IEEE Robotics and Automation Letters (RAL) paper presented at the
2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

End-to-end Tactile Feedback Loop: From Soft Sensor Skin over Deep GRU-Autoencoders to Tactile Stimulation*

Andreas Geier[1,2], Rawleigh Tucker[1], Sophon Somlor[1],
Hideyuki Sawada[3], and Shigeki Sugano[1]

***Abstract*— Tactile feedback is a key sensory channel that contributes to our ability to perform precise manipulations. In this regard, sensor skin provides robots with the sense of touch making them increasingly capable of dexterous object manipulation. However, in applications like teleoperation, the complex sensory input of an infinite number of different textures must be projected to the human user's skin in a meaningful manner. In addressing this issue, a deep gated recurrent unit-based autoencoder (GRU-AE) that captured the perceptual dimensions of tactile textures in latent space was deployed to implicitly understand unseen textures. The expression of unknown textures in this latent space allowed for the definition of a control law to effectively drive tactile displays and to convey tactile feedback in a psycho-physically meaningful manner. The approach was experimentally verified by evaluating the prediction performance of the GRU-AE on seen and unseen data that were gathered during active tactile exploration of objects commonly encountered in daily living. A user study on a custom-made tactile display was conducted in which real tactile perceptions in response to active tactile object exploration were compared to the emulated tactile feedback using the proposed tactile feedback loop. The results suggest that the deep GRU-AE for tactile display control offers an efficient and intuitive method for efficient end-to-end tactile feedback during active tactile texture exploration.**

I. INTRODUCTION

Besides visual perception, tactile sensing constitutes a major sensory channel in the perception of physical properties. The physical contact enables direct transmission of forces and vibrations, which inform us not only about macro-geometric object and micro-geometric surface properties but also about physical quantities like weight distributions in otherwise unstructured environments. Human tactile perception is made possible by the combination of proprioceptive (intrinsic) and cutaneous (extrinsic) mechanoreceptors. The latter are distributed sets of sensory cells that are directly located in the human skin. They provide direct tactile feedback by perceiving forces, pressure, vibrations, and temperatures across the skin.

*This research was supported by the JSPS Grant-in-Aid for Young Scientists (B) No.17K18183 and (B) No. 19K14948 and the Grant-in-Aid for Scientific Research No. 19H02116 and No. 19H01130. Additional financial support was provided by the Ministry of Education, Science, Sports and Culture of Japan (Monbukagakusho).

[1]The authors are with Faculty of Science and Engineering, Department of Modern Mechanical Engineering, Waseda University, 169 Tokyo, Japan. a_geier@sugano.mech.waseda.ac.jp

[2]The author is with Rostock Medical Center, Department of Orthopaedics, 18057 Rostock, Germany

[3]The author is with Faculty of Science and Engineering, Department of Applied Physics, Waseda University, 169 Tokyo, Japan.

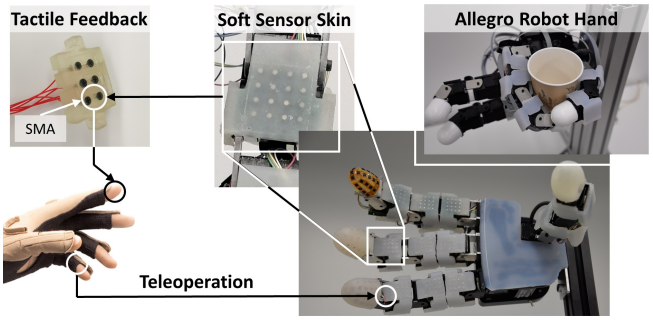

Fig. 1. *Allegro* robot hand with implemented soft sensor skin manipulates typical objects of daily living via teleoperation using the CyberGlove II [3]. Tactile feedback from anthropomorphic robots interacting with their environment is useful to human operators as it facilitates human-robot interaction.

In modern industrial applications as well as in emerging fields like social and service robotics, anthropomorphic dexterous robot hands are increasingly deployed, because they exhibit higher manipulation capabilities and therefore a higher flexibility [1], [2]. In this context, it has been shown that autonomous robot operation tremendously benefits from tactile sensing as it provides intrinsic and extrinsic tactile feedback that facilitates safe operation and robust, dexterous object manipulation [1], [2].

The implementation of a wearable tactile feedback system that can transmit the tactile sensor data from a robot onto a human operator's skin would establish the tactile sensory information flow between robots and humans, thus, enable powerful and immersive teleoperation. One core challenge in developing such a tactile feedback loop is the transformation of the tactile sensor data that arises from the robot's active exploration of a potentially infinite number of different contact surfaces into an actuator driving signal that provides physiologically meaningful stimulation of the cutaneous mechanoreceptors by a tactile display. An end-to-end mapping from tactile sensor data to a tactile actuator driving signal, which results in convincing tactile stimuli with respect to the technical design and transduction principles of both the deployed tactile sensor and the tactile actuator would be desirable.

In this paper, we introduce an end-to-end tactile feedback loop for teleoperation applications using anthropomorphic robot hands. We used a *uSkin* sensor module to actively explore textures and deployed a deep gated recurrent unit-based autoencoder (GRU-AE) to encode the retrieved tactile sensor readings into perceptual coordinates of a psychophys-

ically meaningful latent space. The auto-compression of tactile sensor data into the psychophysical layer of human material perception enabled an end-to-end mapping from tactile sensor data to tactile actuator driving signal. In this manner, we were able to drive a custom-made tactile display module based on shape-memory alloy actuators that can generate convincing tactile stimuli even for new, unseen textures alone from the skin sensor readouts.

