

Article

High-Speed Dynamic Projection Mapping onto Human Arm with Realistic Skin Deformation [†]

Hao-Lun Peng * and Yoshihiro Watanabe

School of Engineering, Tokyo Institute of Technology, Kanagawa 226-8503, Japan; watanabe.y.cl@m.titech.ac.jp

* Correspondence: peng.h.ab@m.titech.ac.jp; Tel.: +81-45-924-5474

[†] This paper is an extended version of Peng, H.-L.; Watanabe, Y. High-Speed Human Arm Projection Mapping with Skin Deformation. In *SIGGRAPH Asia 2020 Emerging Technologies*; Association for Computing Machinery: New York, NY, USA, 2020; article 13, pp. 1–2, doi:10.1145/3415255.3422887.

Abstract: Dynamic projection mapping for a moving object according to its position and shape is fundamental for augmented reality to resemble changes on a target surface. For instance, augmenting the human arm surface via dynamic projection mapping can enhance applications in fashion, user interfaces, prototyping, education, medical assistance, and other fields. For such applications, however, conventional methods neglect skin deformation and have a high latency between motion and projection, causing noticeable misalignment between the target arm surface and projected images. These problems degrade the user experience and limit the development of more applications. We propose a system for high-speed dynamic projection mapping onto a rapidly moving human arm with realistic skin deformation. With the developed system, the user does not perceive any misalignment between the arm surface and projected images. First, we combine a state-of-the-art parametric deformable surface model with efficient regression-based accuracy compensation to represent skin deformation. Through compensation, we modify the texture coordinates to achieve fast and accurate image generation for projection mapping based on joint tracking. Second, we develop a high-speed system that provides a latency between motion and projection below 10 ms, which is generally imperceptible by human vision. Compared with conventional methods, the proposed system provides more realistic experiences and increases the applicability of dynamic projection mapping.

Keywords: projection mapping; augmented reality; human-centered computing; human–computer interface; skin deformation



Citation: Peng, H.-L.; Watanabe, Y. High-Speed Dynamic Projection Mapping onto Human Arm with Realistic Skin Deformation. *Appl. Sci.* **2021**, *11*, 3753. <https://doi.org/10.3390/app11093753>

Academic Editor: Chang-Hun Kim

Received: 31 March 2021

Accepted: 20 April 2021

Published: 21 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Augmented reality is being rapidly developed to enhance the user experience in the real world and has attracted much attention in research and industry. A fundamental approach to realize augmented reality is called projection mapping or spatial augmented reality [1], which has been widely used for a variety of applications. Projection mapping aims to create the perception of changes in the materials and shape of a target surface by overlaying images according to the target position and shape. Compared with other augmented reality techniques relying on handheld devices or head-mounted displays, projection mapping omits using or wearing devices, and the information presented via projection mapping can be shared across multiple users. Various applications have demonstrated the applicability of projection mapping. For instance, theme parks and other entertainment environments have been enhanced and energized by adopting projection mapping [2].

Existing projection mapping techniques can be classified considering the dynamics and shapes of the target objects. With the proposal by Raskar et al. [3] of shader lamps, the targets of traditional projection mapping systems were limited to static and rigid objects.

To overcome these limitations, studies on dynamic projection mapping have been subsequently conducted considering moving and non-rigid targets to achieve highly immersive visual experiences. Siegl et al. [4] proposed a dynamic projection mapping system using a depth sensor. Narita et al. [5] proposed a non-rigid projection mapping system using a deformable dot cluster marker. Miyashita et al. [6] introduced the MIDAS Projection, a marker-less and model-less non-rigid dynamic projection system to represent the appearance of materials using real-time measurements in the infrared region. Nomoto et al. [7] addressed limitations of dynamic projection mapping using multiple projectors and a pixel-parallel algorithm.

Projection mapping can also be performed on the human body, which is a complex non-rigid surface. For instance, accurate face projection mapping systems have been proposed [8,9]. Bermano et al. [8] achieved projection mapping onto a human face based on non-rigid marker-less face tracking. Sigel et al. [9] proposed the FaceForge system to improve the applicability and quality of face projection mapping using multiple projectors. Similarly, other parts of the human body can be augmented by projection mapping. Augmenting the human arm surface via dynamic projection mapping can notably improve applications related to activities of daily living. In addition, projection mapping has found applications in a variety of fields such as fashion [10,11], user interfaces [12–15], prototyping [16,17], education [18], and medical assistance [19]. For instance, Gannon et al. [16] presented ExoSkin, a hybrid fabrication system for designing and projecting digital objects directly on the body through projection mapping and marker-based joint tracking. Xiao et al. [12] introduced LumiWatch, a skin-projected wristwatch, to enable coarse touch screen capabilities on the arm surface. However, conventional projection mapping systems for the human arm surface mostly neglect skin deformation and have a high latency between motion and projection, consequently causing noticeable misalignments between the arm surface and projected images. These problems degrade the user experience and limit the development of more applications.

We focus on high-speed dynamic projection mapping onto the human arm surface with realistic skin deformation for the user not to perceive any misalignment between the arm and projected images. We address two main challenges: reducing the overall latency of the system and reconstructing realistic skin deformation. Ng et al. [20] developed a projection-based touch input device with a response latency of 1 ms and determined that users cannot perceive an average projection latency below 6.04 ms. Therefore, a projection mapping system should achieve a latency of a few milliseconds to avoid perceivable misalignments.

To provide an accurate and realistic experience, we propose a dynamic projection mapping system on the human arm surface with the following two contributions [21]. First, we combine a state-of-the-art parametric deformable surface model with an efficient regression-based accuracy compensation method to describe skin deformation. The compensation method modifies the texture coordinates parallel to the 3D shape reconstruction to achieve fast and accurate image generation for projection mapping based on joint tracking. Second, we develop a high-speed system that provides a latency between motion and projection below 10 ms, which can hardly be perceived by human vision. Compared with conventional methods, the proposed system provides more realistic experiences and broadens the applicability of dynamic projection mapping.

2. Related Work

In this section, we discuss a variety of studies on human skin surface reconstruction.

2.1. Non-Rigid Registration

Non-rigid registration has been widely studied in recent years and is a suitable method for human skin surface reconstruction. Basically, it deforms the vertices of a model to fit the observations at each time step using a non-rigid deformation [22–24]. The as-rigid-as-possible method [22] is commonly used to regularize deformable surfaces [25–31]. Zollhöfer et al. [26] proposed a novel GPU (graphics processing unit) pipeline for non-rigid

registration of live RGB-D (color and depth) data to a smooth template using an extended nonlinear as-rigid-as-possible framework. Gao et al. [30] proposed SurfelWarp for real-time reconstruction using the as-rigid-as-possible method.

However, pure non-rigid deformation suffers from error accumulation and usually fails when tracking long motion sequences [31]. Guo et al. [28] introduced robust tracking of complex human bodies using non-rigid motion tracking techniques. Li et al. [29] developed a robust template-based non-rigid registration method for the human body using local-coordinate regularization, thus improving the registration speed. Nonetheless, these non-rigid registration methods for the human body cannot achieve real-time performance, which usually remains below 1 fps.

2.2. Learned Human Body Model

In computer graphics, the generation of realistic representations of human bodies has been pursued to describe different body shapes and natural deformations according to the pose. For instance, blend skinning [32,33] and blend shapes [34,35] are widely used in animation. In recent years, higher-quality models based on blend skinning and blend shapes have been proposed, such as SCAPE [36], BlendSCAPE [37], and SMPL [38]. SMPL, the skinned multi-person linear model, is a realistic 3D model of the human body learned from thousands of 3D body scans. The SMPL model is suitable to represent multiple persons and provides fast rendering and ease of use.

