Ultrasound-Guided Assistive Robots for Scoliosis Assessment with Optimization-Based Control and Variable Impedance

Anqing Duan, Maria Victorova, Jingyuan Zhao, Yuxiang Sun, *Member, IEEE*
Yongping Zheng, *Senior Member, IEEE*, and David Navarro-Alarcon, *Senior Member, IEEE*

***Abstract*—Assistive robots for healthcare have witnessed a growing demand over the past decades. In this paper, we investigate the development of an optimization-based control framework with variable impedance for an assistive robot to perform ultrasound-guided scoliosis assessment. The conventional procedure for scoliosis assessment using ultrasound imaging typically requires a medical practitioner to slide an ultrasound probe along a patient's back while maintaining a certain magnitude of the contact force. To automate such a procedure, we need to consider multiple objectives, such as contact force, position, orientation, energy, posture, etc. To coordinate different objectives, we propose to formulate the control framework as a quadratic programming problem with each objective weighted by a tunable task priority, subject to a set of equality and inequality constraints. As the procedure requires the robot to establish a constant contact force with the patient during scanning, we incorporate variable impedance regulation of the end-effector to enhance safety and stability during the physical human-robot interaction; The variable impedance gains are then retrieved by learning from medical expert's demonstrations. The proposed methodology is evaluated with a robotic system performing autonomous scoliosis assessment with multiple human subjects involved. The effectiveness of our approach is verified by the coronal spinal images obtained with the robot.**

***Index Terms*—Medical Robots and Systems, Physical Human-Robot Interaction, Task and Motion Planning, Optimization and Optimal Control, Learning from Demonstration.**

I. Introduction

HEALTHCARE and medical assistive robots have received increasing attention due to their growing demand in the market over the past decades [1]. As a typical modality of healthcare and medical assistive robots, ultrasound-guided navigator for physical body examination has numerous applications in clinical practice [2]. Compared with X-rays, ultrasound has several advantages, such as no ionizing radiation and affordable cost. Therefore, it is favored in a wide spectrum of areas such as cardiology, urology, gynecology, etc [3].

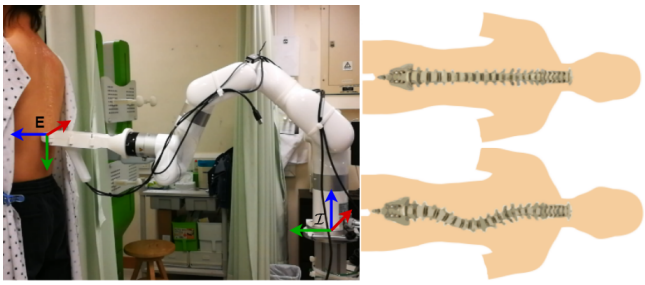

Fig. 1. Illustration of autonomous scoliosis assessment with the end-effector frame E and the base coordinate frame $\mathcal{I}$ labeled where x, y, and $z-$axes are depicted in red, green, and black, respectively (*left*). Comparison between a normal spine (*top right*) and a spine with scoliosis (*bottom right*).

In this paper, our focus lies in automating an ultrasound-guided robotic arm for *scoliosis assessment*, as shown in Fig. 1. Scoliosis assessment is a type of physical examination aimed at screening of scoliosis progression for timely spine correction and intervention [4]. To reduce exposure to ionizing radiation caused by X-ray scans, 3D spinal reconstruction with ultrasound has been gradually gaining popularity for progression monitoring [4]. The process of spinal image generation usually requires a human operator to carefully scan the spine's profile by sliding an ultrasound probe along a patient's back. To ensure that the quality of the ultrasound image is satisfactory, the operator needs to adjust the pose of the probe in real-time during the scan procedure such that the vertebrae (whose presence is normally indicated by the spinous process) are located in the middle of the image, as illustrated in Fig. 2. Therefore, the key to success of 3D spine image reconstruction is to locate and follow the spinous process closely while sweeping the probe along the spine's curvature. The coronal slice of the resulting 3D reconstruction is used for lateral scoliosis curvature measurement. From a robotic point of view, we identify that automating scoliosis assessment heavily relies on technical support from two main research areas, namely, ultrasound-guided navigation [5] and physical human-robot interaction (pHRI) with *variable impedance control* [6].

Ultrasound-based strategies for navigation have been applied to a variety of medical devices and surgical instruments due to their minimally invasive property and high portability. Until recently, several attempts towards physical examination through an ultrasound-guided robot manipulator have emerged.

Manuscript received: February 24, 2022; Revised May 17, 2022; Accepted June 16, 2022.

This paper was recommended for publication by Editor Jens Kober upon evaluation of the Associate Editor and Reviewers' comments. This research is funded in part by the Research Impact Fund (RIF) of the HK Research Grants Council under grant R5017-18F, and in part by PolyU through the Intra-Faculty Interdisciplinary Project under grant ZVVR.

Anqing Duan, Yuxiang Sun, and David Navarro-Alarcon are with the Department of Mechanical Engineering, The Hong Kong Polytechnic University, KLN, Hong Kong (e-mail: aduan@polyu.edu.hk, yx.sun@polyu.edu.hk, dna@ieee.org).

Maria Victorova, Jingyuan Zhao, and Yongping Zheng are with the Department of Biomedical Engineering, The Hong Kong Polytechnic University, KLN, Hong Kong (e-mail: maria.victorova@connect.polyu.hk, jing.y.zhao@connect.polyu.hk, yongping.zheng@polyu.edu.hk).

Digital Object Identifier (DOI): see top of this page.

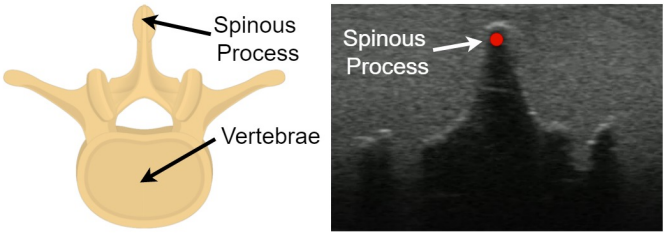

Fig. 2. Anatomy illustration of vertebrae and its spinous process (*left*) as well as the corresponding appearance in ultrasound image (*right*) where the red dot denotes the detected spinous process.

For example, in order to require less expertise from an operator, a probe guidance system is developed for obstetric ultrasound scanning [7]. The system takes ultrasound and motion signals as input and outputs a guidance signal. By contrast with [7], our proposed system can operate fully autonomously and does not require an operator to work in the loop. To capture the complex mapping between the ultrasound image and the probe motion, reinforcement learning has been employed in several works for autonomous probe navigation, such as location of the sacrum [5], ultrasound imaging of a soft and moving target [8], and standard view acquisition in robotic spinal sonography [9]. Although reinforcement learning-based strategies have achieved promising results, data efficiency issue remains a big challenge. To this end, a human-demonstration strategy is adopted for acquiring robotic ultrasound scanning skills [10], yet interpretability issue of learning-based motion controllers draws a big question mark under the context of medical applications. On the contrary, our control framework is devised from an optimization-based perspective, which has shown superior performance in many safety-critical applications, such as humanoid locomotion [11], aerial manipulation [12], and flying humanoids [13].

