

Function Based Sim-to-Real Learning for Shape Control of Deformable Free-form Surfaces

Yingjun Tian[†], Guoxin Fang[‡], Renbo Su[†], Weiming Wang[†], Simeon Gill[†],
Andrew Weightman[†] and Charlie C.L. Wang^{†*}

[†]The University of Manchester

[‡]The Chinese University of Hong Kong

(*Corresponding Author; Email: changling.wang@manchester.ac.uk)

Abstract—For the shape control of deformable free-form surfaces, simulation plays a crucial role in establishing the mapping between the actuation parameters and the deformed shapes. The differentiation of this forward kinematic mapping is usually employed to solve the inverse kinematic problem for determining the actuation parameters that can realize a target shape. However, the free-form surfaces obtained from simulators are always different from the physically deformed shapes due to the errors introduced by hardware and the simplification adopted in physical simulation. To fill the gap, we propose a novel deformation function based sim-to-real learning method that can map the geometric shape of a simulated model into its corresponding shape of the physical model. Unlike the existing sim-to-real learning methods that rely on completely acquired dense markers, our method accommodates sparsely distributed markers and can resiliently use all captured frames – even for those in the presence of missing markers. To demonstrate its effectiveness, our sim-to-real method has been integrated into a neural network-based computational pipeline designed to tackle the inverse kinematic problem on a pneumatically actuated deformable mannequin.

I. INTRODUCTION

Soft robotic systems have demonstrated the ability of dynamically adapting their surfaces in response to programmed actuation or environmental stimuli. Existing prototypes were mainly developed for haptic interfaces (ref. [18, 38, 35, 28, 27]). For the fabrication of customized garments, a soft robotic mannequin with deformable free-form surface has been developed in [39], which is driven by pneumatic actuators to mimic the front shapes of different human bodies on the same mannequin. The research work presented in this paper is to realize a similar function of soft deformable robot, which is to serve as a reusable mould with programmable shape for sustainable garment fabrication (e.g., [32, 34]). Specifically, by giving a target shape of an individual as S^t , the deformable mannequin is expected to deform into a shape S that minimizes the shape difference (see Fig.1(a) for examples). The geometric errors are evaluated by capturing the deformed shapes of the robotic mannequin with the help of a structure-light based 3D scanner (see Fig.1(b)). Shape control of a deformable free-form surface is challenging, and the reasons for this are analyzed below with a review of related work in the literature.

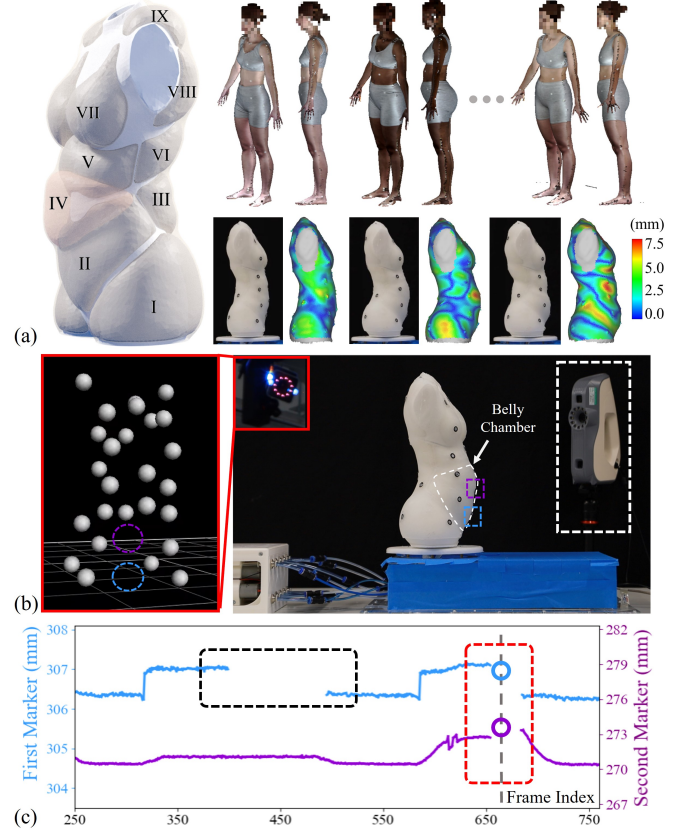


Fig. 1: Motivation and challenges. (a) A soft robotic mannequin with nine air chambers can be deformed by pneumatic actuation to realize the target body shape of individual customers. Shape approximation errors are visualized by colors. (b) The hardware setup includes a 3D scanner and a motion capture system. The zoom-in view shows the captured markers, where two markers on the belly were missed. (c) Positions of two markers captured while deforming the mannequin.

A. Kinematics for Shape Control

An essential shape control problem is to compute the actuation (denoted by a vector $\mathbf{a}$) that can drive the robotic

free-form surface $\mathcal{S}$ to approach a target shape $\mathcal{S}^t$. This is an inverse computing problem, and the shape control of free-form surfaces mainly focused on the configuration space (e.g., [22]). It presents much higher *degrees-of-freedom* (DoFs) compared to the *inverse kinematics* (IK) problems for articulated robots that generally focus on the task space. This leads to high complexity in both the forward shape estimation (i.e., predict the shape $\mathcal{S}(\mathbf{a})$) and the IK computation to determine the value of $\mathbf{a}$. An intuitive solution is to apply an iteration-based IK solver with Jacobian physically evaluated on hardware setup [39, 43]. However, this approach suffers from high costs in both computation and data acquisition, and its performance heavily relies on the robustness of data acquisition and the accuracy of shapes captured on the physical setup.

In the existing research works of kinematics computing for deformable objects, the high DoFs in the configuration space are typically modeled either by reduced analytical models or by discrete numerical models. Comprehensive reviews can be found in the papers of Armanini *et al.* [7] and Arriola-Rios *et al.* [8]. Numerical simulation is commonly employed as a general solution for modeling deformable objects with complex shapes, and it can be adapted in both the cases with internal actuation [14] and those with external interactions [17]. However, high-precision numerical simulation in general cannot be computed efficiently enough to realize a fast IK solver. Moreover, numerical simulation invariably exhibits a prediction error compared to physical performance, a discrepancy often referred to as the ‘reality gap’ [24].

A widely adopted strategy to mitigate the issue of efficiency is to first collect a dataset using the numerical solution and train a *neural network* (NN) for the forward shape estimation of the deformable object [21, 9]. Subsequently, NN-based sim-to-real transfer [13, 44, 26, 16] is applied to achieve high precision while maintaining reasonable efficiency in data acquisition – i.e., only a small set of physical data is needed. When both the forward kinematic computing and the sim-to-real transfer are realized by neural networks, the gradients of IK objectives can be efficiently and effectively computed by differentiating the networks [12, 25]. In short, fast IK can be achieved.

B. Challenges of Sim-to-Real Transfer

Sim-to-real approaches have been developed for controlling robots with compliance [26, 37] and for deformable object manipulation [31, 30]. These works generally adopt rich physical information obtained from well-captured frames with constant DOFs as training data. Less effort has been focused on the sim-to-real problem on free-form surfaces using sparse and incomplete sets of markers. There is an effort to use a 3D scanner to directly obtain the surface shapes as point clouds [19], which however suffers from the problems of inefficiency and unstable boundary regions. Moreover, the scanned shapes at different time currents have no correspondence with each other, which needs a lot of extra effort for non-rigid registration [11].

Using a motion capture system can precisely capture the position of markers attached to a free-form surface [6, 36]. However, only sparse data can be obtained, representing a reduced order of information on the shape of the free-form surface. The motion capture system can also suffer from the problems of marker missing due to camera view occlusion – especially when large deformation is presented (see Fig.1(b, c) for an example). In this paper, we propose a function-based learning method to address the sim-to-real challenge, and thereby also solve the fast IK problem for deformable free-form surfaces.

