AniCraft: Crafting Everyday Objects as Physical Proxies for Prototyping 3D Character Animation in Mixed Reality

Boyu Li
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
blibr@connect.hkust-gz.edu.cn

Linping Yuan
The Hong Kong University of Science and Technology
Hong Kong, China
lyuanaa@connect.ust.hk

Zhe Yan
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
zyan698@connect.hkust-gz.edu.cn

Qianxi Liu
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
qliu930@connect.hkust-gz.edu.cn

Yulin Shen
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
yshen654@connect.hkust-gz.edu.cn

Zeyu Wang*
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
The Hong Kong University of Science and Technology
Hong Kong, China
zeyuwang@ust.hk

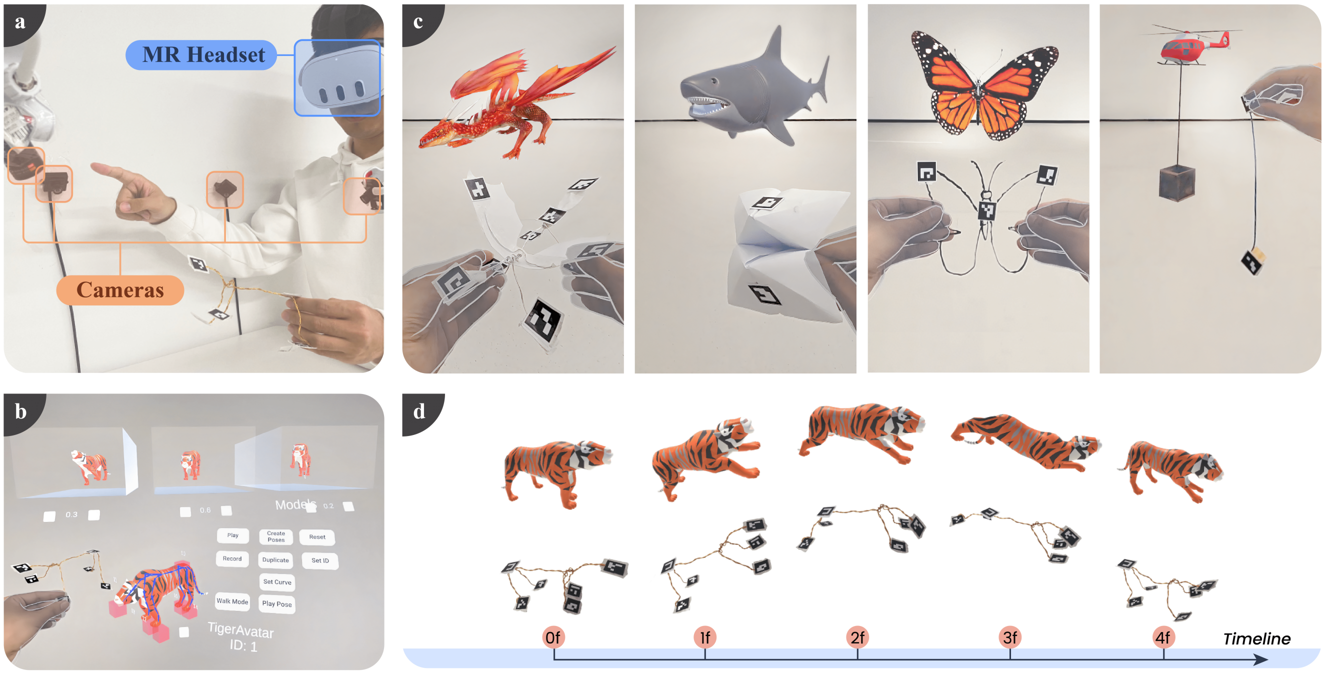


Figure 1: The AniCraft system. AniCraft (a) utilizes affordable cameras for real-time tracking to (b) enable immersive animation prototyping in mixed reality. (c) It empowers creators to craft arbitrary physical proxies to animate diverse characters with (d) keyframe-based or performance-based approaches, accommodating different manipulation types and mapping strategies.

ABSTRACT

We introduce AniCraft, a mixed reality system for prototyping 3D character animation using physical proxies crafted from everyday objects. Unlike existing methods that require specialized equipment

---

*Corresponding author.

UIST '24, October 13–16, 2024, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0628-8/24/10
https://doi.org/10.1145/3654777.3676325

to support the use of physical proxies, AniCraft only requires affordable markers, webcams, and daily accessible objects and materials. AniCraft allows creators to prototype character animations through three key stages: selection of virtual characters, fabrication of physical proxies, and manipulation of these proxies to animate the characters. This authoring workflow is underpinned by diverse physical proxies, manipulation types, and mapping strategies, which ease the process of posing virtual characters and mapping user interactions with physical proxies to animated movements of virtual characters. We provide a range of cases and potential applications to demonstrate how diverse physical proxies can inspire user creativity. User experiments show that our system can outperform traditional animation methods for rapid prototyping. Furthermore, we provide insights into the benefits and usage patterns of different materials, which lead to design implications for future research.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**.

KEYWORDS

Physical Proxy, Rigged Character, 3D Animation, Mixed Reality

ACM Reference Format:
Boyu Li, Linping Yuan, Zhe Yan, Qianxi Liu, Yulin Shen, and Zeyu Wang. 2024. AniCraft: Crafting Everyday Objects as Physical Proxies for Prototyping 3D Character Animation in Mixed Reality. In *The 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24), October 13–16, 2024, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3654777.3676325

1 INTRODUCTION

Rapid prototyping of 3D character animations is crucial in producing films, games, and mixed reality (MR) experiences, as it can help creators test and communicate their design ideas before committing significant resources to final production. Due to the steep learning curve in traditional animation authoring software, there is an increasing interest in leveraging physical proxies and MR technologies to animate rigged characters [6, 30]. This is because physical proxies provide a tangible and intuitive way to manipulate character skeletons [12], and MR allows immediate visualizations of these manipulations on virtual characters within real 3D spaces [21]. Specifically, MR can enhance creators' spatial understanding of character movements and deformations. Unlike desktop setups that involve cumbersome switching between physical proxies and keyboard/mouse, MR enables direct hand-based interaction with the 3D interface. Moreover, while desktop setups disrupt animation creation by requiring users to put down physical proxies to adjust perspectives with a mouse, MR devices allow seamless viewpoint changes by simply moving the head [15].

Previous research has explored the use of human body parts and objects as physical proxies. Although effective, these methods also present several limitations. Specifically, although human faces, hands, or full bodies can drive the movements of humanoid characters naturally and expressively, applying them to non-humanoid characters introduces non-linear and non-intuitive mapping due to the discrepancy in topological structures. To avoid the non-linear mapping, some studies [6, 12, 22] have made use of specialized objects to craft physical proxies mirroring the topological structures of target virtual characters. However, these systems require advanced electronic sensors and mechanical facilities, making them inaccessible to general creators. To lower the barrier, other approaches [13, 15, 33] have adopted everyday objects to control animations, but they primarily focus on simple rigid virtual objects rather than rigged characters.

To address these limitations, we propose AniCraft, a novel MR system for prototyping rigged character animation with physical proxies crafted from everyday objects and materials. AniCraft allows creators to prototype character animations through three key stages: 1) selection of virtual characters, 2) fabrication of physical proxies, and 3) manipulation of these proxies to animate the characters. There are two challenges in designing AniCraft. First, it is non-trivial to support a variety of humanoid and non-humanoid characters with diverse topological structures. To address this, we allow creators to craft custom physical proxies of any shape with diverse objects and materials. We then employ affordable markers and four webcams in AniCraft, offering flexible tracking for changes in any physical proxies, irrespective of their shapes. The use of readily available cameras and MR devices makes AniCraft more accessible than the previous systems [6, 12, 22] that require custom-made electromechanical devices. Second, accommodating the wide variety of motion patterns exhibited by characters is difficult. The system needs to be expressive enough to animate characters from basic rigid transformations to detailed rigged deformations, such as limb movements and gestures. We carefully design diverse manipulation types and mapping strategies, which translate the creator's interactions with the proxies into animated movements of the virtual characters. Then, we incorporate them to support performance-based and keyframe-based animation approaches, which are used for dynamic movements and complex pose changes, respectively.

To evaluate AniCraft, we conducted a comparative user study by asking participants to create several animations with AniCraft and traditional desktop software. The results indicate that our system not only allows users to realize character movement ideas faster but also enhances their creativity through the enjoyable crafting and animating process. Furthermore, we explored the characteristics of various everyday materials in creating physical proxies and character animations through open-explored sessions, providing a reference for future research on developing character animations with tangible interfaces. Finally, we demonstrate three applications of AniCraft. These include incorporating camera movement for rapid previs prototyping, showcasing interactions between characters, and highlighting integration between characters and the physical environment.

This paper makes the following contributions.

- We propose AniCraft, a system leveraging physical proxies crafted from everyday objects and MR technologies for rapid prototyping of 3D character animations.
- We design mapping strategies for various manipulation types and develop physical proxy-based animation examples and potential applications supported by AniCraft.
- Our user study provides insights into system effectiveness, physical proxy preferences, and tangible interface design for character animation prototyping.

2 RELATED WORK

2.1 Physical Proxies for Character Animation

There are two types of physical proxies to animate digital characters: human body parts and general objects.