The formalism of our application, a hybrid motion/force control on a human body, is also analyzed to fall into a category of pHRI. Notably, a rich literature on pHRI is dedicated to enhancing safety and flexibility [14]. Specifically, we aim at endowing the robot with variable impedance control, which has better safety guarantees for contact-sensitive tasks [15], [16] or rehabilitation tasks [17] than high-gain position control. Compared with stochastic searching-based techniques such as evolution strategies [18] or inverse reinforcement learning [19], transferring variable impedance skills via human demonstrations provides a more intuitive and straightforward fashion [20], [21].

Previously, we devised a novel algorithm for trajectory planning based on spinous process localization in ultrasound images [22]. Although the preliminary results have shown that a robotic approach is promising for spinal image reconstruction, there is a lack of a principled control theoretic framework for prioritizing different tasks, such as contact force, end-effector pose, robot configuration, etc, under external kinematic or control constraints. In view of that, in this paper, we propose a new optimization-based control architecture capable of considering multiple objectives simultaneously while respecting a set of equality and inequality constraints. Similar to [11], our controller has two loops and is implemented via quadratic programming (QP). Furthermore, regarding acquisition of variable impedance gains, we follow the general paradigm of learning by demonstration [23]. While most existing works on impedance regulation revolve around either robot joint space or end-effector position [24], it is usually overlooked to incorporate impedance regulation for orientation. To incorporate the impedance regulation of the probe's orientation, we propose to parameterize the demonstrated rotation matrix trajectory such that the issue of exploiting the covariance matrix for a variable in a form other than a vector can be alleviated. Table I summarizes the main properties of our approach to autonomous scoliosis assessment with respect to the state-of-the-art. Compared with the majority of related works that use reinforcement learning, our approach is distinguished by its efficient optimization-based control method.

TABLE I
COMPARISON WITH STATE-OF-THE-ART APPROACHES.

	Contact Awareness	Control Constraints	Variable Impedance	From Demos
Force-ultrasound fusion (Tirindelli et el. [2])	✓	–	–	–
Reinforcement learning (Ning et el. [8])	✓	–	–	–
Dual-agent framework (Li et el. [9])	–	✓	–	–
Guided explorations (Deng et el. [10])	✓	–	–	✓
Follow the curve (Victorova et el. [22])	✓	–	–	–
Our approach	✓	✓	✓	✓

To the best of the authors' knowledge, this is the first time that the proposed methodology has been used to automate the probe's manipulation during scoliosis assessment. The main contributions of this work can be summarized as follows:

- Development of a new optimization-based control architecture for autonomous scoliosis assessment;
- Incorporation of impedance regulation via learning by demonstrations of the ultrasound probe's manipulation;
- Experimental validation of the proposed theory with a spine phantom model and multiple human subjects[1].

The rest of the paper is organized as follows: Section II presents the mathematical models; Section III presents the controller design; Section IV describes the results; Section V presents discussions and gives final conclusions.

II. CONTROL ARCHITECTURE DESIGN

A. Dynamics Modeling

The joint space dynamics for a fixed-base serial robot manipulator whose configuration is characterized by the joint angles $\mathbf{q} \in \mathbb{R}^n$ can be modeled as [25]:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{c}(\mathbf{q}, \dot{\mathbf{q}}) + \boldsymbol{g}(\mathbf{q}) = \boldsymbol{\tau} + \mathbf{J}(\mathbf{q})^{\intercal}\boldsymbol{f}, \tag{1}$$

where $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{n\times n}$ is the mass matrix, $\mathbf{c}(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^n$ accounts for the Coriolis and centrifugal effects, $\boldsymbol{g}(\mathbf{q}) \in \mathbb{R}^n$

[1]Ethical approval HSEARS20210417002 was given by the Departmental Research Committee on behalf of PolyU Institutional Review Board.

is the gravity vector, $\boldsymbol{\tau} \in \mathbb{R}^n$ are the joint actuation torques. $\boldsymbol{f} \in \mathbb{R}^6$ denotes the external wrench (expressed in a frame parallel to the world frame $\mathcal{I}$) that is mapped to joint coordinates through the Jacobian matrix $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{6\times n}$.

When the robot sweeps the ultrasound probe along the subject's back, it exhibits a hybrid motion/force behavior. We consider to model the interaction between the robot end-effector and the human back with a set of holonomic constraints. By explicitly considering the interaction force resulting from the constraints of the end-effector motion, the robot shall possess higher safety level during the diagnostic procedure. Specifically, the time derivatives of these constraints are expressed as

$$\mathbf{B}_c^\intercal[\mathbf{v}_b^\intercal \;\; \boldsymbol{\omega}_b^\intercal]^\intercal = \mathbf{G}^\intercal[\mathbf{v}_E^\intercal \;\; \boldsymbol{\omega}_E^\intercal]^\intercal = \mathbf{J}_c\dot{\mathbf{q}} = \mathbf{0}_{n_c}, \tag{2}$$

where we define $\mathbf{J}_c = \mathbf{G}^\intercal\mathbf{J}$ with $\mathbf{G} = \mathbf{Ad}_g\mathbf{B}_c \in \mathbb{R}^{6\times n_c}$ as the contact map. The matrix $\mathbf{Ad}_g = \texttt{blkdiag}({}^{\mathcal{I}}\mathbf{R}_E,\ {}^{\mathcal{I}}\mathbf{R}_E)$ is the adjoint transformation[2] that maps the robot end-effector *body* velocity to the linear and angular velocities expressed in $\mathcal{I}$ where $\texttt{blkdiag}(\cdot,\cdot)$ denotes the function that returns a block diagonal matrix by sequentially arranging the input matrices along the diagonal. In addition, ${}^{\mathcal{I}}\mathbf{R}_E \in SO(3)$ is the rotation matrix expressing the orientation of the end-effector frame E with respect to $\mathcal{I}$. $\mathbf{v}_b$ and $\boldsymbol{\omega}_b$ are the linear and angular body velocities that are expressed in the instantaneous body frame. $\mathbf{B}_c \in \mathbb{R}^{6\times n_c}$ is the wrench basis with n_c indicating the number of constrained directions of motion (or equivalently, independent forces) of the robot end-effector. $\boldsymbol{\omega}_E$ is the spatial angular velocity of the end-effector such that $\dot{\mathbf{R}}_E = S(\boldsymbol{\omega}_E)\mathbf{R}_E$ where S is the skew operator[3].