C. Our Method

To facilitate the neural network-based fast IK computing, we propose a resilient sim-to-real learning method, which can fully leverage the entire dataset of sparse marker positions obtained from the physical setup. It is a new learning architecture that learns the function space of shape deformation represented by *radial basis functions* (RBF). Specifically, each simulated free-form surface is approximated by a B-spline surface, the control points of which are employed as the shape descriptor in a lower dimension. The input of our sim-to-real network is the control points of a simulated body shape, and the output is the coefficients of RBFs that define a spatial warping function $\Phi(\mathbf{p})$ by using the positions of ‘virtual’ markers as kernel centers. The sim-to-real network can be trained from sparse and even incompletely captured sets of markers.

The space warping function $\Phi(\mathbf{p})$ is in a closed-form and gives a continuous mapping from the shape of a simulated model to the shape of a fixed model that matches the reality. Working together with the forward kinematics prediction that is trained by simulation, this forms an end-to-end NN-based pipeline that can precisely predict the physical shapes of robotic deformable surfaces. This computation pipeline is analytically differentiable so that the gradient-based iteration for IK can be computed efficiently. The effectiveness of our fast IK solver has been verified on a physical setup as shown in Fig.1(b). When testing on a dataset with 100 individual shapes, our IK solver takes about 7.98 seconds in average to generate the actuation parameters. The quality of IK solutions has been verified by scanning the physical shape $\mathcal{S}^p$ to compare with the input target shape $\mathcal{S}^t$. The resultant error color maps for different individuals are shown in Fig.1(a).

The technique contributions of our work are summarized as follows:

- A resilient sim-to-real learning method for deformable free-form surfaces that can employ sparse and incomplete sets of markers to bridge the reality gap by RBF-based spatial warping functions;
- A neural network-based method that can efficiently solve the inverse kinematics problem with its performance physically verified on a pneumatically actuated deformable mannequin.

Experimental tests were conducted on a variety of individual shapes presented in the CAESAR dataset [1]. The shape approximation errors were physically evaluated on a deformable

TABLE I: List of Notations

Symbol	Description	Symbol	Description
$\mathbf{a}$	Actuation parameters (i.e., chamber pressures)	$\mathcal{S}^t$	The target body shape to be achieved by IK
$\mathcal{S}$	The simulated shape of the deformable surface	$\mathcal{S}^*$	The fixed shape after applying our <i>sim-to-real</i> method
$\mathcal{S}^c$	The compact shape descriptor (i.e., control points) of $\mathcal{S}$	$\mathcal{S}^p$	The scanned 3D shape on the physical setup
$\mathcal{N}_{fk}$	Forward kinematics network that predicts $\mathcal{S}^c$ from $\mathbf{a}$	$\mathcal{N}_{mk}$	The baseline network predicting positions of markers
$\mathcal{N}_{rbf}$	Function prediction network that predicts γ by $\mathcal{S}^c$	$\mathcal{N}_{s2r}$	Combination of $\mathcal{N}_{rbf}$ and $\Phi(\cdot)$
$\Phi(\cdot)$	The space warping function used to deform $\mathcal{S}$ to $\mathcal{S}^*$	$\gamma = [\alpha_0, \alpha_1, \dots, \beta_i]$	Variables to determine the function space
$\{\mathbf{q}_i\}$	Centers of Gaussian kernels (also markers on $\mathcal{S}$) in $\Phi(\mathbf{p})$	$\mathbf{p}, \mathbf{p}^* \in \mathbb{R}^3$	The surface point on $\mathcal{S}$ and $\mathcal{S}^*$ respectively
$(\mathbf{R}, \mathbf{t})$	Rotation matrix and translation vector to re-pose $\mathcal{S}^t$	$\mathbf{c}^*$	Closet point of $\mathbf{p}^*$ on $\mathcal{S}^t$
(u_p, v_p)	A point in the parametric domain of B-spline surface	$\mathbf{B}(\cdot)$	B-Spline function as decoder mapping (u_p, v_p) to $\mathbf{p}$

mannequin with 1:2 ratio to the full size. Detail statistics are given in the supplementary document.

II. LEARNING FOR SIM-TO-REAL TRANSFER

This section introduces our function based sim-to-real learning for shape prediction. Given a set of points $\{\mathbf{p} \in \mathbb{R}^3\}$ on the free-form surface of a simulated model $\mathcal{S}$, we propose a method to learn a network $\mathcal{N}_{s2r}$ that can predict a deformation function mapping these points onto a fixed model $\mathcal{S}^*$ (i.e., the fixed shape that has eliminated the sim-to-real gap). All notations used in this paper are summarized in Table I. The shape difference between $\mathcal{S}^*$ and the physical shape $\mathcal{S}^p$ should be minimized by the sim-to-real transfer. $\mathcal{N}_{s2r}$ is expected to be a general network that can adaptively handle different simulated shapes as input.

A. RBF-based Spatial Warping

Considering the gap between the simulated shape $\mathcal{S}$ and the physical shape $\mathcal{S}^p$, we propose to model it as a continuous spatial warping by the *radial basis functions* (RBF) as follows.

$$\mathbf{p}^* = \Phi(\mathbf{p}) = \alpha_0 + A\mathbf{p} + \sum_{i=1}^N \beta_i e^{-c\|\mathbf{p}-\mathbf{q}_i\|^2} \quad (1)$$

with N being the number of kernels employed for the space warping and $A = [\alpha_1, \alpha_2, \alpha_3]$. The centers of Gaussian kernels, $\{\mathbf{q}_i\}$, are sparsely selected to cover the entire surface of a simulated model. The value of coefficient c controlling the width of the Gaussian kernel is chosen as $(3.0e - 5)$ by experiments. In our practice, we use the locations of markers for motion capture as the kernel centers. We remark that $\gamma = [\alpha_0, \alpha_1, \dots, \beta_i]$ are variables of function space, which can be determined by solving a linear system [42]. Note that $\{\mathbf{q}_i\}$ are virtual markers defined on the simulated models, where the free-form surface of a simulated model has been parameterized into u, v -domain and represented by a B-spline surface. The same marker on different simulated models will have the same (u, v) parameters.

With a determined warping function $\Phi(\mathbf{p})$, we are able to map any point $\mathbf{p} \in \mathcal{S}$ to its corresponding point $\mathbf{p}^* \in \mathcal{S}^*$ as demonstrated in Fig.2. However, different warping functions need to be employed for different simulated shapes. A straight-forward solution is to train a network $\mathcal{N}_{mk}$ for predicting the positions of physical markers by using the set of virtual marker positions as $\{\mathbf{q}^*\} = \mathcal{N}_{mk}(\{\mathbf{q}\})$. As a result, the predicted

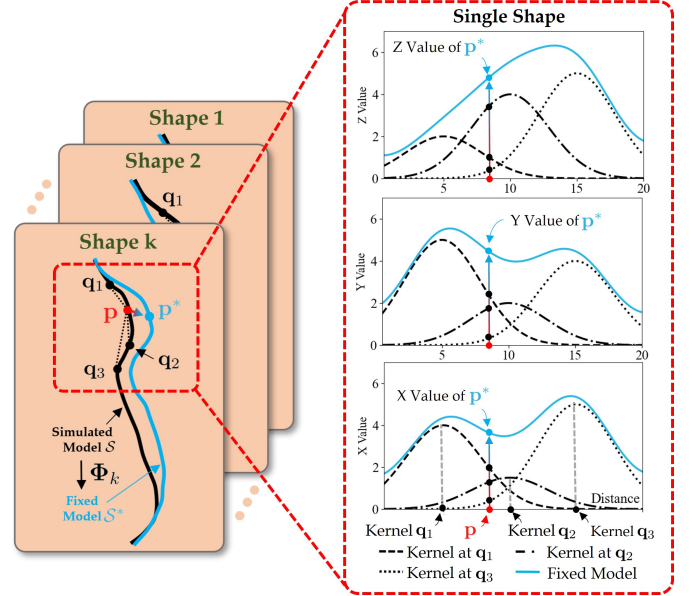


Fig. 2: The space warping $\Phi(\cdot)$ is built on the *radial basis functions* (RBF) with kernel centers located on the free-form surface of a simulated model. Different warping functions need to be determined for different simulated shapes.

positions of physical markers and the positions of virtual markers can be used to determine the value of γ therefore also the warping function – see Fig.3(a) for an illustration of this marker-prediction based pipeline. There are two major problems with this sim-to-real strategy:

- This is not an end-to-end approach as it needs to solve a linear equation system when establishing a warping function for each simulated shape;
- Training the marker prediction network $\mathcal{N}_{mk}$ needs sets of fully captured markers, which can be challenging on largely deformed models as discussed above.