The most common form of using humans as physical proxies is motion capture, which involves mapping human movements onto virtual humanoid characters [8, 25, 32]. Besides animating virtual humanoid characters with human bodies, a lot of work [7, 11, 19, 23, 29, 43] has focused on enabling humans to control various non-humanoid characters. For instance, KinEtre [7] utilizes Kinect to link different parts of the human body with corresponding parts of a virtual character, generating animated effects through physical movement. BodyAvatar [45] treats bodies as physical proxies of the character, allowing users to create and animate 3d avatars by body gestures. Rhodin et al. [29] implemented real-time character control based on the wave properties of gestures. HandAvatar [19] employs optimization algorithms to automatically map the user's hand movements to any virtual model, while FingerPuppet [16] adopted the finger-walking technique to control virtual character. However, these methods may suffer from unnatural control issues because human body parts have different topological structures from non-humanoid characters [19, 34]. Besides, humans are limited in the range of motions they can perform and thus may fail to express some animations, such as the complex flips often seen in animal movements.

To provide natural control with linear mapping between proxies and characters, many methods [12, 14, 18, 24, 30] have explored utilizing an intermediary object, a "puppet," as a physical proxy of a virtual character for animation. Mechanical devices equipped with sensors [12, 22, 30] can also serve as physical proxies to enable users to easily assemble them into a skeleton for posing rigged characters. For example, Glauser et al. [12] developed a modular input device, comprising joints and splitter parts, with custom circuit and Hall effect sensors. Lamberti et al. [22] used devices like servo motors, gyro sensors, and ultrasonic sensors to capture the pose of the physical proxy. Tangible Avatar [30] tracks a wooden doll through a prop-based controller consisting of capacitive sensors, IMU stations, and a Vive tracker. However, these implementations incorporating electromechanical sensing systems present challenges in reproducibility for individuals lacking expertise in electrical engineering. Therefore, other research [13, 15, 33, 36] has explored the use of common objects as alternative proxies to make the technology more accessible. For instance, 3D Puppetry [15] and MotionMontage [13] incorporate a Kinect-based performance capture to track toys, facilitating the generation of animations within virtual environments. Marker-based systems such as Mirror Puppeteering [33] can also be used to manipulate virtual characters, requiring only a single webcam for tracking. However, these studies primarily focus on simple rigid virtual objects rather than rigged characters, which inherently present a higher level of complexity.

To address these limitations, AniCraft aims to support using diverse and readily available materials, such as paper or metal wire, to create physical proxies for rigged characters. Our system only requires several cheap webcams and printed markers, allowing the mapping of the physical proxies' poses onto virtual handles to control the characters.

2.2 Authoring Tools for 3D Character Animation in Mixed Reality

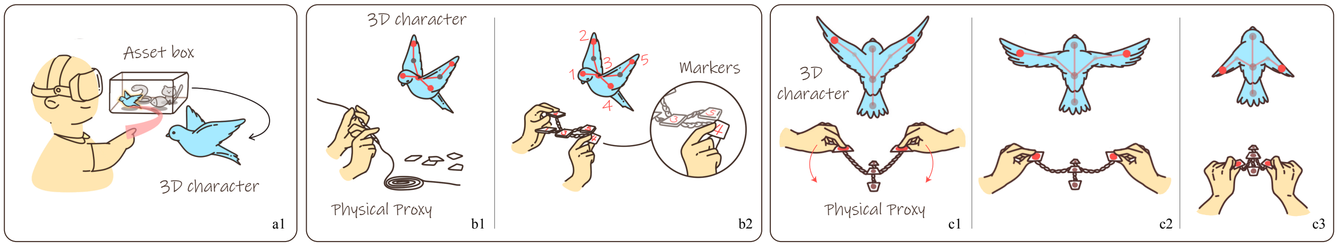
Animation authoring tools in MR blend the physical and digital worlds, allowing creators to design, animate, and interact with 3D content within a real-world context. Commercial VR applications such as Quill [1], Tilt Brush [2], and VR plugins for traditional animation software like Blender and Unreal Engine provide an immersive platform for animation creation. Enhanced spatial understanding and intuitive operations in these systems can reduce the learning curve [37] or inspire animators [10].

In the context of character animation within virtual environments, several systems [5, 27] have explored the enhancement of 3D interaction to facilitate specific stages of the animation creation workflow. For example, PoseMMR [27] attempts to facilitate the collaborative creation of rigged character animations in MR. Compared to a 2D interface, the interaction of manually dragging to pose characters in a 3D environment, although more intuitive, still requires the creation of keyframes for a multitude of control points individually. Therefore, some research has investigated more convenient and affordable methods by utilizing performance-based creation to record movements. For example, AnimationVR [39] is a cost-effective animation system for humanoid characters using the HTC Vive Tracker for tracking body movement.

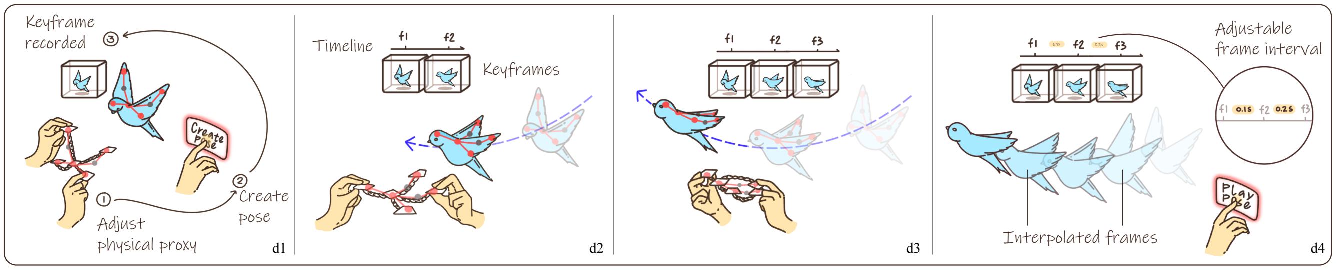
While these studies offer immersive experiences, they often isolate users from the real world, limiting their ability to interact directly with physical elements during the animation process. In response, many animation systems have been proposed to establish connections between physical and virtual objects, thereby enabling the generation of digital animations through physical manipulation. Examples of these systems include MechARspace [47], Teachable Reality [26], Sketched Reality [20], RealityCanvas [41], RealitySketch [35], and HoloBots [17]. Among them, ARAnimator [44] is the system most relevant to us. It uses smartphones as physical proxies for animating characters within physical environments. However, ARAnimator limits its functionality to mapping only the overall displacement of the character, relying on pre-defined character motions to generate character animation. In contrast, AniCraft enhances this concept by supporting more nuanced user-defined adjustments of character pose. This enhancement allows users to control not only the character's overall movement but also the motion of various body parts in an MR setting, offering a more detailed and interactive experience.

3 DESIGN CONSIDERATIONS

We aim to facilitate creators to quickly prototype character animations in the early stages of their projects using self-crafted physical proxies within an MR environment. Toward this, we conducted semi-structured interviews with five creators (C1-C5) for approximately 45 minutes each. The five creators had varying experiences of making animations. Specifically, C1 specialized in video game animations, C2 focused on conceptual design for CG shorts, C3 had experience with animations in virtual environments, C4 excelled in physics simulation animations, and C5 was dedicated to educational video animations. The goals of the interviews included understanding their current practices and challenges when animating characters in the early stages, thereby assessing their needs for



(a) Character Selection (b) Fabrication of Physical Proxy (c) Manipulation and Animation by Performance



(d) Manipulation and Animation by Keyframing

Figure 2: Authoring workflow of AniCraft. To animated a virtual bird with metal wire, (a) a user first selects the bird character from the asset box. (b) Then, the user crafts the physical proxy of the bird with metal wire and attaches markers to it. (c) The user animates the bird through performance, manipulating the proxy's wings for flight. (d) For detailed poses, such as lifting the bird's head, the user switches to the keyframe approach, setting poses and timing for a realistic animation.

a novel animation workflow and system based on physical proxies within an MR environment. Based on our interview results, we have distilled the following design considerations:

Provide a unified authoring process while accommodating to character and animation differences (R1). Participants expressed a desire to create animations of varying complexity for different characters. C2 suggested, *"Aside from the humanoid characters, many other characters, such as animals and non-realistic but fantastic creatures, should be supported together."* C3 mentioned, *"Creating animations involves multiple stages. Sometimes I want to experiment with different movements, and sometimes I focus on pose adjustment."* To meet these needs, AniCraft should provide a unified creation process that allows creators to customize the physics proxy based on the skeletal structure of a character. AniCraft integrates a performance-based animation approach to allow creators to quickly express their intent, and a keyframe-based approach to support detailed and complex poses.

Support physical proxies crafted with daily accessible, easy-to-fabricate, and durable objects and materials (R2). Our participants expressed their desire for a physical proxy to facilitate animating rigged characters, but they preferred not to involve specialized and expensive devices or sensors. For example, C3 stated, *"There is a gap between daily accessible resources and the existing animation creation process. It would be great if everyday stuff could serve as proxies in this process."* C5 expressed a desire for convenience, saying, *"I wish I do not have to master additional skills just for crafting proxies, and I do have to remake proxies frequently."* Reflecting these insights, AniCraft should support the physical proxies crafted from objects and materials that are accessible, easy to fabricate, and durable. Specifically, the materials must be ubiquitous and readily available, enabling creators to obtain the necessary components with minimal effort and financial outlay. Furthermore, these materials should enable quick and straightforward manual construction without necessitating specialized skills or tools. Lastly, the chosen materials need to be robust, ensuring they can endure repeated handling throughout the animation process.

Support intuitive manipulation of multiple handles while utilizing the laws of physics (R3). Our participants expressed frustration with the need to individually adjust control handles when manipulating characters. As stated by C2, *"It is tedious to manually align characters with a ground surface with Gizmos to ensure they move on the ground, rather than sometimes being in the air and sometimes underground."* Furthermore, animations resulting from such adjustments often deviate from physical laws. As mentioned by C4, *"Motion processes like ropes or snakes often require complex physics simulations to achieve realistic movement effects."* To alleviate these difficulties and enhance the realism and engagement of the animations, AniCraft should allow creators to efficiently manipulate multiple handles and capture the real-world behaviors of physical proxies.