2.3. 3D Human Surface Estimation

In computer vision, remarkable results have been achieved in the reconstruction of the human body surface based on parametric deformable surface models, such as SMPL [38], without using markers.

Early works have been aimed to fit parametric models to 2D image observations through iterative optimization, implying time-consuming computations and requiring careful initialization. Bogo et al. [39] proposed the SMPLify framework to automatically estimate the SMPL parameters using human landmarks. Similarly, Lassner et al. [40] introduced a method for estimating the SMPL parameters using silhouettes, and Joo et al. [41] proposed the Total Capture system to estimate the SMPL parameters using a multicamera system.

Convolutional neural networks have been widely studied with the development of deep learning. Kanazawa et al. [42] presented an end-to-end adversarial learning method, HMR, to estimate the SMPL parameters in approximately 40 ms using a single GPU (NVIDIA GeForce GTX 1080 Ti; NVIDIA, Santa Clara, CA, USA). Similarly, Pavlakos et al. [43] introduced an end-to-end framework to estimate the SMPL parameters in 50 ms using a single GPU (NVIDIA Titan X). In these methods, the process is decomposed into two stages—first, regression of some types of 2D representations such as joint heatmap, mask, or 2D segmentation; second, estimation of model parameters from the intermediate results.

As the performance of two-stage methods relies heavily on the accuracy of the intermediate results, input information is not fully utilized. Recent studies have reported the higher efficiency and effectiveness of estimating meshes representing human bodies in a single stage instead of two stages [44–46]. The concept of DenseBody proposed by Yao et al. [46] achieves state-of-the-art performance in speed by introducing new representations for 3D objects, being more suitable for convolutional neural networks. Its runtime is approximately 5 ms for images of 256×256 pixels using a single GPU (NVIDIA GeForce GTX 1080 Ti). Although a runtime of 5 ms may be appropriate, the accuracy of DenseBody is still a critical problem, especially for representing the limbs. The main reason for the limited accuracy of DenseBody is its low resolution, and the limbs only account for a small portion of input images of the whole body. Thus, the skin deformation accuracy on the limbs, including the arms, is low for dynamic projection mapping. Moreover, the reconstruction of realistic skin deformation without markers while achieving high projection accuracy and speed remains to be solved.

3. Proposed Projection Mapping System

3.1. Overview

Figure 1 shows the system configuration and method pipeline. The proposed system is composed of three main parts: tracking, rendering, and projection.

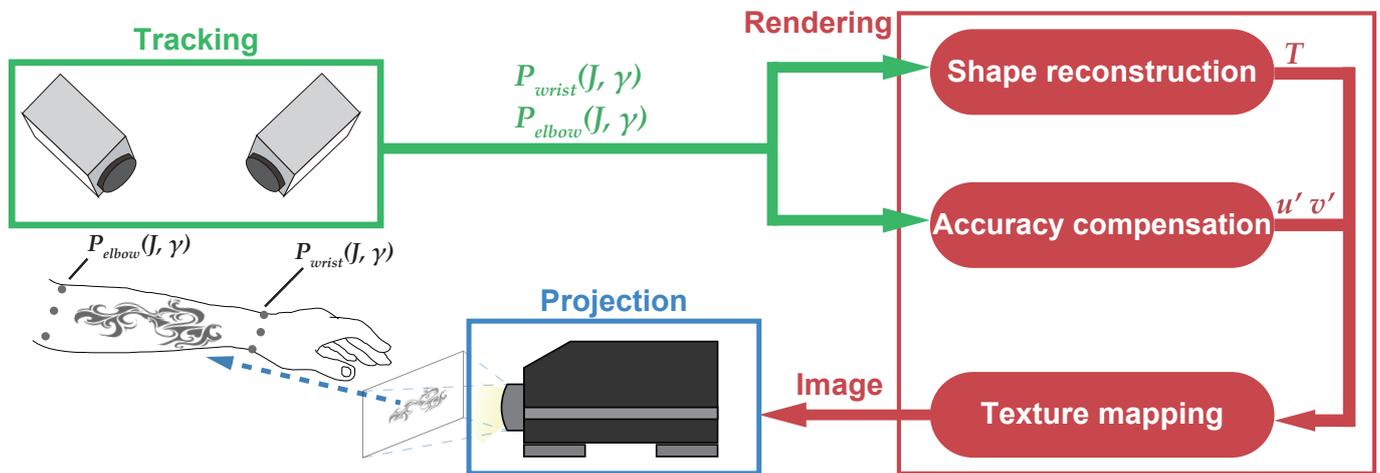


Figure 1. System configuration and method pipeline.

During tracking, arm poses including wrist pose P_{wrist} and elbow pose P_{elbow} are obtained by a marker-based motion tracking system in real time and used for rendering. Gray dots on the arm in Figure 1 indicate markers. During rendering, shape reconstruction outputs a 3D mesh of the arm surface based on the acquired real-time joint poses and predefined shape parameters, while accuracy compensation outputs the 2D texture coordinates of the arm model based on the joint poses only. These two processes are performed in parallel to accelerate rendering. The outputs are then combined by texture mapping to generate an image for projection. During projection, a high-speed projector shows the image on the arm of a user.

3.2. High-Speed Motion Tracking

As mentioned in Section 2, it is difficult to reconstruct skin deformation with high accuracy and speed using marker-less methods. However, fast and accurate methods are required for applications such as prototyping and medical assistance. Although it may increase the burden in setup and cause some discomfort to the user, marker-based methods are more suitable than their marker-less counterparts to track joint poses and reconstruct the arm surface using a parametric deformable surface model, such as SMPL [38]. For instance, the system in [16] used marker-based tracking to obtain accurate arm poses.

The SMPL model represents a realistic human body using 10-dimensional body shape parameter β and 72-dimensional pose parameter θ [38]. As the user's body shape does not vary in a short time, β can be determined in advance using marker-less methods, such as those in [39–43]. In contrast, the user's pose changes rapidly and notably contributes to skin deformation. We consider the forearm (i.e., the region between the elbow and wrist) for the proposed system. According to the definition of blend skinning, each vertex of the mesh is transformed by the weight of its neighboring joints. Hence, only the poses of the wrist and elbow are required to represent the forearm.

Commercial marker-based motion tracking systems, such as OptiTrack (NaturalPoint, Corvallis, OR, USA) [47] and Vicon (Oxford Metrics, Yarnton, UK) [48], can offer robust pose information of rigid bodies with a very short latency. In the proposed system, wrist pose P_{wrist} and elbow pose P_{elbow} are obtained from a motion tracking system. Joint pose P is a six-dimensional vector that includes the joint position, $J \in \mathbb{R}^3$, and joint orientation, $\gamma \in \mathbb{R}^3$.

3.3. High-Speed Rendering

The arm surface, which is represented by 3D mesh $T \in \mathbb{R}^{3N}$ with N vertices, is reconstructed using the SMPL model [38]. Although the SMPL model provides high-quality reconstruction, it suffers from various problems. For instance, despite the accurate shape representation [41], small details of skin deformation are missing in the SMPL model due to its linearity [49].

The first row of Figure 2 illustrates the accuracy problems of the SMPL model. The misalignment between the projected patterns and black crosses is clear. This is a critical problem in projection mapping because users are sensitive to misalignments between the deformed skin and projected images. This problem is more noticeable because the users observe the projected image directly on their arms instead of on a separate display.