A function space learning-based method is introduced below to solve these issues.

B. NN-based Deformation Function Learning

Only learning to predicting the positions of markers does not provide good shape-control in the regions between markers, which relies on dense and complete sets of captured markers

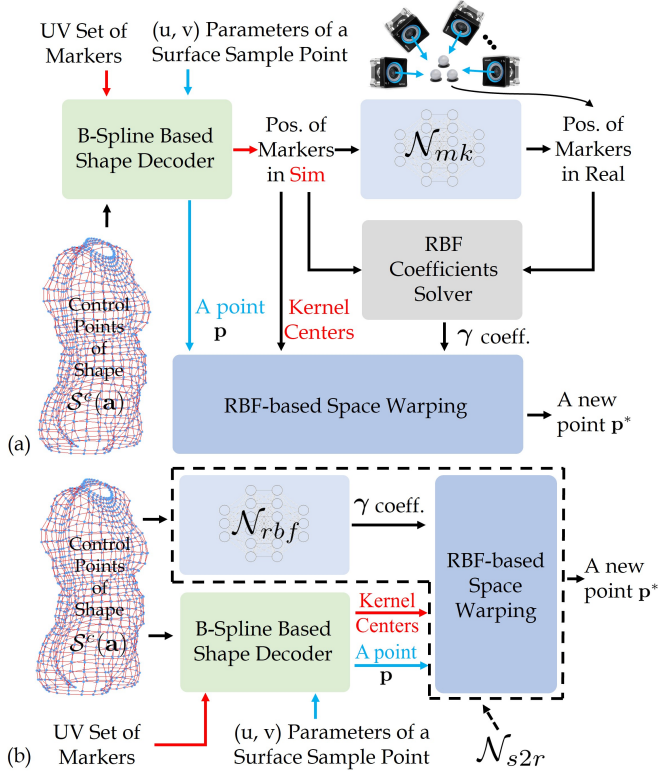


Fig. 3: Different sim-to-real strategies by using RBF-based spatial warping – (a) a marker-prediction based pipeline (the baseline method that can only use frames with complete set of markers) that needs to solve different linear systems for different simulated shapes and (b) our function-prediction based end-to-end pipeline (that can resiliently use all frames – even for those in the presence of missing markers) where the components circled by dash lines form a sim-to-real network $\mathcal{N}_{s2r}$.

as input to train $\mathcal{N}_{mk}$. Differently, we propose a function-prediction based pipeline, which is end-to-end after training. Moreover, all frames of motion capture data – even those in the presence of missing markers can be employed for training.

A network $\mathcal{N}_{rbf}$ is introduced to predict the value of the function variable γ according to the simulated shape $\mathcal{S}$. However, using all sample points on the surface of $\mathcal{S}$ may bring in redundant information, and therefore require a complex network for the prediction. We employ the control points of B-spline surface (denoted by $\mathcal{S}^c$) as the shape descriptor of $\mathcal{S}$, with which any point in the parametric domain with the parameter (u_p, v_p) can be mapped to a point $\mathbf{p} \in \mathbb{R}^3$ as $\mathbf{p} = \mathbf{B}(u_p, v_p, \mathcal{S}^c)$. This function $\mathbf{B}(\cdot)$ is named as the B-spline based shape decoder in our computational pipeline (see the green block in Fig.3 for an example).

Our function-prediction based sim-to-real pipeline is formed by three building blocks: 1) the network $\mathcal{N}_{rbf}$, 2) the B-spline based shape decoder and 3) the RBF-based space warping (i.e., Eq.(1)), where the diagram has been shown in Fig.3(b). When actuating the deformable mannequin into different shapes $\{\mathcal{S}_j\}$

by $\{\mathbf{a}_j\}$, the function-prediction network $\mathcal{N}_{rbf}$ can be trained by the position of all captured markers $\mathbf{q}_{i,j}$ together with their corresponding parameters (u_i, v_i) and the control points $\mathcal{S}_j^c$. Here $\mathbf{q}_{i,j}$ denotes the position of the i -th marker captured on the j -th frame. The loss function for learning $\mathcal{N}_{rbf}(\cdot)$ is defined as

$$\mathcal{L}_{s2r} = \sum_j \sum_i \|\Phi_j(\mathbf{B}(u_i, v_i, \mathcal{S}_j^c)) - \mathbf{q}_{i,j}\|^2 \quad (2)$$

with the function variable γ of Φ_j predicted by the network as $\gamma = \mathcal{N}_{rbf}(\mathcal{S}_j^c)$ and the kernel centers of Φ_j given as $\{\mathbf{B}(u_i, v_i, \mathcal{S}_j^c)\}$. Only captured markers are included to evaluate the loss $\mathcal{L}_{s2r}$ defined in Eq.(2). This pipeline supports resilient sim-to-real learning that can use all frames of captured data by motion camera.

Given a well trained network $\mathcal{N}_{rbf}(\cdot)$, the sim-to-real transfer can be evaluated by the pipeline of Fig.3(b) in the inference phase as a network $\mathcal{N}_{s2r}(\cdot)$ that gives

$$\mathbf{p}^* = \mathcal{N}_{s2r}(\mathbf{B}(u_p, v_p, \mathcal{S}^c(\mathbf{a})), \mathcal{S}^c(\mathbf{a}), \{\mathbf{B}(u_q, v_q, \mathcal{S}^c(\mathbf{a}))\}) \quad (3)$$

(u_p, v_p) denotes a point sampled in the parametric domain of the free-form surface. $\{\mathbf{B}(u_q, v_q, \mathcal{S}^c(\mathbf{a}))\}$ denotes the set of marker points, which are defined by their parameters (u_q, v_q) and are chosen as kernel centers for the RBF-based warping function. All operators in this sim-to-real pipeline is differentiable w.r.t. the actuation parameter $\mathbf{a}$. In the following section, we present the detail of NN-based forward kinematics and the IK computing for shape control, which is based on minimizing the distance between $\mathbf{p}^*$ and its closest point on the target shape $\mathcal{S}^t$.

III. NEURAL NETWORKS BASED KINEMATIC COMPUTING

This section presents the neural network-based computational pipeline for kinematics. Shape control of deformable free-form surfaces can be realized by our fast IK solution.

A. NN-based Forward Kinematics

While the recently developed fast simulator provides an efficient solution to predict the deformed shape of soft robots [15], applying it to estimate gradients of the IK objective by numerical differentiation remains time-consuming [14]. Here we propose to employ an NN-based computational pipeline for forward kinematics. Unlike those work interested in the status of end-effectors in the task space (often in a very low dimension), the output of our network as free-form surfaces can have a very high dimension if the shapes are simply represented as sample points. It's very challenging to train such a network, as a massive number of simulation results are needed.

In this work, we apply B-spline surface representation as a compact shape descriptor for each free-form surface $\mathcal{S}$ generated by simulation. Specifically, the triangular mesh of $\mathcal{S}$ can be approximated by a B-spline surface, with its control points $\mathcal{S}^c$ determined through parameterization and fitting steps (see Fig. 4 for an illustration). The dimension of $\mathcal{S}^c$ using B-spline representation is lower than mesh representation. In

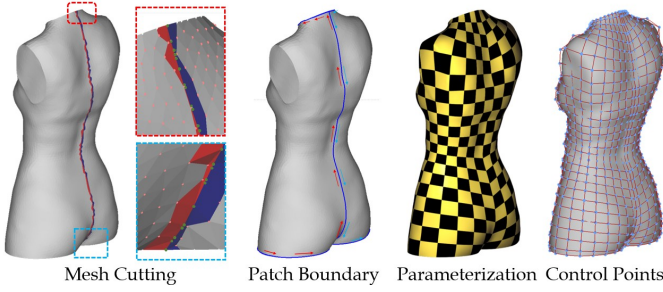


Fig. 4: Steps of generating the compact B-spline representation for the free-form surface of a deformable mannequin.