The remainder of the paper is organized as follows: In **Section II** we review related works of recent tactile sensor and state-of-the-art tactile display developments. Following this, in **Section III**, we present the system architecture. Then, **Section IV** describes the experimental setup and procedure that were used to evaluate our system on experimental data with users. **Section V** shows the results of our study and evaluates them in the light of human tactile recognition. Finally, **Section VI** draws conclusions and gives directions for future works.

II. RELATED WORKS

A. Tactile sensors and skin sensors

Recently, the development and potential applications of tactile sensors attracted much interest leading to the development of a variety of tactile sensors with different transduction principles and design criteria [2], [4]. For example, the tactile sensor *GelSlim* [5] is a high-resolving optical tactile sensor. Even though it has a comparatively low frequency response, it is capable of detecting very detailed macro-geometric and fine micro-geometric object features. The sensor uses a clear elastomer with a reflective coating, colored lights, and a camera to capture highly detailed images of objects. The BioTac sensor [6] is a multi-modal sensor combining three transduction principles. Normal forces are measured with an array of 24 electrodes with a resolution of $0.01N$ at up to $100Hz$. Moreover, vibrations ranging from 10-$1040Hz$ can be measured using the integrated pressure sensor that senses the oscillations of a fluid contained in the sensor's finger tip. These sensors have been integrated into several anthropomorphic robot hands, e.g. the *Shadow Hand* [7] , the *BarretHand* [8] and the *Allegro Hand* [9]. In this regard, our lab has been developing *uSkin*: a compact and soft 3-axis skin sensor [10] for the implementation into anthropomorphic robot hands [11]. Each tactile sensor element consists of a small magnet that is suspended on top of a Hall-effect sensor with an air gap in between allowing for the sensing of 3D force vectors. One *uSkin* sensor module carries an array of 16 3-axis taxels in an area of 26x27mm and is only 4mm thick.

B. Tactile texture recognition

Tactile sensing allows for the discrimination of objects and, moreover, the estimation of physical properties for which reason it is of great interest to the robotics community [2], [4]. Commonly, texture recognition requires sensors with high sampling frequency and fast frequency response for identifying textures: Fishel *et al.* deployed their *BioTac* sensor in combination with an approach termed Bayesian exploration for direct identification of a total of 117 textures by analyzing reaction forces and occurring vibration signals. Kaboli and Cheng [12] proposed a set of defined tactile descriptors inspired by *Hjorth* parameters deployed in real-time electroencephalography signal analysis. In terms of texture recognition, this approach allowed for the robust recognition of 120 materials by analyzing the vibro-tactile signals while executing human-like exploratory movements using a plurality of different and multi-modal tactile sensors. Chathuranga *et al.* [13] combined the calculation of a set of descriptors with artificial neural networks to discriminate seven wooden textures with their biomimetic soft fingertip sensor. Takahashi and Tan [14] employed a convolutional autoencoder technique to implicitly relate the tactile sensor data that was gathered during experiments with a Sawyer robot stroking across 25 surfaces of different materials to the respective images of these material surfaces in order to enable the visual perception of tactile properties. In this context, Polic and co-workers [15] employed a convolutional autoencoder to perform unsupervised feature extraction from tactile sensor data on the basis of an optical-based tactile sensor in order to identify a universally useful set of features that could be used for a variety of manipulation tasks. Clearly, the complementation of robotic manipulators with tactile sensors has demonstrated great potential, as the availability of physical information about the interactions with the environment across the robot hand greatly enhances automation and manipulation capabilities [1], [2], [4].

C. Tactile displays

In contrast to tactile sensors, tactile displays are arrays of actuators that stimulate the cutaneous receptors in order to resemble the extrinsic sense of touch by the projection of forces, pressure, vibration or temperatures onto the human skin. The realization of tactile displays yields the potential of rich and direct cutaneous tactile feedback in the form of wearable devices, however, poses severe challenges in regards to the actuator technology [16], [17]. Mizukami [18], Fukuyama [19] and co-workers investigated the use of an array of up to nine shape-memory alloy wires that produce micro-vibrations to transmit sensations of moving patterns and surface properties. Wang *et al.* [20] developed a biomechanically optimized tactile display that relies on lateral skin deformation. It implements 60 laterally moving skin contactors with a spatial resolution of $1.8\times1.2mm^2$. Kajimoto *et al.* [21] presented an electro-tactile display with 512 electrodes that displayed tactile images based on camera data of a smartphone by electrical stimulation. Mun and co-workers [22] developed a soft tactile actuator based on an electro-active polymer that activates with an electrical voltage and moves vertically to push the skin of the wearer. The maximum protrusion is $650\mu m$ with a maximum force of $255mN$. One actuator has a diameter of around $15mm$ and can be worn on the finger phalanges and the forearm. Young and Kuchenbecker [23] presented a 6-DOF parallel manipulator for tactile stimulation of the fingertip. This device uses a single end-effector to display normal and shear forces at any

location on a fingertip. A more comprehensive review on tactile displays and related technology can be found in [16], [17], [24].

Yet, the projection of tactile sensor data onto the human skin by the use of tactile displays, i.e., the generation of convincing tactile stimuli with respect to the technical design and transduction principles of both the deployed tactile sensor and the tactile actuator constitutes a major technical challenge. The implementation of a vastly applicable end-to-end mapping from tactile sensor data to tactile actuator driving signal to establish a powerful tactile feedback loop is addressed in the present study.

III. End-to-end Tactile Feedback Loop

A. Overview of the system architecture

The components of the complete system architecture to achieve end-to-end tactile feedback, i.e., the **uSkin sensor module** for the deployment on robot hands, our custom-made **SMA-based tactile display module** for the generation of micro-vibrations, the **Deep GRU-AE** for implicit texture understanding and tactile display control, as well as the **experimental studies** are elaborated on in the following sections.