Enable flexible mapping from real to virtual for different animation prototyping tasks (R4). One challenge was that prototyping for character animation often involves various focus points, with participants expressing the need for prototypes of various types of animations. C1 mentioned, *"For Prototyping, sometimes I want to show changes in character posture, and sometimes I just need a rough displacement or rotation."* Participants wanted to achieve

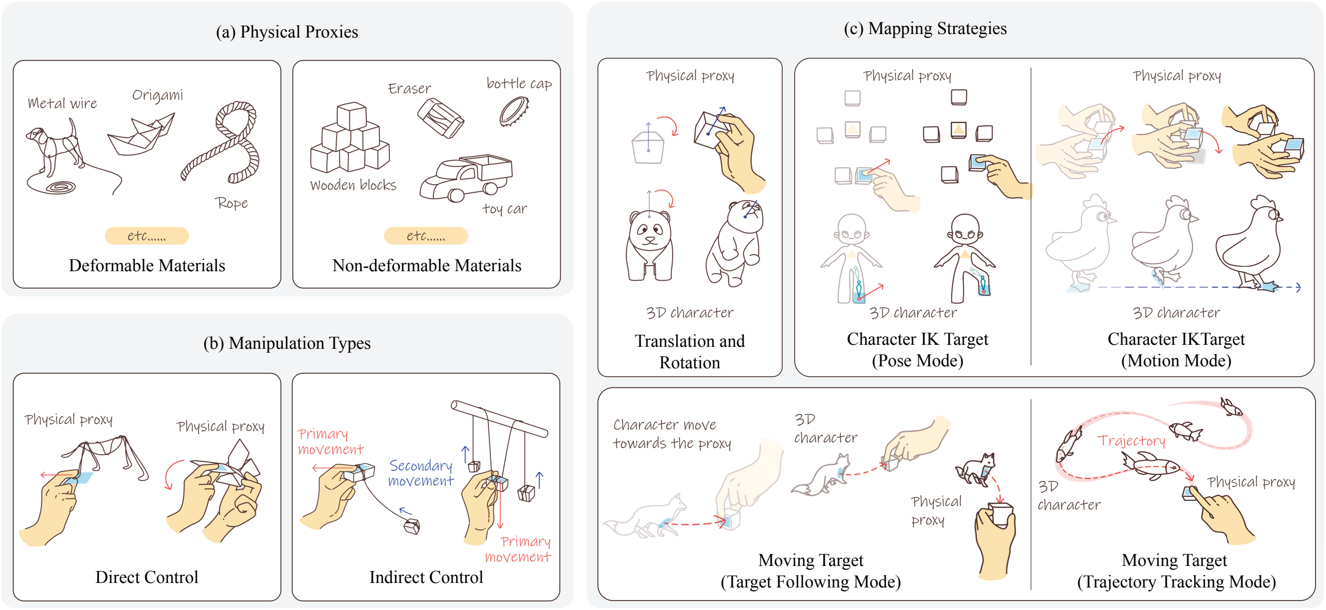


Figure 3: System design of AniCraft. (a) Users can choose from deformable or non-deformable materials to craft physical proxies, (b) manipulate these proxies directly or indirectly, and (c) select proper mapping strategies to translate their physical interactions into character animation.

various types of animations through as simple operations as possible. For example, C5 said, *"I would like to use tools that can adapt character movements to different levels of detail."* Thus, the system should allow for various mapping from physical proxies to virtual characters that can accommodate various animation prototyping tasks and levels of detail.

4 THE ANICRAFT SYSTEM

Informed by the design considerations, we introduce AniCraft, a novel MR authoring system designed for quick and intuitive prototyping of 3D animations. This section begins with a demonstrative example of how creators can utilize AniCraft (Fig. 2), and then elaborates on the system's components (Fig. 3), including physical proxies, manipulation types, and mapping strategies.

4.1 System Walkthrough

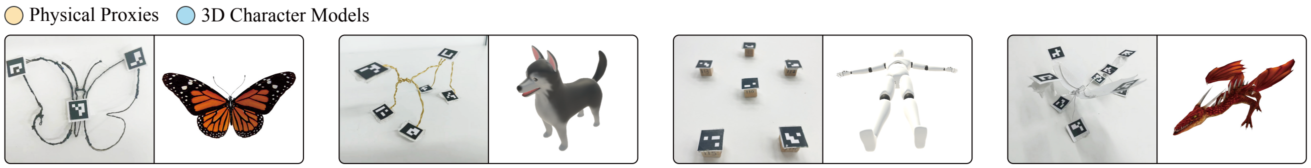
Figure 2 depicts the unified three-stage workflow creators go through to create character animations (**R1**). Imagine a director, Selina, who wants to brainstorm ideas for her upcoming animated film and communicate with her colleagues with tangible animations. She opts to use AniCraft, whose physical proxies allow her to fluidly explore various character movements and combinations. She quickly sets up a workstation with four cameras, connects the cameras and a Quest 3 to her laptop, and puts on the Quest 3 (Fig. 1-a).

Selection of virtual characters (Fig. 2-a). Upon launching AniCraft, Selina steps into an MR creation space. She starts by browsing the 3D character models within an asset box. After reviewing several models, she selects a bird model by dragging it out with her hand. Then, the bird model is enlarged and associated with a menu panel consisting of several buttons for later use.

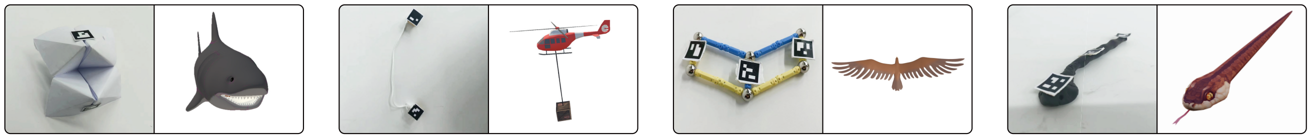
Fabrication of physical proxies (Fig. 2-b). Selina selects aluminum wire for its flexibility. With the visual aids of the virtual bird's skeleton (Fig. 2-b1), she crafts a physical proxy with a similar skeleton in three minutes. Then, Selina notices that AniCraft annotates the control points of the virtual bird with numbers (e.g., 1-5 in Fig. 2-b2). Following these instructions, she attaches numbered markers (e.g., 1-5) to the matching points on the crafted physical proxy (Fig. 2-b2).

Manipulation of proxies to animate characters (Fig. 2-c,d). Selina now can bring her imaginative animations to life using either *performance-based* or *keyframe-based* approaches. Both approaches rely on a sequence of direct or indirect manipulations and proper mapping strategies. Specifically, manipulation allows Selina to change the spatial attributes (e.g., positions and rotations) of the markers, and mapping strategies allow her to define how these marker changes translate into the virtual birds' actual movements.

Initially, Selina wants to animate a bird in flight. In the menu panel, she selects the [Translate] and [IK Target] options as her mapping strategies. Preferring the quicker performance-based method (Fig. 2-c), she hits the [record] button. By continuously bending the wings of the physical proxy with both hands and moving it through the air, she effortlessly creates a virtual bird that flaps its wings in real-time. After reviewing the animation, Selina thinks the virtual bird would look more lifelike if it could lift its head. However, manipulating the wings, head, and proxy position simultaneously is challenging with the performance-based approach. Thus, she switches to the keyframe-based approach (Fig. 2-d), keeping the same mapping strategies. She selects the [Create Poses] button, sets



(a) 3D Printing Pen for Butterfly (b) Metal Wires for Dog (c) Separated Cubes for Human (d) Origami + Metal Wires for Dragon



(e) Origami for Shark (f) Cube + Rope for Helicopter with Box (g) Magnetic Balls + Rods Set for Bird (h) Clay for Snake

Figure 4: Examples of physical proxies for virtual characters. (a) The butterfly model by Rukh3D, (f) helicopter by antonmoek, rope by xenosmashgames, wooden box by Alexey Stepanov, and (h) snake by ChristineDesign, are obtained from Sketchfab under the CC BY 4.0 license. Other models are free assets from Unity Assets Store.

several static poses iteratively at different positions as keyframes, adjusts the timing, and completes the animation by selecting one of the smooth motion interpolation buttons (e.g., [Linear], [EaseIn], [EaseOut], and [EaseInOut]). Satisfied with the animation after hitting the [Play poses] button, she proceeds to animate other characters, returning to the selection of virtual characters and repeating the process until her story is fully animated.

4.2 Physical Proxies

To empower users in the rapid prototyping of 3D character animations through our system easily, our system supports the control of characters through any form of physical proxy by attaching markers (Fig. 2-b). The system can support physical proxies made from readily available materials (**R2**), such as metal wire, origami paper, wooden blocks, rope, and 3D printing pens. Figure 4 demonstrates various physical proxies and their corresponding characters. We primarily categorize the materials used for creating physical proxies into two categories: deformable and non-deformable (Fig. 3-a).

Deformable materials are defined by their capacity to bend and undergo shape transformations. Within the context of our system, users are tasked with the straightforward process of crafting these materials into physical proxies of the target character. This approach enables the mimicry of character posture changes through physical deformation. The intrinsic properties of different materials lend themselves to varied outcomes when utilized for animation within our system. For instance, metal wires (Fig. 4-b) offer plasticity and shape retention, ideal for complex structures and posture maintenance in detailed models. On the other hand, origami (Fig. 4-e) allows for preserving geometric relationships and articulations that can be critical when mimicking certain characters, such as the folded intricacies needed to represent the shark mouth.

