A Fast Forest Reverberator Using Single Scattering Cylinders

1st Shoken Kaneko*
Department of Computer Science
University of Maryland, College Park
College Park, MD, USA
kaneko60@umd.edu

2nd Hannes Gamper
Audio and Acoustics Research Group
Microsoft Research
Redmond, WA, USA
hannes.gamper@microsoft.com

***Abstract*—Simulating forest acoustics has important applications for rendering forest sound scenes in mixed and virtual reality, developing wildlife monitoring systems that use microphone arrays distributed in a forest, or as an artistic effect. Previously proposed methods for forest impulse response (IR) synthesis are limited to small or sparse forests because of their cubic asymptotic complexity with respect to the number of trees. Here we propose a simple and efficient parametric forest IR generation algorithm that relies on a multitude of single scattering cylinders to approximate scattering caused by tree trunks. The proposed method was compared to measured forest IRs in terms of the IR echo density, energy decay, reverberation time (T60), and clarity (C50). Experimental results indicate that the proposed algorithm generates forest reverb with acoustic characteristics similar to real forest IRs at a low computational cost.**

***Index Terms*—Artificial reverberation, forest acoustics, outdoor impulse response, single scattering**

I. Introduction

Modeling forest acoustic impulse responses enables applications including designing or training forest wildlife monitoring systems based on microphone arrays and audio signal processing, synthesizing artificial forest sound scenes to generate or augment training data for machine learning systems for sound event detection, creating reverberation sound effects for artistic expression and multimedia, or enabling plausible auditory rendering of virtual forests in mixed reality (MR) and immersive gaming. Studies related to forest acoustics date back to the 1940s [1], [2]. Prior work on forest acoustics or systems of multiple scatterers focused on computing the attenuation of a sound wave propagating through a forest, which can be represented by the bulk effective wavenumber [3]–[5]. Under the assumption that trees can be approximated by ideal cylinders, previous works considered multiple scattering by trees [3]–[5]. However, the bulk effective wavenumber is a result of statistical averaging and does not take into account geometric details including tree positions. Therefore, a bulk effective wavenumber alone is not sufficient to reproduce the echo patterns due to individual scattering events that depend on source, receiver, and tree positions.

Another line of research related to forest acoustics is recent work in the context of audio effects and outdoor reverberation synthesis. Spratt et al. devised a forest reverb synthesis algorithm based on a digital waveguide approach, referred to as the treeverb algorithm [6]. Stevens et al. introduced a more general scheme to model outdoor reverberation which is referred to as the waveguide web algorithm [7]. It too is based on waveguide techniques but not limited to forests. Although both of these algorithms consider multiple scattering, their asymptotic complexity is cubic with respect to the number of trees. This computational cost severely limits the size and density of the forest that can be modeled. Only results for small or sparse forests consisting of 25 to 30 trees have been reported using these algorithms [6], [7]. As shown later in this paper, experimental results indicate that 30 or fewer trees, even if multiple scattering is considered, may not be sufficient to generate acoustically plausible forest IRs.

Here we propose a simple, efficient, and scalable algorithm for parametric forest impulse response synthesis. Source code and sample audio of the proposed method are available online.[1] The proposed algorithm models the effect of tree scattering analytically via scattering from rigid cylinders, and approximates multiple scattering in a forest by single scattering from a multitude of virtual *single scattering cylinders* (SSCs). Since the SSCs do not interact with each other, this approximation results in an algorithm which can be easily parallelized and which has linear asymptotic complexity with respect to the number of SSCs, allowing the algorithm to generate forest IRs using hundreds of thousands of SSCs within a few seconds. Rather than aiming at a physically accurate simulation of a particular forest scene with given tree positions, the proposed algorithm produces acoustically plausible forest IRs. For many practical applications, including sound effects, audio rendering in MR, or data augmentation for training machine learning systems, physical accuracy may be of lower importance than subjective plausibility, computational cost, and parametric tunability of the reverberator. Experimental results indicate that the proposed algorithm produces synthetic forest IRs that exhibit acoustic and sonic qualities similar to real measured forest IRs [8], given a sufficiently

*This work was done while Shoken Kaneko was an intern at Microsoft Research Labs in Redmond, WA, USA.

