Interact. Tech. **8**(4) (2025) https://doi.org/10.1145/3747862

[94] Tonge, R., Benevolenski, F., Voroshilov, A.: Mass splitting for jitter-free parallel rigid body simulation. ACM Trans. Graph. **31**(4) (2012) https://doi.org/10.1145/2185520.2185601

[95] Stewart, D., Trinkle, J.C.: An implicit time-stepping scheme for rigid body dynamics with coulomb friction. In: Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065), vol. 1, pp. 162–1691 (2000). https://doi.org/10.1109/ROBOT.2000.844054

[96] Castro, A.M., Qu, A., Kuppuswamy, N., Alspach, A., Sherman, M.: A transition-aware method for the simulation of compliant contact with regularized friction. IEEE Robotics and Automation Letters **5**(2), 1859–1866 (2020) https://doi.org/10.1109/LRA.2020.2969933

[97] Müller, M., Heidelberger, B., Hennix, M., Ratcliff, J.: Position based dynamics. Journal of Visual Communication and Image Representation **18**(2), 109–118 (2007) https://doi.org/10.1016/j.jvcir.2007.01.005

[98] Klosowski, J.T., Held, M., Mitchell, J.S.B., Sowizral, H., Zikan, K.: Efficient collision detection using bounding volume hierarchies of k-dops. IEEE Transactions on Visualization and Computer Graphics **4**(1), 21–36 (1998) https://doi.org/10.1109/2945.675649

[99] Li, M., Ferguson, Z., Schneider, T., Langlois, T., Zorin, D., Panozzo, D., Jiang, C., Kaufman, D.: Incremental potential contact: Intersection-and inversion-free, large-deformation dynamics. ACM Transactions on Graphics **39**, 20 (2020) https://doi.org/10.1145/3386569.3392425

[100] Jin, W., Posa, M.: Task-driven hybrid model reduction for dexterous manipulation. IEEE Transactions on Robotics **40**, 1774–1794 (2024) https://doi.org/10.1109/TRO.2024.3359531

[101] Doerner, R., Broll, W., Jung, B., Grimm, P., Göbel, M., Kruse, R.: Introduction to virtual and augmented reality. In: Doerner, R., Broll, W., Grimm, P., Jung, B. (eds.) Virtual and Augmented Reality (VR/AR): Foundations and Methods of Extended Realities (XR), pp. 1–37. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-79062-2_1

[102] Wallet, G., Sauzéon, H., Pala, P.A., Larrue, F., Zheng, X., N'Kaoua, B.: Virtual/real transfer of spatial knowledge: benefit from visual fidelity provided in a virtual environment and impact of active navigation. Cyberpsychology, Behavior, and Social Networking **14**(7-8), 417–423 (2011) https://doi.org/10.1089/cyber.2009.0187

[103] Klein, L., Maiello, G., Fleming, R., Voudouris, D.: Friction is preferred over grasp configuration in precision grip grasping. Journal of Neurophysiology **125** (2021) https://doi.org/10.1152/jn.00021.2021

[104] Bergmann Tiest, W.M., Kappers, A.M.L.: The influence of visual and haptic material information on early grasping force. Royal Society Open Science **6**(1), 181563 (2019) https://doi.org/10.1098/rsos.181563

[105] Frisoli, A., Rocchi, F., Marcheschi, S., Dettori, A., Salsedo, F., Bergamasco, M.: A new force-feedback arm exoskeleton for haptic interaction in virtual environments. In: First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference, pp. 195–201 (2005). https://doi.org/10.1109/WHC.2005.15

[106] Giri, G.S., Maddahi, Y., Zareinia, K.: An application-based review of haptics technology. Robotics **10**, 29 (2021) https://doi.org/10.3390/robotics10010029

[107] Kourtesis, P., Argelaguet, F., Vizcay, S., Marchal, M., Pacchierotti, C.: Electrotactile feedback applications for hand and arm interactions: A systematic review, meta-analysis, and future directions. IEEE Transactions on Haptics **15**(3), 479–496 (2022) https://doi.org/10.1109/TOH.2022.3189866

[108] Perret, J., Vander Poorten, E.: Touching virtual reality: A review of haptic gloves. In: ACTUATOR 2018; 16th International Conference on New Actuators, pp. 1–5 (2018)

[109] Triantafyllidis, E., Hu, W., McGreavy, C., Li, Z.: Metrics for 3d object pointing and manipulation in virtual reality: The introduction and validation of a novel approach in measuring human performance. IEEE Robotics & Automation Magazine **29**, 76–91 (2021)

