MODELING HEAD-RELATED TRANSFER FUNCTIONS VIA SPATIAL-TEMPORAL GAUSSIAN PROCESS

Tatsuya Komatsu[1], Takanori Nishino[2], Gareth W. Peters[3], Tomoko Matsui[4] and Kazuya Takeda[1]

[1]Graduate School of Information Science, Nagoya University, Japan
[2]Graduate School of Engineering, Mie University, Japan
[3]Department of Statistical Science, University College London, UK
[4]Department of Statistical Modeling, Institute of Statistical Mathematics, Japan

ABSTRACT

We propose a novel application of a family of non-parametric statistical models to estimate head-related transfer functions (HRTFs) using spatial-temporal Gaussian processes (GPs). In this approach, we model the head-related impulse response (HRIR) utilizing non-parametric regression via a GP. The challenge posed by this problem involves accurate modeling of the spatial correlation structure jointly with the temporal correlation structure at each spatial location for the HRIR. We solve this problem by constructing a joint spatial-temporal kernel characterizing the GP regression model. To perform inference, we estimate the hyper-parameters of the GP regression kernel via maximum signal-to-deviation-ratio estimation on the basis of a real experimental setup in which we collected observations of the HRIR using two head-and-torso simulators (HATSs): KEMAR and B&K. We also perform cross validation of the model by training on the KEMAR system and assessing the generalization of our model and its out-of-sample predictive power for HRIRs at any locations that we predict by the model assessed on the B&K system. The corresponding HRTFs are obtained as the Fourier transform of the HRIRs. In the experiments, we show that our method is robust against variation in the azimuth interval needed to perform high-accuracy interpolation and has the expressive power to handle the individual characteristics of each HATS.

Index Terms— Head-Related Transfer Function, Head-related Impulse Response, Interpolation, Gaussian Process, Kernel Methods

1. INTRODUCTION

There is strong demand for the development of advanced telecommunications and for surround-sound audio systems that can artificially simulate the feeling of being immersed in a particular spatial environment, such as a room or concert hall when one is physically not located in that environment. For example, this may be useful when creating the atmosphere of a concert hall in a surround-sound system in a home theater or when creating the sense of direct conversation between people in a common spatial environment when they are physically in separate locations and communicating through (wireless) devices.

The current interest in this regard involves the development of the sensation of immersion in alternative acoustic environments, which are simulated artificially for a user by headphones or surround-sound audio devices. To this end, there has been research into what is known in frequency-domain sound modeling as an HRTF and in time-domain sound modeling as an HRIR. The development of statistical models to describe and parameterize the HRIR and HRTF that are accurate and robust and characterize both the spatial and temporal features will improve the synthesis of the acoustic characteristics of alternative environments for users and hence improve the synthesized feeling of immersion in such an environment, e.g., a concert hall.

More formally, HRTF models capture the response in the frequency domain that characterizes how an ear receives a sound from a point in space. Therefore, they are influenced by the shape of the human head and ear and the angles at which the sound wave is incident. In this regard, HRTF models should be flexible at adapting to particular unique characteristics of different anatomies. In general, in order to control a sound image localization accurately, one must measure omnidirectional HRTFs for each human *a priori*, and this clearly cannot be done in practice. Parameterizing all of these characteristics in an HRTF model will be non-generalizable, meaning that the model will not perform well in prediction for both in-sample and out-of-sample forecasting of the HRTF for angles for which the incident interference was not observed.

To meet this challenge, previous HRTF model design methods have involved interpolation from HRTFs measured over a discrete grid (mesh) of angles in the range $[0, 2\pi]$. Several HRTF interpolation methods that have been explored, and ones utilizing linear interpolation [1, 2], filter bank models[3, 4], principal component analysis (PCA) [5, 6, 7], Karhunen-Loeve expansion (KLE)[8], reconstruction via basis functions[9], spherical harmonics expansion and ambisonic sounds[10, 11] have been reported. However, the HRTF interpolation models in the conventional methods are simple and limited and the performance is neither sufficient nor generalizable.