[1] https://github.com/microsoft/Forest_IR_synthesis

large number of SSCs. This is achieved at a significantly lower computational cost compared to previous cubic order methods that try to model multiple scattering accurately. Using recorded real forest IRs as a baseline, the synthetic IRs produced by the proposed algorithm compared favorably to IRs synthesized by previously proposed algorithms in terms of energy decay curves, reverberation time (T60), clarity (C50) [9], and a recently proposed echo density measure [10].

II. Proposed Approach

A. Algorithm overview

Assuming linearity and time-invariance, the acoustic path from a sound source to a receiver position is described by an impulse response (IR). A forest IR h_{forest} is modeled here as:

$$h_{\mathrm{forest}} = h_{\mathrm{direct}} + h_{\mathrm{ground}} + h_{\mathrm{tree}}, \tag{1}$$

where h_{direct}, h_{ground}, and h_{tree} are IRs describing the direct path, ground reflection, and tree scattering components, respectively. The direct-path component is given as:

$$h_{\mathrm{direct}} = h_{\mathrm{air}}(|\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{s}}|) * h^{(1)}_{\mathrm{dist}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{s}}), \tag{2}$$

where $h_{\mathrm{air}}(x) = \mathcal{F}^{-1}(e^{-bx})$ denotes the air dissipation filter given the frequency-dependent sound attenuation factor b and the propagation distance x. $\mathcal{F}^{-1}$ denotes the inverse discrete Fourier transform (DFT). Note that the IRs are functions of time t while the variable t is omitted for clarity. The source and receiver positions are denoted by $\mathbf{r}_{\mathrm{s}}$ and $\mathbf{r}_{\mathrm{r}}$, respectively, and $*$ denotes the convolution operator. $h^{(\alpha)}_{\mathrm{dist}}$ describes a generalized distance attenuation and delay filter defined as:

$$h^{(\alpha)}_{\mathrm{dist}}(\mathbf{r}_1, \mathbf{r}_2) = \delta(t - |\mathbf{r}_1 - \mathbf{r}_2|/c)|\mathbf{r}_1 - \mathbf{r}_2|^{-\alpha}, \tag{3}$$

where δ is Dirac's delta function, c is the speed of sound, and α a distance decay parameter. This parameter allows the spatial spreading of the scattered wave to be interpolated between ideal cylindrical spreading ($\alpha = 0.5$) and spherical spreading ($\alpha = 1$). In the proposed model, the sound wave emitted from the source is modeled as a spherical wave with $\alpha = 1$, while the scattered waves traveling from the tree trunks to the receiver are modeled with a parameter $0.5 \leq \alpha_{\mathrm{c}} \leq 1$. This is to account for the fact that real trees are not ideal cylinders with infinite length. The ground reflection component is modeled as:

$$h_{\mathrm{ground}} = g(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{s}}) \left(h_{\mathrm{air}}(|\mathbf{r}_{\mathrm{r}} - \check{\mathbf{r}}_{\mathrm{s}}|) * h^{(1)}_{\mathrm{dist}}(\mathbf{r}_{\mathrm{r}}, \check{\mathbf{r}}_{\mathrm{s}}) \right), \tag{4}$$

where $\check{\mathbf{r}}_{\mathrm{s}}$ denotes the source position reflected by the ground plane and g the ground reflection coefficient which is implemented as a frequency-independent reflection constant. If the frequency dependence of a forest ground reflection coefficient is known, that can be used here instead of a constant coefficient. The frequency spectrum of the reflection filter could be used to model different boundary conditions, e.g., the presence or absence of snow on the ground, which we leave for future studies. The tree scattering component is defined as:

$$h_{\mathrm{tree}} = \sum_{m=1}^{N_{\mathrm{c}}} h_{\mathrm{air}}(r_{\mathrm{r,c}_m,\mathrm{s}}) * h_{\mathrm{scat}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{c}_m}, \mathbf{r}_{\mathrm{s}}) * h^{(1)}_{\mathrm{dist}}(\mathbf{r}_{\mathrm{c}_m}, \mathbf{r}_{\mathrm{s}}), \tag{5}$$

where m denotes the SSC index, N_{c} the total number of SSCs, $\mathbf{r}_{\mathrm{c}_m}$ the position of the mth SSC, $r_{\mathrm{r,c}_m,\mathrm{s}} = |\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{c}_m}| + |\mathbf{r}_{\mathrm{c}_m} - \mathbf{r}_{\mathrm{s}}|$ the total path length of the wave scattered by the mth SSC, and h_{scat} the scattering-angle dependent tree-scattering filter. The path lengths $|\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{c}_m}|$ and $|\mathbf{r}_{\mathrm{c}_m} - \mathbf{r}_{\mathrm{s}}|$ are evaluated in two dimensions, while the lengths of the direct path and the ground reflection path is evaluated in three-dimensional coordinates. The effect of tree scattering is modeled as the scattering of a plane wave by a rigid cylinder [11]:

$$h_{\mathrm{scat}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{c}_m}, \mathbf{r}_{\mathrm{s}}) = \beta h^{(\alpha_{\mathrm{c}})}_{\mathrm{dist}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{c}_m}) * h_{\mathrm{angle}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{c}_m}, \mathbf{r}_{\mathrm{s}}), \tag{6}$$

with β denoting a parameter controlling the scattering amplitude. The scattering-angle dependent component h_{angle} is defined as:

$$h_{\mathrm{angle}}(\mathbf{r}_{\mathrm{r}}, \mathbf{r}_{\mathrm{c}_m}, \mathbf{r}_{\mathrm{s}}) = \mathcal{F}^{-1}(H_{\mathrm{angle}}), \tag{7}$$

where H_{angle} denotes a frequency-domain filter modeling the frequency and scattering-angle dependent attenuation to an incoming wave scattered by a rigid cylinder:

$$H_{\mathrm{angle}} = e^{i\frac{\pi}{4}} \sqrt{\frac{2}{\pi k}} \sum_{n=0}^{N_{\max}} (2 - \delta_{n,0}) \sin(\gamma_n) \cos(n\varphi) e^{i\gamma_n}, \tag{8}$$

$$\gamma_n = \arctan \frac{J_{n-1}(ka) - J_{n+1}(ka)}{N_{n-1}(ka) - N_{n+1}(ka)}, \tag{9}$$

and

$$\gamma_0 = \arctan \frac{J_1(ka)}{N_1(ka)}, \tag{10}$$

with $\varphi = \arccos(\frac{(\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{c}_m}) \cdot (\mathbf{r}_{\mathrm{c}_m} - \mathbf{r}_{\mathrm{s}})}{|\mathbf{r}_{\mathrm{r}} - \mathbf{r}_{\mathrm{c}_m}||\mathbf{r}_{\mathrm{c}_m} - \mathbf{r}_{\mathrm{s}}|})$ denoting the scattering angle, k the wave number, a the tree radius, and $N_{\max}$ the truncation number of the series expansion. J_n and N_n are the Bessel functions of the first and second kind, respectively. To remove the tree–receiver distance dependency on the spectrum, the far-field approximation is used here [11].

As can be observed from (1), (2), (4) and (5), the asymptotic complexity of the proposed forest IR generation algorithm is linear with respect to the number of SSCs. In addition, since all scattering paths can be evaluated in parallel, the proposed algorithm allows efficient synthesis of forest IRs with dense scattering patterns. In our experiments, we compute IRs for a forest of 1 km $\times$ 1 km size with up to 500 000 SSCs.