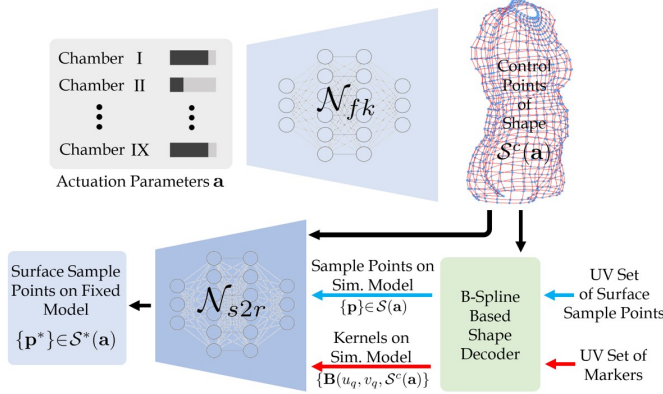


Fig. 5: NN-based pipeline for forward kinematic computing.

our implementation, we use 30×30 control points for a mannequin model as shown in Fig.5. Note that the parameterization is consistent among all models as they are generated by numerical simulation using the mesh representation with the same set of vertices and elements – only the positions of vertices are changed.

Given a set of randomly generated actuation parameters as $\{a_k\}$, we can run the numerical simulators to obtain the deformed shapes $\{S^c\}$ and generate their corresponding control points as $\{S^c_k\}$. A training dataset with M such pairs of results $\{a_k : S^c_k\}_{k=1, \dots, M}$ can be obtained by the simulator using a reduced physical model (i.e., [15]). It is important to generate a dataset that spans the entire space of shape variation. The dataset is then employed to train a network $\mathcal{N}_{fk}$ that can give the fast simulation result with an input actuation as $S^c = \mathcal{N}_{fk}(a)$.

The gap between the simplified simulation and the physical reality can be well captured and fixed by our sim-to-real network. As a result, for any point (u_p, v_p) sampled in the u, v -parametric domain, its position on the physical model can be predicted by

$$\mathbf{p}^* = \mathcal{N}_{s2r}(\mathbf{B}(u_p, v_p, \mathcal{N}_{fk}(a)), \mathcal{N}_{fk}(a), \{\mathbf{B}(u_q, v_q, \mathcal{N}_{fk}(a))\}). \quad (4)$$

This NN-based computational pipeline for forward kinematics was illustrated in Fig.5. The predicted shape $S^*(a)$ can be obtained as a set of points $\{\mathbf{p}^*\}$, which is a differentiable function in terms of the actuation.

B. Gradient Iteration Based Inverse Kinematics

The IK computation is formulated as an optimization problem that minimize the shape difference between the predicted shape S^* and the target shape S^t . That is

$$\arg \min_{\mathbf{a}, \mathbf{R}, \mathbf{t}} D(S^*(\mathbf{a}), S^t) = \sum_{\mathbf{p}^* \in S^*(\mathbf{a})} \|\mathbf{p}^* - (\mathbf{R}\mathbf{c}^* + \mathbf{t})\|^2 \quad (5)$$

where the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$ are applied to the target model to eliminate the influence of orientation variation. $\mathbf{c}^*$ is the closet point of $\mathbf{p}^*$ on the target shape S^t .

The objective function can be minimized using the gradient descent method with linear search. After each iteration of $\mathbf{a}$ update, we apply an ICP-based rigid registration to determine a new pair of $(\mathbf{R}, \mathbf{t})$. In other words, the optimization process alternates between updating the actuation parameter $\mathbf{a}$ and adjusting the pose of target model $(\mathbf{R}, \mathbf{t})$. Thanks to the formulation of NN-based forward kinematics and our function based sim-to-real learning, the gradients of $D(\cdot)$ can be computed analytically. That is

$$\frac{\partial D}{\partial \mathbf{a}} = \sum_{\mathbf{p}^* \in S^*(\mathbf{a})} 2 \left(\frac{\partial \mathbf{p}^*}{\partial \mathbf{a}} \right)^T (\mathbf{p}^* - (\mathbf{R}\mathbf{c}^* + \mathbf{t})), \quad (6)$$

where

$$\begin{aligned} \frac{\partial \mathbf{p}^*}{\partial \mathbf{a}} &= \frac{\partial \mathcal{N}_{s2r}}{\partial \mathbf{B}} \frac{\partial \mathbf{B}}{\partial \mathbf{a}} + \frac{\partial \mathcal{N}_{s2r}}{\partial \mathcal{N}_{fk}} \frac{\partial \mathcal{N}_{fk}}{\partial \mathbf{a}} + \frac{\partial \mathcal{N}_{s2r}}{\partial \{\mathbf{B}\}} \frac{\partial \{\mathbf{B}\}}{\partial \mathbf{a}} \\ &= \left(\frac{\partial \mathcal{N}_{s2r}}{\partial \mathbf{B}} \frac{\partial \mathbf{B}}{\partial \mathcal{N}_{fk}} + \frac{\partial \mathcal{N}_{s2r}}{\partial \mathcal{N}_{fk}} + \frac{\partial \mathcal{N}_{s2r}}{\partial \{\mathbf{B}\}} \frac{\partial \{\mathbf{B}\}}{\partial \mathcal{N}_{fk}} \right) \frac{\partial \mathcal{N}_{fk}}{\partial \mathbf{a}}. \end{aligned} \quad (7)$$

Note that $\{\mathbf{B}\}$ denotes those terms contributed by the set of virtual markers (as kernels of RBF-base warping) while $\mathbf{B}$ is for the surface point $\mathbf{p}$.

For realizing a fast iterative-based IK solver, the analytical evaluation of the gradient using network-based forward kinematics is crucial. It avoids the time-consuming and unstable steps of numerical difference taken in the simulation [14] or on the physical setup [39, 43].

IV. IMPLEMENTATION DETAILS AND RESULTS

We have implemented the proposed approach in C++ and Python. The networks training part are implemented in the PyTouch platform with the batch size as 32, the learning rate at 0.001 and the maximum number of epochs as 150. For the inference phase, we transferred the trained networks to be integrated with our IK solver in C++ running on CPU. The analytical gradients of the networks are employed in our IK solver. All the training and computational tests are conducted on a PC with Intel i7-12700H CPU, RTX 3060 GPU and 32 GB RAM. The results of sim-to-real and fast IK have been verified on a physical setup of soft deformable mannequin.

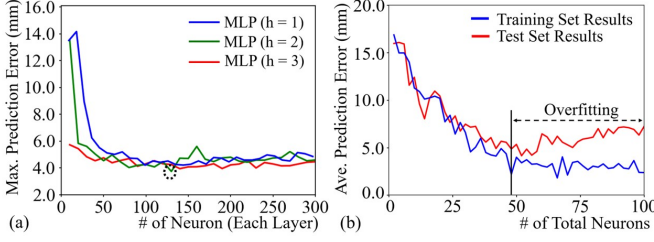


Fig. 6: Study of network parameters – (a) the maximal shape prediction error w.r.t. different numbers of layers (h) and numbers of neurons per layer (l) are evaluated to choose the ‘best’ values of h and l for $\mathcal{N}_{fk}$ (circled by black dash lines), and (b) the average shape prediction error w.r.t. the number of neurons in $\mathcal{N}_{rbf}$ is studied to avoid overfitting while training $\mathcal{N}_{s2r}$ using a limited number of samples. Note that $\mathcal{N}_{s2r}$ contains $\mathcal{N}_{rbf}$ plus an additional space warping module as illustrated in Fig.3.