In this paper, we demonstrate that one can outperform the simple-interpolation-based basis regression models previously proposed, in terms of predictive performance and model generalization by instead considering the flexible class of non-parametric statistical regression models based on GPs[12, 13] to perform such inference modeling for HRTF estimation. There are many ways that one can develop such a class of non-parametric models in the modeling of HRTFs and HRIRs. The key components of these models involve the specification of a distribution over a smooth random function (surface) with its mean surface representing the unknown HRTFs or HRIRs over a mesh of incident angles for a given anatomy configuration. In approaching the design of such a model, one must consider the spatial and temporal dependence features of the response: it is in this respect that we differ from the standard approach to such a problem, which involves separating the spatial and temporal dependences through a product space formulation, which is often common in machine learning and various applications of

GP modeling, e.g., motion tracking modeling[14], modeling gas distribution[15], environmental surveillance[16], modeling MRI brain images[17], transcriptional landscape estimation[18], clustering gene expression[19], inter atomic potential models[20], and modeling of wire-cut electrical discharge machining(WEDM)[21] as discussed in [22, 23]. In this paper, we consider the temporal and spatial features (co-variates) jointly in the covariance and mean functions. To simplify the design and estimation efficiency, we demonstrate the performance of the GP model under the isotropic spatial dependence assumption.

The characteristics of a GP model are ideal for an HRTF since it is a flexible class of models that can adapt to the varied anatomies of humans and all possible sound wave incident directions without the need to make an explicit parametric model. Instead, what we develop is a distribution over a random function that is uniquely characterized by the spatial and temporal dependence structures utilized to define the GP. These attributes can be learnt efficiently from data through specification of mean and covariance functions in the GP, as detailed below.

In short, the goal of this paper is to investigate the utilization of GPs in HRIR/HRTF interpolation and devise an effective method of HRTF interpolation. We assess the performance of this GP regression model by calibrating it with training data from KEMAR and then testing its generalization out-of-sample predictive performance spatially on test data from B&K.

2. HRTF NON-PARAMETRIC REGRESSION VIA A SPATIAL-TEMPORAL GAUSSIAN PROCESS

Here, we overview briefly how to define a spatial-temporal GP regression model and then describe how to model the HRIR using spatial-temporal GPs. Finally, the HRTF is obtained as the Fourier transform of the HRIR.

2.1. Gaussian Process Definition

A GP defines a distribution over a space of functions and it is completely specified by the equivalent of sufficient statistics for such a process. More formally, it is defined as follows [13].

Definition (Gaussian process): *Let $\mathcal{X} \subset \mathbb{R}^D$ be some bounded domain of a d-dimensional real-valued vector space. Denote by $f(\mathbf{x}) : \mathcal{X} \mapsto \mathbb{R}$ a stochastic process parametrized by $\mathbf{x} \in \mathcal{X}$. Then, the random function $f(\mathbf{x})$ is a Gaussian process if all its finite dimensional distributions are Gaussian, where for any $m \in \mathbb{N}$, the random variables $(f(\mathbf{x}_1), \cdots, f(\mathbf{x}_m))$ are normally distributed.*

A GP is formally defined by the following class of random functions:

$$\mathcal{F} := \{f(\cdot) : \mathcal{X} \mapsto \mathbb{R} \text{ s.t. } f(\cdot) \sim \mathcal{GP}(\mu(\cdot;\boldsymbol{\phi}), k(\cdot,\cdot;\boldsymbol{\Psi}))\},$$
$$\mu(\mathbf{x};\boldsymbol{\phi}) := \mathbb{E}[f(\mathbf{x})],$$
$$k(\cdot,\cdot;\boldsymbol{\Psi}) := \mathbb{E}[(f(\mathbf{x}_i) - \mu(\mathbf{x}_i;\boldsymbol{\phi}))(f(\mathbf{x}_j) - \mu(\mathbf{x}_j;\boldsymbol{\phi}))],$$

where at each point the mean of the function is $\mu(\cdot;\phi) : \mathcal{X} \mapsto \mathbb{R}$, which is parameterized by ϕ, and the spatial dependence between any two points is given by the covariance function (Mercer kernel) $k(\cdot,\cdot;\boldsymbol{\Psi}) : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}^+$, parameterized by unknown vector $\boldsymbol{\Psi}$, see details in [13].