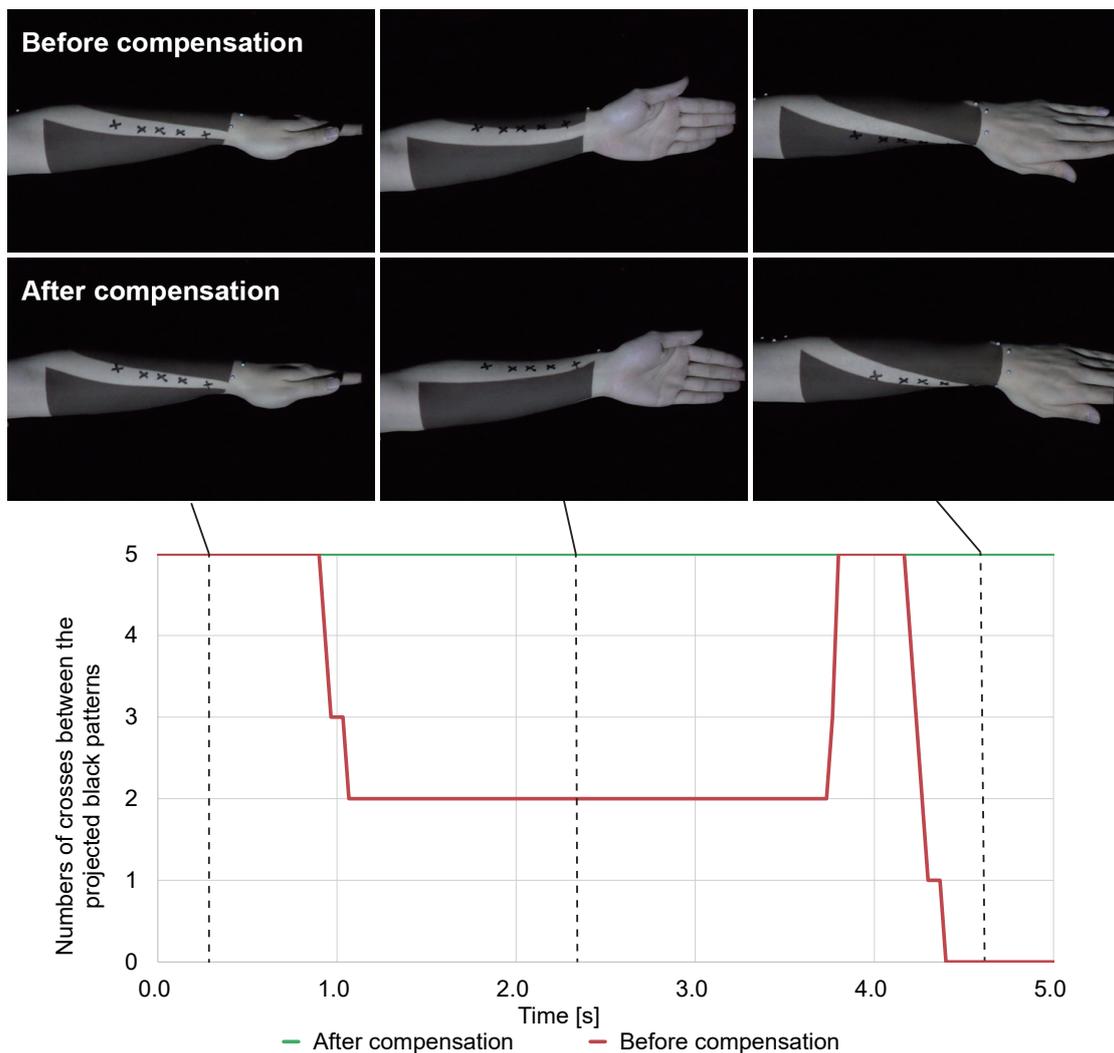


Figure 2. Photographs before (first row) and after (second row) accuracy compensation. The black patterns are projected images. The black crosses drawn on the arm using an ink marker indicate the ground truth for skin deformation. Numbers of crosses between the projected black patterns in the time sequence before and after accuracy compensation (bottom).

Regarding the arm, wrist pronation, and supination tend to cause the most obvious skin deformations that cause the most severe misalignments, as shown in the first row of Figure 2. Pronation and supination are rotations around the axis of the forearm. Thus, we assume that the misalignments caused by these motions are mainly perpendicular to the axis of the forearm. To improve the accuracy of describing skin deformation in the SMPL model for dynamic projection mapping, we propose an efficient regression-based

compensation method. The method modifies the texture coordinates of the model in real time as follows:

$$\begin{cases} \mathbf{u}' = \mathbf{g}(\mathbf{v}^*, \gamma_{wrist}) + \mathbf{u}^* \in \mathbb{R}^N, \\ \mathbf{v}' = \mathbf{v}^* \in \mathbb{R}^N, \end{cases} \quad (1)$$

where $(\mathbf{u}^*, \mathbf{v}^*)$ are the texture coordinates provided by the arm-surface model initially before compensation. Respectively, u -axis is perpendicular to the axis of the forearm and v -axis is parallel to the axis of the forearm. Regression model $\mathbf{g}(\mathbf{v}^*, \gamma_{wrist})$ is prepared in advance using polynomial curve fitting of user-specific data. The compensation method provides \mathbf{u}' and \mathbf{v}' , the texture coordinates of the model that are consistent with the reconstruction of the arm surface, T , obtained by the SMPL model.

In the proposed system, the SMPL model serves as shape reconstruction method, and the compensation method can run in parallel to accelerate processing, as shown in Figure 3. Then, the reconstruction and compensation methods are combined by texture mapping, which maps an image onto a 3D shape. The required processing time of the proposed compensation method is shorter than that of the arm-surface reconstruction due to the simple polynomial regression. Thus, the proposed compensation method does not increase the overall system latency and improves the accuracy of dynamic projection mapping.

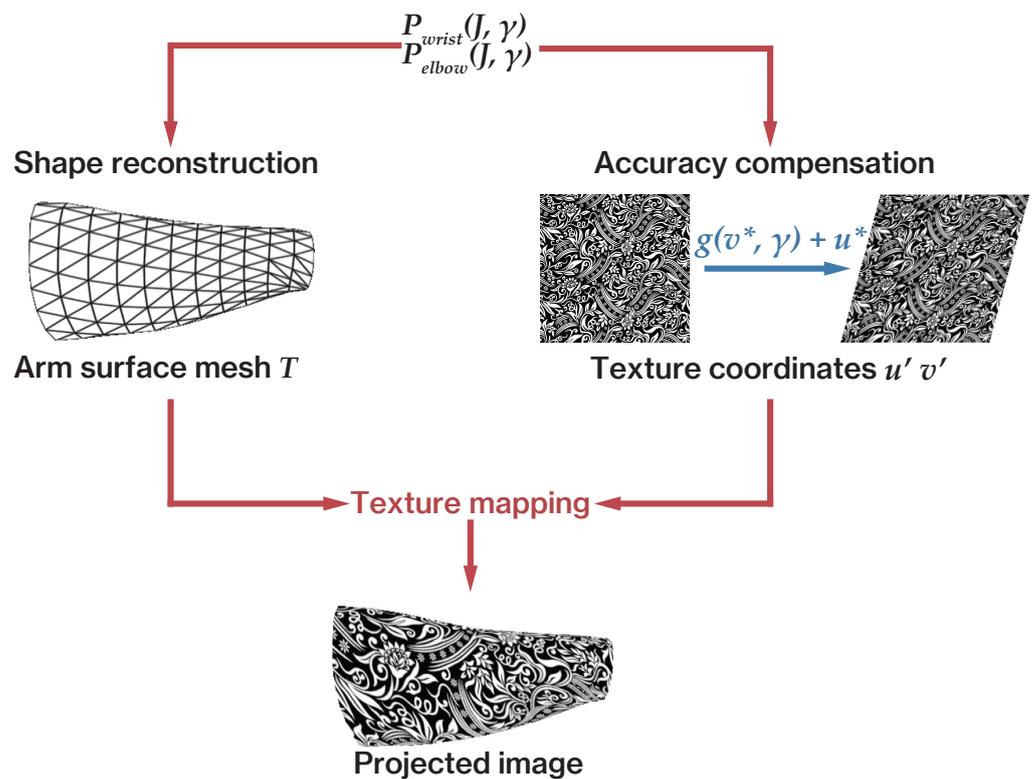


Figure 3. Shape reconstruction and compensation executed in parallel using joint tracking results.