A. Details of Networks and Training

For the training of forward kinematics network $\mathcal{N}_{fk}$, a dataset with 1,000 different pairs of actuation and shapes is collected from the simulation. This dataset contains shapes with the minimal and the maximal pressures applied to every chamber - in total 512 shapes (i.e., 2^9 with 9 chambers) were generated. We also randomly sampled other 488 actuation parameters using the Halton sequence [23]. All these 1,000 samples of actuation are applied in the simulation system to obtain their simulated shapes to explore the whole deformation space. Considering the large non-linearity between actuation and shape, we non-uniformly resample the actuation space by the amount of chamber inflation to enhance the training accuracy. The average time used for completing each simulation is around 150 seconds. This dataset is separated in the ratio of 7 : 3 for training and testing.

We carefully select the network structure, which includes the number of hidden layers (h), and the number of neurons in each hidden layer (l). The experimental results of the maximal shape prediction error according to different network structures are studied and illustrated in Fig. 6(a), and we choose two hidden layers with each layer contains 128 neurons as the final network structure to balance the quality of training result and the computational cost. In our implementation, ReLU is selected as the activation function and batch normalization is applied to improve stability in training. Note that our experiment finds that prediction with better quality can be achieved when learning the translation vectors applied to the positions of control points on an average model of all training shapes.

For the function prediction network $\mathcal{N}_{rbf}$ of sim-to-real transfer, we employ a network architecture with 2 hidden layers where each layer has a certain number of neurons using the ReLU activation function. The input layer contains the positions of 30×30 control points, and the output layer is the coefficients for the RBF-based warping function as

$\gamma \in \mathbb{R}^{3(N+4)}$ with $N = 34$ being the number of markers used in motion capture. The training of $\mathcal{N}_{rbf}$ is taken on 40 frames of markers obtained from the motion capture system while randomly varying the pneumatic actuation within the working range of the chambers. Among these 40 frames, around 18 frames have missing markers mainly due to camera occlusion caused by large deformation on the physical mannequin. However, all captured markers (even for those from frames with missing markers) are employed as samples to train $\mathcal{N}_{rbf}$. In total, 1330 points are involved in the training. By using this dataset, it can be observed from the study as shown in Fig. 6(b) that the prediction error starts to increase when the total number of neurons exceeds 50 – i.e., overfitting happens. Based on this analysis, we select 24 neurons for each layer in $\mathcal{N}_{rbf}$.

In our implementation, we first train $\mathcal{N}_{fk}$ and then train $\mathcal{N}_{rbf}$ by fixing $\mathcal{N}_{fk}$. The RBF-based space warping function Φ and the B-spline shape decoder $\mathbf{B}$ are both implemented as differentiable layers, enabling seamless integration into the standard training process of $\mathcal{N}_{rbf}$. Having a surrogate model $\mathcal{N}_{fk}$ that is trained by physics-based simulation is essential to achieve satisfied results of shape prediction. As an alternative, researchers might directly train a network to predict the positions of markers $\{\mathbf{p}^*\}$ by $\mathbf{a}$ and then deform the rest shape of human body by the RBF-based warping function $\Phi(\cdot)$ [42]. However, large shape approximation errors will be generated by this alternative without $\mathcal{N}_{fk}$ (see the middle of Fig. 7(a)). The other interesting study is about the output of $\mathcal{N}_{fk}$. As can be observed from the comparison given in Fig. 7(b), the B-spline surface representation used in our pipeline results in much smaller shape approximation error than directly predicting the positions of mesh vertices. The average shape approximate error is evaluated on each model of a test set with 300 individuals, and the distribution of errors among all individuals are plotted as a histogram to conduct the comparison. The mean of error distribution can be reduced from 1.82 mm to 0.63 mm when learning B-spline control points instead of mesh vertices by $\mathcal{N}_{fk}$.

B. Hardware for Verification

A physical setup has been developed to test the effectiveness of our sim-to-real and fast IK methods. The hardware contains a soft robotic mannequin representing both the front and back sides of a human body at a 1:2 scale, a closed-loop pressure control system, and a vision system to capture the positions of markers for training. The deformed shapes of the mannequin are captured by a 3D scanner for verification.

Our soft robotic mannequin can be deformed by pressurized air and features a design with three layers: an inner soft chamber layer attached to the base solid core (for chambers I-III & V-IX as shown in Fig.1(a)); a middle soft chamber layer attached on top of the inner layer (as chamber IV in Fig.1(a)); and a top membrane layer (passively morphed) covering the base layers to form the overall smooth shape. By pumping air at varying pressures into the chambers, the soft mannequin can be deformed into different shapes. An open-source pneumatic

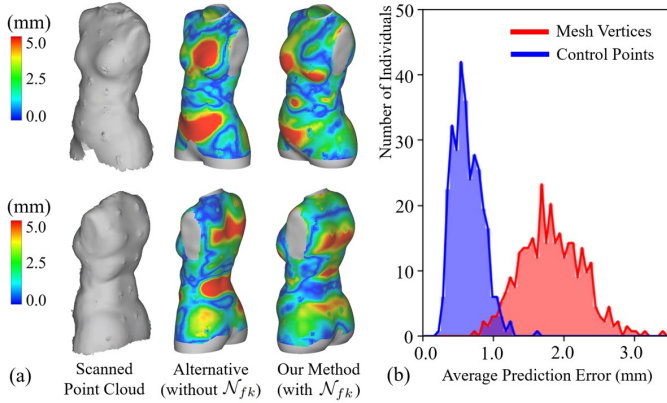


Fig. 7: Study of different alternatives. (a) Much larger of shape approximation errors are generated by using an alternative method (without the simulation-based $\mathcal{N}_{fk}$ network) that first predicts the positions of markers from actuation and then deform the body shape by marker-based RBF deformation [42] – see results shown in the middle column. (b) Comparison of the shape prediction errors (i.e., the average error of 18,000 sample points on each model) among a test set with 300 individuals generated by $\mathcal{N}_{fk}$ when outputting the B-spline control points as compact shape descriptor (blue) vs. the mesh vertices (red).

platform, OpenPneu [40], is employed to control the actuation with well-controlled pressures.

A Vicon motion capture system [5] with 8 cameras is employed to collect the positions of markers on the deformed surfaces, and a structured-light based 3D scanner Artec Eva Lite [4] is used to capture the surface geometry of a deformed mannequin as point clouds. Shape errors between the scanned mannequin and the target shape are measured after applying an ICP-based rigid registration to the target shape. Note that as the robotic mannequin is half size, the target shape is also scaled down to the half size as input in all our experimental tests.

C. Results of Sim-to-Real Learning

Experimental tests have been conducted to verify the performance of our sim-to-real method. First of all, we randomly change all the nine actuation parameters of $\mathbf{a}$ and use our forward kinematic network to predict free-form surface shapes. Four instances are selected to apply the actuation $\mathbf{a}$ onto the physical setup. The resultant shapes on the soft robotic mannequin are scanned and compared with the simulated models (i.e., without sim-to-real transfer) and the fixed models (i.e., using the full pipeline with the sim-to-real network). The shape approximate errors are evaluated by first correcting the pose of scanned model using the ICP-based registration and then compute the distances between every surface sample points to their closest points on the scanned model. As can be found from the results shown in Fig.8, the shape approximation errors on the fixed models were significantly reduced. The reductions are 30.4%, 53.5%, 32.1% and 33.5% for the

average error and by 6.7%, 44.0%, 50.6% and 60.6% for the maximal error. Result can also be found in the supplemental video.

We then study the effectiveness of our function-prediction based sim-to-real pipeline by comparing it with the marker-prediction based pipeline as discussed in Sec.II-A, which is used as a baseline. While randomly changing the pneumatic actuation, markers attached on the surface of the physical mannequin are progressively captured by the Vicon system and employed to train the sim-to-real networks. The trained networks are then employed to work together with $\mathcal{N}_{fk}$ to predict the free-form surfaces generated by a random selected actuation $\mathbf{a}_r$. The shape approximation errors are measured between the predicted shapes and the real surface, which is captured by scanning the physical model deformed by applying $\mathbf{a}_r$. The baseline method can only use those frames with full set captured markers for training. As shown in Fig.9, our method can predict a surface shape with smaller error. This mainly because that all captured markers can contribute to the sim-to-real training of our function-prediction based model even when they are on frames with missing markers. Differently, among all 40 frames involved in the tests shown in Fig.9, only about 55% can be employed by the marker-prediction based method. The other 45% frames all have 1 to 8 missing markers due to the occlusion caused by large deformation on chambers.