The equation expressing the unconstrained direction of motion $\mathbf{v}_u$ can be written as

$$\mathbf{B}_u^\intercal[\mathbf{v}_b^\intercal \;\; \boldsymbol{\omega}_b^\intercal]^\intercal = \mathbf{G}_u^\intercal[\mathbf{v}_E^\intercal \;\; \boldsymbol{\omega}_E^\intercal] = \mathbf{J}_u\dot{\mathbf{q}} = \mathbf{v}_u, \tag{3}$$

where $\mathbf{v}_u$ denotes the velocity of the unconstrained direction of motion expressed in the body frame and we define $\mathbf{J}_u = \mathbf{G}_u^\intercal\mathbf{J}$ with $\mathbf{G}_u = \mathbf{Ad}_{g^{-1}}^\intercal\mathbf{B}_u$ and $\mathbf{B}_u \in \mathbb{R}^{6\times 6-n_c}$ being complementary to $\mathbf{B}_c$. We then rewrite the robot dynamics equation (1) by additionally including the contact constraints and decoupling the external wrench $\boldsymbol{f}$ as a result of non-holonomic constraints as well as free motion[4]:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{c} + \boldsymbol{g} = \boldsymbol{\tau} + \mathbf{J}_c^\intercal\boldsymbol{f}_c + \mathbf{J}_u^\intercal\boldsymbol{f}_u, \tag{4a}$$

$$\mathbf{J}_c\dot{\mathbf{q}} + \dot{\mathbf{J}}_c\ddot{\mathbf{q}} = \mathbf{0}, \tag{4b}$$

where $\boldsymbol{f}_c \in \mathbb{R}^{n_c}$ denote the forces due to the existence of the constraints and $\boldsymbol{f}_u \in \mathbb{R}^{6-n_c}$ represent the forces that are caused by allowable directions of motion such as frictions. The constraints on the acceleration of the end-effector motion (4b) is obtained by differentiating (2) with respect to time.

B. Task Specifications

In order to make the robot automatically perform scoliosis assessment, it is necessary to properly determine a vector of quantities of interest $\boldsymbol{\zeta}$ such that by tracking the corresponding desired values denoted by $\boldsymbol{\zeta}^d$, the robot is able to function and achieve the goal. Each term of the output vector $\boldsymbol{\zeta}$ is typically called a *task* in the robot control terminology. The selected tasks that are relevant to our application are listed as follows:

- Joints configuration of the robotic manipulator $\mathbf{q}$;
- Linear position due to the unconstrained motion $\mathbf{p}_u$;
- Angular position due to the unconstrained motion $\mathbf{R}_E$;
- Contact forces arising from the constraints $\boldsymbol{f}_c$.

Thus, these tasks are simply grouped into the vector

$$\boldsymbol{\zeta} = [\mathbf{q}^\intercal \;\; \mathbf{p}_u^\intercal \;\; \mathbf{R}_E^\intercal \;\; \boldsymbol{f}_c^\intercal]^\intercal. \tag{5}$$

The key to success of the spinal image reconstruction lies in fine motion of the probe's x-direction in the coordinate frame E, depicted in Fig. 1. To this end, we have developed an efficient neural network built upon ResNet that can automatically locate spinous processes [22]. The proposed network takes as input raw ultrasound images and outputs a spatial heatmap that indicates the location of the spinous process as well as the confidence probability. The distance between the detected spinous location and the image center is then used to guide the movement of x coordinate of the probe frame. Specifically, the probe moves in a way such that the vertebrae will remain in the center of the ultrasound image, as shown in Fig. 3. Furthermore, in order to drive the ultrasound probe along the spine in a waist-to-neck direction, the y coordinate of the ultrasound probe is set to move with a constant velocity. For the probe's orientation control, the desired rotation matrix is determined such that the probe is pointing towards normal direction of the back's surface. The reference force is specified based on the subject's body mass index (BMI) to ensure a good acoustic coupling with the skin [2].

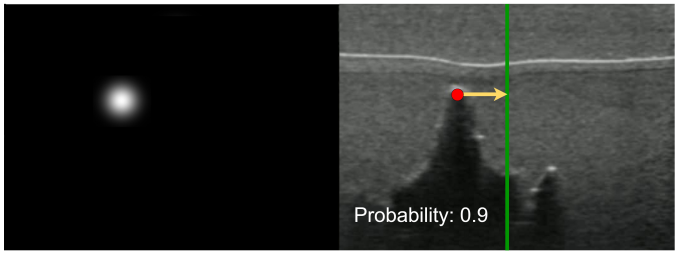

Fig. 3. Illustration of the adjustment strategy for the ultrasound probe. From the heatmap (*left*) of the original ultrasound image of the spinous process (*right*), the deviation (yellow arrow) away from the middle line (green) is used to laterally guide the probe [22].

C. Controller Design

Next, we consider the design of a controller for dynamical system (4) to achieve the aforementioned tasks. The goal here is to make the quantities of interest $\boldsymbol{\zeta}$ track the desired trajectory $\boldsymbol{\zeta}^d$ that is specified by our ultrasound image processor. In view of high complexity of the multi-input multi-output system in addition to several constraints that could emerge, such as control bounds, joint limits, kinematic constraints etc., we formulate our control problem from an optimization perspective. Compared with analytic control law design, optimization-based strategies exhibit great potential for

[2]As an abuse of notation, it is not exactly the mapping from body velocity to spatial velocity.

[3]Reference frame $\mathcal{I}$ is omitted for brevity from now on.

[4]Dependence on the robot states is dropped for brevity from now on.

customization of different requirements [26] and are better at handling constraints [27]. Specifically, our control method (akin to feedback linearization) is composed of two loops [12]. In the outer loop (i.e., the task planner block), it is assumed that higher-order derivatives of the tasks $\boldsymbol{\zeta}$ defined as

$$\mathbf{v} = [\ddot{\mathbf{q}}^\intercal \quad \ddot{\mathbf{p}}_u^\intercal \quad \dot{\boldsymbol{\omega}}_E^\intercal \quad \dot{\boldsymbol{f}}_c^\intercal]^\intercal \tag{6}$$

are directly controllable by a so-called virtual input $\mathbf{v}^*$. Notably, the time derivative of the corresponding virtual input for the contact force task $\boldsymbol{f}_c$ in (6) is zero. The input-output asymptotic stability is guaranteed if there exists a control signal $\mathbf{u}^*$ such that $\mathbf{v}^*$ can be exactly achieved, i.e. $\mathbf{v} = \mathbf{v}^*$. In principle, any forms of stabilizing controller that can make $\boldsymbol{\zeta}$ smoothly track $\boldsymbol{\zeta}^d$ is valid.

Here, the chosen stabilizing controller is designed in the form of a PID-like controller due to its simplicity. The virtual input is then designed as:

$$\mathbf{v}^* = \begin{bmatrix} \ddot{\mathbf{q}}^d - \mathbf{K}_q^D(\dot{\mathbf{q}} - \dot{\mathbf{q}}^d) - \mathbf{K}_q^P(\mathbf{q} - \mathbf{q}^d) \\ \ddot{\mathbf{p}}_u^d - \mathbf{K}_u^D(\dot{\mathbf{p}}_u - \dot{\mathbf{p}}_u^d) - \mathbf{K}_u^P(\mathbf{p}_u - \mathbf{p}_u^d) \\ \dot{\boldsymbol{\omega}}_E^d - \mathbf{K}_R^D(\boldsymbol{\omega}_E - \boldsymbol{\omega}_E^d) - \mathbf{K}_R^P\mathbf{e}_R \\ \boldsymbol{f}_c^d - \mathbf{K}_c^P(\boldsymbol{f}_c - \boldsymbol{f}_c^d) - \mathbf{K}_c^I\int_0^t(\boldsymbol{f}_c - \boldsymbol{f}_c^d)dt \end{bmatrix}. \tag{7}$$

The rotation error $\mathbf{e}_R$ between the desired end-effector rotation $\mathbf{R}_E^d$ and the current rotation $\mathbf{R}_E$ is given by [28]

$$\mathbf{e}_R = \frac{1}{2}(\mathbf{R}_E^\intercal\mathbf{R}_E^d - \mathbf{R}_E^{d\intercal}\mathbf{R}_E)^\vee \in \mathbb{R}^3, \tag{8}$$

where $(\cdot)^\vee$ denotes the inverse of the skew operator.

