

MONTE CARLO EXPLORATION FOR ACTIVE BINAURAL LOCALIZATION

Christopher Schymura, Juan Diego Rios Grajales and Dorothea Kolossa

Institute of Communication Acoustics, Ruhr-Universität Bochum, Germany

ABSTRACT

This study introduces a machine hearing system for robot audition, which enables a robotic agent to pro-actively minimize the uncertainty of sound source location estimates through motion. The proposed system is based on an active exploration approach, providing a means to model and predict effects of the agent's future motions on localization uncertainty in a probabilistic manner. Particle filtering is used to estimate the posterior probability density function of the source position from binaural measurements, enabling to jointly assess azimuth and distance of the source. The framework allows to infer and refine a policy to select appropriate actions via a Monte Carlo exploration approach. Experiments in simulated reverberant conditions are conducted, showing that active exploration and the incorporation of distance estimation significantly improve localization performance.

Index Terms— sound source localization, robot audition, active listening, particle filters, Monte Carlo exploration

1. INTRODUCTION

An essential part of auditory scene analysis (ASA) is the localization of sound sources in the environment [1]. Recently, this task has been addressed by various studies in the context of robot audition. A significant advantage of robotic agents over static acoustic sensors is the ability to move and actively explore the environment. This has triggered research on algorithms for active listening which incorporate feedback into the audition process. A prominent area of research in this context is active localization. Inspired by the abilities of human listeners to improve the assessment of auditory scenes through head and body motions [2], this has led to a variety of approaches for different applications.

An early study on how head movements are utilized by humans to resolve front-back ambiguities in sound localization was introduced in [3]. Based on these findings, several computational localization models incorporating head-movements have recently been introduced [4, 5, 6]. Additionally, methods based on whole-robot motion were proposed, using either microphone arrays [7] for auditory simultaneous

localization and mapping (SLAM) or binaural sensors [8, 9] for active localization of sound sources.

An important question in this context is, how robot motion can optimally support localization. For instance, the framework introduced in [9] describes an information-based feedback control scheme. It selects controls to maximize information gain of the estimated posterior probability density function (PDF) representing the assumed source location. Similar approaches have also been proposed in the broader context of active exploration, aiming at maximizing the robot's knowledge about the environment [10, 11]. The framework proposed in this study follows similar ideas, but differs from the previously described approaches in two important aspects:

Firstly, an extended binaural localization model is introduced, which incorporates azimuth and distance information into the localization process. Binaural models for sound distance estimation based on the direct-to-reverberant energy ratio (DRR) [12] or statistical signal parameters [13] have already been proposed. In this work, sound distance is modeled using the interaural coherence (IC) of reverberant binaural signals, which was originally described in [14]. It was reported that a decreasing DRR results in a decrease of correlation between both binaural channels, which can be approximately represented by the IC. The current study shows that incorporating distance information improves the localization abilities of the robotic agent in reverberant environments, compared to a conventional bearing-only observation model.

Secondly, a closed-loop feedback control scheme is proposed, aiming at minimizing the entropy of the belief state while approaching a specific goal position. The robot motion will be chosen from a set of pre-defined actions based on a Monte Carlo exploration (MCE) approach [11, 15]. This allows for the selection of movements by minimizing the predicted localization uncertainty. This approach extends previously proposed methods [8, 9] with the possibility for trade-offs between exploratory and goal-direction motions.

2. SYSTEM OVERVIEW

The proposed framework follows a modular approach, consisting of individual building blocks which constitute a non-linear, closed-loop feedback control system. Detailed descriptions of all system components will be given below.

This research has been supported by European Union FET grant TWO!EARS, ICT-618075, www.twoears.eu.

2.1. Robot platform and binaural head

A binaural Knowles Electronics Manikin for Acoustic Research (KEMAR) dummy head, mounted on a two-wheel differential drive robotic platform, is considered throughout this study. The acoustic properties of the dummy head are captured by head-related impulse responses (HRIRs) from the CIPIC HRIR database [16].