D. Results of Inverse Kinematics

We have tested the performance of our NN-based fast IK solver on a variety of individual shapes presented in the CAESAR dataset. Figure 10 shows the progressive results when applying our IK solver to realize the target shape on the physical setup of pneumatic driven deformable mannequin. It can be observed that our method converges very fast – i.e., around 15 iterations where each iteration takes less than 0.5 second. The distribution of shape approximation errors are generated by scanning the soft mannequin that has been deformed using the resultant actuation determined by our IK solver.

The effectiveness of our fast IK solver has been tested on other models randomly selected from the CAESAR dataset – see the results shown in Fig.11. It can be find that the shapes realized on the robotic deformable mannequin can well mimic the shape of target models. This can also be verified by the quantitative analysis results generated by the 3D scanner, where the errors are displayed as color maps and reported in the bottom row of Fig.11.

E. Discussion

1) *Computational efficiency:* In this work, the NN-based IK computing pipeline enables fast and robust shape approximation for a deformable free-form surface. We have achieved a computing speed of less than 10 seconds per target model, which is significantly faster than existing methods and aligns well with the physical actuation time of the proposed setup. The computational complexity of our pipeline is influenced

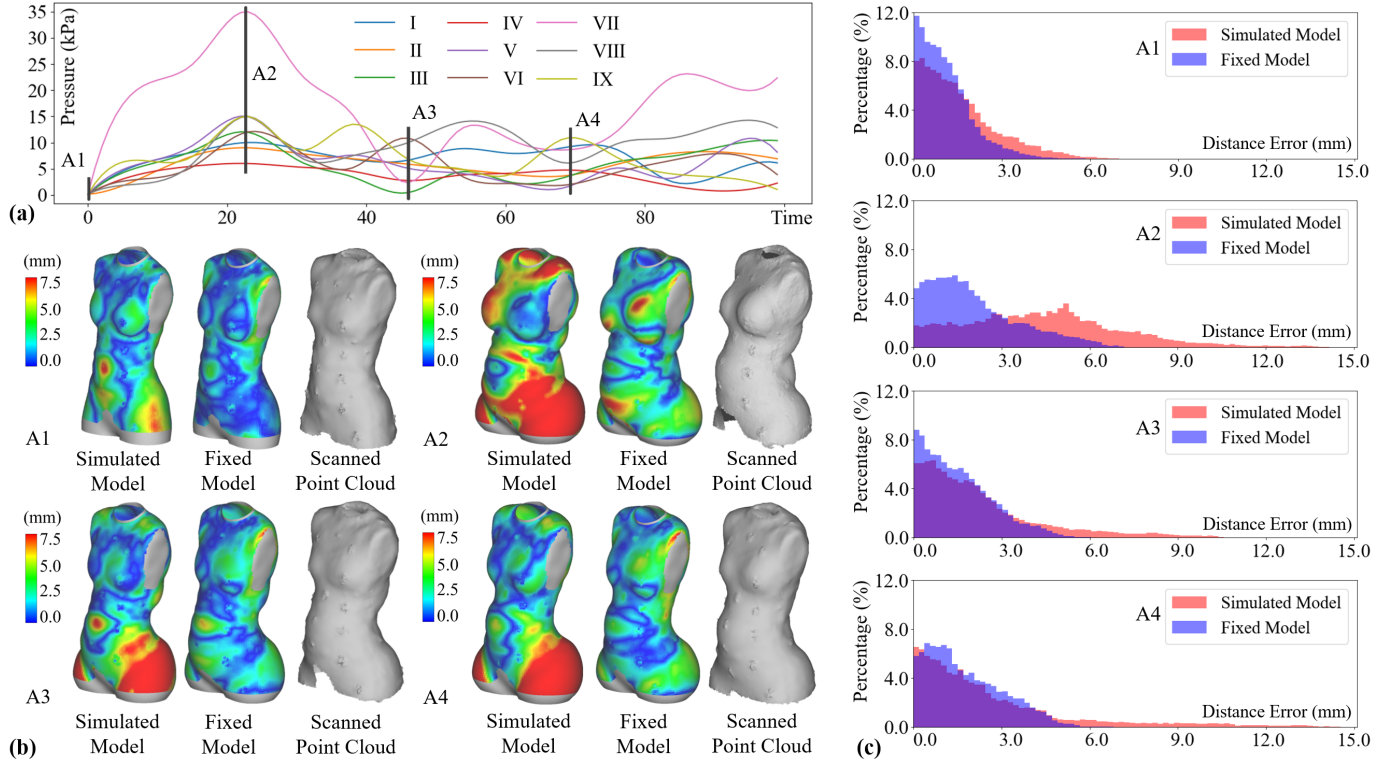


Fig. 8: Result of different free-form surfaces predicted by our forward kinematics pipeline (i.e., Fig.5 and Eq.(4)) while changing the actuation parameters (a). Four instants (A1-A4) are selected to apply onto the physical setup with the resultant shapes scanned and compared with the predicted shapes, where the shape approximation errors are visualized as colors. The results with and without sim-to-real transfer (denoted by fixed and simulated models respectively) are compared in both the color maps (b) and the error histograms (c).

by various factors, including the network architecture, the number of control points on the B-spline surface, and the kernel number of the RBF morphing function. The current selection is mainly based on experimental tests to find the balance between performance and efficiency. We can further accelerate the computation by GPU-based parallel computing.

2) *Applicability to other deformable robots:* Our pipeline is also applicable to other soft and deformable robotic systems. In principle, our function based sim-to-real learning method can be applied to a variety of deformable robots when their freeform shapes can be represented by B-spline surfaces (e.g., Chen *et al.* [10]). We will further explore the potential of our method in other robotic setups, particularly for addressing data collection challenges with learning-based solutions.

3) *Different setups for data acquisition:* In this work, a well-calibrated motion capture system (with position tracking error less than 0.1 mm [41]) is applied to capture data from the physical setup. While our pipeline is designed for sparse yet accurate data from motion capture, it has the potential to be extended to cover the systems with LiDAR [3] or RGBD cameras [2, 29] that capture dense but less accurate point clouds. On the other hand, since the material of the deformable surface exhibits hysteresis, the positions of markers are captured in a quasi-static manner. As such, we have to wait for 3 seconds

after applying the actuation to make the system stable and then capture frames of markers. This represents a limitation of the proposed method when applied to robotic systems that require high dynamic performance. One possible solution is incorporating other techniques that can handle sequential data inputs in the learning pipeline (e.g., [20]).

4) *Design of deformable mannequin:* The design of the proposed deformable surface can be further enhanced. This includes optimizing the shape and location of pneumatic-driven chambers, as well as the geometry of the rigid core. In our current design, we employed a neural distance field based method [33] to determine a volumetric region that is commonly ‘inside’ the body shapes of all individuals. We have not considered the poses of individuals when determining the shape of the core. We plan to address this limitation in our future work.