B. Implementation and optimization details

The tree scattering filter h_{scat} is implemented as a 128-tap finite impulse response (FIR) filter. The scattering-angle dependent tree scattering amplitude spectra are shown in Fig. 1. The scattered wave has an amplitude drop of approximately 30 dB relative to the incident wave for scattering angles larger than about 20 degrees, implying that the majority of the multiple scattering paths will experience a quick decay of amplitude while traveling from the source to the receiver.

The air dissipation filter h_{air} is modeled based on the ISO9613-1 standard [12], and was implemented as a 128-tap FIR filter. We found the incorporation of air dissipation

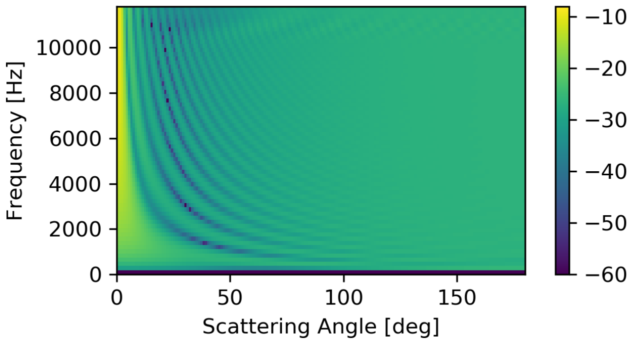

Fig. 1. The amplitude spectrum of the angle-dependent tree scattering filter H_{angle} [11].

to play an important role for obtaining a realistic timbre in the late reverberation. The ground reflection coefficient g was set to 0.8. For faster computation, all tree scattering and air dissipation filter coefficients are pre-computed for predefined discrete angles and distances. The processing cost is dominated by the computation of the tree scattering component h_{tree}. For better performance, the convolution is performed in the frequency-domain. Note that the frequency-domain filter banks for air dissipation and cylindrical scattering can be pre-computed. A Python implementation of the proposed algorithm was able to compute IRs of a forest with 1 km $\times$ 1 km size and 100 000 SSCs at a sampling rate of 24 kHz in about 1.2 seconds on a laptop PC with an Intel Core i7-8665U CPU. In contrast, a Matlab implementation [13] of the algorithm by Spratt et al. [6] took about 31 minutes on the same machine to compute a forest IR considering multiple scattering for a forest with 50 trees. Note that this significant difference of computation cost, which is a consequence of the cubic complexity of the previously proposed method, prevents direct comparison of the methods using the same number of scatterers since the proposed method uses large numbers of scatterers by design.

III. Experimental Evaluation

A. Echo density of synthetic and real forest IRs

The echo density measure [10] of forest IRs for two existing algorithms [6], [7], forest IRs synthesized by the synthesis algorithm proposed here, and real forest IRs measured in the Koli national park in Finland [8], was evaluated. All reference forest IRs were downloaded from Openair [14]–[17]. The IRs using the treeverb and waveguide web algorithms were computed for a forest of 25 trees and source-receiver distance D of 5.8 m. These IRs generated by previous algorithms, which are computed for a forest of a significantly smaller number of scatterers compared to the proposed method, should be considered only as references and not as a direct comparison since their cubic complexity only allows computation of small forests within a realistic amount of time. The real forest IRs were measured with $D = 9$ m. For the proposed algorithm, we set $\alpha_{\mathrm{c}} = 0.7$, $a = 0.25$ m, source and receiver height to 1.5 m, $D = 10$ m, and $\beta = 2$. The SSC positions were randomly sampled using a uniform distribution within a square-shaped forest of size 1 km $\times$ 1 km. A constant tree radius was used here to allow reusing a single scattering coefficient table for all SSCs for the sake of computational efficiency, although this too can be randomized with an additional computational cost linear with respect to the number of SSCs. It was found by informal listening that randomness of the SSC positions is important to synthesize perceptually natural IRs. Configurations with regularity, e.g. regular Cartesian grids or two-dimensional quasi-random sequences with approximate regularity, tend to result in IRs with noticeable unnatural sound.

