

## 審査の結果の要旨

氏 名 劉 玉宇

本論文は「Synchrony-based Audiovisual Analysis」（同期性に基づく音と映像の統合解析）と題し、映像中の話者領域抽出、移動音源の位置特定、音と映像間の同期復元の3つの問題に対し、音と映像の間の同期性に基づき統一的な解決の枠組みを示したものであり、英文で記され全体で6章により構成されている。

第1章「Introduction」（はじめに）では、本研究の背景と目的について論じた後、本研究における2つの主要な論点、すなわち音情報と映像情報との関係を如何にして解析するかという点と、その解析にもとづき映像中の音源特定や音と映像間の同期復元の問題を如何にして解決するかという点についてまとめている。

第2章「Preliminaries」（準備）では、当該研究分野における本研究の位置づけについて整理した上で、音と映像の相関解析に関してこれまでに提案されている特徴量および相関指標について詳説し、本研究における相関解析手法の特長についてまとめている。

第3章「Face region segmentation of a stationary speaker」（映像中の静止話者領域の抽出）では、音と映像の相関解析により映像中の話者領域を自動で切り出す手法を紹介している。まず、従来の音と映像の相関分析にもとづく音源位置推定技術に共通する、断片化された領域しか得られないという課題を指摘した上で、グラフカット最適化による領域分割に音と映像の相関分析を組み入れるという着想にもとづいた解決方法を提案している。複雑な動きを伴う背景中の人物の映像を用いた実験により提案手法の有効が示されている。

第4章「Visual localization of a non-stationary sound source」（移動音源位置の特定）では、音源が移動する場合への拡張について紹介している。音源が移動する場合、音と映像の相関を正しく計算することができないため、既存の音と映像の相関分析にもとづく音源位置推定技術では移動音源を扱えないという問題が存在した。これに対し、本研究では音と映像の相互情報量の増減を手掛かりに対象の移動経路を探索するという新たなアプローチを提案し、移動する音源であっても精度良く位置を推定することを可能としている。また、相互情報量を逐次的に計算する方法を導入することにより、経路探索の計算コストを大幅に削減することを実現している。

第5章「Recovery of audio-to-video synchronization」（音と映像の同期の復元）では、映像の圧縮符号化やネットワークによる伝送により音と映像の同期が一旦ずれてしまうと同期を復元するのは容易ではないという問題に対し、音と映像の間の同期を自動で復元する手法を提案している。一般に、音と映像の同期ずれを生じた場合、音の増減変化と映像の動きのタイミングがずれるため音と映像の相互情報量が低下するという点に着目し、音と映像の相互情報量を最大化する同期ずれ量を推定することにより、音と映像の同期ずれの補正を実現している。さまざまな映像を用いた実験により、提案手法により映像規格として定められている同期誤差範囲内の同期復元が可能であることが確認されている。

第6章「Conclusions」（まとめ）では、全体を総括し、今後の課題と展望について述べている。

以上これを要するに、本論文では、映像中の話者領域抽出、移動音源の位置特定、音と映像間の同期復元というマルチメディア情報処理における3つの重要な課題に対し、音と映像という異なるモダリティ間の相関分析にもとづいた解決方法を提案し、シミュレーションおよび実データを用いた実験により各手法の有効性を示したものであり、電子情報学上貢献するところが少なくない。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。