【11】


| | |
|---|---|
| 氏　　　　　名 | 岡　田　雅　司 |
| 博士の専攻分野の名称 | 博　士（情報科学） |
| 学　位　記　番　号 | 第　２５８４８　号 |
| 学位授与年月日 | 平成 25 年 3 月 25 日 |
| 学位授与の要件 | 学位規則第４条第１項該当<br>情報科学研究科情報システム工学専攻 |
| 学　位　論　文　名 | A Study on 3D Sound Field Rendering System based on Geometrical Acoustic Analysis and Binaural Signal Processing<br>**（幾何音響解析とバイノーラル信号処理に基づく三次元音場生成システムに関する研究）** |
| 論　文　審　査　委　員 | （主査）<br>教　授　尾上　孝雄<br>（副査）<br>教　授　竹村　治雄　　准教授　武内　良典　　准教授　橋本　昌宜 |




## 論 文 内 容 の 要 旨


This dissertation discusses the problems of 3D sound field rendering. Rendering generates graphics or sounds in a virtual scene from the descriptor of the scene. Realistic ones can be rendered by analyzing optical or acoustical wave behaviors in the scene. Interactive rendering of dynamic scenes allows a user to walk around the virtual environment. The synergy of realistic and interactive rendering would immerse a person in a virtual world visually and auditorily. In such virtual environments, various kinds of activities are available without practical environments: professional trainings, simulations, conventions, acoustical designing and testing of prototype buildings, and so on.

3D sound field rendering is realized by simulating acoustical wave phenomena caused by environments and a listener's body. Waves emitted from a sound source reflect at walls and/or diffract at barriers and then arrive at the listener from various directions. The arrival waves reflect and/or diffract again at the torso, shoulders, and outer ears of the listener. Such wave phenomena can be analyzed by numerical means, or numerical acoustics (NA); it numerically solves the wave equation. Although NA is able to involve all of the above wave effects, the analysis is strictly limited to small cases because it requires massive computer resources. In addition, fast analysis for interactive rendering is impractical due to the high computational cost. Geometrical means, or geometrical acoustics (GA), approximately analyzes wave propagations with rays. The procedure is simple and much faster than NA, which is better suited to interactive rendering. On the other hand the ray based analysis cannot handle two kinds of wave effects by itself: diffraction and head-related effects. They can be involved in GA by introducing a diffraction model and a binaural signal processing technique. However, the introduction induces computational difficulties. While coping with the difficulties, this dissertation tries to realize an interactive system of 3D sound field rendering.

This dissertation first discusses diffraction modeling in GA. Diffraction should be involved in sound rendering to give listeners enough cues to perceive and localize a hidden sound object. Moreover, in several applications such as acoustical designing, diffraction fields must be accurately estimated. For this purpose, this dissertation focuses on the analytic secondary source model which is the exact model of diffraction. This model gives accurate results but requires us to solve numerically challenging integrations: multiple integrations and integrations with singularities. Because of this computational complexity, this exact model has rarely been applied to rendering applications. The novel method, proposed in this dissertation, solves the integrations by means of ray-tracing-like procedure and the Monte Carlo method. The similarity between the proposed method and general GA analysis makes it possible to utilize the various acceleration approaches developed for ray tracing such as many-core parallelization. In order to address the difficult integrations, two importance sampling are derived. The proposed method, complemented the importance sampling, can solve the integrations efficiently and accurately. The accuracy of the proposed method is demonstrated by comparing its estimates with the ones calculated by reference software. An analysis of signal-to-noise ratios is performed objectively and subjectively in order to evaluate the error characteristics and perceptual quality. The experiments also confirm that the proposed method is superior over the conventional approximate diffraction model, i.e., the uniform theory of diffraction.

