

氏名(本籍)	なかのともやす (新潟県) 中野倫靖	
学位の種類	博士(情報学)	
学位記番号	博甲第4812号	
学位授与年月日	平成20年3月25日	
学位授与の要件	学位規則第4条第1項該当	
審査研究科	図書館情報メディア研究科	
学位論文題目	歌唱理解及び歌唱表現の解明とその応用システム構築に関する研究	
主査	筑波大学教授 小高和己	
副査	筑波大学教授 田中和世	
副査	筑波大学教授 平賀譲	
副査	筑波大学教授 水谷孝一	
副査	筑波大学准教授(連携大学院) 後藤真孝	

論文の内容の要旨

本論文は広い意味での「歌唱」、つまり音声による音楽表現について著者が行った2つの研究をまとめたものである。論文構成としては、第1章の序論で研究動機や目的、背景となる関連研究が記されたのに続き：

第Ⅰ部：歌唱理解の解明とその応用システム構築（第2～4章）

第Ⅱ部：歌唱表現の解明とその応用システム構築（第5～7章）

がそれぞれの研究を述べたものとして全体の中核を成し、第8・9章の総合考察・結論でまとめの考察が行われている。

「2つの研究」と書いたが、実際にはそれぞれにおいて、共通のテーマに対し、著者の言う「HMIアプローチ」に基づき以下のような個別の研究（アプローチ）が統合されている。

- ・人間による歌唱やそれに関する判断・評価について、心理実験を中心とした手法による調査・分析（Human分野）
- ・得られた知見に基づき、音響・音声信号処理技術を用いて目的とする分析・認識機能を実現するための基礎的研究、及びその計算機上への実装（Machine分野）
- ・その有効性を確認した上で、それを応用した実用性のあるユーザ向けシステムの開発及び評価（Interface, Interaction分野）

これら H, M, I の各分野における研究は、それぞれが単独の研究としての完結性を持つとともに、相互に有機的に関連し合っている。それらの総体として、対象となるテーマについての多面的なアプローチによる検討・解明を目指した研究となっている。

第Ⅰ部の「歌唱理解」では、歌う人の「歌唱力」について、人間はどのような観点から評価を行うか、それを踏まえて計算機上で「歌唱力評価システム」をどのように構築するかがテーマである。ここでの特徴は、楽譜情報のような、比較対象としての典拠になる「正解データ」を用いずに評価を行うという点にある。これは人間が歌唱力の良否を直観的に判断する状況と合致しており、カラオケ評価システムなど、楽譜情報を利用した評価手法とは一線を画する。

第Ⅱ部の「歌唱表現」では、「歌唱」としてはやや特異な「口ドラム」が対象となっている。口ドラムとはドラムなどの打楽器音を音声により模倣・演ずることを指す。これには音響的にリアルな再現を目指すものから、「ドンタン」のような擬音語表現でいわば「記号的」に表現しようというものまで、様々な種類がある。本研究は後者の擬音語的表現について、音声認識技術を使って認識し、その応用を目指したものである。

どちらの研究においても、人間による判断・評価にベースをおいているため、個人の特質や、好み等の判断の違いによる個人差にどう対処するかが、研究やシステム開発にあたっての中心的課題となる。

第Ⅰ部「歌唱理解」の第2章では歌唱者の歌唱力を、原曲を知らない聴取者が評価する評価実験が述べられている。評価用データには、同一曲を10人が歌唱したもの（うち1つはプロのオリジナル歌唱、他はそれを一般人が反復歌唱したもの）8曲分を用い、順位法による歌唱力評価実験を行った。評価は原曲の歌唱とともに、 F_0 成分のみを抽出した F_0 歌唱についても行い、合わせて判断基準のプロトコルも採取した。結果として、原曲歌唱については評価者の判断に高い相関が得られ、評価者個人に拠らない共通する判断基準の存在が示された。 F_0 歌唱についてはそれほど高い相関は得られず、捨象された要因の重要性が示唆された。

第3章ではこの結果に基づき、歌唱力評価に寄与する音響的特徴量の抽出と、それによる評価実験について述べられている。評価は各歌唱データに対し、「うまい／へた」のラベルを設定した2分識別により行った。予備実験では24種の特徴量を設定し、寄与度の高いものとして相対音高（いわゆる「音程」）に関するものと、ヴィブラートの有無に関するものが抽出された。それを用いた本実験では、83.5%の識別率が得られた。