Non-deformable materials, prevalent in everyday life, are characterized by their rigidity and inability to alter shape, offering their own unique advantages for animating 3D characters. They are often chosen for their stability and ease of assembly, allowing users to quickly construct physical proxies by piecing together various rigid parts. This straightforward method can be particularly effective for characters with distinct structural features, such as a human figure represented by separated cubes (Fig. 4-c).

4.3 Manipulation Types

Manipulation refers to how users interact with and manipulate the physical proxy within the physical space. AniCraft supports direct and indirect manipulation (Fig. 3-b) to deal with multiple handles and utilize the laws of physics (**R3**).

4.3.1 *Direct control.*

Direct control is the straightforward act of manually adjusting the proxy, whether by bending and shaping deformable materials (e.g., Fig. 6-top) or repositioning markers on non-deformable objects (e.g., Fig. 6-bottom). Direct control allows for detailed adjustment of each marker's position, offering precision in the manipulation process.

4.3.2 *Indirect control.*

Indirect control, however, leverages the natural laws of physics to manipulate movement, such as pulling a rope to simulate a snake's slither or altering part of an origami structure to induce overall motion. Indirect control capitalizes on the inherent properties of physical laws to inspire more creative and dynamic movements from the user.

Figure 5 illustrates indirect control with two examples. For the upper eagle, we designed a simple pulley system as its physical proxy. Users can simulate the eagle flapping its wings by pulling the wooden block representing the eagle's body, causing the counterweights on either side, which represent the wings, to move accordingly. For the bottom example, the helicopter and the hanging box form a single rigged model, depicted by two wooden blocks connected by a rope as physical proxies. The upper block mimics the helicopter's motion, and the lower block the box's movement. As the helicopter moves, the box sways accordingly, achieving natural motion simulation through real-world physics, effectively replicating physical dynamics without digital physics engines.

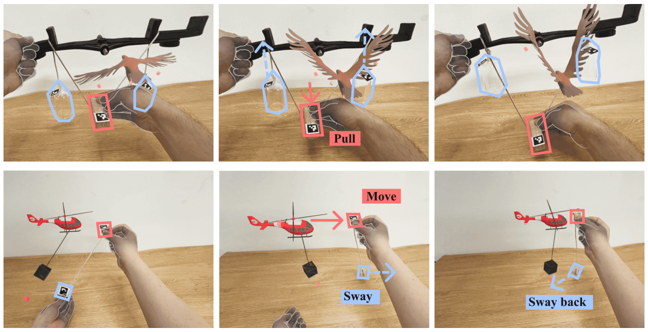


Figure 5: Indirect control examples: (top) an eagle flapping its wings and (bottom) a helicopter with a swinging box. Physical proxies are outlined, where red denotes primary movement and blue denotes secondary movement.

4.4 Mapping Strategies

Mapping strategies define how changes in markers are translated into the characters' movements. We design several mapping strategies (Fig. 3-c) to support various animation prototyping tasks (**R4**).

4.4.1 *Translation and Rotation.* Upon conducting pose estimation on a marker, we can acquire its 6-degrees-of-freedom (DOF) parameters, encompassing both translation and rotation aspects. These parameters are directly mapped to the transformation properties of the corresponding virtual object. It provides a simple way to indicate the entire character's movement.

4.4.2 *Character IK Target.* In order to facilitate the manipulation of skeletal dynamics within rigged characters and minimize the requisite quantity of markers, we employ inverse kinematics (IK) algorithms for two modes: pose mode and motion mode.

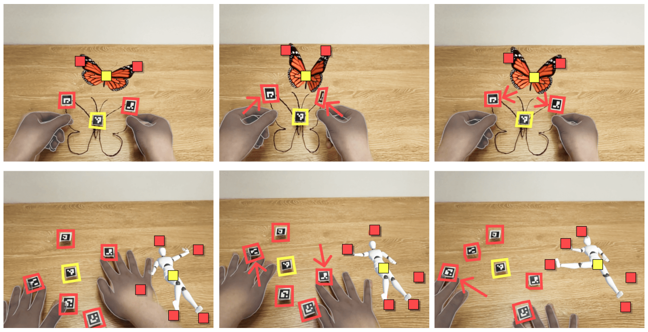


Figure 6: Pose mode examples: (top) a butterfly flapping wings and (bottom) a human performing kung fu. Model origins are denoted by yellow squares, and IK targets are denoted by red squares. Each origin or IK target is associated with a marker.

Pose mode. Pose mode focuses on detailed character posing. It involves placing a primary marker on the origin point to represent the translation and rotation of the character. To further adjust the character's pose, additional markers indicate the relative displacement of IK targets. This method is especially useful for intricate animations that require precise control over each part of the character. Figure 6 illustrates two examples of pose modes. In the top row, there is a progression from left to right where a user manipulates a 3D-printed butterfly to make a virtual butterfly flap its wings. The bottom row shows a user manipulating several wooden blocks with markers to make a virtual person perform kung fu actions. Handles used to control virtual models are visualized as squares, corresponding one-to-one with markers. Yellow squares denote the origin, managing overall character movement, whereas red squares represent the IK target, directing limb movement.

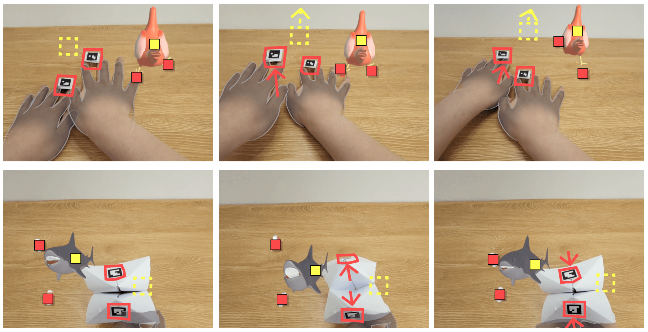


Figure 7: Motion mode examples: (top) a chicken walking and (bottom) a shark opening its mouth. Yellow dashed boxes denote the inferred markers corresponding to model origins.

Motion mode. Motion mode is designed for broader control with fewer markers, ideal for animating simple movements or when minimal markers are desired. In this technique, markers only correspond with IK targets on the character. The process involves calculating the character's potential movement for each IK target based on the initial offsets to the model origin. These individual movements are then averaged to get the character's estimated position and rotation. Figure 7 shows two examples. The upper example demonstrates a chicken walking, requiring only two wooden cubes for the chicken's feet. The full body movement is inferred from the foot's IK position and rotation, allowing users to control the chicken's walking animation by simply moving its feet forward. Similarly, the bottom example shows that controlling a shark's upper and lower jaws can manipulate both the opening of the mouth and the shark's overall movement.

4.4.3 *Moving Target.* This mapping strategy is designed for animating character locomotion, utilizing a looping animation of the character's walking in place and a physical proxy to influence movement directionality rather than exact positioning. Two distinct operational modes are employed: **target following mode** (Fig. 8) propels the character towards the proxy's direction for straightforward target movement, while **trajectory tracking mode** (Fig. 9) guides the character along a path marked by the proxy's movement, ideal for complex navigation through routes or obstacles.

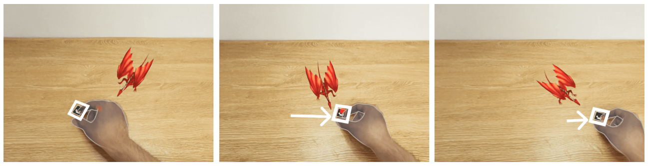


Figure 8: Target following mode example: a dragon moving towards a target.

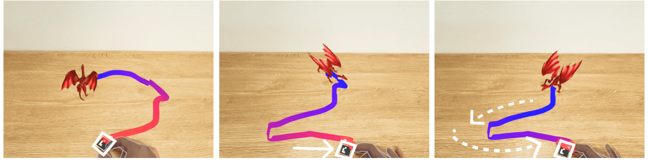


Figure 9: Trajectory tracking mode example: a dragon following a trajectory drawn by a block. The trajectory is represented by a line with a blue-red gradient from the start to the end point

5 IMPLEMENTATION

Our system is developed by Unity and tested on the Meta Quest 3, using four driverless USB 1080p webcams for tracking. For computational support, we deployed two laptops with identical configurations (an Intel Core i5-12450H CPU with NVIDIA GeForce RTX 4050 Laptop GPU and 16GB of memory). One laptop serves as a Python server dedicated to processing the multi-camera data and detecting/tracking marker poses, while the other runs the Unity client interface application and manages streaming to the Quest 3. We will open source the Python and Unity project for reproducibility.

Tracking markers with multiple webcams. To achieve comprehensive tracking coverage, the cameras are positioned to form a 90-degree angle between each other, with a 45-degree downward inclination. This arrangement maximizes the visibility of markers from various angles. We utilize Aruco markers, each assigned a unique ID for precise identification and tracking. A 4x4 Aruco board establishes the world coordinate origin, eliminating the need for external camera calibration. The position and orientation (6DoF) of additional markers are estimated relative to this origin, allowing for accurate spatial tracking. Marker positions recognized by multiple cameras undergo Kalman filtering to merge and smooth the calculated poses, though this process introduces an acceptable latency between marker movements and the corresponding virtual object animations.

