

論文 / 著書情報
Article / Book Information

題目(和文)	2次元画像からの後部残響特性推定に関する研究
Title(English)	
著者(和文)	今誉
Author(English)	Homare Kon
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11481号, 授与年月日:2020年3月26日, 学位の種別:課程博士, 審査員:小池 英樹,篠田 浩一,徳永 健伸,三宅 美博,齋藤 豪
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11481号, Conferred date:2020/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報 コース	系 コース	申請学位 (専攻分野)： 博士 (工学) Academic Degree Requested Doctor of
学生氏名： Student's Name	今 誉		指導教員 (主)： 小池 英樹 Academic Supervisor(main)
			指導教員 (副)： Academic Supervisor(sub)

要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters)

本論文は「2次元画像からの後部残響特性推定に関する研究」と題し、複合現実における現実空間に適合した後部残響特性を1枚の2次元画像から推定する手法と推定性能の改善手法、およびその性能を論じるもので、全7章からなる。

第1章「序論」では、本研究の背景である拡張現実(AR)と複合現実(MR)で再現すべき空間に適合した聴覚刺激となる残響について述べ、1枚の画像からも後部残響特性が推定できる可能性を示唆し、本研究の立ち位置を明確にしている。従来は物理的アプローチとして空間の幾何学構造と境界の反射特性から残響特性を演算していたが、演算量が膨大であり、AR/MRを想定したアプリケーションではリアルタイムに実現することが困難であった。本論文では人間の脳が視覚刺激から聴覚刺激を想起できることに着目し、空間画像からの残響特性を推定する低演算な感覚的アプローチとして、脳機能の特性に類似した数理モデルであるニューラルネットワークでの実現手法を提案している。

第2章「畳み込みニューラルネットワークによる残響特性の推定」では、本研究で扱う残響特性を定義し、畳み込みニューラルネットワーク(CNN)を用いて1枚の2次元画像から残響特性を推定するネットワークの実装と評価をし、改善すべき課題を提示している。後部残響特性の時間周波数特性であるEDR(Energy Decay Relief)の初期エネルギーとディケイタイムのオクターブバンド周波数での値をそれぞれ10次元の推定パラメータとしてCNNで対になる空間画像を入力として学習させたモデルによって残響特性の推定にかかる演算時間は150ms程度と物理的アプローチより優れていることを述べている。一方で推定結果の評価から、推定パラメータの汎化性能が低いこと、聴覚特性を考慮したパラメータのスケールリングと学習の重み付けが必要であること、入力画像の方向変化で推定結果が安定しないことを改善すべき課題として挙げている。

第3章「主成分分析による推定残響パラメータの次元削減」では、ニューラルネットワークの汎化性能を向上させるための適切な残響パラメータの次元数を主成分分析にて主観評価と客観値の観点から検討を行った。次元数が3未満の場合、主観評価としてDMOS(Degradation Mean Opinion Score)を用いた弁別が可能であることを示し、次元数が3の場合には、誤差が初期エネルギーで3~5dB、ディケイタイムで20~30%以内となり、関連研究の知見と類似することから、3次元の圧縮が妥当であると論じている。

第4章「聴覚心理を応用した残響パラメータのスケールリングと学習手法」では、人間の聴覚心理の知見を参考に残響パラメータのスケールリングと学習の重み付けをすることで推定誤差が改善し、主観評価にて再現精度が向上することを述べている。パラメータのスケールリングにはマスキング効果、等ラウドネス曲線やヴェーバー・フェヒナーの法則を用い、周波数毎に最小可聴域から疼痛域までのダイナミックレンジに応じた重み付け学習を用いることでCNNの推定精度が有意に改善することを記している。また、MUSHRA(Multi Stimulus test with Hidden Reference and Anchor)を応用した主観評価を行い、スコアも改善したことを述べている。

第5章「空間画像の連続的な特徴量変化による推定値のロバスト化」では、同一空間の画像で撮影方向が変わった場合に推定値が安定しない課題に対して、ニューラルネットワークに確率分布モデルを用いることを提案し、連続的に撮影方向が変化する画像に対して推定値の変化が少なくなることを述べている。潜在変数部分に確率分布を導入した変分自己符号化器(VAE)の連続的に潜在変数が変化すると出力も連続的に変化するという特徴に着目し、これをネットワークに組み込んだ例を記している。残響パラメータを出力とするネットワークに加え、入出力対になる学習データセットが少ない場合を考慮し、空間画像のみと残響パラメータのみのVAEをそれぞれ独立に対になっていないデータセットで学習させた後にそれぞれの潜在変数への写像を学習させるネットワークを提案している。それぞれの学習モデルをパノラマ画像から方角を徐々に変えて切り出した画像を入力し、出力の変化を階差で評価すると、それぞれ確率モデルが導入されたネットワークで出力が有意に安定化することを記している。

第 6 章「考察」では、本論文での提案手法の効果を簡便な従来手法と比較し、有効性を示すとともに、さらに今後の課題として提案手法では再現できなかった残響パラメタについての再現手法について議論している。

第 7 章「結論」では、第 1 章から第 6 章までの研究内容を総括している。

以上のように、本論文では、人間の脳が視覚から聴覚刺激を想起できることに着目した 2 次元画像から後部残響特性を推定するニューラルネットワークを提案し、実現への課題を明確にし、課題の解決手法を提案、実施するとともに客観的な側面と主観的な側面からの評価を行い、推定精度を向上させている。本研究は、AR/MR で空間に適合した後部残響をリアルタイムに再現する低演算手法を提案し、その効果を示したものである。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報 系 コース	申請学位 (専攻分野)： 博士 Academic Degree Requested	(工学) Doctor of
学生氏名： Student's Name	今 誉	指導教員 (主)： Academic Supervisor(main)	小池 英樹
		指導教員 (副)： Academic Supervisor(sub)	

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

The reproduction of effects adapted to real space is essential for creating a high-presence experience in augmented reality (AR) and mixed reality (MR) applications. In a visual example, reproducing realistic shadow and shade of the virtual objects can make them look real and allow users to feel they sharing the space. The element of sound affected by the environment is reverberation. Therefore, the audio component of an AR/MR experience needs to simulate the acoustics of the environment that the users are experiencing in order to create the feeling of sharing the space with virtual objects.

Earlier, in virtual reality (VR) applications, sound engineers could program all the reverberation parameters for a particular scene in advance or when the user is at a fixed position. However, it is not easy to adjust the reverberation parameters using conventional procedures because of the inability to program the unlimited range of such parameters for AR/MR applications. Meanwhile, the reverberation synthesis method with physical simulation requires an enormous amount of calculation and is not realistic for mobile AR/MR application. Therefore, it is necessary to estimate the reverberation characteristics at low computational cost based on the environment of the users' location dynamically. Considering that the human brain can recall auditory stimuli from visual stimuli, I assumed that it is possible to simulate the ability using the mathematical model of neural network, which is a network or circuit of neurons.

In this thesis, by focusing on late reverberation, which is not easily affected by the spatial position, I develop an estimation method of late reverberation characteristics from a single two-dimensional environmental image using neural networks, clarify the problems, propose the method to solve them, discuss based on the objective and subjective evaluations, and improve the estimation accuracy. In this study, I propose a low computation method for real-time reverberation adaptation to space for AR/MR application and show its effects. The results of this study contribute to realizing high-presence audio for AR/MR applications.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).