The waveforms and frequency responses of each forest IR are shown in Fig. 2. The comb-filter effect visible in the spectrum of the synthetic IRs obtained with the proposed method is a result of the interference of the direct path and the ground reflection. This effect can be controlled by tuning the amplitude, or, if frequency-dependence is considered, the spectrum, of the ground reflection coefficient g. While this comb-filtering is noticeable in the spectra, it did not result in unnatural sound in informal listening.

Fig. 3 shows the echo density profiles for all evaluated IRs. It is notable that real forest IRs have nearly constant echo density profiles. We can see that this property is present in forest IRs synthesized with the proposed algorithm using a large number of SSCs ($N_{\mathrm{c}} = 500\mathrm{k}$), especially at low source–receiver distances D. However, this is not the case when using fewer SSCs ($N_{\mathrm{c}} \leq 200\mathrm{k}$) or for synthetic IRs obtained with previously proposed algorithms. While the reason for this near-constant echo density profile in real forest IRs is unclear, the results suggest that having a large number of SSCs may be necessary to produce comparable synthetic IRs.

Informal listening indicates that for $N_{\mathrm{c}} \leq 200\mathrm{k}$, the reduced echo density in the first 200 ms compared to real forest IRs is audible. It should be noted that $N_{\mathrm{c}} = 500\mathrm{k}$ corresponds to an SSC density about five to ten times higher than tree densities observed in typical forests [18]. We hypothesize that the high SSC density used here leads to denser echo profiles similar to real forest IRs by compensating for the fact that the SSCs are modeled as simple cylinders rather than complex geometric objects and that multiple scattering is not considered.

B. Energy decay curves of synthetic and real forest IRs

Fig. 4 shows the energy decay curves (EDCs) of the reference and synthetic forest IRs with parameters described in Sec. III-A. As can be seen, the synthetic forest IRs obtained with the proposed method exhibit energy decay characteristics similar to real forest IRs. On the other hand, the EDCs of the previous algorithms exhibit characteristics which are far from real forest IRs. This difference was clearly audible in informal listening. To study the source–receiver distance (D) dependence of the EDCs, we measured forest IRs in a real forest (Volunteer Park, Seattle, USA) at various distances D. As can be seen in Fig. 4 (bottom), the energy decays more slowly with increasing D, a property present in both the synthetic and real measured forest IRs.