3.4. High-Speed Projection

To perform dynamic projection mapping within milliseconds [20], it is necessary to reduce the latency between image transmission and projection. Commercial off-the-shelf projectors usually have a latency above 10 ms after the projected image is delivered by the computer, being unsuitable for the proposed system. Alternatively, projectors based on the Digital Light Processing (DLP) chipsets have been proposed to achieve extremely high projection speeds. For instance, Watanabe et al. [50] developed a single-chip DLP high-speed monochrome projector with a rate of 1000 fps and a latency of 3 ms. Such high performance is achieved by the synchronized operation of a digital micromirror device and an LED (light emitting diode) by using a specialized image transmission module. To project

RGB images using DLP projectors, three-chip DLP architectures have been combined to project the red, green, and blue image channels [8,51]. For instance, a customized 24-bit RGB projector at 480 fps has been devised using a three-chip DLP configuration [8]. However, this configuration requires complicated optical systems that are costly and bulky. Watanabe et al. [52] then developed a single-chip DLP 24-bit RGB projector with fast response and high brightness.

In the proposed system, considering the system configuration and latency, we use a state-of-the-art high-speed single-chip projector that displays 24-bit XGA (Extended Graphics Array) images at a maximum rate of 947 fps [52]. As a result, the latency in image transmission from computer generation until projection can be reduced to less than 3 ms.

4. Experiments

4.1. Hardware Setup

Figure 4 shows the system configuration used for experiments to test the proposed system. Eight OptiTrack Prime 17 W cameras (360 fps, 1664×1088 resolution, 70° field of view) were used to determine the joint poses within 4 ms [47]. In addition, a 24-bit high-speed projector was used. The projector achieves a maximum rate of 947 fps and 3 ms latency [52]. For the experiments, we used a computer equipped with a dual-core Intel Xeon Gold 6136 processor (3.00 GHz, 24 cores; Intel, Santa Clara, CA, USA), an NVIDIA GeForce RTX 2080 Ti (VRAM 11.0 GB) GPU, and 80.0 GB (2666 MHz) memory.

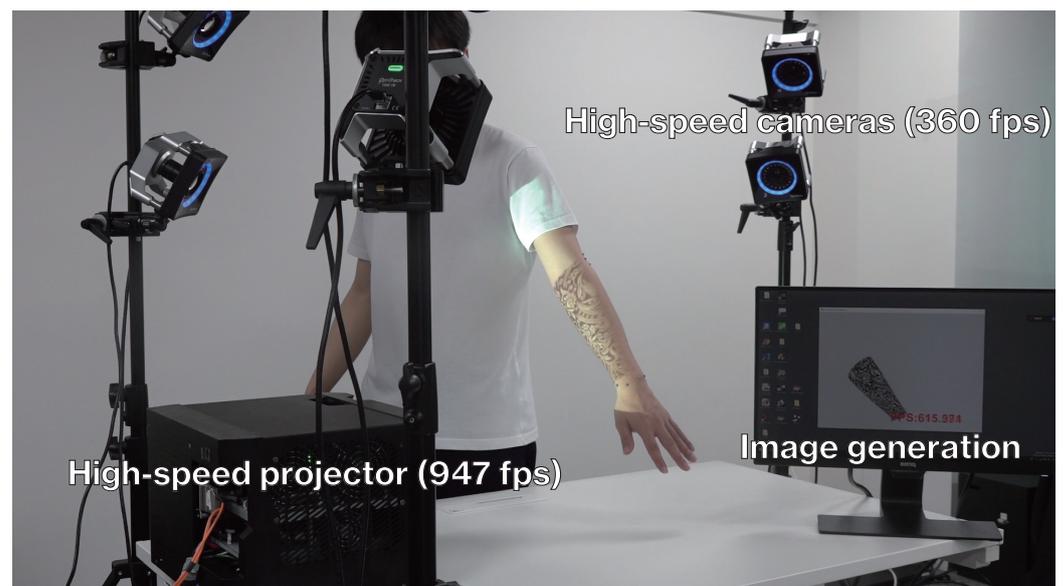


Figure 4. Experimental system configuration.

The stereo calibration between the projector and tracking system was performed in advance by collecting several pairs of corresponding 2D and 3D data points from the projector and tracking system, respectively.

The motion tracking system can use both passive and active markers to obtain the 3D position and orientation. For simplicity in the setup on the skin surface, we used passive markers to test the proposed system. We directly attached the markers to the skin surface around the wrist and elbow, using four markers per body part. The markers were registered as two rigid bodies in the motion tracking system to accurately determine wrist pose P_{wrist} and elbow pose P_{elbow} in real time. The origin points and local coordinate system of each rigid body are fixed to be the same as the corresponding joint by manual operation.

4.2. Model Preparation

The SMPL model is described by pose parameter θ and shape parameter β [38]. Compared with real-time tracking of pose parameter θ , shape parameter β can be obtained

in advance for each user. We used the open-source frameworks SMPLify [39] and OpenPose [53] to estimate shape parameter β . OpenPose can estimate a 2D human pose from a single image. Then, SMPLify can estimate the SMPL parameters from a single image with a 2D human pose.

The proposed accuracy compensation method was applied in the 2D space of the texture. However, the default UV map of the SMPL model hinders compensation [54]. The default UV map of the SMPL model for a whole human body including the arm is shown in the left graph of Figure 5. We used Blender [55], an open-source 3D creation software, to modify the whole-body UV map and extract the forearm region. The extracted UV map was modified such that u -axis and v -axis were respectively perpendicular and parallel to the axis of the forearm, and Euclidean distances in 2D texture coordinate space between all vertices were proportional to their Euclidean distances in 3D world space, obtaining the result shown in the right graph of Figure 5. The vertices $(u_{wrist}^*, v_{wrist}^*) \in (u^*, v^*)$ are around the wrist and also closest to the wrist among (u^*, v^*) . In the experiment, we set the v_{wrist}^* to be the zero vector.