For the inner loop, it is responsible to generate control signals $\boldsymbol{\tau}$ for the tracking of the virtual input $\mathbf{v}^*$. Typically, it is usually very difficult to design an analytical control law for a multi-input (namely $\boldsymbol{\tau}$) multi-output (namely $\mathbf{v}$) system, it is thus considered to leverage optimization-based techniques as they require less manual efforts compared with analytical control law design. To do so, we first write the system dynamics equation of the inner loop in a control-affine style as follows

$$\mathbf{v} = \mathbf{A}(\mathbf{q})\mathbf{u} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}), \quad \text{with} \quad \mathbf{u} = [\boldsymbol{\tau}^\intercal \quad \boldsymbol{f}_c^\intercal]^\intercal. \tag{9}$$

It should be noted here that contact force $\boldsymbol{f}_c$ is also included as a part of the control signals $\mathbf{u}$ in addition to joint torques $\boldsymbol{\tau}$, which implies that it is both a task vector and a control input. In the system's dynamics, the matrix $\mathbf{A}(\mathbf{q})$ and the bias vector $\mathbf{b}(\mathbf{q}, \dot{\mathbf{q}})$ that contain all the terms independent of the extended control input $\mathbf{u}$ are given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{M}^{-1} & \mathbf{M}^{-1}\mathbf{J}_c^\intercal \\ \mathbf{J}_u\mathbf{M}^{-1} & \mathbf{J}_u\mathbf{M}^{-1}\mathbf{J}_c^\intercal \\ \mathbf{0}_{n_c\times n} & \mathbf{I}_{n_c} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \boldsymbol{\Gamma} \\ \mathbf{J}_u\boldsymbol{\Gamma} + \dot{\mathbf{J}}_u\dot{\mathbf{q}} \\ \mathbf{0}_{n_c} \end{bmatrix}, \tag{10}$$

where we denote $\boldsymbol{\Gamma} = \mathbf{M}^{-1}(\mathbf{J}_u^\intercal\boldsymbol{f}_u - \mathbf{c} - \boldsymbol{g})$. The optimization problem to be solved in the inner loop is formulated as:

$$\min_{\mathbf{u}} \quad \|\mathbf{v} - \mathbf{v}^*\|_{\mathbf{W}}^2 \tag{11a}$$

$$\texttt{s.t.} \quad \underline{\mathbf{u}} \le \mathbf{u} \le \overline{\mathbf{u}} \tag{11b}$$

$$\mathbf{C}\boldsymbol{f}_c \le \mathbf{0} \tag{11c}$$

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{c} + \boldsymbol{g} = \boldsymbol{\tau} + \mathbf{J}_c^\intercal\boldsymbol{f}_c + \mathbf{J}_u^\intercal\boldsymbol{f}_u \tag{11d}$$

$$\mathbf{J}_c\dot{\mathbf{q}} + \dot{\mathbf{J}}_c\ddot{\mathbf{q}} = \mathbf{0} \tag{11e}$$

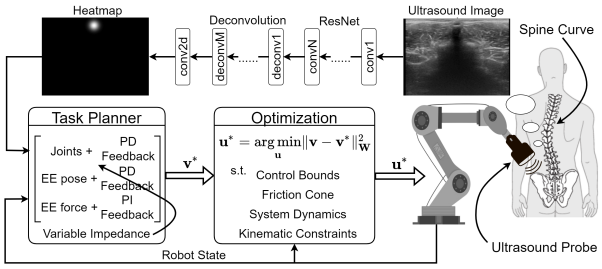

Fig. 4. Illustration of the overall proposed control framework for autonomous scoliosis assessment. First, the ultrasound images are processed with a fully connected neural network, which outputs the predicted location of the spinous process. Then the outer loop of the controller (task planner) sends the higher-order derivatives of the desired values of quantities of interest while the inner loop (optimization) solves an optimization problem for control commands.

The control objective (11a) is designed as a squared weighted Euclidean norm $\|\cdot\|_{\mathbf{W}}^2$ where each task priorities can be specified by $\mathbf{W}$. The constraint (11b) specifies the bounds on the control input such as joint torques limits and contact force magnitude, which are critical when establishing interactions with humans for medical purposes. The specified upper bound and lower bound are denoted by $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$, respectively. The constraint (11c) requires that the contact friction cones $\mathcal{FC}$ are satisfied by approximating $\boldsymbol{f}_c \in \mathcal{FC}$ with a linear inequality. (11d) impose constraints due to system's dynamics while (11e) correspond to the kinematic constraints as a result of contact with the environment.

It can be verified that (11) in fact belongs to a Quadratic Programming (QP) problem. In general, the standard formulation of a QP problem is composed of a quadratic-form objective function of the design variables and a set of linear equality and inequality constraints:

$$\min_{\mathbf{u}} \frac{1}{2}\mathbf{u}^\intercal\mathbf{P}_c\mathbf{u} + \mathbf{q}_c^\intercal\mathbf{u} \quad \texttt{s.t.} \quad \mathbf{L}\mathbf{u} \le \mathbf{h} \text{ and } \mathbf{D}\mathbf{u} = \mathbf{z}. \tag{12}$$

By re-arranging the objectives and constraint terms of (11) towards the general QP formalism of (12), the original optimization problem (11) can be equivalently written as

$$\begin{aligned} \min_{\mathbf{u}} \quad & \frac{1}{2}\mathbf{u}^\intercal\mathbf{A}^\intercal\mathbf{W}\mathbf{A}\mathbf{u} + (\mathbf{b} - \mathbf{v}^*)^\intercal\mathbf{W}\mathbf{A}\mathbf{u} \\ \texttt{s.t.} \quad & \begin{bmatrix} \mathbf{I}_{n+n_c} \\ -\mathbf{I}_{n+n_c} \\ \mathbf{0}_n^\intercal \ \mathbf{C} \end{bmatrix}\mathbf{u} \le \begin{bmatrix} \overline{\mathbf{u}} \\ -\underline{\mathbf{u}} \\ \mathbf{0} \end{bmatrix} \\ & \begin{bmatrix} \dot{\mathbf{J}}_c\mathbf{M}^{-1} & \dot{\mathbf{J}}_c\mathbf{M}^{-1}\mathbf{J}_c^\intercal \end{bmatrix}\mathbf{u} = -\mathbf{J}_c\dot{\mathbf{q}} - \dot{\mathbf{J}}_c\boldsymbol{\Gamma}. \end{aligned} \tag{13}$$

The correspondence relationship between (11) and (12) is thus yielded explicitly with the help of (13). As a result, the inner-loop optimization can be solved fast enough due to the merits of QP and it is possible to implement real-time optimization-based robot control. A diagram of the overall robotic system for scoliosis assessment is illustrated in Fig. 4.