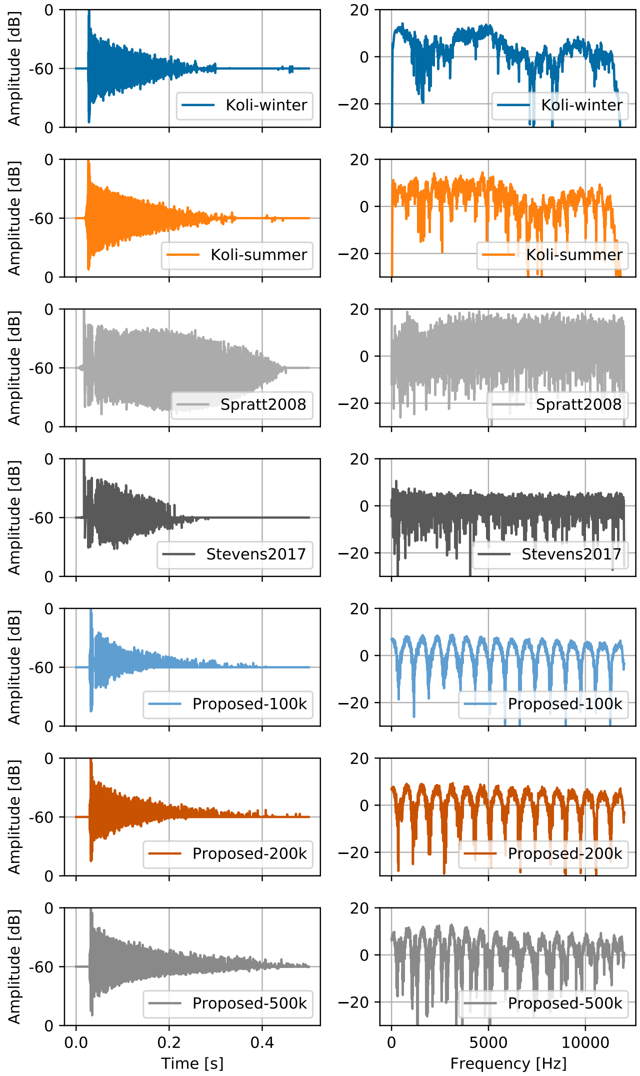

Fig. 2. The waveforms (left) and frequency responses (right) of various forest IRs. From top to bottom: Koli national park in winter and summer [8], [16], [17], treeverb algorithm [6], [15], waveguide web algorithm [7], [15], proposed method with $N_c = 100$k, 200k, 500k.

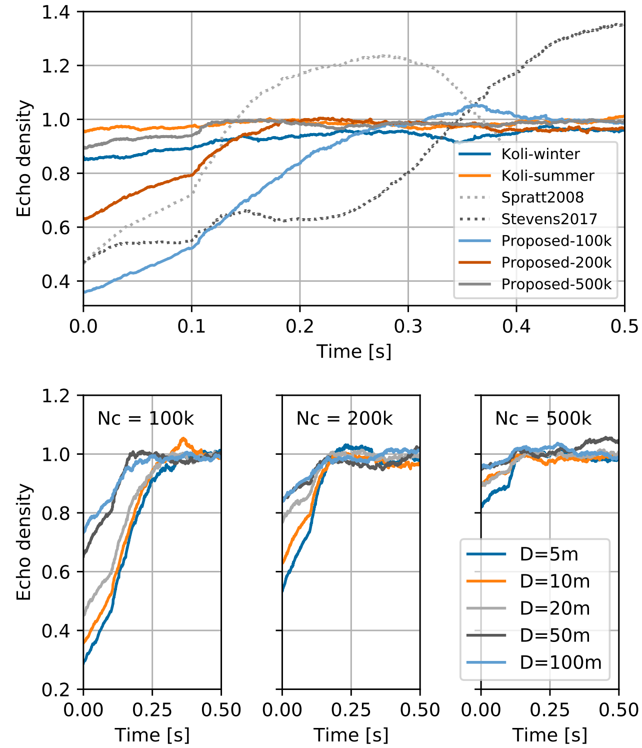

Fig. 3. The echo density profiles of forest IRs. Comparison with reference forest IRs (top), and comparison among forest IRs generated by the proposed method for a square-shaped forest of 1 km $\times$ 1 km, with various D and N_c (bottom). Each subplot in the bottom shows results for $N_c = 100$k (bottom left), 200k (bottom center), and 500k (bottom right), respectively.

C. *Reverberation time and clarity*

Fig. 5 shows the T60 and C50 of the synthetic forest IRs and the reference forest IRs introduced in Sec. III-A and Sec. III-B. In the proposed method, T60 was able to be controlled over a wide range by varying N_c. Clear frequency dependence of T60 was only observed in real forest IRs and in synthesized forest IRs by the proposed algorithm. The synthetic forest IRs obtained with the proposed method exhibit a decrease of C50 as a function of D, which was also observed in measured forest IRs.