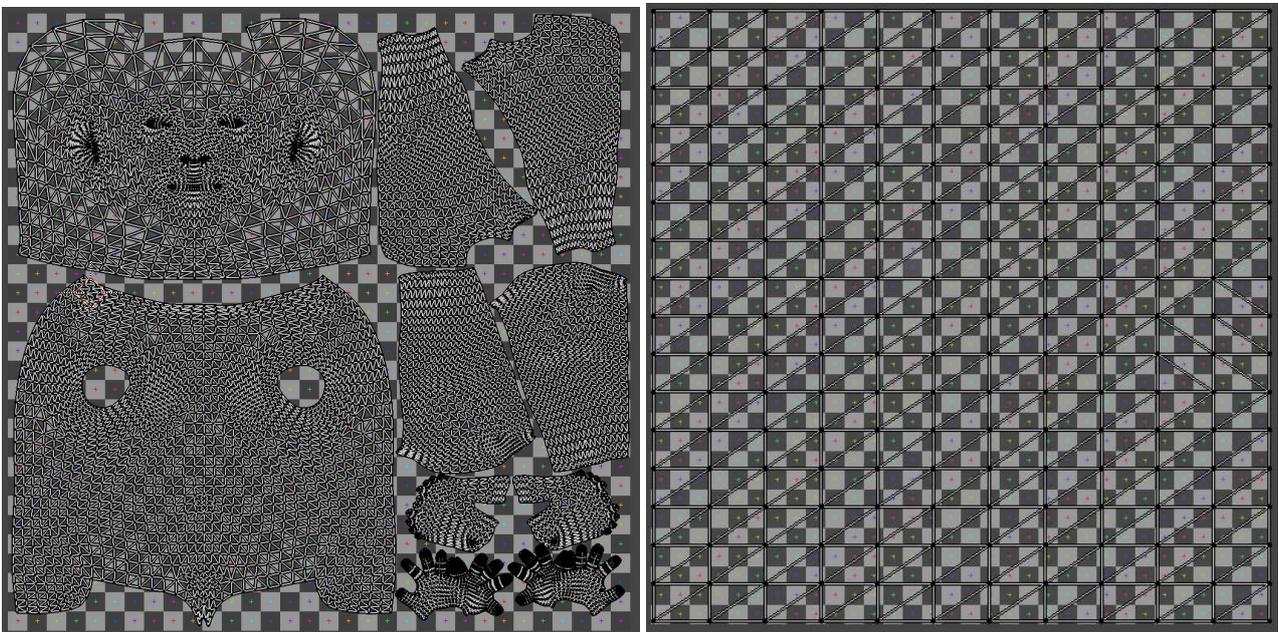


Figure 5. Default UV map of SMPL model for whole body [38] (left). UV map for the forearm after modification (right).

4.3. Data Preparation

We determined Equation (1) in advance using polynomial curve fitting of user-specific data. Six markers for data collection were placed along the forearm discretely, as shown in Figure 6. The 3D position of i -th marker M_i^t at time t was obtained by the motion tracking system. The data were processed as follows:

$$\begin{cases} o_i^t = |M_i^t - J_{wrist}^t - M_i^{t-1} + J_{wrist}^{t-1}| + o_i^{t-1}, \\ d_i^t = |M_i^t - J_{wrist}^t|, \end{cases} \quad (2)$$

where o_i^t is the cumulative offset that is perpendicular to the axis of the forearm, of i -th marker at time t , and d_i^t is the Euclidean distance between i -th marker and wrist at time t .

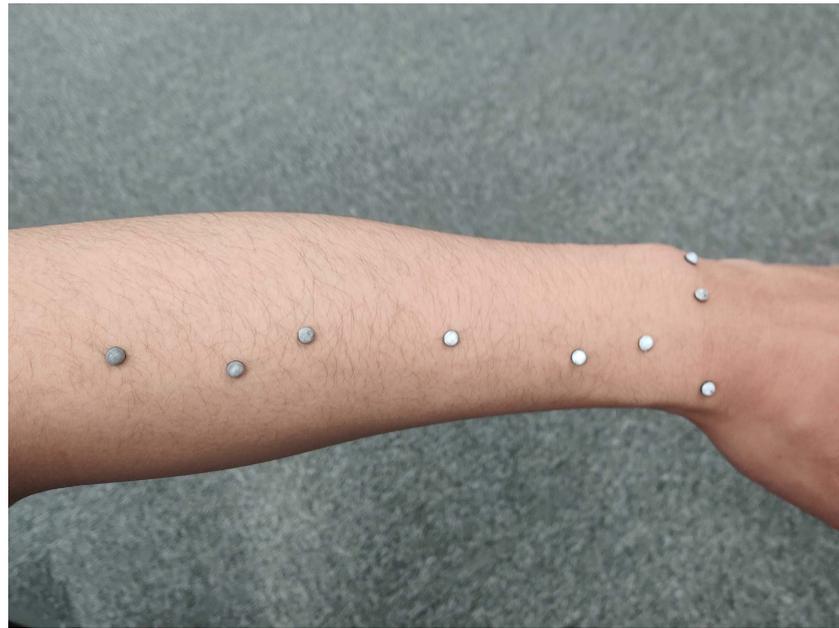


Figure 6. Markers attached to the forearm for data collection.

Over 10,000 datasets of (o_i^t, d_i^t) were collected for a user in the experiments, and the results of the following sixth-degree polynomial curve fitting were obtained:

$$o' = \begin{bmatrix} 3E^{-12} \\ 2E^{-10} \\ -2E^{-8} \\ -1E^{-6} \\ 7E^{-5} \\ 0.0014 \\ 0.0227 \end{bmatrix}^T \begin{bmatrix} \gamma_{wrist}^x 6 \\ \gamma_{wrist}^x 5 \\ \gamma_{wrist}^x 4 \\ \gamma_{wrist}^x 3 \\ \gamma_{wrist}^x 2 \\ \gamma_{wrist}^x \\ \gamma_{wrist}^x 1 \end{bmatrix} d + o^*, \tag{3}$$

where the x -axis is parallel to the axis of the forearm in wrist coordinate; $d \in \mathbb{R}^N$ is a vector that contains the Euclidean distances in 3D world space from the arm-surface model's vertices to the wrist; $o^* \in \mathbb{R}^N$ is a vector that contains the initial value before calculating cumulative offset of the arm-surface model's vertices; and $o' \in \mathbb{R}^N$ is a vector that contains the cumulative offset of the arm-surface model's vertices according to corresponding γ_{wrist}^x , d and o^* .

Since o' and o^* are perpendicular to the axis of the forearm and d is approximately parallel to the axis of the forearm, we accordingly obtained the conversion relationship as follows:

$$\begin{cases} o' - o^* = k(u' - u^*), \\ d = k(v^* - v_0^*), \end{cases} \tag{4}$$

where k is the parameter that converts pixels to millimeters; and $v_0^* \in \mathbb{R}^N$ is a vector that contains the v -axis value of the vertex $(u_{wrist}^*, v_{wrist}^*) \in (u_{wrist}^*, v_{wrist}^*)$ for each vertex $(u^*, v^*) \in (u^*, v^*)$, where $(u_{wrist}^*, v_{wrist}^*)$ and (u^*, v^*) have the minimum Euclidean distance in 2D texture coordinate space.

Due to the modification of the UV map in Section 4.2, v_0^* is the zero vector and Equation (4) can be simplified as:

$$\begin{cases} o' - o^* = k(u' - u^*), \\ d = kv^*. \end{cases} \tag{5}$$

Hence, after combining Equation (3) and Equation (5), the regression model $g(v^*, \gamma_{wrist})$ was obtained:

$$u' = \begin{bmatrix} 3E^{-12} \\ 2E^{-10} \\ -2E^{-8} \\ -1E^{-6} \\ 7E^{-5} \\ 0.0014 \\ 0.0227 \end{bmatrix}^T \begin{bmatrix} \gamma_{wrist}^x 6 \\ \gamma_{wrist}^x 5 \\ \gamma_{wrist}^x 4 \\ \gamma_{wrist}^x 3 \\ \gamma_{wrist}^x 2 \\ \gamma_{wrist}^x \\ \gamma_{wrist}^x 1 \end{bmatrix} v^* + u^*. \quad (6)$$

4.4. Results

Figure 2 (first and second rows) presents the projection results obtained before and after accuracy compensation. Black patterns were projected on the forearm surface, which contained five black crosses drawn using an ink marker as the ground truth of skin deformation. The misalignment between the projected patterns and black crosses in the first row of the figure is clear. The second row of the figure shows the improvement in projection accuracy achieved by applying the proposed compensation method, which suppresses the misalignment. We evaluated the accuracy before and after accuracy compensation by counting the numbers of crosses between the projected black patterns in the time sequence. The evaluation was conducted for 5 s, with 150 frames. The user performed wrist supination and pronation once during the evaluation. The more the crosses between the projected black patterns, the higher the accuracy. Figure 2 (bottom) presents the evaluation results. All of the five black crosses stay between the projected black patterns before wrist supination and pronation. Before accuracy compensation, the number of crosses between the projected black patterns drops to 2 and 0 after wrist supination and pronation, respectively. After accuracy compensation, the number of crosses remains 5 in both movements of wrist supination and pronation.