2.2. Binaural front-end

An auditory front-end as proposed in [17] is used to extract binaural cues from the dummy head's ear signals, sampled with a rate of $f_s = 16$ kHz. Each channel of the ear signals is decomposed into $L = 32$ auditory channels using a phase-compensated gammatone filterbank. Binaural cues, namely interaural time differences (ITDs), interaural level differences (ILDs) and IC, are computed and integrated across all frequency channels at the filterbank output, using non-overlapping time frames with a length of 25 ms. All binaural cues are composed into a 3-dimensional observation vector $\hat{\mathbf{y}}_k$ at each time step k .

2.3. System dynamics

The localization model used in this study assumes a generic nonlinear state space representation

$$\begin{aligned} \mathbf{x}_k &= \begin{bmatrix} \mathbf{x}_{S,k} \\ \mathbf{x}_{R,k} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{S,k-1} \\ f(\mathbf{x}_{R,k-1}, \mathbf{u}_k) \end{bmatrix} + \mathbf{v}_k \\ \mathbf{y}_k &= g(\mathbf{x}_k) + \mathbf{n}_k, \end{aligned} \quad (1)$$

where \mathbf{x}_k and \mathbf{y}_k denote the state and observation vectors and $\mathbf{u}_k = [u_{L,k} \ u_{R,k}]$, $u_{\{L,R\},k} \in [-1, 1]$ represents the control input at the left and right wheel actuators. The system dynamics in Eq. (1) is composed of two augmented state vectors $\mathbf{x}_{S,k} = [m_{x,k} \ m_{y,k}]^T$ and $\mathbf{x}_{R,k} = [p_{x,k} \ p_{y,k} \ \theta_k]^T$, representing the position of the sound source in Cartesian coordinates and the robots pose including its heading direction θ_k , respectively. The sound source is assumed to be static up to state noise, whereas the robot dynamics is governed by a nonlinear motion model $f(\mathbf{x}_{R,k-1}, \mathbf{u}_k)$ [11]. Observations are predicted using a nonlinear mapping function $g(\mathbf{x}_k)$, which will be discussed in detail in Sec. 2.4. Both state and measurement noise characteristics are modeled as additive, zero-mean Gaussian random variables \mathbf{v}_k and \mathbf{n}_k with corresponding covariance matrices \mathbf{Q} and \mathbf{R} .

2.4. Measurement model

The measurement model in Eq. (2) introduces a nonlinear mapping function $g(\mathbf{x}_k)$ from states \mathbf{x}_k to observations \mathbf{y}_k . As described previously, ITDs, ILDs and IC are used as primary binaural cues for the proposed localization model. ITDs and ILDs are cues that correspond to the relative angle ϕ_k between the sound source and the heading direction of the robot.

In addition to that, the IC is used here to model the sound-to-receiver distance d_k . Hence, the systems state is first mapped from Cartesian to polar coordinates. The mapped state is subsequently used to predict binaural observations by a regression model

$$g(\mathbf{x}_k) = \mathbf{W}^T \Phi(\phi_k(\mathbf{x}_k), d_k(\mathbf{x}_k)), \quad (3)$$

where $\Phi(\phi_k(\mathbf{x}_k), d_k(\mathbf{x}_k))$ represents the regressors and \mathbf{W} is a matrix of regression coefficients. The latter are computed via multivariate linear regression [18, Chap. 3] using rendered binaural room impulse responses (BRIRs) and white noise as stimulus signal. A finite Fourier-series representation [19] is used to model angle-dependent regressors in $\Phi(\phi_k(\mathbf{x}_k), d_k(\mathbf{x}_k))$, whereas the distance-related regressors are modeled via polynomials. Both representations are computed up to an order of 4.

The residuals obtained after training are used to estimate the measurement noise covariance matrix \mathbf{R} . This model extends the approach from [6], which was restricted to azimuth prediction based on a spherical head model. The use of a regression function according to Eq. (3) yields a more flexible framework, which can be trained on both measured or simulated BRIRs.