Binding markers to virtual handles. Compared to having users specify the correspondence among markers and virtual handles one by one, we simplify this step by only requiring users to assign character IDs (Fig. 2-b), which also considers ID conflicts in multi-character scenarios. After users have assigned IDs to characters, our system interface provides intuitive visual indicators to guide users on which specific part of the character should be manipulated with each assigned marker ID. This feature is designed to help users to efficiently attach markers to physical proxies. This guidance ensures that users can easily select the correct marker ID and attach it to the corresponding part of the physical proxy. The system categorizes Aruco marker IDs into three groups to facilitate the organization and control of virtual entities: 1) Marker IDs 0-99 are designated for rigged characters. Each group of 10 Marker IDs is assigned to one character. This allows for up to 10 markers to control a single character, with support for up to 10 characters in total. 2) IDs 100-199 are allocated for rigid objects, with each ID representing a distinct object. This range supports the inclusion of various non-deformable items within the animation environment. 3) The remaining IDs, 200-215, are reserved for the Aruco board that defines the world coordinate origin, ensuring a consistent reference point for all spatial calculations.

6 APPLICATIONS

To demonstrate the usefulness of AniCraft, this section presents three applications that AniCraft can facilitate.

6.1 Rapid Previs Prototyping

To create film or animation, it is common practice to first create a rough 3D animation as a prototype, known as Previsualization (Previs), which acts like a 3D version of a storyboard [28, 46]. It allows directors and technical teams to explore and communicate ideas, object movement, character actions, camera angles, scene layouts, and other elements before investing a significant amount of time and resources into producing the final product. However, the creation of Previs currently demands considerable effort and time from professional animators, which can hinder rapid discussion and iterative creation. Our system, AniCraft, addresses this bottleneck by enabling users to quickly prototype 3D animations using various physical proxies. By simply attaching markers to objects, users gain the ability to freely manipulate characters and even the camera within the virtual space.

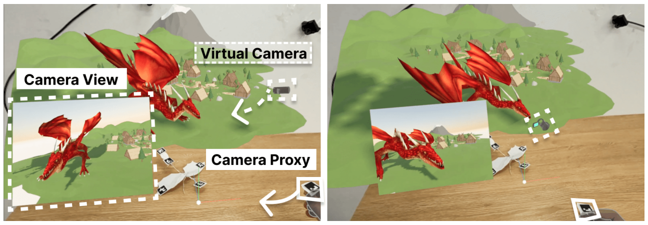


Figure 10: Previs example. A user prototypes a scene where a dragon lowers its wings and head. The animation is made by keyframe in advance, and the user moves a wooden block while playing the animation, which is the physical proxy of the camera. The virtual dragon and the captured scene by the camera are visualized in the real environment.

To refine the previs process further, as shown in Fig. 10, AniCraft allows users to control camera motion using physical proxies, providing real-time previews of the camera's view. Additionally, we have integrated virtual scenes into the system, enabling the visualization of character movements within these settings in mixed environments. This application is designed with the intent of empowering creators who may not have extensive expertise in animation techniques or photography. By moving physical proxies, these users can swiftly communicate their visions for character and camera movements, facilitating clearer and more effective communication with professionals.

6.2 Interactions between Characters

Most tangible user interface systems for character control are usually limited to single-character animations. In contrast, our system allows users to control multiple characters simultaneously using multiple physical proxies. This advancement not only breaks through previous limitations, making character animations more comprehensive but also enables interactions between characters. Specifically, it allows for richer and more realistic virtual character interactions by simulating physical laws of the real world, such as collisions or connections.

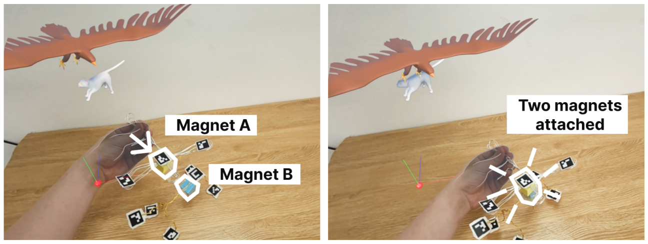


Figure 11: Character interaction example of an eagle snatching a cat. A user employs physical proxies made of aluminum wire for both the eagle and the cat, equipped with magnets. The magnets simulate the action of the eagle capturing the cat through attraction.

As an example, we present an animation case of an eagle snatching a cat (Fig. 11). In this case, users can animate in real time by manipulating two different physical proxies, representing the eagle and the cat, respectively. To simulate the action of the eagle grabbing the cat in the real world, we attached magnets to both physical proxies. When the eagle proxy is moved close to the cat proxy, the magnetic force causes the two proxies to attract each other, vividly simulating the scene of the cat being snatched by the eagle.

6.3 Integration with Physical Environment

AniCraft empowers users to manipulate characters through physical proxies, offering the significant advantage of establishing a direct connection between virtual characters and the real-world environment. In the MR setting, it enables the creation of animations that depict characters moving within real-world environments. Furthermore, inspired by HandAvatar [19], we use other physical objects as proxies for virtual environmental elements, addressing the issue of lacking tactile feedback when controlling character movement in virtual settings.

Figure 12 are two scenarios where wooden blocks are used to control the movement of a chicken's feet, facilitating interaction with the physical environment. The first scenario involves creating an animation of a character moving in the real world. Users can align the virtual character with its physical proxy, guiding the virtual chicken up the steps of the book in the real world. The second scenario focuses on animating a character within a virtual environment, where a box acts as a physical proxy for a bench in the virtual space. By having the chicken step from the tabletop onto the box, users simulate the action of the chicken stepping from grass onto a bench, thereby bridging the gap between virtual movements and physical feedback.

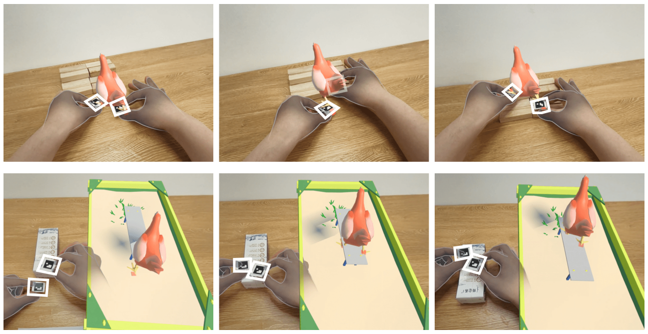


Figure 12: Physical environment example. The top images depict two wooden blocks controlling a virtual chicken stepping onto a staircase constructed from wooden blocks in the real world. The bottom images show a milk box being used as a physical proxy for a bench in a virtual scene, where a user manipulates wooden blocks to step on the virtual box, simulating the animation of a chicken walking onto a bench.

7 USER STUDY

In this section, we evaluate how AniCraft facilitates 3D character animation prototyping compared to the desktop-based system, and how different materials affect the quality and user experience of physical proxies in animation. We conducted a user study and collected quantitative and qualitative results to gain these insights.

7.1 Study Setup and Design

7.1.1 Participants. We recruited 12 participants from our university, including six males and six females aged between 22 and 29 years. Among them, four participants had prior experience in 3D animation production ranging from 3 to 48 months, and one participant had 36 months of professional expertise in character animation and motion capture. The remaining participants were novices in this domain. Nine participants had experience in using AR/VR devices.

7.1.2 Study Setup. As shown in Fig. 2, the workflow consists of three stages. In the manipulation stage, we provided a table of 1.2m by 0.6m as a workstation for crafting proxies. We provided four types of everyday materials, including 1) three different gauges of metal wires, 2) A4 hard and soft paper, 3) clay, and 4) wooden cubes. Essential tools such as scissors, tape, and double-sided tape were also provided. In the mapping stage, a separate table of 0.5m by 0.5m served as a workstation for manipulating proxies and creating animations. To capture markers attached to the proxies, we positioned four cameras around this table (Fig. 1-a). We used two laptops to support computing: one as a server for processing multi-camera data and tracking marker poses, and the other to run the client interface application and stream with Quest 3. These laptops were interconnected through a local network. Participants wore a Quest 3 headset during the whole workflow.

7.1.3 Procedure. We conducted the study with each participant individually. Each study consisted of the following four sessions.

Introduction and training session (15 mins). After filling in a consent form and getting familiar with the basic usage of Quest 3, participants received a brief introduction about the project background and a tutorial on the system with a walk-through example: animating a tiger with manipulation of wire (Fig. 1-a). Participants were asked to follow the tutorial to familiarize themselves with the system, including proxy fabrication, marker and skeletal correspondence, manipulation techniques, and buttons for selecting mapping strategies and triggering keyframe-based or performance-based creation. They could freely ask questions and explore every function until they felt comfortable using the system.

Replication tasks with AniCraft and a desktop-based system (30 mins). Participants were asked to perform replication tasks, which included a total of four tasks with AniCraft: making a chicken walk with blocks (Fig. 7), guiding a butterfly with a 3D printing pen creation (Fig. 6), controlling a dragon to move with blocks (Fig. 9), and directing an eagle with a pulley system (Fig. 5). These tasks were designed to cover most functions in our system design, including proxies, manipulation types, and mapping strategies. At the start of this phase, participants watched example videos of these four tasks and received pre-made physical proxies corresponding to each task. Finally, participants performed a comparative task where they removed the headset and manipulated the character in a desktop system (Unity) using a keyboard and mouse.

Freely exploration with various materials (50 mins). Participants were asked to create physical proxies for characters, including humans, birds, and dogs, using various materials provided, including metal wire, origami, clay, and wooden cubes. We did not provide more materials to prevent users from aimless exploration, and these materials were the most common ones we found in daily life when conducting a Google search for crafting. They were encouraged to use at least two different materials or combinations. After creating each proxy, participants created animations using our system.

Post-study survey and interview (15 mins). Participants finally filled out a seven-point Likert questionnaire based on their previous experimental experience. For the replication tasks, the survey investigated the intuitiveness, comfort, engagement, and overall preference of AniCraft and the baseline. For free exploration, the study evaluated the performance of four materials: metal art, clay, and origami, across seven criteria: learning curve, comfort, flexibility, fun, precision, durability, and adaptability. Subsequently, we conducted semi-structured interviews to collect qualitative feedback. We asked about their opinions of physical proxies, manipulation, and mapping strategies. We also asked about potential application scenarios for AniCraft that users might consider.

