THE VIRTUAL REALITY SOCIETY OF JAPAN

第 27 回日本バーチャルリアリティ学会大会論文集（2022 年 9 月）

Sound Pressure Field Reconstruction for Airborne Ultrasound Tactile Display by Neural Network

陳剣宇 [1)]，鈴木颯 [1)]，牧野泰才 [1)2)] ，篠田裕之 [1)2)]
Jianyu CHEN, Shun SUZUKI, Yasutoshi MAKINO, and Hiroyuki SHINODA

1) 東京大学 新領域創成科学研究科複雑理工学専攻 （〒277-8561 千葉県 柏市 柏の葉 5-1-5, 7707189878@edu.k.u-tokyo.ac.jp）
2) 東京大学 情報理工学系研究科システム情報学専攻 （〒113-8656 東京都文京区本郷 7-3-1）

Abstract: Airborne Ultrasound Tactile Display (AUTD) is usually used to generate a specific foci sound field to provide different tactile experiences. As a result, how to generate a specific foci field is widely studied. Here, we propose to use a neural network to train the data of phase sets and the coordinates of foci, and developed a model that only needs to input the foci coordinates which can obtain the corresponding phase set. Our proposed method can ignore the calculation of the transmission matrix, so we expect to apply this method even in the presence of time-invariant nonlinear effects between the field and the AUTD.
Keyword: Ultrasound, Tactile Display, Neural network

1. Introduction

The Airborne Ultrasound Tactile Display (AUTD) is usually used to present tactile sensations. The ultrasound transducers of AUTD whose phases can be individually controlled. The ultrasound waves emitted from these transducers interfere with each other and form a sound field in space [1]. By controlling the phases appropriately, the AUTD can present various sensations by generating a specific field. Monnai et al. used focus to develop a mid-air interaction system that allows users to touch a floating virtual screen with non-contact tactile feedback [2]. Inoue et al. formed an 11 foci sound field to present a 3D shape of a five-point star tactile sensation [3]. Similarly, Long et al. optimized the phases to achieve a 3D tactile sensation like virtual objects in the air [4].

All the applications above are developed by calculating the transmission matrix to obtain the objective phases. In general, optimizing the phase by the transmission matrix can provide a better experience. The fastest way to form a foci field is to use Linear Synthesis Scheme (LSS). As its name implies, generates the foci field by linearly synthesizing each focus signal without doing the optimization. However, the field generated by this method sometimes cannot form the foci clearly. Thus, some methods are developed to solve the problem. GS-PAT, which can optimize the transmission matrix obtained from LSS using GPU, greatly reduce the delay in generating the sound field [5]. Suzuki et al. proposed to use the Greedy Algorithm with Brute-Force Search to obtain the optimal phases. They treated each transducer in the AUTD as a separate problem, and searches for the best solution (i.e., strongest sound pressure in foci position) for each one to arrive at the overall best solution for the entire problem. Similarly, it still has to compute the transmission matrix within each exploration [6].

Currently, all methods are based on the transmission matrix to obtain the objective phase. If there is an obstacle between the target field and the AUTD, or if there are nonlinear factors, the transmission matrix will become difficult to calculate. Therefore, we propose to use a neural network to obtain the phase. We only need to train the phase data and the corresponding foci coordinates, which can ignore the transmission matrix.

At the present stage, we only consider the linear condition, and use the multilayer perceptron (MLP) model to learn the single-focus and two-focus sound fields. As long as the two-dimensional coordinates (x-y) of each focus are input, the model can obtain the corresponding phase. All experiments are run in simulation.

2. Methods

In this chapter, we will first describe how to obtain the dataset for our neural network. We will use the GS-PAT algorithm to obtain the data set. Therefore, we will briefly describe how GSPAT works. Finally, we will introduce the MLP model.

2.1 GS-PAT

GS-PAT uses an iterative method to optimize the phase set. It needs to calculate the phases from each focus to transducers for forward propagation first. Then the phases of back-propagation from the transducers to each focus need to be calculated as well. And normalize the amplitude of the transducers. By fixing the above three steps under the iterations, the optimal phase can be obtained.

2.2 Multilayer perceptron

A multilayer perceptron (MLP) is a fully connected class of feedforward artificial neural network. The basic structure of multi-layer perception consists of three parts: the first input layer, the hidden layers, and the last output layer. Our MLP has 5 hidden layers. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. This time, we used ReLU as activation function.

The structure of the MLP model is shown in Fig.1.

This time, we have four-dimensional inputs, the x-y coordinates of two foci. And we simulate the AUTD has 18×14 transducers, so the output is a set of 252 phases.