2.5. State estimation

As this study focuses on source localization and does not consider the full SLAM problem, the robot position is assumed to be known and deterministic, depending only on the applied control input. Hence, state estimation reduces to recursively compute the posterior PDF or belief of the sound source position $p(\mathbf{x}_{S,k} | \mathbf{y}_{1:k}, \mathbf{u}_{1:k})$. This distribution might have multimodal characteristics, as front-back confusions can occur due to the ambiguous nature of binaural cues [2]. This renders Bayesian filtering methods based on unimodal Gaussian assumptions inappropriate for this task.

To overcome these limitations, a Gaussian mixture sigma-point particle filter (GMSPPF) [20] is used, which represents the posterior PDF as a Gaussian mixture model (GMM). However, an analytic evaluation of the entropy is not possible for GMMs. To obtain a measure of uncertainty, an approximation of the belief states entropy is necessary. As described in [20], the conditional mean state estimate $\hat{\mathbf{x}}_{S,k} = E\{\mathbf{x}_{S,k} | \mathbf{y}_{1:k}\}$ and the error covariance matrix $\hat{\mathbf{P}}_k = E\{(\mathbf{x}_{S,k} - \hat{\mathbf{x}}_{S,k})(\mathbf{x}_{S,k} - \hat{\mathbf{x}}_{S,k})^T\}$ can be utilized to approximate the resulting GMM representation of the belief state by a unimodal Gaussian distribution. Hence, the entropy of the corresponding PDF is defined as

$$H(\mathbf{x}_{S,k}) = \frac{1}{2} \log \left((2\pi e)^D \cdot |\hat{\mathbf{P}}_k| \right), \quad (4)$$

where $D = 2$ is the dimensionality of the state vector $\hat{\mathbf{x}}_{S,k}$.

3. MONTE CARLO EXPLORATION

MCE has been widely used in the context of robotics as a means to actively gain information about uncertain entities in the environment [11]. A prominent application for exploration techniques is SLAM, where the robot tries to actively explore its environment to reduce uncertainty in the map-building process [21]. However, unrestricted exploration is usually not desired in many applications, for instance, if the robot must reach a specific goal position.

Hence, this study focuses on using MCE to construct a policy $\pi(\mathbf{x}_k)$, which allows the robot to select appropriate actions that maximize a specific reward function which trades off exploration and goal-directed movements. Thereby, exploration serves two purposes: using rotational movements to resolve front-back ambiguities and translatory movements to support distance estimation via triangulation.

3.1. Algorithm description

A generic MCE algorithm as described in [11, Chap. 17] is adopted in this work. It aims at finding the action \mathbf{u}_{k+1} that maximizes the expected reward at the subsequent time step. To assure computational tractability, the continuous-valued controls described in Sec. 2.3 are discretized, yielding a set of N_u actions $\mathcal{U} = \{\mathbf{u}_{k+1}^{(1)}, \dots, \mathbf{u}_{k+1}^{(N_u)}\}$.

A control policy is obtained by running Monte Carlo simulations that predict the immediate reward of all actions in the set \mathcal{U} , using the system dynamics and measurement model as a black-box simulator. The procedure is initialized by drawing a set of N samples $\tilde{\mathbf{x}}_{S,k}^{(i)} \sim p(\mathbf{x}_{S,k} | \mathbf{y}_{1:k}, \mathbf{u}_{1:k})$, $i = 1, \dots, N$ from the belief distribution at the current time step. As the robot pose is assumed to be deterministic, the full state samples are represented as $\tilde{\mathbf{x}}_k^{(i)} = [\tilde{\mathbf{x}}_{S,k}^{(i)} \quad \mathbf{x}_{R,k}]^T$. Subsequently, a particle filter update step is conducted for all available actions by sampling observations using the measurement model $\tilde{\mathbf{y}}_{k+1} \sim p(\mathbf{y}_{k+1} | \tilde{\mathbf{x}}_k^{(i)})$. The update steps generate a set of predicted posterior PDFs $p(\tilde{\mathbf{x}}_{S,k+1} | \tilde{\mathbf{y}}_{1:k+1}, \mathbf{u}_{1:k+1})$ along with their corresponding entropies, using the approximation introduced in Eq. (4). This allows for the calculation of the immediate reward $r(\tilde{\mathbf{x}}_{S,k+1}^{(i)}, \mathbf{x}_{R,k+1})$ based on the negative entropy, which will be explained in Sec. 3.2.