Furthermore, we make the following notational definitions:

$$k(\mathbf{x}_*, \mathbf{x}_{1:N}) := \mathbb{E}[f(\mathbf{x}_*)\ f(\mathbf{x}_{1:N})] \in \mathbb{R}^N,$$
$$\mathcal{K}(\mathbf{x}_{1:N}, \mathbf{x}_{1:N}) := \begin{bmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & \cdots & k(\mathbf{x}_1,\mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N,\mathbf{x}_1) & \cdots & k(\mathbf{x}_N,\mathbf{x}_N) \end{bmatrix} \in \mathcal{S}^+\left(\mathbb{R}^d\right),$$

where $\mathcal{K}$ is known as the Gramm matrix defined on $\mathcal{S}^+\left(\mathbb{R}^d\right)$ corresponding to the manifold of symmetric positive definite matrices. Having formally specified the semi-parametric class of Gaussian process models, we proceed with presenting the system model for the HRIR under the GP framework.

2.2. HRIR Modeling via Spatial-Temporal Gaussian Process

Suppose that we observe N observations over a spatial-temporal domain utilizing a given selection of a HATS to produce training data, $\left\{\left(\mathbf{x}^{(i)}, z^{(i)}\right)\right\}_{i=1}^N$. In the model that we develop, element $\boldsymbol{x}^{(i)}$ will contain the location information (encoded by angle) and the temporal information (in ms) and the observations $z^{(i)}$ correspond to the HRIR corresponding to a given $x^{(i)}$. Then, the standard GP regression can be expressed as

$$\boldsymbol{Z}_{1:N} = \mathbf{h}_{1:N} + \boldsymbol{W}_{1:N}, \tag{1}$$

with the random function corresponding to the noisily observed HRIR evaluated at a set of points $\mathbf{x}_{1:N}$, denoted by the random vector $\mathbf{h}_{1:N} := [h(\mathbf{x}_1), \cdots, h(\mathbf{x}_N)]^T$ (here $h(\cdot)$ corresponds to $f(\cdot)$ in section 2.1), and $\boldsymbol{W}_{1:N}$ is a zero-mean i.i.d vector of Gaussian noise with covariance $\sigma_{\mathrm{w}}^2\mathbf{I}$. Furthermore, we consider zero mean GP models in this paper.

Having specified the generic GP framework in Section 2.1 and equation (1), we now make explicit the spatial-temporal GP model used for the HRIR formulation, given by

$$Z(t,\theta) = h(t,\theta) + W(t,\theta), \tag{2}$$

where $h(t,\theta)$ indicates the unknown random function with time t and location θ observed in noise, which is be assumed to be distributed according to a GP non-parametric distribution over smooth functions. Here, θ is the azimuth from the front of the head, and we assume $W_n(t) \sim \mathcal{N}(0,\sigma^2)$, which is independent of the random function $h(t,\theta)$. The diagram in Fig. 1 describes these quantities.

In specification of the GP model, the choice of kernel is an important consideration because it dictates the extend to which the joint spatial-temporal correlation structures of the HRIR will be adequately modeled by the GP. In this paper, we make the challenging choice of not separating space and time; instead, we incorporate both into a common kernel $k(\mathbf{x},\mathbf{x}')$, which is defined as follows for two vectors (angle,time) given by

$$\begin{aligned} k(\mathbf{x},\mathbf{x}') &= \exp\left(-|x-x'|\Sigma^{-1}|x-x'|\right) \\ \Sigma_{ij} &= \begin{cases} \sigma_i^2, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

The hyper-parameters σ_i are estimated so as to maximize the signal-to-deviation-ratio(SDR) as mentioned below. Therefore, the two stages involve in modeling the HRIR via a GP framework are as follows:

Azimuth interval	Number of HRIRs
5° and 10°	5
15°	4
30°	3
45° or larger	2

Table 1. The numbers of HRIRs used in interpolation for the azimuth intervals of training data measured.

Stage 1: Estimation of GP Hyper-Parameters for Spatial-Temporal Kernel.