7.2 Results and Findings: Effectiveness and Usability

Users can pose characters more naturally and intuitively with 3D physical interaction than with traditional 2D interfaces. From the results shown in Fig. 13, it is evident that AniCraft receives higher scores than the traditional 2D animation UI. For **intuitiveness**, both systems were rated similarly by some participants, but the results for AniCraft were better, with 6 out of 12 participants rating 7. P8, a novice user said, *"It was challenging to quickly understand the complex 2D interface to control characters, but directly manipulating physical proxies is intuitive."* Therefore, 3D physical interaction significantly lowers the entry barrier for animation, making it accessible even to novices. For **comfort**, while we still achieved better results, one notable point was that there was one participant who disagreed with the comfort of the system, with this dissent primarily stemming from experiences of VR-induced motion sickness. Regarding **engagement**, AniCraft again had a modest edge, with the most common score being 7 from 10 out of 12 participants, compared to a more even distribution of scores for the 2D animation UI. P3 mentioned, *"In a 2D interface, you have first to select the correct handle and then drag the 3D object using an awkward set of axes, which is annoying."* Though we have proposed detailed mapping strategies, users do not need to understand the theory behind them. They only need to familiarize themselves with the operation of the mappings through simple practice where even 10 minutes of training is sufficient for a novice. As P4 mentioned, "The 3D spatial interface of AniCraft is much easier to use compared to traditional software." Nearly all users agreed that 3D interactions are more fitting for prototyping character animations. In the end, AniCraft gets a higher preference.

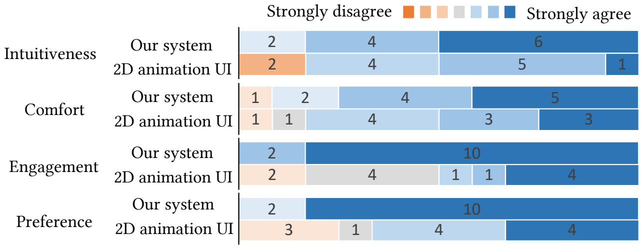


Figure 13: User ratings for AniCraft and a 2D user interface.

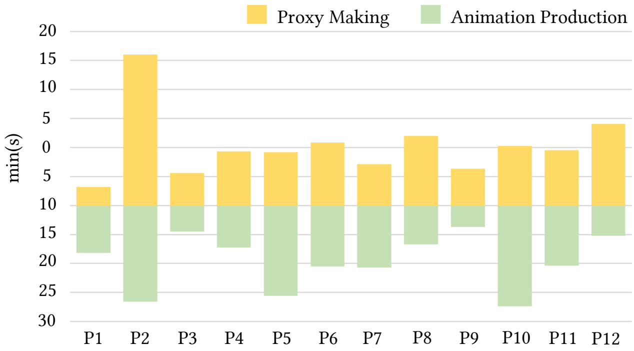


Figure 14: Time allocation of 12 participants on proxy making and animation production.

Our animation workflow with physical crafting enhances creativity, enjoyment, and efficiency. Incorporating physical crafting into the animation workflow significantly alters the creative process, as evidenced by a time-use analysis depicted in Fig 14.

Participants spent their time across two phases: crafting physical proxies and manipulating for animation creation. This time allocation varies from person to person. For example, P1 emphasizes efficiency, allocating just three minutes to create the physical proxy and dedicating the remaining seven minutes to perfecting the animation. P1 stated, *"I'm not very good at handicrafts, so I just focused on making a physical proxy that could support me in completing the animation to the minimum extent necessary."* Conversely, P2 took 26 minutes to create an elaborate proxy, expressing that *"Making the physical proxy was super interesting, especially because it wasn't just for good-looking but could be used for animation."* The result reveals that crafting physical proxies not only made the process more tangible compared to the abstract manipulation seen in 2D software but also transformed it into a therapeutic activity that stimulated creativity. P6 suggested, *"Seeing my handmade creations come to life as virtual models feels incredibly satisfying."* This shift from experiencing animation as a laborious task in 2D to engaging in a pleasurable, stress-relieving exercise in a tactile environment underscores the method's efficacy in enhancing the creative workflow. P2 commented, *"Making animations in 2D software always made me feel frustrated, but this system is a game changer. It lets me enjoy the process and gives me the freedom to create whatever I want."* In summary, our system was instrumental in fostering a deeper connection to the creative work and opening new avenues for innovative thought.

7.3 Results and Findings: Proxies

Material choice for physical proxies crucially impacts animation quality, balancing user ease and creative expression. The utilization of various materials, including origami paper, wooden blocks, metal wire, and clay, demonstrates unique capabilities in simulating animal movements and the dynamics of characters. This insight is supported by user evaluations across seven metrics (Fig. 15), underscoring the nuanced role materials play in animation:

Material selection crucially shapes the user experience and quality of crafted physical proxies. Different materials have their unique characteristics, which make the crafting of physical proxies quite distinct. For instance, **origami** received a lower score for the learning curve because participants found it challenging to create shapes corresponding to the characters. P11 pointed out, *"Although origami paper offers incredible versatility, getting the precise shape and maintaining it can be tricky, especially for complex poses."* **Metal art** overall received the highest ratings, indicating that it is the most suitable among the materials we selected. P6 said *"The metal wire itself is particularly well-suited for creating skeletons, and it's very easy to handle, allowing for the creation of some very attractive metal art."* **Clay** received negative feedback on durability, mainly because over time clay becomes dry and hard, making it unsuitable for reshaping. **Wood blocks** generally refer to non-deformable material commonly found in everyday life. This type of proxy, in most cases, does not require special creation; it only needs to be placed in the correct position.

Physical proxies crafted from various materials feature distinct manipulation ways and movement behaviors. Participants noted that different materials not only facilitated a better emulation of animal movements but also exhibited unique movement patterns for identical characters. For example, no one disagreed with the precision of metal art, the reason is that wire is relatively easy to fix in place, making it especially convenient for posing characters, particularly when combined with keyframe animation. Non-deformable materials like wood cubes, appreciated for their ease of handling and stability, were also noted for their potential incorrect pose due to the lack of restrictions. P4 mentioned, *"Wooden blocks make maintaining poses easier, but they can't capture the subtlety of a character's gesture as well as softer materials can."*

Combining different materials could complement their strengths and weaknesses, allowing for more expressive physical proxies. The strategic combination of materials emerged as a creative solution to harness the strengths of different materials. For example, P10 creatively combined wooded blocks with metal wire to construct a human character proxy that could stand. The wooden blocks acted as a solid base for the proxy, ensuring stability, whereas the wire functioned as a malleable backbone, offering both form and flexibility for adjusting poses. Additionally, for birds, some users used origami for the wings and clay for the body, allowing the bird's wings to flutter in the breeze. P7 stated, *"Combining wire and paper brought the best of both worlds—stability, and flexibility, allowing for a broader range of expressions."*

Reflecting on these findings, it becomes clear that the selection of materials directly influences the animation's fidelity and the user's ability to convey intended movements. The insights from user feedback emphasize the importance of material characteristics in achieving desired animation outcomes.

7.4 Design Implications

Based on the above findings, we distill several implications for the future design of animation prototyping systems utilizing physical proxies.

Prioritize intuitive 3D interaction for novice-friendly animation prototyping. While acknowledging the potential limitations of 3D interaction in precision compared to 2D interfaces, the preference for 3D interaction in character animation prototyping underscores its suitability for tasks that do not require granular detail. Future prototyping tools could leverage this insight to prioritize 3D interaction paradigms, particularly for novice users, as they offer greater ease of understanding. By emphasizing intuitive 3D interaction, future tools can simplify the prototyping process, lower the barrier to entry for animation novices, and ultimately create a more accessible and user-friendly animation creation environment.

Integrate tangible interactions into the animation prototyping workflow to enhance efficiency and creativity. By incorporating physical crafts into the animation creation process, creators can leverage tactile feedback and intuitive manipulation of materials to explore more expressive and innovative design options. This hands-on approach makes the prototyping process more engaging and enjoyable and introduces a more iterative and experimental workflow. Such tangible interactions provide immediate visual and physical feedback, facilitating quicker adjustments and improvements. Hence, research on prototyping may consider embedding physical proxies as a significant factor.