[110] Kimble, K., Albrecht, J., Zimmerman, M., Falco, J.: Performance measures to benchmark the grasping, manipulation, and assembly of deformable objects typical to manufacturing applications. Front. Robot. AI **9**, 999348 (2022)

[111] Chen, S., Xu, Y., Yu, C., Li, L., Ma, X., Xu, Z., Hsu, D.: Daxbench: Benchmarking deformable object manipulation with differentiable physics. In: The Eleventh International Conference on Learning Representations (2023). https://openreview.net/forum?id=1NAzMofMnWl

[112] Liu, H., Zhang, Z., Xie, X., Zhu, Y., Liu, Y., Wang, Y., Zhu, S.-C.: High-fidelity

grasping in virtual reality using a glove-based system. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 5180–5186 (2019). https://doi.org/10.1109/ICRA.2019.8794230

[113] Georgii, J., Westermann, R.: Mass-spring systems on the gpu. Simulation Modelling Practice and Theory **13**(8), 693–702 (2005) https://doi.org/10.1016/j.simpat.2005.08.004 . Programmable Graphics Hardware

[114] Nealen, A., Müller, M., Keiser, R., Boxerman, E., Carlson, M.: Physically based deformable models in computer graphics. Computer Graphics Forum **25**(4), 809–836 (2006) https://doi.org/10.1111/j.1467-8659.2006.01000.x

[115] Moreno-Guerra, M.R., Martínez-Romero, O., Palacios-Pineda, L.M., Olvera-Trejo, D., Diaz-Elizondo, J.A., Flores-Villalba, E., Silva, J.V.L., Elías-Zúñiga, A., Rodriguez, C.A.: Soft tissue hybrid model for real-time simulations. Polymers **14**(7) (2022) https://doi.org/10.3390/polym14071407

[116] Höll, M., Oberweger, M., Arth, C., Lepetit, V.: Efficient physics-based implementation for realistic hand-object interaction in virtual reality. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 175–182 (2018). https://doi.org/10.1109/VR.2018.8448284

[117] Mendes, D., Caputo, F.M., Giachetti, A., Ferreira, A., Jorge, J.: A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. Computer Graphics Forum **38**(1), 21–45 (2019) https://doi.org/10.1111/cgf.13390

[118] Bouzbib, E., Pacchierotti, C., Lécuyer, A.: When tangibles become deformable: Studying pseudo-stiffness perceptual thresholds in a vr grasping task. IEEE Transactions on Visualization and Computer Graphics **29**(5), 2743–2752 (2023) https://doi.org/10.1109/TVCG.2023.3247083

[119] Misra, S., Ramesh, K.T., Okamura, A.M.: Modeling of tool-tissue interactions for computer-based surgical simulation: A literature review. Presence: Teleoper. Virtual Environ. **17**(5), 463–491 (2008) https://doi.org/10.1162/pres.17.5.463

[120] Hulin, T., Albu-Schäffer, A., Hirzinger, G.: Passivity and stability boundaries for haptic systems with time delay. IEEE Transactions on Control Systems Technology **22**(4), 1297–1309 (2014) https://doi.org/10.1109/TCST.2013.2283372

[121] Diolaiti, N., Niemeyer, G., Barbagli, F., Salisbury, J.K.: Stability of haptic rendering: Discretization, quantization, time delay, and coulomb effects. IEEE Transactions on Robotics **22**(2), 256–268 (2006) https://doi.org/10.1109/TRO.2005.862487

[122] Gil, J.J., Sánchez, E., Hulin, T., Preusche, C., Hirzinger, G.: Stability boundary for haptic rendering: Influence of damping and delay. Journal of Computing

and Information Science in Engineering **9**(1), 011005 (2009) https://doi.org/10.1115/1.3074283

[123] Allin, S., Matsuoka, Y., Klatzky, R.: Measuring just noticeable differences for haptic force feedback: implications for rehabilitation. In: Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002, pp. 299–302 (2002). https://doi.org/10.1109/HAPTIC.2002.998972

[124] Zadeh, M.H., Wang, D., Kubica, E.: Human factors for designing a haptic interface for interaction with a virtual environment. In: 2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games, pp. 21–26 (2007). https://doi.org/10.1109/HAVE.2007.4371580

[125] Zadeh, M.H., Wang, D., Kubica, E.: Perception-based lossy haptic compression considerations for velocity-based interactions. Multimedia Systems **13**(4), 275–282 (2008) https://doi.org/10.1007/s00530-007-0106-9