The first involves estimation of the hyper-parameters $\mathbf{\Psi}$ composing the spatial-temporal kernel $k(\cdot,\cdot;\mathbf{\Psi})$. Note that we estimate them via a maximum SDR procedure with a grid search. This procedure was found to be more robust to over-fitting of the GP model that can occur when compared to the maximum-likelihood approach. It involved using SDR with the hyper parameter $\mathbf{\Psi}$, $d(\mathbf{\Psi})$ between a measured HRIR value $\hat{h}(t,\theta)$ and the GP model mean value $\overline{h}(t,\theta;\mathbf{\Psi})$ which is estimated using the spatial-temporal kernel. Therefore the estimate was iteratively found as the solution to equation (4).

$$\begin{aligned} d(\mathbf{\Psi}) &= \frac{1}{M}\sum_m 10\log\frac{|\hat{h}(t,\theta_m)|^2}{|\hat{h}(t,\theta_m)-\overline{h}(t,\theta_m;\mathbf{\Psi})|^2} \\ \widehat{\mathbf{\Psi}}^{SDR} &= \arg\max_{\mathbf{\Psi}} d(\mathbf{\Psi}). \end{aligned} \tag{4}$$

where M is the number of all interpolated azimuths and θ_m is the *m*-th azimuth for interpolation. Here, we make explicit in the notation for $\overline{h}(t,\theta;\mathbf{\Psi})$ the dependence of the mean of the GP model on the hyper-parameters $\mathbf{\Psi}$ in the GP HRIR model. We selectively used a few HRIRs with the closest degrees according to the condition in interpolation. Table 1 lists the numbers of HRIRs for the azimuth intervals of training data measured for interpolation. $\hat{h}(t,\theta)$ and $\overline{h}(t,\theta;\mathbf{\Psi})$ are scaled by the maximum amplitude value.

Stage 2: GP Spatial-Temporal Prediction for the HRIR.

The second stage involves predicting at new space and time points for which observations were not made. Under this model, we can express the mean prediction function, at any 'location' $\mathbf{x}_*$, having observed $\boldsymbol{Z}_{1:N}$ spatial observations, as:

$$\begin{aligned} \overline{f}(\mathbf{x}_*) &= \mathbb{E}\left[h(\mathbf{x}_*)\,|\mathbf{x}_*,\boldsymbol{Z}_{1:N}\right] \\ &= k(\mathbf{x}_*,\mathbf{x}_{1:N})\left(\mathcal{K}(\mathbf{x}_{1:N},\mathbf{x}_{1:N})+\sigma_{\mathrm{w}}^2\mathbf{I}\right)^{-1}\boldsymbol{Z}_{1:N}. \end{aligned} \tag{5}$$

The prediction error variance (uncertainty) function can be expressed as:

$$\begin{aligned} &\sigma^2(\mathbf{x}_*) = \mathbb{E}\left[h(\mathbf{x}_*)^2\,|\mathbf{x}_*,\boldsymbol{Z}_{1:N}\right]-\mathbb{E}\left[h(\mathbf{x}_*)\,|\mathbf{x}_*,\boldsymbol{Z}_{1:N}\right]^2 \\ &= k(\mathbf{x}_*,\mathbf{x}_*)-k(\mathbf{x}_*,\mathbf{x}_{1:N})\left(\mathcal{K}(\mathbf{x}_{1:N},\mathbf{x}_{1:N})+\sigma_{\mathrm{w}}^2\mathbf{I}\right)^{-1}k(\mathbf{x}_{1:N},\mathbf{x}_*). \end{aligned} \tag{6}$$

Interpreting these generic GP model results in the case of the HRIR model we propose in equation (2) involves considering the estimation of the sufficient statistics of the distribution of the prediction $h(\boldsymbol{x}_*)$ for an input $\mathbf{x}_*$ when inputs $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n]^{\mathrm{T}}$ and the random function at the observation locations $\mathbf{h}=[h(\mathbf{x}_1),h(\mathbf{x}_2),...,h(\mathbf{x}_n)]^{\mathrm{T}}$ (in short, $\mathbf{x}=(t,\theta)$) are given. Obtaining the results in equations (5) and (6) is achieved by considering the joint distribution of $\mathbf{h}$ and $h(\mathbf{x}_*)$. This is also a Gaussian distribution according to the GP definition. When the mean values of $\mathbf{h}$ and $h(\mathbf{x}_*)$ are 0, the joint distribution is formulated as follows.

$$\begin{bmatrix} \mathbf{h} \\ h(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathcal{K}(\mathbf{x}_{1:n},\mathbf{x}_{1:n})+\sigma^2 I & k(\mathbf{x}_*,\mathbf{x}_{1:n}) \\ k(\mathbf{x}_{1:n},\mathbf{x}_*) & k(\mathbf{x}_*,\mathbf{x}_*)) \end{bmatrix}\right). \tag{7}$$

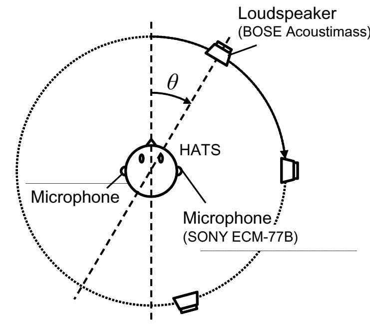

Fig. 1. Experimental setup for HRIR measurement.