All obtained immediate rewards are averaged across all Monte Carlo simulations, yielding an approximation of the expected reward $R(\mathbf{x}_k, \mathbf{u}_{k+1})$. The policy obtained by MCE is a greedy policy, as it exclusively considers the expected reward at the next time step. Hence, action selection can be conducted by evaluating $\pi(\mathbf{x}_k) = \arg \max_{\mathbf{u}_{k+1}} R(\mathbf{x}_k, \mathbf{u}_{k+1})$.

3.2. Reward function

The approach proposed in this study relies on an immediate reward function that constitutes two possibly conflicting

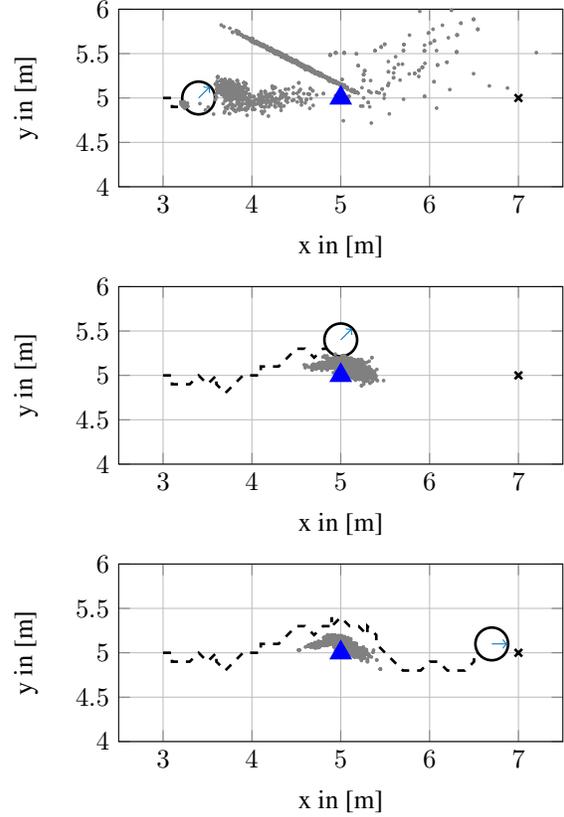


Fig. 1: Active exploration in reverberant conditions with $T_{60} = 250$ ms. The source position (blue triangle) is located between the initial robot position and the goal position (black cross). MCE is performed with $\lambda = 0.5$, generating a trajectory (dashed black line) towards the goal which helps reducing the uncertainty of the particle set (gray dots).

goals: minimizing localization uncertainty by exploratory movements and reaching a specified goal position \mathbf{x}_G . Hence, a trade-off has to be found which balances exploration and goal-directed actions. This is expressed via the function

$$r(\tilde{\mathbf{x}}_{S,k+1}^{(i)}, \mathbf{x}_{R,k+1}) = - \left(\lambda \|\mathbf{x}_{R,k+1} - \mathbf{x}_G\|_2 + (1 - \lambda) H(\tilde{\mathbf{x}}_{S,k+1}^{(i)}) \right), \quad (5)$$

where $\lambda \in [0, 1]$ is a trade-off parameter that balances the minimization of the Euclidean distance from the robot to the goal position and exploration achieved by considering the entropy predicted for the next time step. The implications of the choice of λ will be discussed in Sec. 4.3.