III. Impedance Gain Retrieval

In this section, we study the problem of regulating the impedance gains, namely the stiffness and damping matrices $\mathbf{K}_u^D$, $\mathbf{K}_u^P$, $\mathbf{K}_R^D$, and $\mathbf{K}_R^P$ of (7), to guarantee safe and reliable interaction with human subjects. To this end, we consider to retrieve the desired impedance gains by resorting to the general

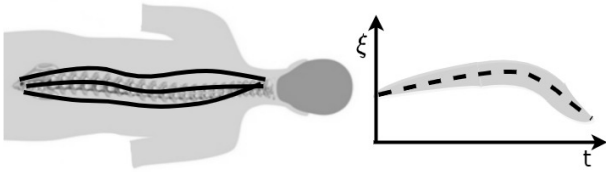

Fig. 5. Illustration of the learning by demonstration approach for the retrieval of impedance gains. The desired impedance gain is obtained by estimation of covariance matrix from multiple demonstrations.

framework of learning by demonstration or imitation learning since it provides a straightforward and intuitive manner to transfer human motion skills to robots [29]. Specifically, we propose to leverage the techniques from probabilistic imitation learning framework, where we can obtain variability from multiple human demonstrations to regulate the robot manipulator's impedance as shown in Fig. 5. To retrieve the desired impedance profile from the demonstrated trajectories, the region where the dispersion of the trajectories is high implies that robots will lower stiffness gains.

One popular approach is to encode variability of the demonstrated trajectories through Gaussian Mixture Regression (GMR) [23]. Without loss of generality, we denote $\boldsymbol{\xi}$ to represent either one of $\mathbf{p}_u$, $\dot{\mathbf{p}}_u$, and $\boldsymbol{\omega}_E$ whereas the retrieval of the orientation stiffness profile will be explained later as the rotation matrix does not follow the Euclidean distance metric.

Given M demonstrations from an expert, we can collect a dataset $\{\{d_n^m, \boldsymbol{\xi}_n^m\}_{n=1}^N\}_{m=1}^M$ with each demonstration length being N. The input $d \in \mathbb{R}$ is chosen as time stamps. For the purpose of generalization or different test speed, we could simply scale it accordingly.

In order to retrieve the variability of the demonstrated quantities with GMR from the dataset, we first need to encode the joint probability distribution of input d and output $\boldsymbol{\xi}$ with the Gaussian Mixture Model (GMM) representation: $\mathcal{P}(d, \boldsymbol{\xi}) = \sum_{k=1}^K \pi_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ where π_k denotes the prior probability of each Gaussian component in total of K Gaussian components with $\sum_{k=1}^K \pi_k = 1$, and

$$\boldsymbol{\mu}_k = \begin{bmatrix} \boldsymbol{\mu}_{d,k} \\ \boldsymbol{\mu}_{\xi,k} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_k = \begin{bmatrix} \boldsymbol{\Sigma}_{dd,k} & \boldsymbol{\Sigma}_{d\xi,k} \\ \boldsymbol{\Sigma}_{\xi d,k} & \boldsymbol{\Sigma}_{\xi\xi,k} \end{bmatrix}. \tag{14}$$

During the reproduction phase, the conditional distribution $\mathcal{P}(\boldsymbol{\xi}(d))$ given a query point d is given as

$$\mathcal{P}(\boldsymbol{\xi}(d)) = \sum_{k=1}^K \eta_k(d) \mathcal{N}(\boldsymbol{\mu}_{k|d}, \boldsymbol{\Sigma}_{k|d}), \tag{15}$$

where the weighting term $\eta_k(d)$, the conditional mean $\boldsymbol{\mu}_{k|d}$ and the covariance of a Gaussian component $\boldsymbol{\Sigma}_{k|d}$ are respectively given by [23]

$$\eta_k(d) = \frac{\pi_k \mathcal{N}(\boldsymbol{\mu}_{d,k}, \boldsymbol{\Sigma}_{dd,k})}{\sum_{j=1}^K \pi_j \mathcal{N}(\boldsymbol{\mu}_{d,j}, \boldsymbol{\Sigma}_{dd,j})}, \tag{16a}$$

$$\boldsymbol{\mu}_{k|d} = \boldsymbol{\mu}_{\xi,k} + \boldsymbol{\Sigma}_{\xi d,k} \boldsymbol{\Sigma}_{dd,k}^{-1} (d - \boldsymbol{\mu}_{d,k}), \tag{16b}$$

$$\boldsymbol{\Sigma}_{k|d} = \boldsymbol{\Sigma}_{\xi\xi,k} - \boldsymbol{\Sigma}_{\xi d,k} \boldsymbol{\Sigma}_{dd,k}^{-1} \boldsymbol{\Sigma}_{d\xi,k}. \tag{16c}$$

Notably, we are especially interested in retrieving the covariance matrix whose inverse serves as a proxy of the impedance gains. Therefore, the single peaked covariance of the multi-modal distribution (15) is approximated with

$$\mathtt{cov}(\boldsymbol{\xi}(d)) = \sum_{k=1}^K \eta_k(d)(\boldsymbol{\Sigma}_{k|d} + \boldsymbol{\mu}_{k|d}\boldsymbol{\mu}_{k|d}^\intercal) - \boldsymbol{\mu}_d \boldsymbol{\mu}_d^\intercal, \tag{17}$$

where the conditional mean value of the approximated single normal distribution is calculated as $\boldsymbol{\mu}_d = \sum_{k=1}^K \eta_k(d) \boldsymbol{\mu}_{k|d}$. Finally, the impedance gains $\mathbf{K}_u^D$, $\mathbf{K}_u^P$, and $\mathbf{K}_R^D$ are then set as the inverse of (17).

It should be noted that special attention shall be paid to the calculation of the stiffness matrix $\mathbf{K}_R^P$ for rotation control. Indeed, the problem is not trivial, if not ill-posed, when directly applying GMM/GMR to quantifying the aleatoric uncertainty of a variable in the form of a matrix rather than a vector, as in our case of rotation matrix. As a workaround, we propose to parameterize a rotation matrix with its exponential coordinate, namely $\mathbf{R} = \exp(S(\boldsymbol{\omega})\theta)$ with $\boldsymbol{\omega} \in \mathbb{R}^3$ being a unit vector and $\theta \in \mathbb{R}$ given by [30]

$$\theta = \cos^{-1}\left(\frac{\mathrm{tr}(\mathbf{R}) - 1}{2}\right), \quad \boldsymbol{\omega} = \frac{1}{2\sin(\theta)}(\mathbf{R} - \mathbf{R}^\intercal)^\vee. \tag{18}$$

Consequently, with such parameterization of the end-effector's orientation, the rotational stiffness matrix $\mathbf{K}_R^P$ is retrieved by instantiating $\boldsymbol{\xi}$ with $\boldsymbol{\omega}\theta$ and then following a similar procedure as per (14)–(17). By doing so, we can avoid employing complicated manifold-based regression algorithms to deal with the manifold-structured rotation matrix data. Arguably, the eligibility of employing the inverse of the covariance matrix of $\boldsymbol{\omega}\theta$ to represent $\mathbf{K}_R^P$ is evidenced by its compatibility with our choice of the rotation error as in (8).