3. EXPERIMENTS

3.1. Experimental Condition

The HRIRs were measured with two kinds of HATSs—KEMAR and B&K4128 (hereafter B&K)—for every 1° azimuth on a horizontal plane in different soundproof chambers whose each reverberation times were less than 150 ms (as shown in Fig. 1). The distance between the sound source and the midpoint between the ears was 1.2 m for KEMAR and 1.5 m for B&K. A swept sine signal [24] was transduced by a loudspeaker (BOSE Acoustimass). Microphones (SONY ECM-77B) were arranged at the entrances of the ear canals to block them. The sampling frequency was 48 kHz. Both HATSs were positioned on a turntable that can be moved at intervals of 1°, with accuracy of 0.3°. The power spectra of the measured HRTFs for the left ears of KEMAR and B&K are shown in Fig. 2. The azimuthal angle of 0° corresponds to the speaker position of the frontal direction of the HATS, and it increases clockwise. The frequency characteristics of the HATSs are different. We used HRIR data with KEMAR for training and conducted open and closed tests using both HATSs. Where an open test is one in which training is done on one HATS apparatus and testing on the other, whilst a closed test involved within HATS training and testing of the GP predictive performance.

From two perspectives—(1) the measurement of the azimuth interval needed to perform high-accuracy interpolation and (2) the robustness for different kinds of HATSs—we compared our method with two conventional methods: linear interpolation (denoted "linear") and interpolation by spatial linear prediction ("spatial linear prediction") [25]. For the linear method, the interpolated HRIR $\hat{h}(t,\theta)$ for angle θ obtained from those measured at θ_a and θ_b is obtained as

$$\hat{h}(t,\theta)=\alpha h(t,\theta_a)+(1-\alpha)h(t,\theta_b) \qquad a\neq b, \tag{8}$$

where

$$\alpha=\frac{\theta_b-\theta}{\theta_b-\theta_a}. \tag{9}$$

We used SDRs in equation (4) averaged over all locations for interpolation to evaluate the performance. A larger SDR score indicates a higher-accuracy interpolation. The initial delays of HRIRs for different azimuths are different. For the linear method, the initial

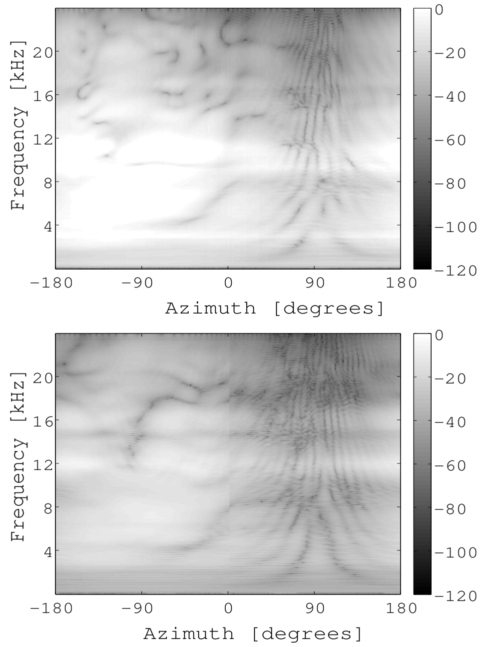

Fig. 2. Power spectra of measured HRTFs for KEMAR (top) and B&K (bottom). The gray scale indicates the relative level in dB.

delay for each HRIR was adjusted in advance. For our method and the spatial linear method, the time gap between adjacent HRIR data was adjusted by using the point having the maximum amplitude as a reference point.