V. CONCLUSION

In this paper, we addressed the challenge of bridging the gap between simulated and physically deformed free-form surfaces caused by hardware errors and simplifications in physical simulation. Our proposed novel approach, based on deformation function-driven sim-to-real learning, offers a promising solution. It effectively maps simulated model geometries to their real-world counterparts, even accommodating

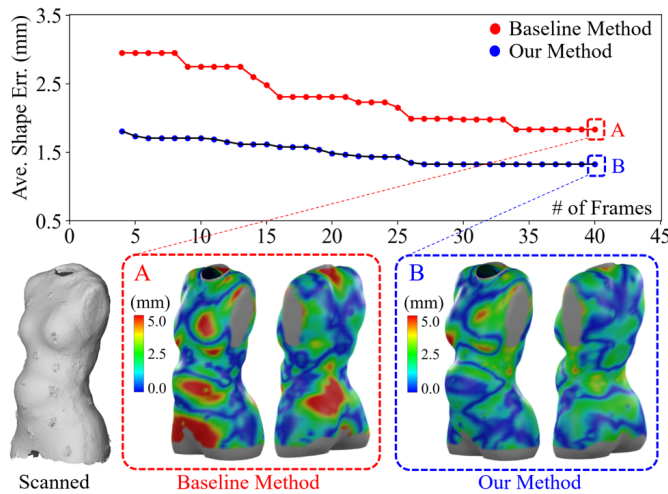


Fig. 9: Comparison of our function-prediction based sim-to-real method vs. the marker-prediction based method (as baseline). (Top) The errors between the scanned shape and the predicted shape when more and more frames of markers are employed for the sim-to-real learning. (Bottom) The scanned surface and the error maps of the predicted surfaces.

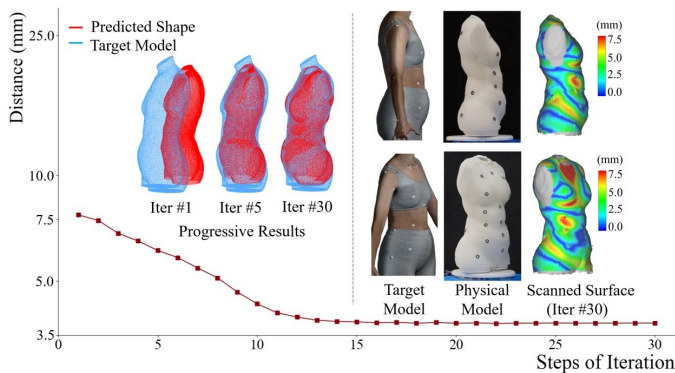


Fig. 10: The progressive results by applying our NN-based fast IK solver, where the curve shows the average shape approximation errors during the iterations of gradient descent. Using the actuation determined by our fast IK solver, a shape similar to the target model can be realized on the soft mannequin where the error analysis is conducted by 3D scanning with its result shown as the color map.

sparingly distributed set of markers with missing data. By integrating this method into a neural network-based computational pipeline, we demonstrate its effectiveness in solving the inverse kinematic problem for deformable mannequins with pneumatic actuation. Through this work, we contribute to advancing the field of shape control for deformable free-form surfaces, paving the way for more accurate and reliable realization in various applications.

ACKNOWLEDGMENTS

The project is partially supported by the chair professorship fund of the University of Manchester and the research fund

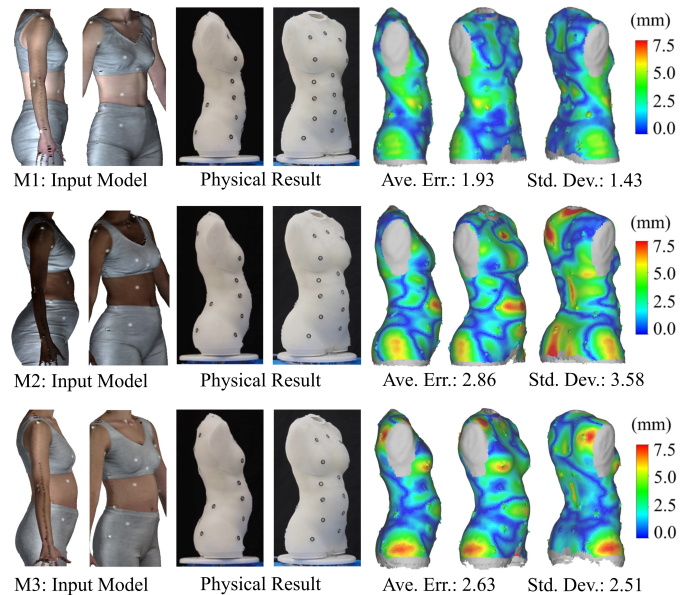


Fig. 11: The physical results of our fast IK solver for three different target models, where the distributions of shape approximation errors are generated by scanning and visualized as the color maps.

of UK Engineering and Physical Sciences Research Council (EPSRC) (Ref.#: EP/W024985/1).

REFERENCES

- [1] CAESAR: Civilian American and European Surface Anthropometry Resource Project. URL <https://www.sae.org/standardsdev/tsb/cooperative/caesar.htm>.
- [2] Intel® realsense depth camera d457. URL <https://www.intelrealsense.com/depth-camera-d457/>.
- [3] Intel® realsense lidar l1515. URL <https://www.intelrealsense.com/lidar-camera-l1515/>.
- [4] Artec eva lite: Structure light based 3d scanner. URL <https://www.artec3d.com/portable-3d-scanners/artec-eva-lite>.
- [5] Vicon: Award winning motion capture systems. URL <https://www.vicon.com/>.
- [6] Mehrnoosh Afshar, Jay Carriere, Hossein Rouhani, Tyler Meyer, Ron S. Sloboda, Siraj Husain, Nawaid Usmani, and Mahdi Tavakoli. Accurate tissue deformation modeling using a kalman filter and admm-based projective dynamics. *IEEE/ASME Transactions on Mechatronics*, 27(4):2194–2203, 2022. doi: 10.1109/TMECH.2022.3174510.
- [7] Costanza Armanini, Frédéric Boyer, Anup Teejo Mathew, Christian Duriez, and Federico Renda. Soft robots modeling: A structured overview. *IEEE Transactions on Robotics*, 39(3):1728–1748, 2023. doi: 10.1109/TRO.2022.3231360.
- [8] Veronica E Arriola-Rios, Puren Guler, Fanny Ficuciello, Danica Kragic, Bruno Siciliano, and Jeremy L Wyatt. Modeling of deformable objects for robotic manipula-