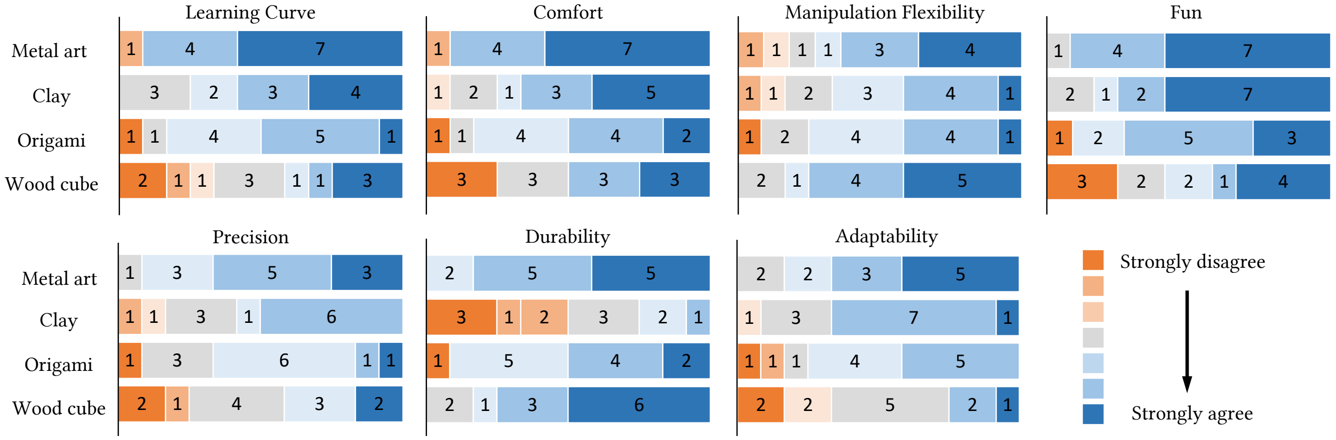


Figure 15: Rating of different metrics on different materials. Users rate for metal art, origami, clay, and wooden cubes on learning curve, comfort, manipulation flexibility, fun, precision, durability, adaptability

Focus on the impact of materials used in making physical proxies on character animation. To accommodate users of varying skill levels, animation systems should consider the complexity differences of materials, providing customized support and guidance to better utilize these materials for creating easily manipulated physical proxies. Furthermore, future research should pay attention to the different movement patterns exhibited by character proxies made of different materials in animations, aiding in selecting materials best suited for specific animation tasks. Additionally, developing algorithms that can detect and correct character pose errors caused by using physical proxies is crucial for maintaining the accuracy and realism of animations. Finally, exploring the potential of combining different materials to leverage their complementary properties is key to their efficient use in the creation of physical proxies, further expanding the possibilities of animation expression.

8 DISCUSSION

8.1 Limitations of Tracking Physical Proxies

While we have utilized a marker-based system to effectively address the need for tracking physical proxies in prototype character animations, we acknowledge this process unavoidably introduces problems common to marker-based systems [26, 36].

Accuracy, range, and angle precision of tracking. In our mixed reality environment, several key factors influence the user experience. Firstly, achieving **accuracy** is paramount. We have adopted markers with a side length of 2cm to ensure better adhesion to physical proxies, although this places significant demands on pose estimation and tracking algorithms. Despite our efforts to improve tracking through algorithmic enhancements, we still struggle to match the precision of 2D interfaces. Secondly, the **range** of interaction is constrained by the fixed camera position on the workbench. Users are confined to the immediate workbench area, roughly within a hemisphere with a radius of 0.3m, limiting their spatial freedom. Thirdly, **orientation** poses a challenge due to our system's reliance on only four desktop cameras. Consequently, markers facing the desktop become untrackable when objects are flipped over, affecting the continuity of interactions. While these limitations could be alleviated by adding more cameras in appropriate locations, our primary focus with AniCraft is concept conveyance rather than addressing these technical constraints directly.

Hand occlusion. When creators manipulate physical proxies, they need to avoid occluding the markers with their hands, as this can result in lost tracking. This issue is particularly critical when employing the performance method for animation creation, where inadvertent occlusion can lead to animation inconsistencies. While keyframes offer the advantage of allowing the user to remove their hands after setting the pose for the physical proxy, ensuring visibility of all markers to the camera, the occlusion problem still impacts the user experience to some degree. Therefore, an algorithm to estimate the position of occluded markers should be proposed. By knowing the starting shape of the character, even if one or two markers are blocked, the pose of the character can be inferred from the position of the detected markers [3].

Although some tangible systems have achieved markerless tracking [4, 9, 26], controlling 3D rig characters still requires tracking multiple small targets with a high level of precision, which remains challenging without using markers. The best way to avoid all issues related to markers is through AI algorithms for 6DoF tracking of deformable objects. We look forward to integrating more advanced algorithms into our system in the future.

8.2 Future Work for Enhancing Usability

Provide an application for 3D-generated models. Currently, our system does not allow users to directly import original 3D models, but rather preset characters that are already rigged and set inversive kinematics. If we could drive arbitrary 3D character models, leveraging methods like auto-rig [38], it would be possible to incorporate existing 3D content generation. This incorporation allows creators to generate characters in mixed reality environments using text-to-model, and to manipulate them directly using real objects. Additionally, integrating reference-driven animationauthoring techniques [42] would allow creators to use pre-existing animations as references, enhancing the creative process and providing more flexibility in animation creation.

Interactions in mixed reality environments offer unique advantages. With the use of physical proxies, avatars can directly interact with the real environment to leverage the natural feedback for better control, as demonstrated in Sec. 6.3. We can further explore integrating 3D scene scanning and style transfer technologies to build immersive environments. There are already some methods [31, 40] that can virtualize real-world scenes and use generative AI to edit them. Through these methods, real-world scenes can be transformed into objects with the same shape but different semantics, such as turning a table into a grassland. Users can directly engage in motion within such scenes using physical proxies.

9 CONCLUSION

In conclusion, AniCraft provides a novel MR platform for intuitive 3D character animation prototyping using everyday materials as physical proxies to streamline the creative process. Our system operates with just a few inexpensive webcams, effectively lowering the threshold for replication and use. We initially conducted a formative study focusing on professionals to gather in-depth design considerations. Additionally, our user study included 7 novices and 5 professionals, ensuring a balanced evaluation of the system's usability across different experience levels. Through user study, AniCraft demonstrates its ability to boost creativity and improve the efficiency of character animation prototyping beyond traditional desktop methods. It also helps provide insights into different materials for crafting physical proxies for virtual characters. Considering its limitations and user feedback, future work includes improving tracking accuracy, expanding material compatibility, and enriching user interactions in MR environments, thus bringing more possibilities for animation prototyping for both novices and professionals.

REFERENCES

[1] 2016. Quill. https://quill.art/.

[2] 2016. Tilt Brush. https://www.tiltbrush.com/.

[3] Andreas Aristidou and Joan Lasenby. 2013. Real-Time Marker Prediction and Cor Estimation in Optical Motion Capture. *The Visual Computer* 29 (01 2013), 7–26. https://doi.org/10.1007/s00371-011-0671-y

[4] Ananta Narayanan Balaji, Clayton Kimber, David Li, Shengzhi Wu, Ruofei Du, and David Kim. 2023. RetroSphere: Self-Contained Passive 3D Controller Tracking for Augmented Reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 157 (jan 2023), 36 pages. https://doi.org/10.1145/3569479

[5] Alberto Cannavò, Claudio Demartini, Lia Morra, and Fabrizio Lamberti. 2019. Immersive Virtual Reality-Based Interfaces for Character Animation. *IEEE Access* 7 (2019), 125463–125480. https://doi.org/10.1109/ACCESS.2019.2939427

[6] A. Cannavò and F. Lamberti. 2018. A Virtual Character Posing System based on Reconfigurable Tangible User Interfaces and Immersive Virtual Reality. In *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference*, Marco Livesu, Gianni Pintore, and Alberto Signoroni (Eds.). The Eurographics Association. https://doi.org/10.2312/stag.20181297

[7] Jiawen Chen, Shahram Izadi, and Andrew Fitzgibbon. 2012. KinEtre: Animating the World with the Human Body. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. 435–444. https://doi.org/10.1145/2380116.2380171

[8] Yann Desmarais, Denis Mottet, Pierre Slangen, and Philippe Montesinos. 2021. A Review of 3D Human Pose Estimation Algorithms for Markerless Motion Capture. *Computer Vision and Image Understanding* 212 (2021), 103275. https://doi.org/10.1016/j.cviu.2021.103275

[9] Ruofei Du, Alex Olwal, Mathieu Le Goc, Shengzhi Wu, Danhang Tang, Yinda Zhang, Jun Zhang, David Joseph Tan, Federico Tombari, and David Kim. 2022. Opportunistic Interfaces for Augmented Reality: Transforming Everyday Objects into Tangible 6DoF Interfaces Using Ad hoc UI. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 183, 4 pages. https://doi.org/10.1145/3491101.3519911

[10] Quentin Galvane, I-Sheng Lin, Marc Christie, and Tsai-Yen Li. 2018. Immersive Previz: VR Authoring for Film Previsualisation. In *ACM SIGGRAPH 2018 Studio* (Vancouver, British Columbia, Canada) *(SIGGRAPH '18)*. Association for Computing Machinery, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3214822.3214831

[11] Adélaïde Genay, Anatole Lécuyer, and Martin Hachet. 2022. Being an Avatar "for Real": A Survey on Virtual Embodiment in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (dec 2022), 5071–5090. https://doi.org/10.1109/TVCG.2021.3099290

[12] Oliver Glauser, Wan-Chun Ma, Daniele Panozzo, Alec Jacobson, Otmar Hilliges, and Olga Sorkine-Hornung. 2016. Rig Animation with a Tangible and Modular Input Device. *ACM Trans. Graph.* 35, 4, Article 144 (jul 2016), 11 pages. https://doi.org/10.1145/2897824.2925909

[13] Ankit Gupta, Maneesh Agrawala, Brian Curless, and Michael Cohen. 2014. Motionmontage: A System to Annotate and Combine Motion Takes for 3D Animations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2017–2026. https://doi.org/10.1145/2556288.2557218

[14] Saikat Gupta, Sujin Jang, and Karthik Ramani. 2014. Puppetx: A Framework for Gestural Interactions with User Constructed Playthings. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (Como, Italy) *(AVI '14)*. Association for Computing Machinery, New York, NY, USA, 73–80. https://doi.org/10.1145/2598153.2598171

[15] Robert Held, Ankit Gupta, Brian Curless, and Maneesh Agrawala. 2012. 3D Puppetry: A Kinect-based Interface for 3D Animation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 423–434. https://doi.org/10.1145/2380116.2380170