3.2. Assessing Model Generalization for Spatial-Temporal GP versus Linear Interpolation

We investigated the measurement of the azimuth interval needed for high-accuracy interpolation by comparing our method and the linear method. Fig. 3 shows the averaged SDRs when using training data measured with azimuth intervals 10°, 15°, 30°, 45°, 60°, and 90° and interpolating HRIRs every azimuth 5° selectively using a few HRIRs with the closest angles according to the azimuth intervals as listed in Table 1. Here, we used KEMAR for both training and testing.

For all azimuth intervals, our method outperformed the linear method. As the interval for training data increased, the difference in the SDR became larger. We think that the assumption of spatial linearity for HRIRs in the linear method tends to fail when the interval is large. Note that the SDR of our method with azimuth 90° is roughly the same as that of the linear method with azimuth 45°. Our method is robust against variable measurement of the azimuth interval for training.

3.3. Predictive Performance of Spatial-Temporal GP versus Spatial Linear Prediction

We compared the performance of our method and the spatial linear prediction method [25] when using the same kind ("closed test") and different kinds ("open test") of HATSs for training and testing. Fig. 4 shows the averaged SDRs when using training data with azimuths 5° and 10° and interpolating HRIRs every azimuth 1°. HRIRs with KEMAR were used for training and HRIRs with B&K were used for

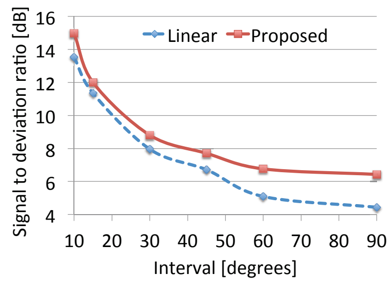

Fig. 3. Comparison of SDRs between our GP regression method and linear method.

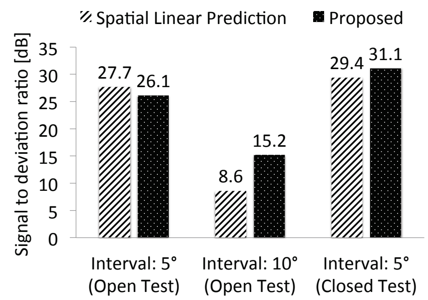

Fig. 4. SDRs in closed and open tests when using the same and different kinds of HATSs.

testing in the open test. For our method, four HRIRs with the closest azimuths were used in interpolation. For the spatial linear prediction method, the number of taps was set to 8 for azimuth interval 5° and 6 for azimuth interval 10° on the basis of preliminary experiments.

When using training data with azimuth interval 10°, our method clearly outperformed the spatial linear prediction method, although when using training data with azimuth interval 5°, it has a slightly lower SDR than the spatial linear prediction method. This result indicates that our method has the expressive power to handle the individual characteristics of multiple HATSs. When training data with azimuth interval 5° was used, we think that the hyper-parameters in our method were excessively over fitted to the data reflecting the characteristics of KEMAR. Evidence for this appeared in the closed test, in which our method outperformed the spatial linear prediction method. Our future work includes determining how to estimate hyper-parameters that are robust for any HATS in any conditions.

4. CONCLUSION

We studied HRIR modeling based on spatial-temporal GPs and HRTF interpolation. By comparing our method with the conventional linear interpolation and interpolation by spatial linear prediction, we found that it has superior expressive power for achieving high-accuracy HRTF interpolation.

Our future work includes investigating a robust method using hyper-parameters for various head and ear shapes. Moreover, we will investigate HRTF interpolation for the angle of elevation and develop a method of estimating HRTFs more flexibly from sound sources in a three-dimensional space. We will also confirm the effectiveness of our method in a subjective evaluation paradigm.

Acknowledgements: This work has been partially supported by the JST/CREST program.

5. REFERENCES

[1] E.M. Wenzel and S.H. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'93)*. IEEE, 1993, pp. 102–105.

[2] T. Nishino, S. Mase, S. Kajita, K. Takeda, and F. Itakura, "Interpolating HRTF for auditory virtual reality," *The Journal of the AcousticalSociety of America*, vol. 100, no. 4, pp. 2602–2602, 1996.

[3] F.L. Wightman and D.J. Kistler, "Headphone simulation of free-field listening. ii: Psychophysical validation," *The Journal of the Acoustical Society of America*, vol. 85, pp. 868–878, 1989.