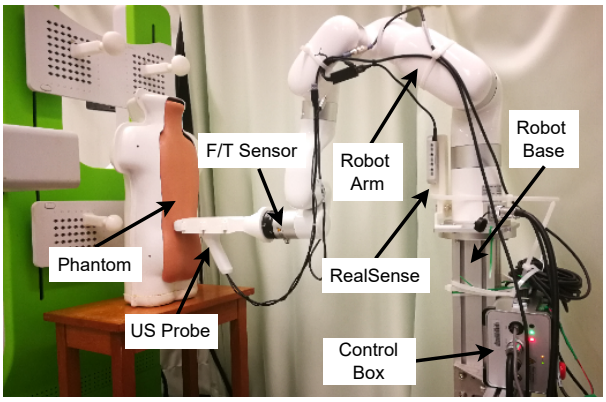

Fig. 6. Illustration of the experimental platform for scoliosis assessment.

IV. Results

A. Experiment Setup

Fig. 6 illustrates the proposed robotic system and its various components for autonomous scoliosis assessment [31]. The core component of the proposed robotic platform is an industrial robot manipulator called UFACTORY xArm 6, which is a fixed-base and serial robot manipulator with six DoFs. At the robot end effector, a USB ultrasound probe Sonoptek is mounted. To enable direct force control, a six-axis Force/Torque sensor Robotiq FT300 is also installed at

Fig. 7. Snapshots of real experiments on scoliosis assessment demonstrated by a medical expert for impedance gains retrieval (*first row*), execution on the phantom (*second row*), and execution on six human subjects (*last two rows*).

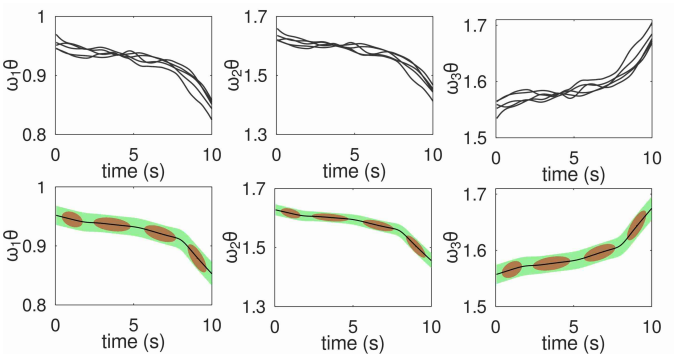

Fig. 8. Multiple demonstrated trajectories of $\boldsymbol{\omega}\theta$ (*top row*) and the retrieved probabilistic trajectories with red ellipses denoting GMM components, green shallow area denoting the covariance and black line denoting the mean value (*bottom row*).

the robot end-effector. To retrieve the normal vectors of the test subjects' back, a depth camera RealSense is attached near the robot base to collect point clouds of the surface.

B. Ultrasound Scanning Experiments

The aim of these tests is to reconstruct the spinal images of a phantom model as well as multiple human subjects. Specifically, to show the generalizability of the proposed approach, six human subjects with varying heights, BMI, and stiffness are used in the tests. Throughout the experiments, the parameters of the dynamic model are obtained from the URDF model provided by PyBullet [32]. The normal vectors of the test subject's back, which guide the orientation of the ultrasound probe, are estimated with the Point Cloud Library [33]. The QP problem that needs to be solved in the inner loop of the control architecture is handled by QP-solvers in Python [34]. See the accompanying multimedia file.

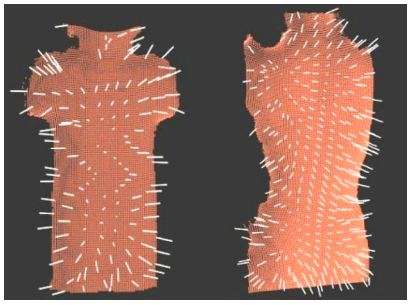

Fig. 9. Illustration of point cloud data (depicted in flesh) as well as the associated normal vectors (marked in white) of the phantom's back (*left*) and a human subject's back (*right*).

To estimate the impedance gains with the inverse of the covariance matrices, we provide multiple demonstrations on the phantom as shown in the first row of Fig. 7. The demonstrated trajectories are conveyed via kinesthetic teaching by an expert practitioner. The back region, where the adipose tissue is thick, is expected to have more concentration of the demonstrated trajectories, implying that higher stiffness of the end-effector behavior would overcome the adipose's effects on the ultrasound image quality. Similarly, the back region with bumpy bones is expected to have more dispersion of the demonstrated trajectories, resulting in a lower stiffness of the end-effector in order to avoid injuring the patient.

In these experiments, the number of the demonstrated trajectories is determined as $M = 5$. During the human demonstration phase, the robot has a sampling period of $0.05\,\mathrm{s}$. The demonstration duration lasts for $10\,\mathrm{s}$, which results in a total trajectory length of $N = 200$ points. The number of GMM components is chosen as $K = 4$, which is determined by Bayesian information criterion for trade-off between data fitting and model complexity [23]. The parameters of the GMM, namely π_k, $\boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$, are iteratively updated by means of the expectation maximization (EM) algorithm until the convergence criteria is satisfied, which is set to be less than $1\mathrm{e}-4$ for the increase of the average log-likelihood value. Fig. 8 shows the probabilistic modeling results of demonstrated orientation trajectories using Octave.

After these steps, we are ready to perform scoliosis assessment of the phantom and the human patients. To do so, we first spread ultrasound gel over the backs as common practice to avoid the air gap and enhance a tight contact between the back and the ultrasound transducer. Then, the normal vectors of the subject's backs are estimated from point cloud data, as shown in Fig. 9. The ultrasound probe is set to move upwards at a speed of $0.003\,\mathrm{m/s}$ starting from the waist towards the neck. The coordinate frame of the probe is set to align with the normal vector of the nearest back point such that the probe can maintain perpendicular to the test subject's back for best acoustic coupling. Furthermore, we employ the point contact model that imposes a constraint on the force f_z along the $z-$direction of the probe frame. The employed contact model, which will help establish a stable contact with the test subject's back, instantiates $\boldsymbol{f}_c$ with f_z and yields $\mathbf{C}$ of (11c) as -1.

Two experiments are conducted with the phantom model with the desired contact force set to $7\,\mathrm{N}$ and $15\,\mathrm{N}$, respectively.

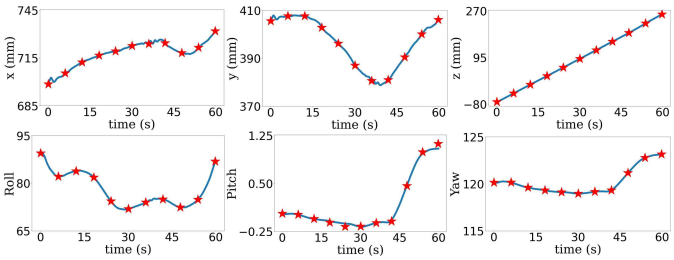

Fig. 10. Tracking performance of end-effector position (*top row*) and orientation (*bottom row*) on the *phantom* with contact force being 7 N where red stars denote the desired value.