[126] Kadleček, P., Kmoch, P., Křivánek, J.: Haptic rendering for under-actuated 6/3-dof haptic devices. In: Auvray, M., Duriez, C. (eds.) Haptics: Neuroscience, Devices, Modeling, and Applications, pp. 63–71. Springer, Berlin, Heidelberg (2014)

[127] Lopes, P., You, S., Ion, A., Baudisch, P.: Adding force feedback to mixed reality experiences and games using electrical muscle stimulation. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, pp. 1–13. Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3174020 . https://doi.org/10.1145/3173574.3174020

[128] Wang, X., Monteiro, D., Lee, L.-H., Hui, P., Liang, H.-N.: Vibroweight: Simulating weight and center of gravity changes of objects in virtual reality for enhanced realism. In: 2022 IEEE Haptics Symposium (HAPTICS), pp. 1–7 (2022). https://doi.org/10.1109/HAPTICS52432.2022.9765609

[129] Choi, I., Culbertson, H., Miller, M.R., Olwal, A., Follmer, S.: Grabity: A wearable haptic interface for simulating weight and grasping in virtual reality. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. UIST '17, pp. 119–130. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3126594.3126599 . https://doi.org/10.1145/3126594.3126599

[130] Wojna, K., Georgiou, O., Beattie, D., Frier, W., Wright, M., Lutteroth, C.: An exploration of just noticeable differences in mid-air haptics. In: 2023 IEEE World Haptics Conference (WHC), pp. 410–416 (2023). https://doi.org/10.1109/WHC56415.2023.10224388

[131] Luo, Q., Xiao, J.: Contact and deformation modeling for interactive environments. IEEE Transactions on Robotics **23**(3), 416–430 (2007) https://doi.org/10.1109/TRO.2007.895058

[132] Macklin, M., Müller, M., Chentanez, N.: Xpbd: position-based simulation of compliant constrained dynamics. In: Proceedings of the 9th International Conference on Motion in Games. MIG '16, pp. 49–54. Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2994258.2994272 . https://doi.org/10.1145/2994258.2994272

[133] Va, H., Choi, M.-H., Hong, M.: Efficient simulation of volumetric deformable objects in unity3d: Gpu-accelerated position-based dynamics. Electronics **12**(10) (2023) https://doi.org/10.3390/electronics12102229

[134] Martins, M., Morais, L., Torchelsen, R., Nedel, L., Maciel, A.: Efficient position-based deformable colon modeling for endoscopic procedures simulation. In: ACM SIGGRAPH 2024 Conference Papers. SIGGRAPH '24, pp. 91–19110. Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3641519.3657454 . https://doi.org/10.1145/3641519.3657454

[135] Ciaurriz, P., Diaz, I., Gil, J.J.: Stable discrete-time impedances for haptic systems with vibration modes and delay. IEEE Transactions on Control Systems Technology **22**(3), 884–895 (2014) https://doi.org/10.1109/TCST.2013.2271931

[136] Sun, J.S., Lee, K.H., Lee, H.P.: Comparison of implicit and explicit finite element methods for dynamic problems. Journal of Materials Processing Technology **105**(1), 110–118 (2000) https://doi.org/10.1016/S0924-0136(00)00580-X

[137] Liu, T., Bargteil, A.W., O'Brien, J.F., Kavan, L.: Fast simulation of mass-spring systems. ACM Trans. Graph. **32**(6) (2013) https://doi.org/10.1145/2508363.2508406

[138] Todorov, E., Erez, T., Tassa, Y.: Mujoco: A physics engine for model-based control. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5026–5033 (2012). https://doi.org/10.1109/IROS.2012.6386109

[139] Buckingham, G.: Hand tracking for immersive virtual reality: Opportunities and challenges. Frontiers in Virtual Reality **2** (2021) https://doi.org/10.3389/frvir.2021.728461

[140] Zhang, J., Ma, F., Pi, Y., Du, H., Pan, X.: Digital twin embodied interactions design: Synchronized and aligned physical sensation in location-based social vr. Frontiers in Virtual Reality **Volume 6 - 2025** (2025) https://doi.org/10.3389/frvir.2025.1499845

[141] Jain, D., Sra, M., Guo, J., Marques, R., Wu, R., Chiu, J., Schmandt, C.: Immersive terrestrial scuba diving using virtual reality. In: Proceedings of the

2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. CHI EA '16, pp. 1563–1569. Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2851581.2892503 . https://doi.org/10.1145/2851581.2892503