[16] Ching-Wen Hung, Chung-Han Liang, and Bing-Yu Chen. 2024. FingerPuppet: Finger-Walking Performance-based Puppetry for Human Avatar. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 163, 6 pages. https://doi.org/10.1145/3613905.3650840

[17] Keiichi Ihara, Mehrad Faridan, Ayumi Ichikawa, Ikkaku Kawaguchi, and Ryo Suzuki. 2023. HoloBots: Augmenting Holographic Telepresence with Mobile Robots for Tangible Remote Collaboration in Mixed Reality. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 119, 12 pages. https://doi.org/10.1145/3586183.3606727

[18] Alec Jacobson, Daniele Panozzo, Oliver Glauser, Cédric Pradalier, Otmar Hilliges, and Olga Sorkine-Hornung. 2014. Tangible and Modular Input Device for Character Articulation. *ACM Trans. Graph.* 33, 4, Article 82 (jul 2014), 12 pages. https://doi.org/10.1145/2601097.2601112

[19] Yu Jiang, Zhipeng Li, Mufei He, David Lindlbauer, and Yukang Yan. 2023. HandAvatar: Embodying Non-Humanoid Virtual Avatars through Hands. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 309, 17 pages. https://doi.org/10.1145/3544548.3581027

[20] Hiroki Kaimoto, Kyzyl Monteiro, Mehrad Faridan, Jiatong Li, Samin Farajian, Yasuaki Kakehi, Ken Nakagaki, and Ryo Suzuki. 2022. Sketched Reality: Sketching Bi-Directional Interactions Between Virtual and Physical Worlds with AR and Actuated Tangible UI. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 1, 12 pages. https://doi.org/10.1145/3526113.3545626

[21] Fabrizio Lamberti, Alberto Cannavo, and Paolo Montuschi. 2020. Is Immersive Virtual Reality the Ultimate Interface for 3D Animators? *Computer* 53, 4 (2020), 36–45. https://doi.org/10.1109/MC.2019.2908871

[22] Fabrizio Lamberti, Gianluca Paravati, Valentina Gatteschi, Alberto Cannavò, and Paolo Montuschi. 2018. Virtual Character Animation Based on Affordable Motion Capture and Reconfigurable Tangible Interfaces. *IEEE Transactions on Visualization and Computer Graphics* 24, 5 (2018), 1742–1755. https://doi.org/10.1109/TVCG.2017.2690433

[23] Dong-Yong Lee, Yong-Hun Cho, and In-Kwon Lee. 2018. Being Them: Presence of Using Non-human Avatars in Immersive Virtual Environment. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 73, 2 pages. https://doi.org/10.1145/3281505.3283384

[24] Ronald Metoyer, Lanyue Xu, and Madhusudhanan Srinivasan. 2003. A Tangible Interface for High-Level Direction of Multiple Animated Characters. *Proceedings - Graphics Interface*, 167–176.

[25] Thomas B. Moeslund, Adrian Hilton, and Volker Krüger. 2006. A Survey of Advances in Vision-based Human Motion Capture and Analysis. *Computer Vision and Image Understanding* 104, 2 (2006), 90–126. https://doi.org/10.1016/j.cviu.2006.08.002 Special Issue on Modeling People: Vision-based understanding of a person's shape, appearance, movement and behaviour.

[26] Kyzyl Monteiro, Ritik Vatsal, Neil Chulpongsatorn, Aman Parnami, and Ryo Suzuki. 2023. Teachable Reality: Prototyping Tangible Augmented Reality with

Everyday Objects by Leveraging Interactive Machine Teaching. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 459, 15 pages. https://doi.org/10.1145/3544548.3581449

[27] Ye Pan and Kenny Mitchell. 2020. PoseMMR: A Collaborative Mixed Reality Authoring Tool for Character Animation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 758–759. https://doi.org/10.1109/VRW50115.2020.00230

[28] Andrew Paquette. 2013. *3D Animation*. Springer London, London, 239–246. https://doi.org/10.1007/978-1-4471-5100-5_13

[29] Helge Rhodin, James Tompkin, Kwang In Kim, Edilson de Aguiar, Hanspeter Pfister, Hans-Peter Seidel, and Christian Theobalt. 2015. Generalizing Wave Gestures from Sparse Examples for Real-time Character Control. *ACM Trans. Graph.* 34, 6, Article 181 (nov 2015), 12 pages. https://doi.org/10.1145/2816795.2818082

[30] J. Saint-Aubert, F. Argelaguet, and A. Lecuyer. 2023. Tangible Avatar : Enhancing Presence and Embodiment During Seated Virtual Experiences with a Prop-Based Controller. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE Computer Society, Los Alamitos, CA, USA, 572–577. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00121

[31] Yulin Shen, Yifei Shen, Jiawen Cheng, Chutian Jiang, Mingming Fan, and Zeyu Wang. 2024. Neural Canvas: Supporting Scenic Design Prototyping by Integrating 3D Sketching and Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1056, 18 pages. https://doi.org/10.1145/3613904.3642096

[32] Soshi Shimada, Vladislav Golyanik, Weipeng Xu, and Christian Theobalt. 2020. PhysCap: Physically Plausible Monocular 3D Motion Capture in Real Time. *ACM Trans. Graph.* 39, 6, Article 235 (nov 2020), 16 pages. https://doi.org/10.1145/3414685.3417877

[33] Ronit Slyper, Guy Hoffman, and Ariel Shamir. 2015. Mirror Puppeteering: Animating Toy Robots in Front of a Webcam. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (Stanford, California, USA) *(TEI '15)*. Association for Computing Machinery, New York, NY, USA, 241–248. https://doi.org/10.1145/2677199.2680548

[34] William Steptoe, Anthony Steed, and Mel Slater. 2013. Human Tails: Ownership and Control of Extended Humanoid Avatars. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (2013), 583–590. https://doi.org/10.1109/TVCG.2013.32

[35] Ryo Suzuki, Rubaiat Habib Kazi, Li-yi Wei, Stephen DiVerdi, Wilmot Li, and Daniel Leithinger. 2020. RealitySketch: Embedding Responsive Graphics and Visualizations in AR through Dynamic Sketching. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 166–181. https://doi.org/10.1145/3379337.3415892

[36] Wai Tong, Zhutian Chen, Meng Xia, Leo Yu-Ho Lo, Linping Yuan, Benjamin Bach, and Huamin Qu. 2022. Exploring interactions with printed data visualizations in augmented reality. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 418–428.

[37] Wai Tong, Kento Shigyo, Linping Yuan, Mingming Fan, Ting-Chuen Pong, Huamin Qu, and Meng Xia. 2024. VisTellAR: Embedding Data Visualization to Short-form Videos Using Mobile Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* (2024).

[38] Ritika Verma, Sarthak Mittal, Siddharth Pawar, Moolchand Sharma, Shalini Goel, and Victor Hugo C. de Albuquerque. 2024. Automatic Rigging of 3D Models With Stacked Hourglass Networks and Descriptors. *AIP Conference Proceedings* 2919, 1 (03 2024), 050006. https://doi.org/10.1063/5.0184393

[39] Daniel Vogel, Paul Lubos, and Frank Steinicke. 2018. AnimationVR - Interactive Controller-based Animating in Virtual Reality. In *2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE)*. 1–6. https://doi.org/10.1109/ANIVAE.2018.8587268

[40] Zeyu Wang, Cuong Nguyen, Paul Asente, and Julie Dorsey. 2023. PointShopAR: Supporting Environmental Design Prototyping Using Point Cloud in Augmented Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, 15 pages. https://doi.org/10.1145/3544548.3580776

[41] Zhijie Xia, Kyzyl Monteiro, Kevin Van, and Ryo Suzuki. 2023. RealityCanvas: Augmented Reality Sketching for Embedded and Responsive Scribble Animation Effects. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 115, 14 pages. https://doi.org/10.1145/3586183.3606716

[42] Liwenhan Xie, Zhaoyu Zhou, Kerun Yu, Yun Wang, Huamin Qu, and Siming Chen. 2023. Wakey-Wakey: Animate Text by Mimicking Characters in a GIF. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 98, 14 pages. https://doi.org/10.1145/3586183.3606813

[43] Katsu Yamane, Yuka Ariki, and Jessica Hodgins. 2010. Animating Non-humanoid Characters with Human Motion Data. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (Madrid, Spain) *(SCA '10)*. Eurographics Association, Goslar, DEU, 169–178.

[44] Hui Ye, Kin Chung Kwan, Wanchao Su, and Hongbo Fu. 2020. ARAnimator: In-situ Character Animation in Mobile AR with User-defined Motion Gestures. *ACM Trans. Graph.* 39, 4, Article 83 (aug 2020), 12 pages. https://doi.org/10.1145/3386569.3392404

[45] Yupeng Zhang, Teng Han, Zhimin Ren, Nobuyuki Umetani, Xin Tong, Yang Liu, Takaaki Shiratori, and Xiang Cao. 2013. BodyAvatar: creating freeform 3D avatars using first-person body gestures. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 387–396. https://doi.org/10.1145/2501988.2502015

[46] Qian Zhu, Linping Yuan, Zian Xu, Leni Yang, Meng Xia, Zhuo Wang, Hai-Ning Liang, and Xiaojuan Ma. 2024. From reader to experiencer: Design and evaluation of a VR data story for promoting the situation awareness of public health threats. *International Journal of Human-Computer Studies* 181 (2024), 103137.

[47] Zhengzhe Zhu, Ziyi Liu, Tianyi Wang, Youyou Zhang, Xun Qian, Pashin Farsak Raja, Ana Villanueva, and Karthik Ramani. 2022. MechARspace: An Authoring System Enabling Bidirectional Binding of Augmented Reality with Toys in Real-time. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 50, 16 pages. https://doi.org/10.1145/3526113.3545668