The developed system achieved a rate of 360 fps, with a bottleneck being caused by the tracking cameras. The tracking part causes a latency of 4 ms, obtained by the OptiTrack software [47]. The rendering part causes a latency of 3 ms, measured directly in the program. The projection part causes a latency of 3 ms, mentioned in [52]. Therefore, by simply adding the latency in each part, we can deduce that the overall latency between motion and projection was approximately 10 ms. At the achieved latency, users could hardly notice the misalignment between the projected image and arm surface, as illustrated in Figure 7. Thus, the proposed system provides fast and accurate projection.



Figure 7. Projection of tattoo images on the forearm. The user can hardly perceive any misalignment between the projected image and the arm even during high-speed rotations (first row) and translations (second row).

Figure 8 presents an example of tattoo image projection on the forearm with varying skin deformation using the developed system. The projected image can be selected by the user freely and is deformed realistically as the wrist and elbow joints rotate. The developed system can also enable flexible on-body user interfaces because both sides of the arm surface can be used. In the left photograph of Figure 9, the inner side of the forearm serves as a menu panel, and, in the right photograph, the outer side of the forearm serves

as a music player interface. Figure 10 presents another example of wound effects using a physical-based rendering texture.

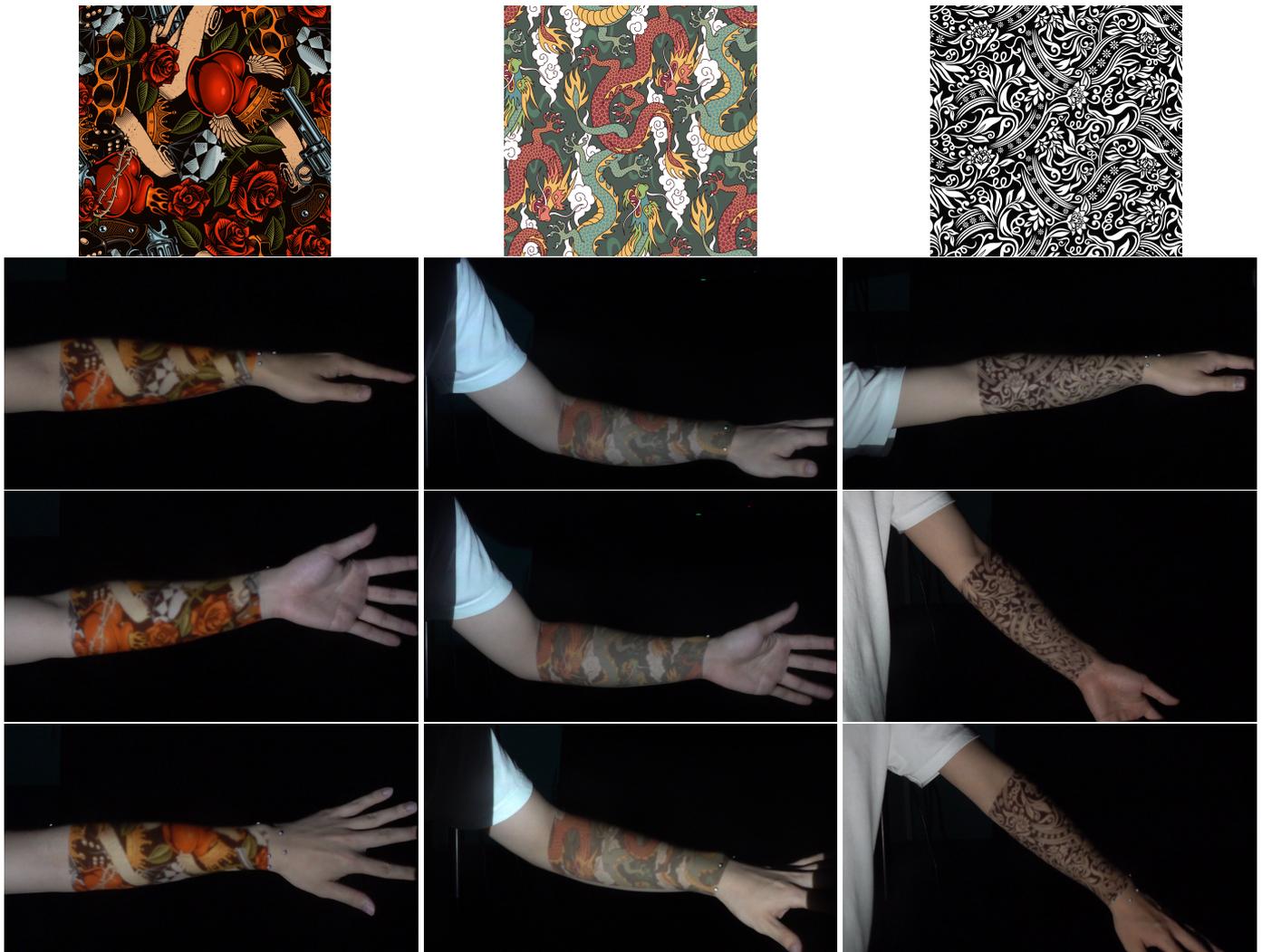


Figure 8. Proposed system projecting tattoo images (first row) onto the forearm surface under skin deformation (second to fourth rows).

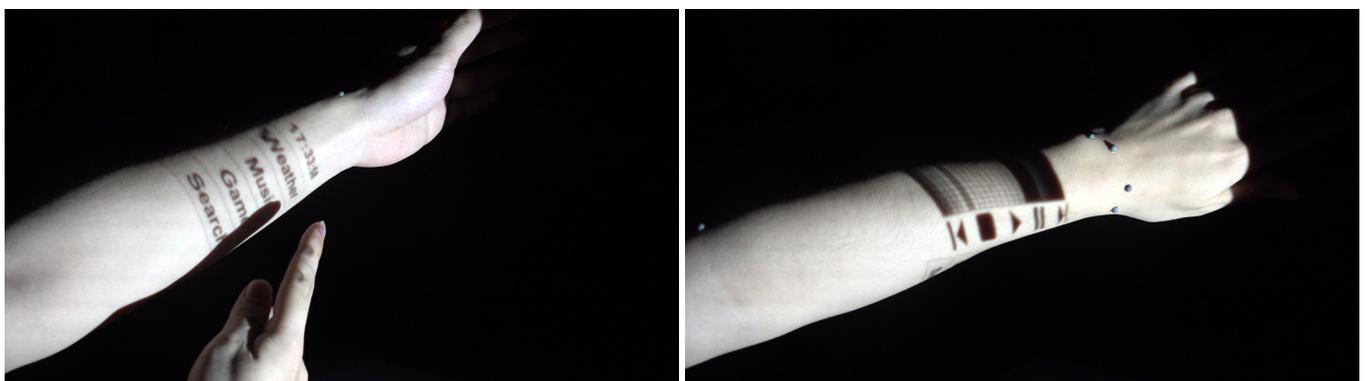


Figure 9. Both sides of the arm surface serving as on-body interfaces. The inner side of the arm serves as a menu panel (left), and the outer side serves as a music player interface (right).



Figure 10. Wound generated with physically based rendering texture (first column), wound effect projected on the forearm (second to fourth columns).