[142] Choi, I., Hawkes, E.W., Christensen, D.L., Ploch, C.J., Follmer, S.: Wolverine: A wearable haptic interface for grasping in virtual reality. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 986–993 (2016). https://doi.org/10.1109/IROS.2016.7759169

[143] Pollard, N.S., Zordan, V.B.: Physically based grasping control from example. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. SCA '05, pp. 311–318. Association for Computing Machinery, New York, NY, USA (2005). https://doi.org/10.1145/1073368.1073413 . https://doi.org/10.1145/1073368.1073413

[144] Nasim, K., Kim, Y.J.: Physics-based interactive virtual grasping. In: Proceedings of HCI Korea. HCIK '16, pp. 114–120. Hanbit Media, Inc., Seoul, KOR (2016). https://doi.org/10.17210/hcik.2016.01.114 . https://doi.org/10.17210/hcik.2016.01.114

[145] Müller, M., Heidelberger, B., Teschner, M., Gross, M.: Meshless deformations based on shape matching. ACM Trans. Graph. **24**(3), 471–478 (2005) https://doi.org/10.1145/1073204.1073216

[146] Chen, L., Huang, J., Sun, H., Bao, H.: Cage-based deformation transfer. Computers & Graphics **34**(2), 107–118 (2010) https://doi.org/10.1016/j.cag.2010.01.003

[147] Casti, S., Livesu, M., Mellado, N., Abu Rumman, N., Scateni, R., Barthe, L., Puppo, E.: Skeleton based cage generation guided by harmonic fields. Computers & Graphics **81**, 140–151 (2019) https://doi.org/10.1016/j.cag.2019.04.004

[148] Le, B.H., Deng, Z.: Interactive cage generation for mesh deformation. In: Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. I3D '17, pp. 3–139. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3023368.3023369 . https://doi.org/10.1145/3023368.3023369

[149] Sacht, L., Vouga, E., Jacobson, A.: Nested cages. ACM Trans. Graph. **34**(6) (2015) https://doi.org/10.1145/2816795.2818093

[150] Liu, H., Zhang, Z., Jiao, Z., Zhang, Z., Li, M., Jiang, C., Zhu, Y., Zhu, S.-C.: A reconfigurable data glove for reconstructing physical and virtual grasps. Engineering **32**, 202–216 (2024) https://doi.org/10.1016/j.eng.2023.01.009

[151] Li, Y., Pollard, N.S.: A shape matching algorithm for synthesizing human-like enveloping grasps. In: 5th IEEE-RAS International Conference on Humanoid Robots, 2005., pp. 442–449 (2005). https://doi.org/10.1109/ICHR.2005.1573607

[152] Li, Y., Fu, J.L., Pollard, N.S.: Data-driven grasp synthesis using shape matching and task-based pruning. IEEE Transactions on Visualization and Computer Graphics **13**(4), 732–747 (2007) https://doi.org/10.1109/TVCG.2007.1033

[153] Taheri, O., Ghorbani, N., Black, M.J., Tzionas, D.: Grab: A dataset of whole-body human grasping of objects. In: European Conference on Computer Vision (2020). https://api.semanticscholar.org/CorpusID:221245720

[154] Fan, Z., Taheri, O., Tzionas, D., Kocabas, M., Kaufmann, M., Black, M.J., Hilliges, O.: Arctic: A dataset for dexterous bimanual hand-object manipulation. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 12943–12954 (2022)

[155] Zhang, Y., He, Q., Wan, Y., Zhang, Y., Deng, X., Ma, C., Wang, H.: Diffgrasp: Whole-body grasping synthesis guided by object motion using a diffusion model. In: AAAI Conference on Artificial Intelligence (2024). https://api.semanticscholar.org/CorpusID:275134319

[156] Zilles, C.B., Salisbury, J.K.: A constraint-based god-object method for haptic display. In: Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots, vol. 3, pp. 146–1513 (1995). https://doi.org/10.1109/IROS.1995.525876

[157] Ericson, C.: Real-Time Collision Detection. Morgan Kaufmann, Boca Raton, FL, USA (2004). https://doi.org/10.1201/b14581

[158] Wang, T.: Collision detection algorithm based on particle system in virtual simulation. In: 2023 4th International Conference for Emerging Technology (INCET), pp. 1–5 (2023). https://doi.org/10.1109/INCET57972.2023.10170181

[159] Macklin, M., Müller, M., Chentanez, N., Kim, T.-Y.: Unified particle physics for real-time applications. ACM Trans. Graph. **33**(4) (2014) https://doi.org/10.1145/2601097.2601152