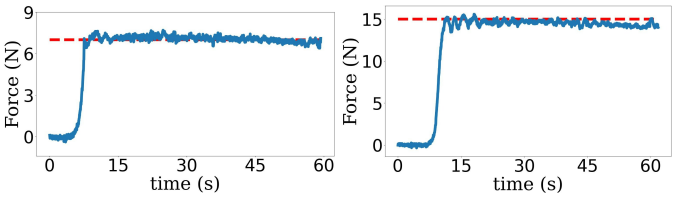

Fig. 11. Tracking performance of the desired contact force on the phantom with red dashed line denoting the reference and black line denoting measured value from the F/T sensor for 7 N (*left*) and 15 N (*right*), respectively.

The weights of the QP for different tasks of $\mathbf{v}$ are experimentally set to $[0.1, 1, 1.5, 2]^\intercal$. The procedure for scanning the spine is shown in the second row of Fig. 7. The tracking performance of different contact forces is shown in Fig. 11. It can be seen that the ultrasound probe can maintain a small tracking force error with an absolute value smaller than 1 N, which is a reasonable allowable performance in practice. The tracking performance of the end-effector pose expressed in the base frame is shown in Fig. 10. The reconstructed spine images of the phantom model are shown in the first two columns of Fig. 14. It is observed that the case with the bigger contact force has a clearer image, which highlights the importance of using for an optimal contact force.

For the experiments with the human subjects, the desired contact force usually depends on the back region and the subject's BMI. Here we experimentally set it to 10 N for all the volunteers, which provides acceptable imaging results. The weights of the QP for different tasks are set to $[0.1, 1, 1.5, 5]^\intercal$. The priority for the contact force task is raised with respect to the phantom case in order to guarantee the force tracking performance for the human back. The procedure for scanning the spines of the human subjects is shown in the last two rows of Fig. 7. The corresponding tracking performance of the end-effector pose for human subject (a) is shown in Fig. 12, where the error is negligible in our application. For statistical analysis, we show the controller's tracking performance among the human subjects using the normalized root mean square error (RMSE), defined as

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\zeta_i - \zeta_i^d)^2} \Big/ (\zeta_{\max} - \zeta_{\min}) \tag{19}$$

where ζ represents the quantity of interest and ζ^d its desired value. Fig. 13 shows the quantitative results. It can be seen that the tracking precision on both the pose tasks and the force task is quite high.

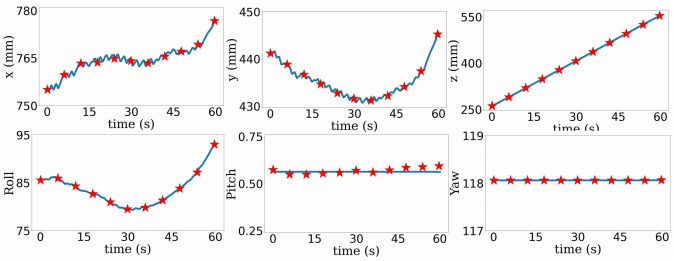

Fig. 12. Tracking performance of end-effector position (*top row*) and orientation (*bottom row*) on human subject (a) with contact force being 10 N where red stars denote the desired value.

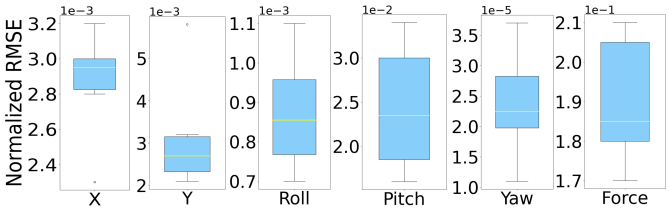

Fig. 13. Statistical analysis of tracking performance among human subjects for x, y-direction, roll, pitch, yaw, and force.

It is observed in Fig. 12 that there are some minor oscillations for trajectory tracking. There are two possible reasons for such phenomena. First, the soft tissue mechanics of the human back is quite different from that of the phantom. The motion of the probe may be disturbed by the compliance of the tissues. Second, human backs have more apparent bumps due to the spine inside. This may also lead to oscillations in the probe's motion. The obtained spine images of six human subjects are also depicted in Fig. 14. The quality of the image is clear and acceptable for medical personnel to assess the curvature of human spines.

V. Discussions and Conclusion

In this paper, we presented an optimization-based control architecture for autonomous scoliosis assessment. Arguably, the convergence issue [35] and unstructured exploration [36] usually pose an arduous challenge for reinforcement learning to be deployed for medical applications. The effectiveness of the proposed approach is verified with comprehensive real-world experiments on a phantom and multiple human subjects for spine curvature detection.

For future work, we would like to investigate system sensitivity to model parameters as it is not considered in the present work. Also, to locate the position of the spinous process more precisely, other forms of sensory information such as thermal images [37], [38] as well as various random image transformations techniques, such as rotation, horizontal and vertical flipping, can be employed to increase the robustness of spinal feature detection [39]. Our team is currently working along these two directions.

References

[1] A. M. Okamura, M. J. Matarić, and H. I. Christensen, "Medical and health-care robotics," *IEEE Robotics & Automation Magazine*, vol. 17, no. 3, pp. 26–37, 2010.

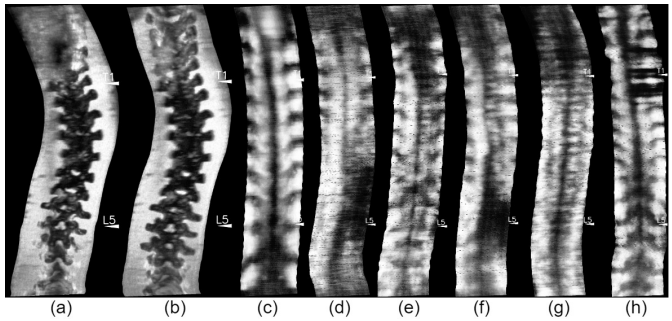

Fig. 14. Spine reconstruction of a *scoliosis* phantom model with contact force set as (a) 7 N and (b) 15 N, respectively; (c-h) Spine reconstruction of six human subjects (no signs of scoliosis) with interaction force being 10 N.

[2] M. Tirindelli, M. Victorova, J. Esteban, S. T. Kim, D. Navarro-Alarcon, Y. P. Zheng, and N. Navab, "Force-ultrasound fusion: Bringing spine robotic-US to the next "level"," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5661–5668, 2020.

[3] K. Li, Y. Xu, and M. Q.-H. Meng, "An overview of systems and techniques for autonomous robotic ultrasound acquisitions," *IEEE Transactions on Medical Robotics and Bionics*, 2021.

[4] Y.-P. Zheng, T. T.-Y. Lee, K. K.-L. Lai, B. H.-K. Yip, G.-Q. Zhou, W.-W. Jiang, J. C.-W. Cheung, M.-S. Wong, B. K.-W. Ng, J. C.-Y. Cheng *et al.*, "A reliability and validity study for Scolioscan: a radiation-free scoliosis assessment system using 3D ultrasound imaging," *Scoliosis and spinal disorders*, vol. 11, no. 1, pp. 1–15, 2016.