An example of a trajectory generated by MCE is depicted in Fig. 1. It shows that the proposed control scheme makes use of translatory and rotational movements to support distance estimation and reduce front-back ambiguities, while approaching the goal position.

4. EVALUATION

4.1. Sound database

A collection of speech and non-speech sounds was obtained using the dataset provided with the “sound event detection in synthetic audio” task of the Detection and Classification of Acoustic Scenes and Events (DCASE) challenge 2016¹. The database consists of isolated recordings from 11 sound categories with 20 samples per category.

4.2. Experimental setup

BRIRs of four rooms, comprising reverberation times (T_{60}) of 250 ms, 500 ms, 750 ms and one anechoic room, were created using HRIRs [16] and the image-source method [22]. Supervised training of the measurement model (2) was conducted using pre-rendered BRIRs on a grid with 25 cm distance and 5° azimuth spacing. During evaluation, BRIRs were rendered online at each time step. The source position was always placed at the center of the room.

The experimental procedure is based on a 4-fold cross-validation approach, where BRIRs of three rooms were used for training. The remaining BRIRs were used for evaluation, by running 50 simulations with random initial robot poses and goal positions. Sound samples were randomly selected from the database and the initial conditions were consistent over all cross-validation folds. Each simulation was restricted to a maximum simulated duration of 60 s. All experiments were repeated for different settings of the trade-off parameter $\lambda \in \{0, 0.25, 0.5, 0.75, 1\}$. An additional experiment with a bearing-only measurement model was also conducted.

The system performance was assessed by the root mean square error (RMSE) of the estimated Cartesian sound source position, averaged over the corresponding sets of simulations. To better account for outliers that sporadically occur due to the stochastic nature of the particle filter, the median of the achieved localization performance is reported in Tab. 1.

4.3. Results and discussion

The results depicted in Tab. 1 show that MCE improves localization performance compared to a fixed trajectory towards the goal ($\lambda = 1$). The best localization accuracy is achieved for $\lambda = 0.25$. However, this comes at the cost of an increased time to reach the goal position, as depicted in Tab. 2. Therefore, λ must be chosen appropriately for specific applications.

An interesting outcome of the conducted experiments is, that localization performance is significantly degraded for $\lambda = 0$. This can be explained by the fact, that the robot is able to move around freely, without being restricted to approach the goal position. The policy obtained in this case tends to steer the robot very close to the assumed source position. This disturbs the observed binaural cues due to the

Table 1: Median localization errors (m) for all investigated acoustic conditions (T_{60} , denoted in the top row). Experiments using a measurement model based on azimuth and distance (AD) were conducted. For comparison, a bearing-only measurement model (A) was evaluated with $\lambda = 0.25$.

Obs.	λ	Anec.	250ms	500ms	750ms	Avg.
AD	0.00	0.58	0.72	1.63	2.82	1.44
AD	0.25	0.51	0.63	0.70	0.77	0.65
AD	0.50	0.63	0.67	1.05	1.29	0.91
AD	0.75	0.87	0.91	1.18	1.44	1.10
AD	1.00	1.09	1.08	1.36	1.37	1.23
A	0.25	0.82	0.84	0.87	0.89	0.86

Table 2: Time-to-goal (TTG) in (s) and percentage of goal positions reached (GPR) within the available simulation time, averaged across all acoustic conditions for four investigated values of λ (top row). As $\lambda = 0$ does not steer the robot towards the goal position, it is not explicitly shown here.

Metric	0.25	0.5	0.75	1.00
TTG	38.31	14.49	9.94	9.44
GPR	60.00	98.00	100.00	100.00

near-field effect of the HRIRs and large jumps in azimuth between consecutive time steps. As this is not explicitly covered by the measurement model, the particle filter is not able to appropriately predict these effects.