B. uSkin sensor module

The *uSkin* sensor is a three-axis *Hall* effect-based skin sensor with a silicone structure that holds a magnet. The sensor module delivers 3-axis raw sensor readouts across the module's surface (x- and y-direction, shear forces) and along the surface normal (z-direction, normal forces). The skin sensor module [25] used in this study was similar to the previously presented compact, soft, and distributed 3-axis sensor module *uSkin* for the deployment to robots [10] and was originally designed to cover the fingers of the *Allegro* robot hand [11] (Wonik Robotics[1]) for tactile sensing during object manipulation, Fig. 1. A microcontroller unit (*MTB3* [26], originally developed by the *Italian Institute of Technology*) served as the master and managed the tactile sensor readings via I2C (Inter-Integrated Circuit) communication. In the experiments, a flat 4x4 *uSkin* module was used and the *MTB3* was connected via a CAN (Controlled Area Network) bus and a CAN/USB (Universal Serial Bus) converter[2] to the host computer. Aiming for an end-to-end design, exclusively uncalibrated sensor readings at a sampling rate of approximately $100Hz$ were used. Note, since the *uSkin* sensor module is designed for the implementation onto robot hands, it comes covered with a gripping tape to protect the silicone structure and to increase the friction coefficient to facilitate in-hand manipulation. However, the high friction made the natural exploration of some textures difficult.

[1]http://www.simlab.co.kr/

[2]https://esd.eu/en

C. SMA-based tactile display

The tactile display module utilized a small shape-memory alloy (SMA) wire to generate micro-vibrations and project tactile stimuli onto the skin [19] [27]. The actuator was electrically controlled by periodic current pulses for the presentation of tactile sensations. Fig. 2-B and Fig. 3-B show the tactile display module with three actuators implementing three $5mm$-long SMA wires with a diameter of $0.1mm$. The SMA wire shrinks around 5% lengthwise when transitioning between the martensite and austenite phases due to an electrically induced temperature change. By driving the SMA wire with a pulse signal current, a periodic micro-vibration of up to $300Hz$ was generated. One pulse had an amplitude of $H[V]$ and a width of $W[ms]$, hence the duty ratio of the pulse determined the heating/ cooling times of the SMA and the electrical energy $H[V] \times W[ms]$ controlled the strength of the vibrations. The SMA actuator had a frequency bandwidth of around $10Hz$ to $300Hz$. Thus, it can stimulate the fast-adapting mechanoreceptors *Meissner* and *Pacini*, yet does not cover the full range of the *Pacinian* frequency bandwidth of up to $1kHz$. Moreover, the SMA actuator design used in this study was not capable of projecting the full variety of mechanical tactile stimuli on the skin: For example, the projection of shear forces due to lateral skin stretch or the conveyance of softness-hardness would require a complementation with other actuators [17].

The implementation of three of these actuators into a small 3D-printed package enabled the generation of stimuli that resemble the occurrence of various moving textures under the resting finger with different speeds. The percept of phantom actuators and seemingly moving stimuli are psychological effects of the human higher-level perception, i.e the phantom sensation (PS) [28] and the apparent movement (AM) [29], respectively. Hence, the actuators were driven independently by current pulse signals $A_i(t)$, which were parameterized depending on the target texture. The actuator activation probability $A_i(t)$ was given by a Gaussian distribution, thereby, completely defined the timely occurrence and density of the micro vibrations for the emulation of different textures:

$$A_n(t) = \alpha + \beta e^{-(t-\mu)^2 \big/ 2\sigma^2} \text{ for } n \in 1,...,N_A, \qquad (1)$$

where t is the time in ms, μ is the mean, σ^2 is the variance, α is the offset, β is the gain, and N_A is the total number of available actuators. Note that $\alpha + \beta < 1$.

Based on the known parameters $c_{i,j}$ in the intervals $c_i \in [\underline{\mu}, \overline{\mu}]$ and $c_j \in [\underline{\sigma}, \overline{\sigma}]$ in control space C that produced a desired textural stimulation [19], a linear relation to the corresponding latent space dimension $z_{l,p}$ was assumed. Thus the driving signal $A_i(t)$ was constructed by calculating the driver parameters from the latent space coordinates as encoded by the GRU-AE:

$$c_{i,j} = \frac{\underline{c}_{i,j} - \overline{c}_{i,j}}{\underline{z}_{l,p} - \overline{z}_{l,p}} (z_{l,p} - \overline{z}_{l,p}) + \overline{c}_{i,j}. \qquad (2)$$