[5] H. Hase, M. F. Azampour, M. Tirindelli, M. Paschali, W. Simson, E. Fatemizadeh, and N. Navab, "Ultrasound-guided robotic navigation with deep reinforcement learning," in *IEEE/RSJ Int Conf on Intelligent Robots and Systems*, 2020, pp. 5534–5541.

[6] F. J. Abu-Dakka and M. Saveriano, "Variable impedance control and learning—A review," *Frontiers in Robotics and AI*, vol. 7, 2020.

[7] R. Droste, L. Drukker, A. T. Papageorghiou, and J. A. Noble, "Automatic probe movement guidance for freehand obstetric ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 583–592.

[8] G. Ning, X. Zhang, and H. Liao, "Autonomic robotic ultrasound imaging system based on reinforcement learning," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 9, pp. 2787–2797, 2021.

[9] K. Li, Y. Xu, J. Wang, D. Ni, L. Liu, and M. Q.-H. Meng, "Image-guided navigation of a robotic ultrasound probe for autonomous spinal sonography using a shadow-aware dual-agent framework," *IEEE Transactions on Medical Robotics and Bionics*, 2021.

[10] X. Deng, Y. Chen, F. Chen, and M. Li, "Learning robotic ultrasound scanning skills via human demonstrations and guided explorations," in *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2021, pp. 372–378.

[11] G. Nava, F. Romano, F. Nori, and D. Pucci, "Stability analysis and design of momentum-based controllers for humanoid robots," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 680–687.

[12] G. Nava, Q. Sablé, M. Tognon, D. Pucci, and A. Franchi, "Direct force feedback control and online multi-task optimization for aerial manipulators," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 331–338, 2019.

[13] D. Pucci, S. Traversaro, and F. Nori, "Momentum control of an underactuated flying humanoid robot," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 195–202, 2017.

[14] Y. Hu, M. Benallegue, G. Venture, and E. Yoshida, "Interact with me: An exploratory study on interaction factors for active physical human-robot interaction," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6764–6771, 2020.

[15] T. Boaventura, J. Buchli, C. Semini, and D. G. Caldwell, "Model-based hydraulic impedance control for dynamic robots," *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1324–1336, 2015.

[16] S. Huo, A. Duan, C. Li, P. Zhou, W. Ma, H. Wang, and D. Navarro-Alarcon, "Keypoint-based planar bimanual shaping of deformable linear objects under environmental constraints with hierarchical action framework," *IEEE Robot Autom Let*, vol. 7, no. 2, pp. 5222–5229, 2022.

[17] J. Lee, H. R. Warren, V. Agarwal, M. E. Huber, and N. Hogan, "Modulating hip stiffness with a robotic exoskeleton immediately changes gait," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 733–739.

[18] Y. Hu, X. Wu, P. Geng, and Z. Li, "Evolution strategies learning with variable impedance control for grasping under uncertainty," *IEEE Trans on Industrial Electronics*, vol. 66, no. 10, pp. 7788–7799, 2018.

[19] X. Zhang, L. Sun, Z. Kuang, and M. Tomizuka, "Learning variable impedance control via inverse reinforcement learning for force-related tasks," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2225–2232, 2021.

[20] C. Yang, C. Zeng, C. Fang, W. He, and Z. Li, "A DMPs-based framework for robot learning and generalization of humanlike variable impedance skills," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 3, pp. 1193–1203, 2018.

[21] A. Duan, R. Camoriano, D. Ferigo, Y. Huang, D. Calandriello, L. Rosasco, and D. Pucci, "Learning to sequence multiple tasks with competing constraints," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 2672–2678.

[22] M. Victorova, M. K.-S. Lee, D. Navarro-Alarcon, and Y. Zheng, "Follow the curve: Robotic ultrasound navigation with learning based localization of spinous processes for scoliosis assessment," *IEEE Access*, 2022.

[23] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intelligent service robotics*, vol. 9, no. 1, pp. 1–29, 2016.

[24] A. Duan, R. Camoriano, D. Ferigo, Y. Huang, D. Calandriello, L. Rosasco, and D. Pucci, "Learning to avoid obstacles with minimal intervention control," *Frontiers in Robotics and AI*, vol. 7, p. 60, 2020.

[25] D. Navarro-Alarcon, Y.-H. Liu, J. G. Romero, and P. Li, "Energy shaping methods for asymptotic force regulation of compliant mechanical systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 6, pp. 2376–2383, 2014.

[26] Y. Li, A. Duan, A. Gratner, and L. Feng, "A geometric programming approach to the optimization of mechatronic systems in early design stages," in *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2016, pp. 1351–1656.

[27] A. Duan, R. Camoriano, D. Ferigo, D. Calandriello, L. Rosasco, and D. Pucci, "Constrained DMPs for feasible skill learning on humanoid robots," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–6.

[28] Y. Zhu, X. Chen, and C. Li, "Some discussions about the error functions on so (3) and se (3) for the guidance of a uav using the screw algebra theory," *Advances in Mathematical Physics*, vol. 2017, 2017.

[29] O. Zahra, S. Tolu, P. Zhou, A. Duan, and D. Navarro-Alarcon, "A bio-inspired mechanism for learning robot motion from mirrored human demonstrations," *Frontiers in Neurorobotics*, vol. 16, 2022.

[30] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[31] K. K.-L. Lai, T. T.-Y. Lee, M. K.-S. Lee, J. C.-H. Hui, and Y.-P. Zheng, "Validation of scolioscan air-portable radiation-free three-dimensional ultrasound imaging assessment system for scoliosis," *Sensors*, vol. 21, no. 8, p. 2858, 2021.

[32] E. Coumans and Y. Bai, "PyBullet, a python module for physics simulation for games, robotics and machine learning," http://pybullet.org, 2016–2021.

[33] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China: IEEE, May 9-13 2011.

[34] "QP solvers for Python," https://github.com/stephane-caron/qpsolvers.

[35] Y. Li, K. H. Johansson, and J. Mårtensson, "Lambda-policy iteration with randomization for contractive models with infinite policies: Well-posedness and convergence," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 540–549.

[36] A. Raffin, J. Kober, and F. Stulp, "Smooth exploration for robotic reinforcement learning," in *Conference on Robot Learning*. PMLR, 2022, pp. 1634–1644.

[37] L. Hu, A. Duan, M. Li, A. Cherubini, L. Li, and D. Navarro-Alarcon, "Paint with the sun: A thermal-vision guided robot to harness solar energy for heliography."

[38] L. Hu, D. Navarro-Alarcon, A. Cherubini, M. Li, and L. Li, "On radiation-based thermal servoing: New models, controls, and experiments," *IEEE Transactions on Robotics*, pp. 1–15, 2021.

[39] Y. Sun, W. Zuo, P. Yun, H. Wang, and M. Liu, "Fuseseg: semantic segmentation of urban scenes based on rgb and thermal data fusion," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 1000–1011, 2020.