IV. Summary and Discussion

We proposed a simple and efficient parametric forest reverberator that approximates multiple scattering in real forests by scattering from a multitude of single scattering cylinders (SSCs). Experimental results show that the proposed algorithm synthesizes forest IRs with plausible acoustic and sonic properties at a low computational cost. The synthetic forest IRs were compared to real forest IRs, captured at two forests with varying seasons or source-receiver distances, in terms of the echo density, energy decay curves, reverberation time, and clarity. When modeling 0.5 million SSCs distributed over one square kilometer, the proposed algorithm produces IRs with an echo density profile similar to real forest IRs. The experimental results also indicate that the proposed algorithm produces plausible energy decay profiles and characteristics similar to real forest IRs, such as the frequency-dependence of T60 and the distance-dependence of C50. The cubic arithmetic complexity of previously proposed algorithms may be too high to model forests with large tree numbers, resulting in synthetic IRs with audibly different echo density profiles compared to real forest IRs.

Note that the exact relation between the model parameter N_c determining the number of simulated SSCs and the number of actual trees in a real forest is unknown. While this relation could potentially be determined by additional measurements and simulation, this was outside the scope of the present work.This parameter can be subjectively tuned by a sound designer to match the desired acoustic characteristics. A practical recommended default value is between $N_c = 200$ k

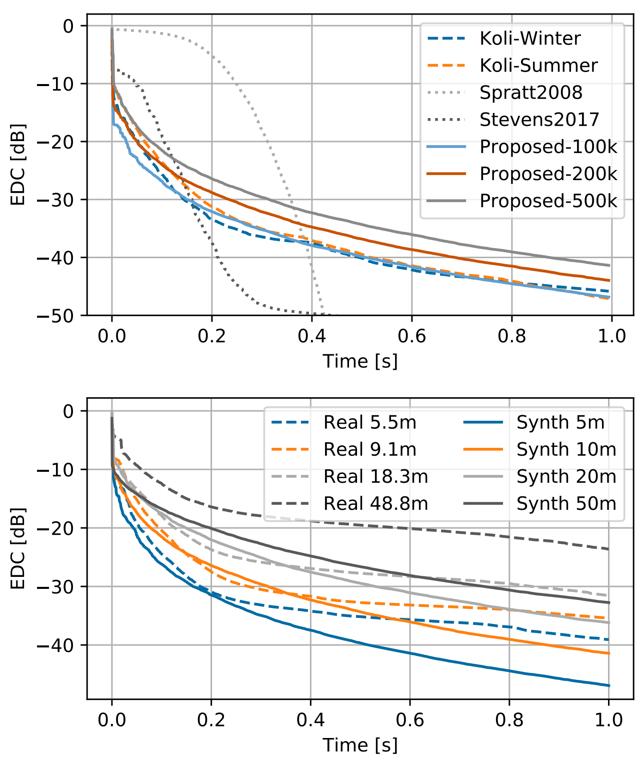

Fig. 4. Top: EDCs for various short distance forest IRs. The set of forest IRs is the same as in Sec. III-A. Bottom: EDCs for synthetic and real forest IRs with varying D. N_c was set to 500k for the proposed method. All EDCs are computed from the IRs after applying a band-pass filter with a passband of 500Hz-8kHz.

and $N_\mathrm{c} = 500$ k for a forest of 1 km^2, which results in a good balance between echo density and energy decay characteristics, as shown in the experiments. A limitation of both previous methods as well as the method proposed here is that scattering from fine structures, e.g., leaves and branches, is not considered. It is possible that the high density of SSCs used in the experiments here is necessary to account for this limitation of the models. Efficient incorporation of multiple scattering and the consideration of scattering from fine structures, which may reduce the number N_c that delivers realistic acoustic profiles, is a subject of future research.