- tion: A tutorial and review. *Frontiers in Robotics and AI*, 7:82, 2020.
- [9] James M Bern, Yannick Schnider, Pol Banzet, Nitish Kumar, and Stelian Coros. Soft robot control with a learned differentiable model. In *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*, pages 417–423. IEEE, 2020.
 - [10] Feifei Chen, Zenan Song, Shitong Chen, Guoying Gu, and Xiangyang Zhu. Morphological design for pneumatic soft actuators and robots with desired deformation behavior. *IEEE Transactions on Robotics*, 2023.
 - [11] Bailin Deng, Yuxin Yao, Roberto M Dyke, and Juyong Zhang. A survey of non-rigid 3d registration. In *Computer Graphics Forum*, volume 41, pages 559–589. Wiley Online Library, 2022.
 - [12] Tao Du, Kui Wu, Pingchuan Ma, Sebastien Wah, Andrew Spielberg, Daniela Rus, and Wojciech Matusik. Diffpd: Differentiable projective dynamics. *ACM Trans. Graph.*, 41(2), nov 2021. ISSN 0730-0301. doi: 10.1145/3490168. URL <https://doi.org/10.1145/3490168>.
 - [13] Mathieu Dubied, Mike Yan Michelis, Andrew Spielberg, and Robert Kevin Katzschnmann. Sim-to-real for soft robots using differentiable fem: Recipes for meshing, damping, and actuation. *IEEE Robotics and Automation Letters*, 7(2):5015–5022, 2022. doi: 10.1109/LRA.2022.3154050.
 - [14] Guoxin Fang, Christopher-Denny Matte, Rob B. N. Scharff, Tsz-Ho Kwok, and Charlie C. L. Wang. Kinematics of soft robots by geometric computing. *IEEE Transactions on Robotics*, 36(4):1272–1286, 2020. doi: 10.1109/TRO.2020.2985583.
 - [15] Guoxin Fang, Yingjun Tian, Andrew Weightman, and Charlie C.L. Wang. Collision-aware fast simulation for soft robots by optimization-based geometric computing. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 12614–12621, 2022. doi: 10.1109/IROS47612.2022.9981870.
 - [16] Guoxin Fang, Yingjun Tian, Zhi-Xin Yang, Jo M. P. Geraedts, and Charlie C. L. Wang. Efficient jacobian-based inverse kinematics with sim-to-real transfer of soft robots by learning. *IEEE/ASME Transactions on Mechatronics*, 27(6):5296–5306, 2022. doi: 10.1109/TMECH.2022.3178303.
 - [17] Fanny Ficuciello, Alessandro Miglione, Eulalie Coevoet, Antoine Petit, and Christian Duriez. Fem-based deformation control for dexterous manipulation of 3d soft objects. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4007–4013. IEEE, 2018.
 - [18] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. Inform: Dynamic physical affordances and constraints through shape and object actuation. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST ’13, page 417–426, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450322683. doi: 10.1145/2501988.2502032. URL <https://doi.org/10.1145/2501988.2502032>.
 - [19] Antonio Elia Forte, Paul Z. Hanakata, Lishuai Jin, Emilia Zari, Ahmad Zareei, Matheus C. Fernandes, Laura Sumner, Jonathan Alvarez, and Katia Bertoldi. Inverse design of inflatable soft membranes through machine learning. *Advanced Functional Materials*, 32(16):2111610, 2022. doi: <https://doi.org/10.1002/adfm.202111610>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.202111610>.
 - [20] Thomas George Thuruthel, Egidio Falotico, Federico Renda, and Cecilia Laschi. Learning dynamic models for open loop predictive control of soft robotic manipulators. *Bioinspiration & Biomimetics*, 12, 08 2017. doi: 10.1088/1748-3190/aa839f.
 - [21] Thomas George Thuruthel, Yasmin Ansari, Egidio Falotico, and Cecilia Laschi. Control strategies for soft robotic manipulators: A survey. *Soft Robotics*, 5, 01 2018. doi: 10.1089/soro.2017.0007.
 - [22] Thomas George Thuruthel, Yasmin Ansari, Egidio Falotico, and Cecilia Laschi. Control strategies for soft robotic manipulators: A survey. *Soft robotics*, 5(2):149–163, 2018.
 - [23] John H Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2:84–90, 1960.
 - [24] Sebastian Höfer, Kostas E. Bekris, Ankur Handa, Juan Camilo Gamboa Higuera, Florian Golemo, Melissa Mozifian, Christopher G. Atkeson, Dieter Fox, Ken Goldberg, John Leonard, C. Karen Liu, Jan Peters, Shuran Song, Peter Welinder, and Martha White. Perspectives on sim2real transfer for robotics: A summary of the RSS 2020 workshop. *CoRR*, abs/2012.03806, 2020. URL <https://arxiv.org/abs/2012.03806>.
 - [25] Yuanming Hu, Jiancheng Liu, Andrew Spielberg, Joshua B Tenenbaum, William T Freeman, Jiajun Wu, Daniela Rus, and Wojciech Matusik. Chainqueen: A real-time differentiable physical simulator for soft robotics. In *2019 International conference on robotics and automation (ICRA)*, pages 6265–6271. IEEE, 2019.
 - [26] Keshav Iyengar, S.M.Hadi Sadati, Christos Bergeles, Sarah Spurgeon, and Danail Stoyanov. Sim2real transfer of reinforcement learning for concentric tube robots. *IEEE Robotics and Automation Letters*, 8(10):6147–6154, 2023. doi: 10.1109/LRA.2023.3303714.
 - [27] Seungwoo Je, Hyunseung Lim, Kongpyung Moon, Shan-Yuan Teng, Jas Brooks, Pedro Lopes, and Andrea Bianchi. Elevate: A walkable pin-array for large shape-changing terrains. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021.
 - [28] Margaret Koehler, Nathan S Usevitch, and Allison M Okamura. Model-based design of a soft 3-d haptic shape display. *IEEE Transactions on Robotics*, 36(3):613–628, 2020.
 - [29] Gregorij Kurillo, Evan Hemingway, Mu-Lin Cheng, and Louis Cheng. Evaluating the accuracy of the azure kinect

- and kinect v2. *Sensors*, 22(7):2469, 2022.
- [30] Xiao Liang, Fei Liu, Yutong Zhang, Yuele Li, Shan Lin, and Michael Yip. Real-to-sim deformable object manipulation: Optimizing physics models with residual mappings for robotic surgery. *arXiv preprint arXiv:2309.11656*, 2023.
 - [31] Fei Liu, Zihan Li, Yunhai Han, Jingpei Lu, Florian Richter, and Michael C. Yip. Real-to-sim registration of deformable soft tissue with position-based dynamics for surgical robot autonomy. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12328–12334, 2021. doi: 10.1109/ICRA48506.2021.9561177.
 - [32] Alice Morby. Anielia Hoitink creates dress from mushroom mycelium. URL <https://www.dezeen.com/2016/04/01/aniela-hoitink-neffa-dress-mushroom-mycelium-textile-materials-fashion/>.
 - [33] Jeong Joon Park, Peter R. Florence, Julian Straub, Richard A. Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 165–174, 2019. doi: 10.1109/CVPR.2019.00025.
 - [34] James Parkes. Non-woven dress sprayed onto models body on coperni runway at paris fashion week, Oct 2022. URL <https://www.dezeen.com/2022/10/04/spray-on-dress-paris-fashion-coperni-bella-hadid/>.
 - [35] Bryan Peele, Shuo Li, Chris Larson, Jason Cortell, Ed Hachtour, and Robert Shepherd. Untethered stretchable displays for tactile interaction. *Soft Robotics*, 6(1):142–149, 2019. doi: 10.1089/soro.2017.0059. URL <https://doi.org/10.1089/soro.2017.0059>. PMID: 30566378.
 - [36] Rob BN Scharff, Guoxin Fang, Yingjun Tian, Jun Wu, Jo MP Geraedts, and Charlie CL Wang. Sensing and reconstruction of 3-d deformation on pneumatic soft robots. *IEEE/ASME Transactions on Mechatronics*, 26(4):1877–1885, 2021.
 - [37] Dylan S Shah, Joshua P Powers, Liana G Tilton, Sam Kriegman, Josh Bongard, and Rebecca Kramer-Bottiglio. A soft robot that adapts to environments through shape change. *Nature Machine Intelligence*, 3(1):51–59, 2021.
 - [38] Andrew A. Stanley and Allison M. Okamura. Controllable surface haptics via particle jamming and pneumatics. *IEEE Transactions on Haptics*, 8(1):20–30, 2015. doi: 10.1109/TOH.2015.2391093.
 - [39] Yingjun Tian, Guoxin Fang, Justas Savinijus Petrusis, Andrew Weightman, and Charlie C. L. Wang. Soft robotic mannequin: Design and algorithm for deformation control. *IEEE/ASME Transactions on Mechatronics*, 27(4):1820–1828, 2022. doi: 10.1109/TMECH.2022.3175759.
 - [40] Yingjun Tian, Renbo Su, Xilong Wang, Nur Banu Altin, Guoxin Fang, and Charlie CL Wang. Openpneu: Compact platform for pneumatic actuation with multi-channels. In *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 765–770. IEEE, 2023.
 - [41] Vicon. Accuracy in motion. URL <https://www.vicon.com/wp-content/uploads/2022/07/Vicon-Metrology-Solutions.pdf>.
 - [42] Charlie C. L. Wang, Kin-Chuen Hui, and Kai-Man Tong. Volume parameterization for design automation of customized free-form products. *IEEE Transactions on Automation Science and Engineering*, 4(1):11–21, 2007. doi: 10.1109/TASE.2006.872112.
 - [43] Michael C. Yip and David B. Camarillo. Model-less feedback control of continuum manipulators in constrained environments. *IEEE Transactions on Robotics*, 30(4):880–889, 2014. doi: 10.1109/TRO.2014.2309194.
 - [44] John Z. Zhang, Yu Zhang, Pingchuan Ma, Elvis Nava, Tao Du, Philip Arm, Wojciech Matusik, and Robert K. Katzschmann. Sim2real for soft robotic fish via differentiable simulation. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 12598–12605, 2022. doi: 10.1109/IROS47612.2022.9981338.