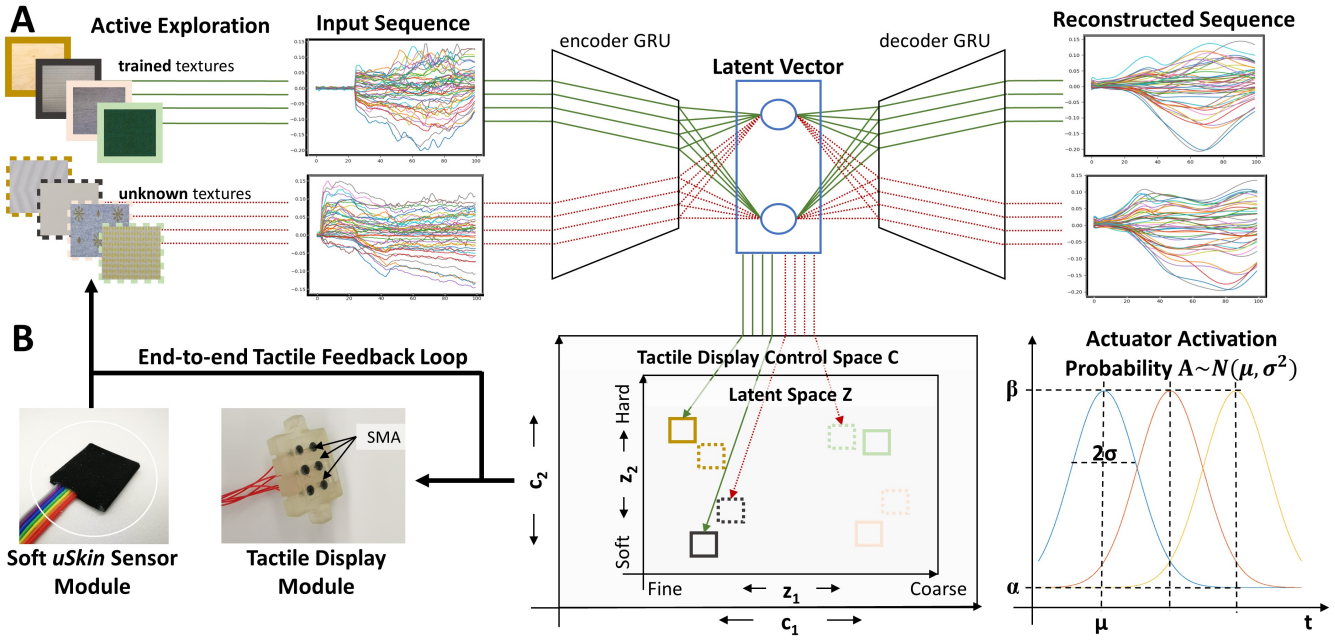

Fig. 2. Active tactile texture recognition with tactile feedback loop. The *uSkin* sensor module, originally designed to be implemented on robot hands, was used to stroke across various objects of daily living. A deep GRU-based autoencoder was deployed to obtain an implicit representation of the textures in latent space (A). The latent space variables $\boldsymbol{Z}$ were used to construct an actuator driving signal $\boldsymbol{C}$ to control a SMA-based tactile display module (B).

D. Deep GRU-AEs for active tactile texture recognition

For safety reasons and to facilitate the manipulation of arbitrary objects, skin sensors for the deployment to robot hands are often designed to be soft and compliant [1], [2], [30]. Since soft material layers act as a lowpass filter, the direct extraction of texture-specific vibrations for texture recognition may become difficult. However, it has been shown [31] that the time-series data may contain frequency-related features that are unique to a type of texture: A deep gated recurrent unit (GRU)-recurrent neural network [32] was deployed to capture abstract temporal features that allowed for accurate texture recognition in a preliminary study on active tactile texture multi-class classification [31]. Likewise, the GRU-AE's encoder comprised three stacked GRU layers with 96 units to capture highly abstract temporal features. However, instead of one dense-layer with soft-max activation for multi-class classification between a limited number of textures, another GRU layer compressed the temporal feature sequence into a point in the latent space $\mathbf{Z}$. No regularization was implemented as the small number of latent space variables acts as a *reconstruction loss*. Table I summarizes the GRU-AE architecture as used in our studies. The GRU units were implemented in their originally introduced version as reported by Cho *et al.* [32]. The flat *uSkin* sensor module contains 16 taxels that measure force quantities in the x,y,z-axes. Due to failure, the measurements of one taxel needed to be discarded for all experimental trials resulting in a maximum feature number of $N_F = 45$. During back-propagation, the network's weights were optimized using Adam [33]. The learning rate was set to α=10^{-5}, the exponential decay rate for the first moment estimates to β_1=0.900, the exponential decay rate for the second moment estimates to β_2=0.999, and the decay to δ=0.0. The loss function was defined as *mean squared error* between measured raw sensor data and their reconstructed counterpart.

TABLE I
GATED-RECURRENT-UNIT AUTOENCODER ARCHITECTURE

Layer	Type	Output Shape	Units	Remarks
0	Input	(_, _, 45)	/	
1	GRU	(_, _, 96)	96	returns seq.
2	GRU	(_, _, 96)	96	returns seq.
3	GRU	(_, _, 96)	96	returns seq
Encoding	GRU	(_, 3)	3	latent vector **z**
4	Repeat	(_, 100, 3)		
5	GRU	(_, 100, 3)	3	returns seq.
6	GRU	(_, 100, 96)	96	returns seq.
7	GRU	(_, 100, 96)	96	returns seq.
8	GRU	(_, 100, 96)	96	returns seq.
9	Dense	(_, 100, 45)	45	Linear Activation

Regularization implicitly included as *reconstruction loss*.
Total trainable parameters: 299,106.

E. Psychophysical space of textures

The psychophysical space is the lower layer of material perception [34], [24], in which a material is described by means of its coordinates in the dimensions *hard - soft*, *rough - fine*, *high-friction - low-friction*, and *cold - warm*. This definition implies that these quantities are, firstly, physical quantities, and, secondly, they can be directly perceived by humans with their mechanoreceptors. This layer is the most primal layer of material perception and is regarded as an interface between human responses and physical stimuli. Hence, the evaluation with respect to these dimensions were regarded as a common ground truth against which the performance of the tactile feedback loop were quantitatively measured during the user studies.