3. Evaluation

We randomly generate a single focus (while the two foci are overlapped in the same position) or two foci on a plane that is 150mm parallel to the AUTD. The center of the AUTD is the origin, the x-axis range is (-50, 50) mm, and the y-axis range is (-50, 50) mm. We recorded the x and y coordinates and the corresponding 252 phases as one sample. In this experiment, we collected a total of 15,000 samples, of which 80% were used as the training set and 20% were used as the test set.

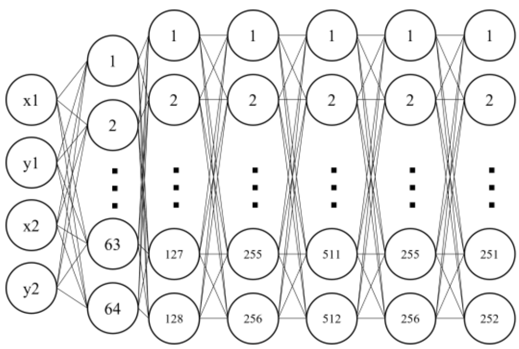

Figure 1: The multilayer perceptron model

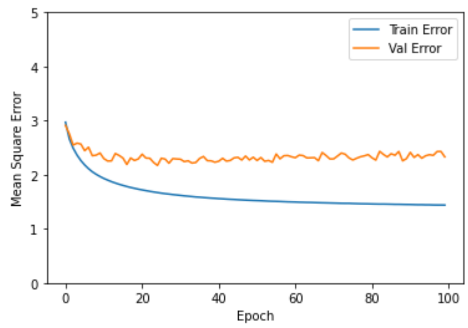

Figure 2: Mean square error

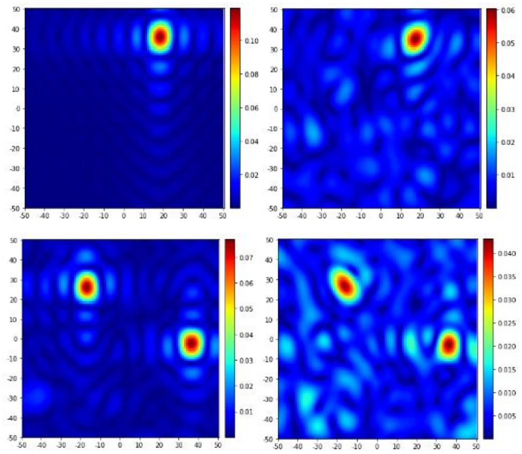
Figure 3: Evaluation of sound field. The fields on the left side are the ground truth generated by the GS-PAT. And the fields on the right side are the field generated by our test set.

The training was run in a desktop PC (Intel Core i9-9900X CPU @3.5GHz) with a single GPU (NVIDIA GeForce RTX 2080Ti). After the training, the evaluation graph of mean square error is shown in Fig.2.

All the phases of test set generate the field successful by the trained model. We chose one of the single focus fields and one of the two foci fields to show the results in Fig.3.

We can see the pressure of the foci generate by our model is weaker than the ground truth, but the foci are still form in the right position.

4. Discussion

The MLP model has generated the foci field successful, but with a weaker sound pressure. This will affect the experience of the tactile sensation, as a result, we will refine the NN model to improve the accuracy of the results.

Moreover, this time we used a method that requires the use of the transmission matrix to collect samples. But when testing the model, we can get the target phase without relying on the calculation of the transmission matrix, which provides more possibilities in the presence of nonlinear factors.

Above all, the NN we proposed which can reconstruct the foci field without relying on the transmission matrix, present a new way to provide various tactile experience.

Acknowledgement

This research is supported by JST PRESTO 17939983

Reference

[1] Tayayuki Hoshi, et al. 'Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound', IEEE Transactions on Haptics, Volume: 3, Issue: 3, July-Sept. 2010

[2] Yasuaki Monnai, et al. 'HaptoMime: mid-air haptic interaction with a floating virtual screen', UIST '14: The 27th Annual ACM Symposium on User Interface Software and Technology Honolulu Hawaii USA October 2014.

[3] S. Inoue and Y. Makino and H. Shinoda: 'Active touch perception produced by airborne ultrasonic haptic hologram'. 2015 IEEE World Haptics Conference (WHC), 362–367 (2015)

[4] B. Long and S. A. Seah and T. Carter and S. Subramanian: 'Rendering volumetric haptic shapes in mid-air using ultrasound'. ACM Transactions on Graphics (TOG), 1–10 (2014)

[5] Diego Martinez Plasencia, Ryuji Hirayama, Roberto Montano-Murillo, Sriram Subramanian: 'GS-PAT: high-speed multi-point sound-fields for phased arrays of transducers'. ACM Transactions on Graphics, Article No.: 138 (2020)

Shun Suzuki and Masahiro Fujiwara and Yasutoshi Makino and Hiroyuki Shinoda: Radiation Pressure Field Reconstruction for Ultrasound Midair Haptics by Greedy Algorithm with Brute-Force Search. IEEE Transactions on Haptics