The comparison of the proposed azimuth and distance-dependent measurement model with the bearing-only approach shows, that modeling distance information using the IC helps to further improve the localization capabilities of the system. The improvements are statistically significant according to a t-test conducted with $p < 0.01$.

5. CONCLUSIONS

A machine hearing framework for active binaural localization on a mobile robot was presented. The experimental results obtained in simulated acoustic scenes indicate that MCE improves localization accuracy, if the trade-off parameter of the reward function is properly chosen. Furthermore, the proposed measurement model, which incorporates predictions of distance-dependent binaural cues, further improves localization performance of the system. The proposed framework provides a starting point for further investigations. For instance, these might focus on adaptively selecting the trade-off parameter λ . Additionally, the proposed measurement model could be improved by incorporating additional cues like DRR, to yield better modeling capabilities of sound distance.

¹<http://www.cs.tut.fi/sgn/arg/dcaset2016/>

6. REFERENCES

- [1] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, Cambridge, MA, 1990.
- [2] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, MA, 1999.
- [3] H. Wallach, "The role of head movements and vestibular and visual cues in sound localization," *Journal of Experimental Psychology*, vol. 27, no. 4, 1940.
- [4] N. Ma, T. May, H. Wierstorf, and G. J. Brown, "A machine-hearing system exploiting head movements for binaural sound localisation in reverberant conditions," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 2699–2703.
- [5] N. Ma, G. J. Brown, and J. A. Gonzalez, "Exploiting deep neural networks and head movements for binaural localisation of multiple speakers in reverberant conditions," in *16th Annual Conference of the International Speech Communication Association, Dresden, Germany, September 6-10, 2015*, pp. 3066–3070.
- [6] C. Schymura, F. Winter, D. Kolossa, and S. Spors, "Binaural sound source localisation and tracking using a dynamic spherical head model," in *16th Annual Conference of the International Speech Communication Association, Dresden, Germany, September 6-10, 2015*, pp. 165–169.
- [7] C. Evers, A. H. Moore, and P. A. Naylor, "Acoustic simultaneous localization and mapping (a-slam) of a moving microphone array and its surrounding speakers," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, Shanghai, China, March 20–25, 2016*, 2016.
- [8] G. Bustamante, A. Portello, and P. Danès, "A three-stage framework to active source localization from a binaural head," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing*, April 2015, pp. 5620–5624.
- [9] G. Bustamante, P. Danès, T. Forgue, and A. Podlubne, "Towards information-based feedback control for binaural active localization," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, Shanghai, China, March 20–25, 2016*, 2016, pp. 6325–6329.
- [10] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, "Information based adaptive robotic exploration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2002, pp. 540–545.
- [11] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, The MIT Press, 2005.
- [12] Y. C. Lu and M. Cooke, "Binaural estimation of sound source distance via the direct-to-reverberant energy ratio for static and moving sources," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 7, pp. 1793–1805, Sept 2010.
- [13] E. Georganti, T. May, S. van de Par, and J. Mourjopoulos, "Sound source distance estimation in rooms based on statistical properties of binaural signals," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 8, pp. 1727–1741, Aug 2013.
- [14] S. Vesa, "Binaural sound source distance learning in rooms," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 8, pp. 1498–1507, 2009.
- [15] H. Nakhost and M. Müller, "Monte-Carlo Exploration for Deterministic Planning," in *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, San Francisco, CA, USA, 2009, IJCAI'09, pp. 1766–1771, Morgan Kaufmann Publishers Inc.
- [16] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, 2001, pp. 99–102.
- [17] T. May, S. van de Par, and A. Kohlrausch, "A probabilistic model for robust localization based on a binaural auditory front-end," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 1–13, 2011.
- [18] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [19] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time Signal Processing*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1999.
- [20] R. van der Merwe and E. Wan, "Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003.
- [21] C. Stachniss, D. Hahnel, and W. Burgard, "Exploration with active loop-closing for FastSLAM," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2004, vol. 2, pp. 1505–1510.
- [22] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.