As these psychophysical coordinates allow humans to rec-

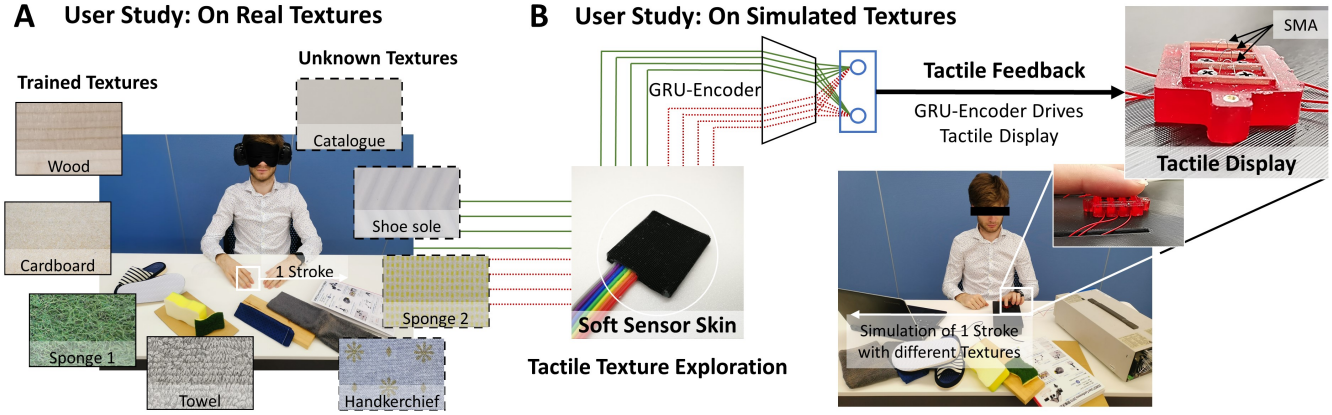

Fig. 3. Experimental setup for the evaluation of the psychophysical dimensions of the texture probes (A) and the proposed tactile feedback loop (B).

ognize different textures, we hypothesized that the previously presented deep GRU-classifier [31] would be able to project the tactile sensor data into a meaningful latent space $\mathbf{Z}$ that coincided with the psychophysical space. Thus the derived AE-architecture of the deep GRU-AE as shown in Table I would encode the tactile sensor readings along a timely trajectory into coordinates related to the psychophysical space. Accordingly, relatively few textures would be necessary to span the latent space across the psychophysical dimensions. This fact motivated the choice of a small training set of only four objects of daily living with mutually different texture properties: a polishing *sponge1*, a cotton *towel*, a board of *wood*, and a piece of *cardboard*. To test this hypothesis, the latent encodings of four completely unknown textures, namely: a softer *sponge2*, a *handkerchief*, a *catalogue* back cover, and a *shoe sole* were evaluated. Finally, the perceived psychophysical coordinates were compared to the coordinates in the GRU-AE's latent space of both the trained as well as the unknown textures by conducting a user study on the very same textures. An overview of the materials is given in Fig. 3-A. Note, the dimension *cold - warm* was not investigated, since the *uSkin* tactile sensor does not allow for determining the heat flow.

IV. Experiments and User Studies

The experiments comprised the data collection of the tactile data using the soft *uSkin* sensor module as well as two user studies on the perception of the above-defined textures and their evaluation in terms of the psychophysical dimensions Fig. 3. Within the first user study, the participants were asked to explore the real textures, while in the second user study, the participants were asked to evaluate the perceived stimulation provided by the proposed end-to-end tactile feedback loop from soft tactile sensor input to the tactile display output. Finally, both the perceptions were compared to draw conclusion on the performance of the end-to-end tactile feedback loop.

The user evaluations are depicted as box plots, since user ratings are ordinal data that are not necessarily normally distributed or symmetric in general. Additionally, *Mann-Whitney-U* tests were conducted on the psychophysical percept evaluations between real textures and the artificial tactile feedback. In this manner, statistically significant discrepancies between a real and an emulated tactile feedback were identified.

A. Experimental setup and GRU-AE implementation

For the experimental evaluation, the flat *uSkin* sensor module similar to the ones implemented onto the *Allegro* robot hand was used to manually explore the surface of the different objects as depicted in Fig. 3-A. The time horizon for stroking across the texture was set to $2s$ comprising the impact on the object surfaces of the objects as well as the exploratory motion across the object. The manual exploration resembled a rather natural motion, in which one freely strokes across a surface in a linear pattern. Accordingly, the sensor module was tightly fixed onto the index finger, then moved across one texture at a time. The GRU-AE was trained exclusively on the textures *sponge1*, *towel*, *wood*, and *cardboard*. Each of these four textures was explored 20 times streaming the tactile data of 200-time steps into 80 *.csv*-files. In regard to the *uSkin* sensor module's sampling frequency of $100Hz$, the sequence length was set to S_L=100, which corresponds to 1s real-time and should enable a relatively responsive texture recognition. To improve the usability during the online deployment of the GRU-AE in conjunction with the tactile display, a sliding window approach was used during training resulting in 80 times 100 sequences of sequence length S_L=100 and covering the complete time horizon of the motion during training. The complete dataset was split into training (6335 sequences), test (1188 sequences), and cross-validation set (397 sequences). The proposed GRU-AE was implemented in the high-level neural networks API KERAS[3] with Tensorflow[4] backend. The calculations were run on an off-the-shelf PC: Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz and 32.0GB RAM, GeForce GTX 1080 Ti with 11 GB frame buffer. The GRU-AE was trained for 500 epochs with a batch size of one.

[3]https://keras.io/
[4]https://www.tensorflow.org/

B. User studies

In the first user study, four subjects were asked to freely stroke across the surface of objects defined in section III-E in a rather linear motion as depicted in Fig. 3-A. The subjects were blindfolded and wearing soundproof earmuffs to eliminate the effect of the other sensory channels on the evaluations. After each trial, the subjects were handed a questionnaire and asked to evaluate the perceived sensation in terms of their psychophysical dimensions with integer numbers from 1 to 6, e.g. 1-very soft or 6-very hard. Only one object at a time was presented to the subjects. All the eight objects were shown in a random order for three times and the evaluations previously made were made unavailable to reduce the training effect. Note, in an introductory briefing, it was made sure that the subjects understood the meaning of the quantities *hard*, *rough*, and *friction*. Moreover, before the very first trial, one very soft, rough, and high-friction texture, as well as a very hard, smooth, and low-friction texture, were presented to the subjects in order to allow the subjects to evaluate over the full range from the first object during the user studies. The second user study was similar, however, the subjects were asked to rest their index finger on top of the tactile display as shown in Fig. 3-B and evaluate the perceived stimulus from the tactile sensor via the tactile display in the same manner as the real textures.