This dissertation then examines a problem with regard to head-related wave effects. A transfer function of a sound signal from a source to a listener's outer ear is referred to as head-related transfer function (HRTF). By simulating HRTFs with digital filters, head-related wave effects can be involved in rendering; this process is referred to as binaural signal processing. However, many HRTF filters must be operated simultaneously for sound field rendering because numerous indirect sounds come from the surroundings of the listener's head in echoic environments. To deal with this problem, an efficient binaural processing method for multiple sources is discussed. The proposed method introduces an existing idea that groups several sources by means of clustering and shares a single HRTF filter with the group, reducing the total number of HRTF filters. Another idea, amplitude-panning, is also introduced and combined to the above scheme by utilizing fuzzy clustering. The proposed method, exploiting the two conventional methods' features, can efficiently render multiple sources with a smaller number of HRTF filters. The proposed method is applied to 3D sound field rendering and evaluated. Objective and subjective evaluations reveal that the proposed method is about 2-7 times more efficient than the conventionals in terms of the number of HRTF filters.

Finally, a 3D sound field rendering system is realized by utilizing the contributions above mentioned. For fast GA analysis including the proposed ray tracing for diffraction, this system employs the OptiX ray tracing engine, which exhibits good acceleration performance on a commercial graphics processor. In order to make it efficient to execute multiple HRTF filters, OpenMP based multi-threading is employed. The system, accelerated by the above techniques, can interactively analyze dynamic auditory scenes and render sound fields in real-time, while supporting the difficult wave effects. The realization of the system significantly contributes to the evolution of various applications which involve sound rendering. The sense of reality and immersion in VR applications and video games will improve with the proposed system. Interactive and accurate rendering of the system can support acoustical designing of rooms. The combination with authoring tools for acoustic content expands the possibility of the content production.


## 論 文 審 査 の 結 果 の 要 旨

本論文は、三次元音場生成技術に関する研究の成果をまとめたものであり、以下の主要な結果を得ている。

**1. レイトレーシングによる回折シミュレーション法の提案**

幾何音響解析法の一つであるレイトレーシングは、音の伝播を高速に解析するうえで有効な方法である。しかしながら、一般のレイトレーシングは重要な音響効果の一つである回折を扱うことができない。本論文では回折の厳密解である解析的二次音源モデルに着目し、本モデルに対する新たな解法を提案している。本手法はレイトレーシングを利用して本モデルにおける多重積分を計算する。その手続きは、既存のレイトレーシングと共通し、従来法における回折の欠如を補完することができる。また、従来法に対する高速化・効率化法を本手法にそのまま適用可能である。また、二種類の重点サンプリング法を提案することで、多重積分や特異点を含む積分の計算容易化を実現している。

**2. 多音源に対する効率的なバイノーラル信号処理法の提案**

任意の仮想環境に対して幾何音響解析を実行し、得られた間接音（反射音、回折音）をバイノーラル信号処理することで、頭部による音響効果を考慮しつつ当該環境の音場を再現できる。しかしながら間接音数は数千・数

万になるため、それら全てに対して信号処理を行うのは現実的ではない。本論文ではファジィクラスタリングを利用して複数音源をいくつかのグループに分割し、各グループに対する信号処理を共有化する手法を提案している。本手法は従来法の課題である定位精度・音像品質劣化を同時に改善することができる。また、客観・主観聴取実験の結果、本手法は1/7～1/2の信号処理負荷で従来手法と同等の品質を達成可能なことを示した。

**3. 三次元音場生成システムの実現**

音の反射に加え、回折および頭部による音響効果を再現可能な世界初の音場生成システムを実現した。本システムは上記の成果を応用することで、回折と頭部効果の計算を行っている。また、ＧＰＵやマルチコアＣＰＵを利用した並列計算を導入し、高速な音響解析および実時間での信号処理を実現している。性能評価実験の結果、本システムは、あるコンサートホールの音場を約28.5msで推定できることを示した。これは、音源・聴取者位置の変化に伴う、動的な音場の変化を約35Hzの速度で更新できることを示しており、仮想現実などの対話的なアプリケーションへの応用が期待できる。

以上のように、本論文で述べた三次元音場生成技術に関する研究は、より現実に近い仮想音場を生成するための要素技術の確立という点で非常に有用である。各種マルチメディアコンテンツの高臨場化が進むなか、本成果は音響分野における技術発展に大きく寄与するものと期待できる。従って、博士（情報科学）の学位論文として価値のあるものと認める。