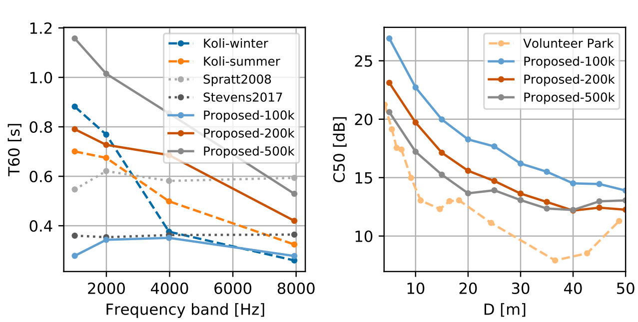

Fig. 5. T60 as a function of octave-band frequency (left) and C50 as a function of D (right). The set of shown forest IRs is the same as in Sec. III-A and Sec. III-B for the left and right sub-plot, respectively. IRs from the Koli forest and previous algorithms are not included in the C50 plot since the available dataset only contained IRs for a fixed distance.

Acknowledgment

This work is part of Microsoft's *AI for Earth* program. The authors thank Professor Justin Kitzes for helpful discussions on acoustic forest wildlife monitoring.

References

[1] C. F. Eyring, "Jungle acoustics," *The Journal of the Acoustical Society of America*, vol. 18, no. 2, pp. 257–270, 1946.

[2] L. L. Foldy, "The multiple scattering of waves. I. general theory of isotropic scattering by randomly distributed scatterers," *Physical review*, vol. 67, no. 3-4, p. 107, 1945.

[3] V. Twersky, "On scattering of waves by random distributions. I. free-space scatterer formalism," *Journal of Mathematical Physics*, vol. 3, no. 4, pp. 700–715, 1962.

[4] S. Bose and A. Mal, "Longitudinal shear waves in a fiber-reinforced composite," *International Journal of Solids and Structures*, vol. 9, no. 9, pp. 1075–1085, 1973.

[5] C. Linton and P. Martin, "Multiple scattering by random configurations of circular cylinders: Second-order corrections for the effective wavenumber," *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3413–3423, 2005.

[6] K. Spratt and J. S. Abel, "A digital reverberator modeled after the scattering of acoustic waves by trees in a forest," in *Audio Engineering Society Convention 125*. Audio Engineering Society, 2008.

[7] F. Stevens, D. T. Murphy, L. Savioja, and V. Välimäki, "Modeling sparsely reflecting outdoor acoustic scenes using the waveguide web," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 8, pp. 1566–1578, 2017.

[8] S. B. Shelley, D. T. Murphy, and A. J. Chadwick, "B-format acoustic impulse response measurement and analysis in the forest at Koli national park, Finland," in *Proceedings of the 16th International Conference on Digital Audio Effects (DAFx13)*. York, 2013, pp. 351–355.

[9] G. A. Soulodre and J. S. Bradley, "Subjective evaluation of new room acoustic measures," *The Journal of the Acoustical Society of America*, vol. 98, no. 1, pp. 294–301, 1995.

[10] H. P. Tukuljac, V. Pulkki, H. Gamper, K. Godin, I. J. Tashev, and N. Raghuvanshi, "A sparsity measure for echo density growth in general environments," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 1–5.

[11] P. M. Morse and K. U. Ingard, *Theoretical acoustics*. Princeton university press, 1986.

[12] ISO Central Secretary, "Acoustics — Attenuation of sound during propagation outdoors — Part 1: Calculation of the absorption of sound by the atmosphere," International Organization for Standardization, Standard ISO 9613-1:1993, June 1993.

[13] T. Wiens, "Forest reverb generator – matlab central file exchange," 2008. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/20764-forest-reverb-generator

[14] D. T. Murphy and S. Shelley, "Openair: An interactive auralization web resource and database," in *Audio Engineering Society Convention 129*. Audio Engineering Society, 2010.

[15] "Waveguide web example audio – Openair," https://openairlib.net/?page_id=782.

[16] "Koli national park – winter – Openair," https://openairlib.net/?page_id=584.

[17] "Koli national park – summer – Openair," https://openairlib.net/?page_id=577.

[18] T. W. Crowther, H. B. Glick, K. R. Covey *et al.*, "Mapping tree density at a global scale," *Nature*, vol. 525, pp. 201—-205, 2015.