V. RESULTS AND DISCUSSION

A. Textures and their psychophysical representation

The results of the user studies on the perceived tactile stimuli during active tactile texture exploration are depicted in Fig. 4. Since the tactile stimuli percepts in all of the dimensions are ordinal scores, results are shown in terms of the maximum, minimum and median values.

Generally, the ratings on the texture properties had a high to very high range. While the intra-subject ratings for one texture over all the trials were relatively constant, i.e. deviations were typically within 1 rating, they varied considerably between subjects and were surely one reason for the high variance in the results. In comparison to the roughness and friction, hardness was more consistently rated between trials and subjects.

As indicated above, the majority of subjects found it difficult to make a clear decision on a texture's friction and roughness. These ratings were considerably more ambiguous with friction having the highest variance among all dimensions for all the trials. Friction is perceived as a quantity related to the force counteracting and the dynamic shear deformations occurring during finger motion. Thus friction is related to the sticky-slippery or the moist-dry factor levels of the individual subject's finger pad. Especially the moist-dry factor could have had an impact on the perceived friction in the experiments and may have caused the high variances in the results. Although roughness and friction are independent dimensions, these two percepts are often understood related [24], because roughness physically influences the friction coefficient and counteracts the finger motion. This effect

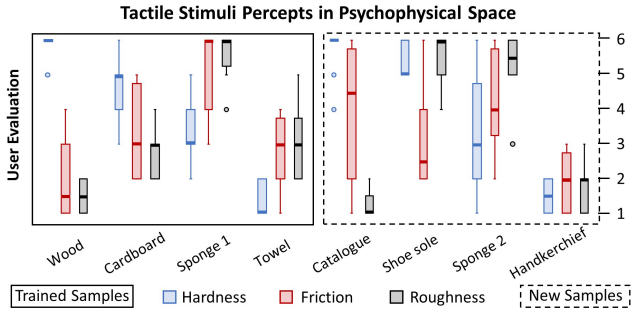

Fig. 4. Results of the user evaluation showing the perceived tactile coordinates for all the texture samples in the psychophysical space. The median is depicted as bold horizontal bar, outliers are depicted as circles.

might explain the higher variance of the ratings of both the roughness and the friction.

As it can be seen from Fig. 4, the textural property profile of the four chosen textures used for GRU-AE training exhibited a good variety in terms of their psychophysical coordinates and represented the psychophysical space well enough. For a good reconstruction of the tactile sensor data via the GRU-AE's bottleneck layer, the GRU-AE is forced to learn a strong and mutually independent representation of the tactile input data. While it was assumed that the chosen four textures force the GRU-AE to learn a good representation of the tactile data, it would be more desirable to sample the psychophysical space in a systematic manner. However, the coordinates of a texture, both in the psychophysical and the correlating latent space, will ultimately depend on the ratings of the subject cohort and might have considerable variations. Regarding the GRU-AE training, the textural property profile of the unknown set of textures was different enough from the textures that were included in the training set and allowed for judging the generalization capabilities of the GRU-AE.

B. Relation between latent and psychophysical space

Fig. 5-A shows the coordinates of the encodings of the tactile sensor readings in latent space and their coincidence with the dimensions of the psychophysical space (Fig. 4). Comparing the perceived hardness to the GRU-AE encodings, the encodings were consistently ordered in accordance with the median hardness percepts alongside the z_2-axis, therefore, coincided with the respective user ratings. Likewise, textures that were rated as rough and frictious were encoded alongside the z_1-axis with the textures exhibiting very high friction/ roughness being located at the right end of the latent space.

As it concerns the encodings of the unseen texture probes, their coordinates in latent space, thus, their relation to the psychophysical space were generally meaningful. The encodings of *catalogue* and *shoe sole* were placed between *wood* and *cardboard* along the z_2-axis. These encodings were in great agreement with the user ratings and in relation to the trained textures. Similarly, alongside the z_1-axis that was assumed to be a measure of friction/ roughness, the compression of *catalogue*, *shoe sole*, and *sponge2* were encoded in accordance with the user ratings and the trained

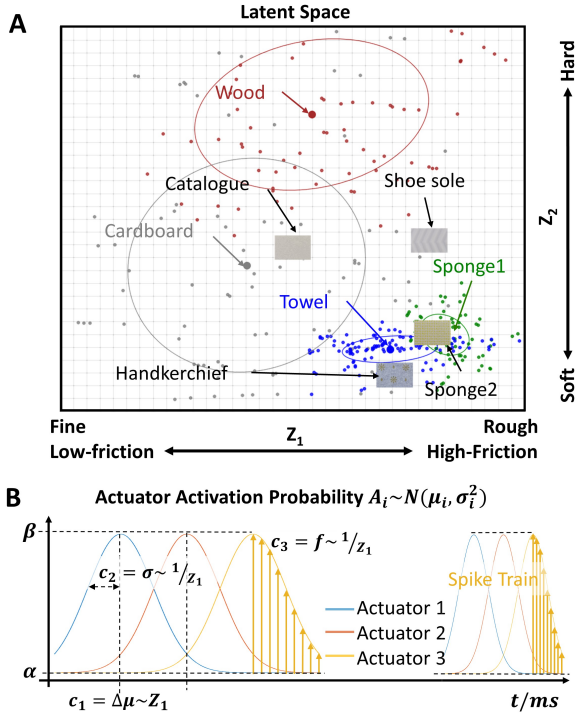

Fig. 5. The latent space of the trained GRU-AE (A) for the parameterization of the driving signal of the SMA actuators (B), 1σ-confidence ellipses.

textures. For example, the roughness of *shoe sole* was rated as high as the roughness of *sponge1* for which reason their z_1-coordinates were close. On the other hand, *sponge1* was rated to be much softer which resulted in a lower z_2-coordinate than *shoe sole*.