[4] A. Kulkarni and H.S. Colburn, "Infinite-impulse-response models of the head-related transfer function," *The Journal of the Acoustical Society of America*, vol. 115, pp. 1714–1728, 2004.

[5] D.J. Kistler and F.L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *The Journal of the Acoustical Society of America*, vol. 91, pp. 1637–1647, 1992.

[6] W.L. Martens, "Principal components analysis and resynthesis of spectral cues to perceived direction," in *Proceedings of the 1987 International Computer Music Conference*, 1987, pp. 274–281.

[7] S. Carlile, C. Jin, and V. Van Raad, "Continuous virtual auditory space using HRTF interpolation: Acoustic and psychophysical errors," in *International Symposium on Multimedia Information Processing, Sydney*, 2000, pp. 220–223.

[8] J. Chen, B.D. Van Veen, and K.E. Hecox, "A spatial feature extraction and regularization model for the head-related transfer function," *The Journal of the Acoustical Society of America*, vol. 97, pp. 439–452, 1995.

[9] W. Zhang, R.A. Kennedy, and T.D. Abhayapala, "Efficient continuous HRTF model using data independent basis functions: Experimentally guided approach," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 17, no. 4, pp. 819–829, 2009.

[10] M.J. Evans, J.A.S. Angus, and A.I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 104, pp. 2400–2411, 1998.

[11] M. Noisternig, T. Musil, A. Sontacchi, and R. Holdrich, "3d binaural sound reproduction using a virtual ambisonic approach," in *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003 IEEE International Symposium on*. IEEE, 2003, pp. 174–178.

[12] C.K.I. Williams, "Prediction with gaussian processes: From linear regression to linear prediction and beyond," *NATO ASI SERIES D BEHAVIOURAL AND SOCIAL SCIENCES*, vol. 89, pp. 599–621, 1998.

[13] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2005.

[14] Manfred Lau, Ziv Bar-Joseph, and James Kuffner, "Modeling spatial and temporal variation in motion data," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 171:1–171:10, Dec. 2009.

[15] C. Stachniss, C. Plagemann, and A.J. Lilienthal, "Learning gas distribution models using sparse gaussian process mixtures," *Autonomous Robots*, vol. 26, no. 2, pp. 187–202, 2009.

[16] A. Singh, F. Ramos, H.D. Whyte, and W.J. Kaiser, "Modeling and decision making in spatio-temporal processes for environmental surveillance," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 5490–5497.

[17] A. Marquand, M. Howard, M. Brammer, C. Chu, S. Coen, and J. Mourão-Miranda, "Quantitative prediction of subjective pain intensity from whole-brain fmri data using gaussian processes," *NeuroImage*, vol. 49, no. 3, pp. 2178–2189, 2010.

[18] P. Nicolas, A. Leduc, S. Robin, S. Rasmussen, H. Jarmer, and P. Bessières, "Transcriptional landscape estimation from tiling array data using a model of signal shift and drift," *Bioinformatics*, vol. 25, no. 18, pp. 2341–2347, 2009.

[19] C.E. Rasmussen, B.J. De la Cruz, Z. Ghahramani, and D.L. Wild, "Modeling and visualizing uncertainty in gene expression clusters using dirichlet process mixtures," *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, vol. 6, no. 4, pp. 615–628, 2009.

[20] A.P. Bartók, M.C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons," *Physical review letters*, vol. 104, no. 13, 136403, 2010.

[21] J. Yuan, C.L. Liu, X. Liu, K. Wang, and T. Yu, "Incorporating prior model into gaussian processes regression for wedm process modeling," *Expert Systems With Applications*, vol. 36, no. 4, pp. 8084–8092, 2009.

[22] T. Hida, "Canonical representations of gaussian processes and their applications," *Kyoto Journal of Mathematics*, vol. 33, no. 1, pp. 109–155, 1960.

[23] J.S. Clark and A. Gelfand, *Hierarchical modeling for the environmental sciences: statistical methods and applications*, Oxford University Press, USA, 2006.

[24] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *The Journal of the Acoustical Society of America*, vol. 69, pp. 1484–1488, 1981.

[25] R. Nishimura, H. Kato, and N. Inoue, "Interpolation of head-related transfer functions by spatial linear prediction," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 1901–1904.