[160] Zou, Y., Liu, X.: Research of collision detection of soft body in virtual environment. In: 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pp. 379–383 (2016). https://doi.org/10.1109/CSCWD.2016.7566018

[161] Fazioli, F., Ficuciello, F., Fontanelli, G.A., Siciliano, B., Villani, L.: Implementation of a soft-rigid collision detection algorithm in an open-source engine for surgical realistic simulation. In: 2016 IEEE International Conference on

Robotics and Biomimetics (ROBIO), pp. 2204–2208 (2016). https://doi.org/10.1109/ROBIO.2016.7866657

[162] Friston, S., Griffith, E., Swapp, D., Marshall, A., Steed, A.: Position-based control of under-constrained haptics: A system for the dexmo glove. IEEE Robotics and Automation Letters **4**(4), 3497–3504 (2019) https://doi.org/10.1109/LRA.2019.2927940

[163] Shi, Y., Shen, G.: Haptic sensing and feedback techniques toward virtual reality. Research **7**, 0333 (2024) https://doi.org/10.34133/research.0333

[164] Grimm, P., Broll, W., Herold, R., Hummel, J., Kruse, R.: Vr/ar input devices and tracking. In: Doerner, R., Broll, W., Grimm, P., Jung, B. (eds.) Virtual and Augmented Reality (VR/AR): Foundations and Methods of Extended Realities (XR), pp. 107–148. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-79062-2_4

[165] Sharp, T., Keskin, C., Robertson, D., Taylor, J., Shotton, J., Kim, D., Rhemann, C., Leichter, I., Vinnikov, A., Wei, Y., Freedman, D., Kohli, P., Krupka, E., Fitzgibbon, A., Izadi, S.: Accurate, robust, and flexible real-time hand tracking. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15, pp. 3633–3642. Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2702123.2702179

[166] Karelin, A., Brazhenko, D., Kliukovkin, G., Chernenko, Y.: Real-time hand tracking and collision detection for immersive mixed-reality boxing training on apple vision pro. Sensors **25**(16) (2025) https://doi.org/10.3390/s25164943

[167] Tütüncü, E.K., Gonzalez-Franco, M., Gonzalez, E.J.: Handover: Enabling precise selection & manipulation of 3d objects with mouse and hand tracking. In: Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology. UIST '25. Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3746059.3747757

[168] Heo, J., Choi, H., Lee, Y., Kim, H., Ji, H., Park, H., Lee, Y., Jung, C., Nguyen, H.-N., Lee, D.: Hand tracking: Survey. International Journal of Control, Automation and Systems **22**(6), 1761–1778 (2024) https://doi.org/10.1007/s12555-024-0298-1

[169] Koubaa, A., Qureshi, B.: Dronetrack: Cloud-based real-time object tracking using unmanned aerial vehicles over the internet. IEEE Access **6**, 13810–13824 (2018) https://doi.org/10.1109/ACCESS.2018.2811762

[170] Clay, V., König, P., König, S.: Eye tracking in virtual reality. Journal of Eye Movement Research **12**(1), 1–18 (2019) https://doi.org/10.16910/jemr.12.1.3

[171] Li, Z., Akkil, D., Raisamo, R.: Gaze-based kinaesthetic interaction for virtual reality. Interacting with Computers **32**(1), 17–32 (2020) https://doi.org/10.1093/iwcomp/iwaa002

[172] Zhang, Y., Wang, Z.: Review of robotic grippers for high-speed handling of fragile foods. Advanced Robotics (2025) https://doi.org/10.1080/01691864.2025.2508785 . Advance online publication

[173] Pacchierotti, C., Sinclair, S., Solazzi, M., Frisoli, A., Hayward, V., Prattichizzo, D.: Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives. IEEE Transactions on Haptics **XX** (2017) https://doi.org/10.1109/TOH.2017.2689006

[174] Okamura, A.M.: Methods for haptic feedback in teleoperated robot-assisted surgery. Industrial Robot: An International Journal **31**(6), 499–508 (2004) https://doi.org/10.1108/01439910410566362

[175] Tang, Y., Xu, J., Liu, Q., Hu, X., Xue, W., Liu, Z., Lin, Z., Lin, H., Zhang, Y., Zhang, Z., Ma, X., Wang, J., Zhong, J., Wang, D., Jiang, H., Ma, Y.: Advancing haptic interfaces for immersive experiences in the metaverse. Device **2**(6), 100365 (2024) https://doi.org/10.1016/j.device.2024.100365