However, the z_1-coordinate of the encoding of *handkerchief* suggests a texture with a rather high surface friction or roughness which was in contrast to the user ratings. From the experiments it became clear that the material pairing of *handkerchief* and the gripping tape of the *uSkin* sensor module resulted in a very high friction coefficient. This was different to the subject's finger sliding easily across the *handkerchief* that was relatively consistently rated as smooth with low friction. Consequently, the modulated or lost tactile information due to the specifics of the deployed sensor limit the capacity of the tactile feedback loop.

Finally, comparing the distribution of the encoded tactile sensor readings in the latent space to their respective user evaluations, it became clear that the driving parameters μ, σ of the actuator's driving signal $A_i(t)$ acc. to equation (1) are strongly linked to the latent space coordinates along the z_1-axis (Fig. 5-B). This allowed to parameterize the control function (1) in dependence on the encoding in latent space. However, due to the inherent non-linearity of the GRU-AE approach, the relations between the encodings were not linear. As a result, the assumed linear mapping (2) between the latent space coordinates z_1 and the driver signal parameters might not be optimal in terms of their conveyed tactile information. Note, the parameter *Hardness* was not further considered, since hardness is mainly perceived as proprioceptive feedback instead of cutaneous feedback. Tactile

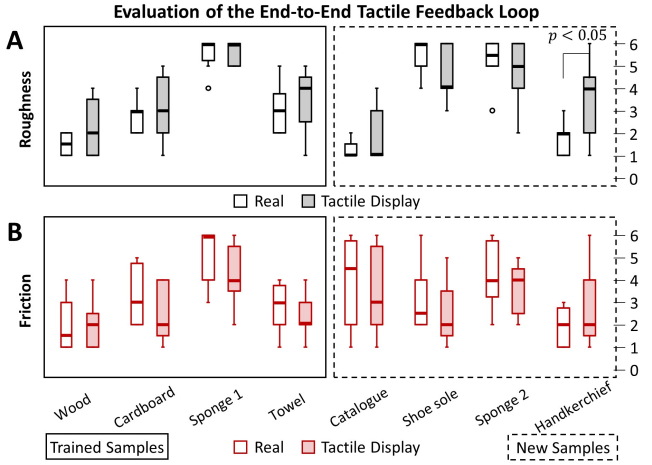

Fig. 6. Evaluation of the end-to-end tactile feedback loop for the dimensions roughness (A) and friction (B). Results are depicted as box plots. Note, a p-value<0.05 indicates a statistically significant difference (Mann-Whitney-U test).

displays are devices that primarily convey information on textural surface properties as cutaneous feedback.

C. Evaluation of the end-to-end tactile feedback loop

Fig. 6 shows the results of the perceived tactile stimuli as generated by the end-to-end tactile feedback loop based on the *uSkin* sensor data in comparison to the real tactile sensations from the exploration with one's own finger as box plots. Generally, the tactile feedback loop was able to generate realistic tactile sensations for all the seen as well as unseen textures. The median perceived psychophysical coordinate in response to the tactile feedback loop's stimulation was roughly similar to the median perceived psychophysical coordinate in response to the actual tactile exploration with one's own finger. Especially the textures with very distinct psychophysical coordinates, e.g. *sponge1* and *catalogue*, were very well perceived. In the case of the *cardboard*, the users even reported feeling the bumps from the corrugated structure inside typical cardboards.

However, the ranges of the percepts in the *Tactile Display*-group were often larger than in the *Real*-group, see Fig. 6-A. These variances indicate that the users found it more difficult to evaluate the roughness of an emulated texture when compared to a real texture. Moreover, a rather huge discrepancy was found between the perceived roughness and friction for the texture *handkerchief*. The stimuli provided by the tactile display were generally evaluated to be rougher and exhibiting a higher friction. Referring back to Fig. 5-A, it became evident that this issue originated from the GRU-AE, because it placed *handkerchief* rather close to *sponge2*. The analysis in the previous subsection would suggest that this was due to the higher relative friction coefficient between the *uSkin* sensor's gripper tape and the fabric of the *handkerchief*. This mismatch of relative friction coefficients between human finger and *uSkin* sensor might have resulted in biased tactile stimuli and emphasizes that hardware specifics can have a considerable effect on the perceived tactile stimuli.

VI. Conclusion

The contribution of our work lies in the introduction of an end-to-end tactile feedback loop for teleoperation applications in combination with anthropomorphic robot hands. We used a *uSkin* sensor module to actively explore textures and deployed a deep gated recurrent unit-based autoencoder to encode the retrieved tactile sensor data into perceptual coordinates of a psychophysically meaningful latent space. In this manner, we were able to drive a tactile display module for the generation of convincing tactile stimuli even for new, unseen textures alone from the skin sensor readouts. We experimentally verified the capabilities of our approach by comparing the perceived tactile sensations during the exploration of real textures to the emulated tactile feedback that was generated by our tactile feedback loop. It is important to note that the tactile sensor and the tactile actuator eventually determine the quality of the tactile stimuli. Hardware limitations in measuring and transmitting surface features impact both the physiological meaningfulness of the encodings into latent space and the projection of tactile stimuli on the human skin as it was demonstrated in our experiments. Future work is necessary to extend this approach to a wearable tactile feedback system to ultimately arrive at a compact and wearable solution for teleoperation applications.

References

[1] C. Piazza, G. Grioli, M. Catalano, and A. Bicchi, "A century of robotic hands," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 1–32, 2019.

[2] Z. Kappassov, J. A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands -— review," *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[3] "Cyber Glove Systems CyberGlove® II Wireless Data Glove," http://www.cyberglovesystems.com/support, accessed: 2020-02-19.