5. Discussion

Herein, we propose a high-speed dynamic projection mapping system for the human arm with realistic skin deformation. The developed system achieves a latency of 10 ms between the motion and projection, almost imperceptible by the human vision even under high-speed rotations and translations, as shown in Figure 7. This is the first projection mapping system that can realistically handle human arm surfaces with a very low latency, although the achieved latency is still a few milliseconds higher than that reported in [20]. We believe that our system can be beneficial for applications in many fields such as fashion, user interfaces, and education. For example, it is difficult to remove a tattoo after getting tattooed; however, by using the proposed system, people can enjoy the realistic tattoo digitally and intuitively find their favorite tattoo design, as shown in Figure 8. On-body interfaces can become more flexible and interactive owing to realistic skin deformation and low latency as shown in Figure 9. Special-effect makeup, typically considered complicated and costly, can be easily applied in the film industry, as shown in Figure 10. Physiotherapy students can gain a better understanding by directly projecting the anatomical structures on the body.

Nevertheless, the proposed system has various limitations that remain to be addressed. For instance, the system uses a marker-based method, which complicates the setup. The registration of the joints requires careful manual operation to ensure that the coordinate systems of the registered rigid bodies and the joints are consistent. The error between the joints and the registered rigid bodies may lead to severe misalignments in the projection and error of the skin deformation. We believe that the problem can be alleviated if more joints are registered in the tracking system. The relationship of the skeletal structure can provide information to improve the registration accuracy.

In addition, we only focused on the arm surface in this study, and the proposed system should be extended to other body parts to widen the applicability. Accordingly, the proposed accuracy compensation method should be carefully designed and modified because the neighboring joints of vertices of other body parts could be different from the forearm. For example, the torso of a body has more than two neighboring joints.

The system described herein is user-specific in that the regression model was prepared in advance for a specific user. This limitation narrows the applications, and the regression model should also be extended to multiple persons. It is known that the SMPL model represents different body shapes through shape parameters [38]. If we collect the data of multiple persons with different shape parameters and regress the collected data with the shape parameters, the regression model can possibly be extended to multiple persons. In future work, we will implement the system using a marker-less method and widen its application to multiple persons simultaneously using learning-based methods.

Author Contributions: Conceptualization, H.-L.P. and Y.W.; methodology, H.-L.P. and Y.W.; software, H.-L.P.; validation, H.-L.P. and Y.W.; formal analysis, H.-L.P. and Y.W.; investigation, H.-L.P. and Y.W.; writing—original draft preparation, H.-L.P.; writing—review and editing, H.-L.P. and Y.W.; supervision, Y.W.; project administration, H.-L.P. and Y.W.; and funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by JST Accelerated Innovation Research Initiative Turning Top Science and Ideas into High-Impact Values (JPMJAC15F1) and JSPS Grant-in-Aid for Transformative Research Areas (20H05959).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bimber, O.; Raskar, R. *Spatial Augmented Reality: Merging Real and Virtual Worlds*; CRC Press: Boca Raton, FL, USA, 2005.
2. Mine, M.R.; van Baar, J.; Grundhofer, A.; Rose, D.; Yang, B. Projection-Based Augmented Reality in Disney Theme Parks. *Computer* **2012**, *45*, 32–40. [CrossRef]
3. Raskar, R.; Welch, G.; Low, K.L.; Bandyopadhyay, D. Shader Lamps: Animating Real Objects With Image-Based Illumination. In Proceedings of the 12th Eurographics Workshop on Rendering Techniques, London, UK, 25–27 June 2001; pp. 89–102.
4. Siegl, C.; Colaïanni, M.; Thies, L.; Zollhöfer, M.; Izadi, S.; Stamminger, M.; Bauer, F. Real-Time Pixel Luminance Optimization for Dynamic Multi-Projection Mapping. *ACM Trans. Graph.* **2015**, *34*, 1–11. [CrossRef]
5. Narita, G.; Watanabe, Y.; Ishikawa, M. Dynamic Projection Mapping onto Deforming Non-Rigid Surface Using Deformable Dot Cluster Marker. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1235–1248. [CrossRef] [PubMed]
6. Miyashita, L.; Watanabe, Y.; Ishikawa, M. MIDAS Projection: Markerless and Modelless Dynamic Projection Mapping for Material Representation. *ACM Trans. Graph.* **2018**, *37*, 1–12. [CrossRef]
7. Nomoto, T.; Li, W.; Peng, H.L.; Watanabe, Y. Dynamic Projection Mapping with Networked Multi-Projectors Based on Pixel-Parallel Intensity Control. In Proceedings of the SIGGRAPH Asia 2020 Emerging Technologies (SA '20), Online, 4–13 December 2020; Virtual Event Republic of Korea.
8. Bermanno, A.H.; Billeter, M.; Iwai, D.; Grundhofer, A. Makeup Lamps: Live Augmentation of Human Faces via Projection. *Comput. Graph. Forum* **2017**, *36*, 311–323. [CrossRef]
9. Siegl, C.; Lange, V.; Stamminger, M.; Bauer, F.; Thies, J. FaceForge: Markerless Non-Rigid Face Multi-Projection Mapping. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 2440–2446. [CrossRef] [PubMed]
10. Ink Mapping: Video Mapping Projection on Tattoos, by Oskar & Gaspar. Available online: <https://vimeo.com/143296099> (accessed on 7 March 2021).
11. Pourjafarian, N.; Koelle, M.; Fruchard, B.; Mavali, S.; Klamka, K.; Groeger, D.; Strohmeier, P.; Steimle, J. BodyStylus: Freehand On-Body Design and Fabrication of Epidermal Interfaces. In Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems (CHI '21), Yokohama, Japan, 8–13 May 2021.
12. Xiao, R.; Cao, T.; Guo, N.; Zhuo, J.; Zhang, Y.; Harrison, C. LumiWatch: On-Arm Projected Graphics and Touch Input. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Montreal, QC, Canada, 21–26 April 2018; pp. 1–11.
13. Harrison, C.; Benko, H.; Wilson, A.D. OmniTouch: Wearable Multitouch Interaction Everywhere. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11), Santa Barbara, CA, USA, 16–19 October 2011; pp. 441–450.
14. Sodhi, R.; Benko, H.; Wilson, A. LightGuide: Projected Visualizations for Hand Movement Guidance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12), Austin, TX, USA, 5–10 May 2012; pp. 179–188.
15. Harrison, C.; Tan, D.; Morris, D. Skinput: Appropriating the Body as an Input Surface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10), Atlanta, GA, USA, 10–15 April 2010; pp. 453–462.
16. Gannon, M.; Grossman, T.; Fitzmaurice, G. ExoSkin: On-Body Fabrication. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16), San Jose, CA, USA, 7–12 May 2016; pp. 5996–6007.
17. Gannon, M.; Grossman, T.; Fitzmaurice, G. Tactum: A Skin-Centric Approach to Digital Design and Fabrication. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), Seoul, Korea, 18–23 April 2015; pp. 1779–1788.
18. Hoang, T.; Reinoso, M.; Joukhadar, Z.; Vetere, F.; Kelly, D. Augmented Studio: Projection Mapping on Moving Body for Physiotherapy Education. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17), Denver, CO, USA, 6–11 May 2017; pp. 1419–1430.
19. Ni, T.; Karlson, A.K.; Wigdor, D. AnatOnMe: Facilitating Doctor-Patient Communication Using a Projection-Based Handheld Device. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), Vancouver, BC, Canada, 7–12 May 2011; pp. 3333–3342.
20. Ng, A.; Lepinski, J.; Wigdor, D.; Sanders, S.; Dietz, P. Designing for Low-Latency Direct-Touch Input. In Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST'12), Vancouver, BC, Canada, 7–12 May 2012; pp. 453–464.
21. Peng, H.L.; Watanabe, Y. High-Speed Human Arm Projection Mapping with Skin Deformation. In Proceedings of the SIGGRAPH Asia 2020 Emerging Technologies (SA '20), Online, 4–13 December 2020; Virtual Event Republic of Korea.
22. Sorkine, O.; Alexa, M. As-Rigid-as-Possible Surface Modeling. In Proceedings of the Fifth Eurographics Symposium on Geometry Processing (SGP '07), Barcelona, Spain, 4–6 July 2007; pp. 109–116.
23. Sumner, R.W.; Schmid, J.; Pauly, M. Embedded Deformation for Shape Manipulation. *ACM Trans. Graph.* **2007**, *26*, 80-es. [CrossRef]

24. Szeliski, R.; Lavalley, S. Matching 3D anatomical surfaces with non-rigid deformations using octree-splines. In Proceedings of the IEEE Workshop on Biomedical Image Analysis, Seattle, WA, USA, 24–25 June 1994; pp. 144–153.
25. Newcombe, R.A.; Fox, D.; Seitz, S.M. DynamicFusion: Reconstruction and tracking of non-rigid scenes in real-time. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 343–352.
26. Zollhöfer, M.; Nießner, M.; Izadi, S.; Rehmann, C.; Zach, C.; Fisher, M.; Wu, C.; Fitzgibbon, A.; Loop, C.; Theobalt, C.; et al. Real-Time Non-Rigid Reconstruction Using an RGB-D Camera. *ACM Trans. Graph.* **2014**, *33*, 1–12. [[CrossRef](#)]
27. Li, C.; Guo, X. Topology-Change-Aware Volumetric Fusion for Dynamic Scene Reconstruction. In Proceedings of the Computer Vision—ECCV 2020, Glasgow, UK, 23–28 August 2020; pp. 258–274.
28. Guo, K.; Xu, F.; Wang, Y.; Liu, Y.; Dai, Q. Robust Non-rigid Motion Tracking and Surface Reconstruction Using L0 Regularization. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 3083–3091.
29. Li, W.; Zhao, S.; Xiao, X.; Hahn, J.K. Robust Template-Based Non-Rigid Motion Tracking Using Local Coordinate Regularization. In Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, 1–5 March 2020; pp. 390–399.
30. Gao, W.; Tedrake, R. SurfelWarp: Efficient Non-Volumetric Single View Dynamic Reconstruction. In Proceedings of the Robotics: Science and Systems (RSS '18), Pittsburgh, PA, USA, 26–30 June 2018.
31. Wei, X.; Zhang, P.; Chai, J. Accurate Realtime Full-Body Motion Capture Using a Single Depth Camera. *ACM Trans. Graph.* **2012**, *31*, 1–12. [[CrossRef](#)]
32. Le, B.H.; Deng, Z. Smooth Skinning Decomposition with Rigid Bones. *ACM Trans. Graph.* **2012**, *31*. [[CrossRef](#)]
33. Kavan, L.; Collins, S.; O'Sullivan, C. Automatic Linearization of Nonlinear Skinning. In Proceedings of the 2009 Symposium on Interactive 3D Graphics and Games (I3D '09), Boston, MA, USA, 27 February–1 March 2009; pp. 49–56.
34. Lewis, J.P.; Corder, M.; Fong, N. Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-Driven Deformation. In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00), New Orleans, LA, USA, 23–28 July 2000; pp. 165–172.
35. Kurihara, T.; Miyata, N. Modeling Deformable Human Hands from Medical Images. In Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '04), Grenoble, France, 27–29 August 2004; pp. 355–363.
36. Anguelov, D.; Srinivasan, P.; Koller, D.; Thrun, S.; Rodgers, J.; Davis, J. SCAPE: Shape Completion and Animation of People. *ACM Trans. Graph.* **2005**, *24*, 408–416. [[CrossRef](#)]
37. Hirshberg, D.; Loper, M.; Rachlin, E.; Black, M. Coregistration: Simultaneous alignment and modeling of articulated 3D shape. In *European Conference on Computer Vision (ECCV)*; LNCS 7577, Part IV; Springer: Berlin/Heidelberg, Germany, 2012; pp. 242–255.
38. Loper, M.; Mahmood, N.; Romero, J.; Pons-Moll, G.; Black, M.J. SMPL: A Skinned Multi-Person Linear Model. *ACM Trans. Graph.* **2015**, *34*. [[CrossRef](#)]
39. Bogo, F.; Kanazawa, A.; Lassner, C.; Gehler, P.; Romero, J.; Black, M.J. Keep It SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image. In Proceedings of the Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 8–16 October 2016; pp. 561–578.
40. Lassner, C.; Romero, J.; Kiefel, M.; Bogo, F.; Black, M.J.; Gehler, P.V. Unite the People: Closing the Loop Between 3D and 2D Human Representations. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4704–4713.
41. Joo, H.; Simon, T.; Sheikh, Y. Total Capture: A 3D Deformation Model for Tracking Faces, Hands, and Bodies. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8320–8329.
42. Kanazawa, A.; Black, M.J.; Jacobs, D.W.; Malik, J. End-to-End Recovery of Human Shape and Pose. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7122–7131.
43. Pavlakos, G.; Zhu, L.; Zhou, X.; Daniilidis, K. Learning to Estimate 3D Human Pose and Shape from a Single Color Image. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 459–468.
44. Güler, R.A.; Neverova, N.; Kokkinos, I. DensePose: Dense Human Pose Estimation in the Wild. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7297–7306.
45. Wei, L.; Huang, Q.; Ceylan, D.; Vouga, E.; Li, H. Dense Human Body Correspondences Using Convolutional Networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1544–1553.
46. Yao, P.; Fang, Z.; Wu, F.; Feng, Y.; Li, J. DenseBody: Directly Regressing Dense 3D Human Pose and Shape From a Single Color Image. *arXiv* **2019**, arXiv:1903.10153.
47. OptiTrack - Motion Capture Systems. Available online: <https://optitrack.com/> (accessed on 18 March 2021).
48. Vicon | Award Winning Motion Capture Systems. Available online: <https://www.vicon.com/> (accessed on 18 March 2021).
49. Murai, A.; Hong, Q.Y.; Yamane, K.; Hodgins, J.K. Dynamic Skin Deformation Simulation Using Musculoskeletal Model and Soft Tissue Dynamics. *Comput. Vis. Media* **2016**, *3*, 49–60. [[CrossRef](#)]

50. Watanabe, Y.; Narita, G.; Tatsuno, S.; Yuasa, T.; Sumino, K.; Ishikawa, M. High-speed 8-bit Image Projector at 1000 fps with 3 ms Delay. In Proceedings of the International Display Workshops IDW '15, Otsu, Japan, 9–11 December 2015; pp. 1064–1065.
51. Panasonic Showcases New Video Technologies and Industry Solutions at InfoComm 2017. Available online: <https://news.panasonic.com/global/topics/2017/48029.html> (accessed on 15 March 2021).
52. Watanabe, Y.; Ishikawa, M. High-Speed and High-Brightness Color Single-Chip DLP Projector Using High-Power LED-Based Light Sources. *IDW* **2019**, *19*, 1350–1352. [[CrossRef](#)]
53. Cao, Z.; Hidalgo Martinez, G.; Simon, T.; Wei, S.; Sheikh, Y.A. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 172–186. [[CrossRef](#)] [[PubMed](#)]
54. SMPL. Available online: <https://smpl.is.tue.mpg.de/> (accessed on 31 March 2021).
55. blender.org—Home of the Blender Project—Free and Open 3D Creation Software. Available online: <https://www.blender.org/> (accessed on 17 March 2021).