[4] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing—from humans to humanoids," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 1–20, 2010.

[5] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1927–1934.

[6] "SynTouch Inc. Resource & Download," https://syntouchinc.com/wp-content/uploads/2017/01/BioTac-Brochure.pdf, accessed: 2020-06-19.

[7] "Shadow Robot Company Shadow Dexterous Hand™," https://www.shadowrobot.com/products/dexterous-hand, accessed: 2020-06-19.

[8] "Barrett™ Advanced Robotics BarrettHand™," https://advanced.barrett.com/barretthand, accessed: 2020-06-19.

[9] "Wonik Robotics Allegro Hand," http://www.simlab.co.kr/Allegro-Hand.htm, accessed: 2020-06-19.

[10] T. P. Tomo, A. Schmitz, W. K. Wong, H. Kristanto, S. Somlor, J. Hwang, L. Jamone, and S. Sugano, "Covering a robot fingertip with uskin: A soft electronic skin with distributed 3-axis force sensitive elements for robot hands," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 124–131, 2018.

[11] S. Funabashi, G. Yan, A. Geier, A. Schmitz, T. Ogata, and S. Sugano, "Morphology-specific convolutional neural networks for tactile object recognition with a multi-fingered hand," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 57–63.

[12] M. Kaboli and G. Cheng, "Robust tactile descriptors for discriminating objects from textural properties via artificial robotic skin," *IEEE Transactions on Robotics*, vol. PP, pp. 1–19, 2018.

[13] D. S. Chathuranga, Z. Wang, V. A. Ho, A. Mitani, and S. Hirai, "A biomimetic soft fingertip applicable to haptic feedback systems for texture identification," in *2013 IEEE International Symposium on Haptic Audio Visual Environments and Games (HAVE)*, 2013, pp. 29–33.

[14] K. Takahashi and J. Tan, "Deep visuo-tactile learning: Estimation of material properties from images," *CoRR*, vol. abs/1803.03435, 2018.

[15] M. Polic, I. Krajacic, N. Lepora, and M. Orsag, "Convolutional autoencoder for feature extraction in tactile sensing," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3671–3678, 2019.

[16] C. Pacchierotti, S. Sinclair, M. Solazzi, A. Frisoli, V. Hayward, and D. Prattichizzo, "Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives," *IEEE Transactions on Haptics*, vol. 10, no. 4, pp. 580–600, 2017.

[17] H. Culbertson, S. Schorr, and A. Okamura, "Haptics: The present and future of artificial touch sensation," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 385–409, 2018.

[18] Y. Mizukami and H. Sawada, "Tactile information transmission by apparent movement phenomenon using shape-memory alloy device," *International Journal on Disability and Human Development*, vol. 5, pp. 277–284, 2006.

[19] K. Fukuyama, N. Takahashi, F. Zhao, and H. Sawada, "Tactile display using the vibration of sma wires and the evaluation of perceived sensations," in *2009 2nd Conference on Human System Interactions*, 2009, pp. 685–690.

[20] Q. Wang and V. Hayward, "Biomechanically optimized distributed tactile transducer based on lateral skin deformation," *The International Journal of Robotics Research*, vol. 29, no. 4, pp. 323–335, 2010.

[21] H. Kajimoto, M. Suzuki, and Y. Kanno, "Hamsatouch: Tactile vision substitution with smartphone and electro-tactile display," in *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2014, pp. 1273–1278.

[22] S. Mun, S. Yun, S. Nam, S. K. Park, S. Park, B. J. Park, J. M. Lim, and K. Kyung, "Electro-active polymer based soft tactile interface for wearable devices," *IEEE Transactions on Haptics*, vol. 11, no. 1, pp. 15–21, 2018.

[23] E. M. Young and K. J. Kuchenbecker, "Implementation of a 6-dof parallel continuum manipulator for delivering fingertip tactile cues," *IEEE Transactions on Haptics*, vol. 12, no. 3, pp. 295–306, 2019.

[24] H. Kajimoto, M. Konyo, and S. Saga, *Pervasive haptics: Science, design, and application*. Japan: Springer, 2016.

[25] T. P. Tomo, M. Regoli, A. Schmitz, L. Natale, H. Kristanto, S. Somlor, L. Jamone, G. Metta, and S. Sugano, "A new silicone structure for uskin—a soft, distributed, digital 3-axis skin sensor and its integration on the humanoid robot icub," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2584–2591, 2018.

[26] G. Cannata, M. Maggiali, G. Metta, and G. Sandini, "An embedded artificial skin for humanoid robots," in *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*. IEEE, 2008, pp. 434–438.

[27] H. Sawada, "Micro-vibration actuators driven by shape-memory alloy wires and its application to tactile displays," in *2017 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, 2017, pp. 1–5.

[28] D. S. Alles, "Information transmission by phantom sensations," *IEEE Transactions on Man-Machine Systems*, vol. 11, no. 1, pp. 85–91, 1970.

[29] G. v. Békésy, "Sensations on the skin similar to directional hearing, beats, and harmonics of the ear," *The Journal of the Acoustical Society of America*, vol. 29, no. 4, pp. 489–501, 1957.

[30] H. Iwata and S. Sugano, "Design of human symbiotic robot twendy-one," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 580–586.

[31] A. Geier, G. Yan, T. P. Tomo, S. Somlor, and S. Sugano, "Deep gru-ensembles for active tactile texture recognition with soft, distributed skin sensors in dynamic contact scenarios," in *2020 IEEE/SICE International Symposium on System Integration (SII)*, 2020, pp. 127–132.

[32] K. Cho, B. van Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. abs/1406.1078, 2014.

[33] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[34] S. Okamoto, H. Nagano, and Y. Yamada, "Psychophysical dimensions of tactile perception of textures," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 81–93, 2013.