第4章では、この相対音高・ヴィブラートの検出機能を組み込んだ応用システム、Mirusingerが紹介されている。MirusingerはCD等のヴォーカルデータを、それをなぞったユーザの歌唱から抽出した相対音高・ヴィブラートデータと合わせてリアルタイムにグラフィック表示することにより、ユーザが自身の歌唱を理解し、向上させることを支援するシステムである。システムはカラオケ感覚で使うことができ、歌唱力向上という直接目的だけでなく、エンターテインメントとしての意義もある。

第Ⅱ部「歌唱表現」の第5章では、バスドラム、スネアドラムという2種類のドラム音の口ドラム表現方法を収集する実験について述べられている。結果として、多くの被験者（一般人）は擬音語的表現を用いる反面、同一音素の音高の違いによる表現方法なども見られた。擬音語表現にも一定の共通性と個別的な多様性が存在している。

これを受けて、第6章では口ドラム表現からドラムパターンを認識する手法の検討及び実装について述べられている。分析には音声認識で用いられる確率モデルの手法などを援用し、ドラムパターンデータベース中のデータと照合・検索する形で認識を実現した。このような処理ではユーザの個人差による影響が大きく、もっとも適応した条件のもとでは92%の認識率が得られた。

第7章ではこの識別手法を利用したドラム譜入力インタフェース、Voice Drummerが紹介されている。Voice Drummerは口ドラム入力によりドラム譜作成や編曲を行うための支援システムで、入力・認識内容がリアルタイムにグラフィック表示される。楽譜入力という直接的用途の他に、「太鼓の達人」のようなエンターテインメント性も備えている。

第8・9章ではこれらの結果を踏まえ、残された研究上の課題や、歌唱表現研究全般にわたる今後の研究への展望が論じられている。

審査の結果の要旨

まず「HMIアプローチ」について述べる。

「論文要旨」でも述べたように、本論文では2つの研究テーマについて、H、M、Iそれぞれのアプローチによる研究が行われており、それぞれ単独の研究としても十分自立性を有する内容になっている。しかしさらに、これらが個々に別々の研究として行われているのではなく、一貫した意図及び計画のもとで行われている点で、相互の関連性による相乗効果を生じており、それが研究全体の意義・価値を高めることにつながっている。

一般論として、このような多面的・総合的アプローチによる研究が極めて有意義であることは言うまでもないだろう。もちろんこれがすべて1人の研究者あるいは研究グループにより遂行される必要性は必ずしもなく、共同研究あるいは研究コミュニティ全体の研究成果として実現されていくのでもかまわない。逆に言えば、それぞれのアプローチは方法論、背景、必要な専門知識などが異なっており、そのすべてを同一人が遂行するというのは難しい。その点において、大学院生という身分でこれらの研究を遂行しえたこと自身、極めて高く評価できるとともに、研究者たる資質・力量において非凡なものを感じさせる。(付言：本論文に関する研究発表は連名共著で行われているが、共著者は研究についての指導・助言・環境提供などを主として担当し、研究の中核部分は著者本人により遂行されている。)

もちろん個々の分野において、専門的見地から見れば不十分な点はいろいろあるかもしれない。しかしそういった不備の如何に関わらず、異分野にわたるアプローチに積極的に取り組んだ意欲と、それによって初めて研究として完結しようという意識・信念は高く評価したい。

具体的な研究内容については、第1章の関連研究の記述でも述べられているように、第I部「歌唱理解」、第II部「歌唱表現」のいずれも、そのまま過去の研究を踏襲したようなものではなく、オリジナリティは極めて高い。

歌唱力評価については少し意外な感がないでもないが、従来の研究はオペラ歌唱のような特定の歌唱の分析や心理学的研究は行われていても音響的解析とは結びついていなかったり、また評価システムの場合、楽譜情報などの「正解データ」との照合を行うものが中心だった。その点、楽譜情報を用いず、歌唱そのものから判断しようという本研究のアプローチは発表時点で斬新性があり、また応用範囲も広い。

口ドラムのほうはテーマとしては少し特異ではあるが、それでも他に類例のない孤立した研究というわけではなく、本研究の発表と前後して同様のテーマを扱う研究が増えてきており、むしろ発展性のある分野と言えよう。

本研究の発表については、情報処理学会論文誌への採択、ISM2007 (9th IEEE International Symposium on Multimedia) などの著名国際会議での採択・発表があり、また歌唱力評価研究については2006年の日本音楽知覚認知学会研究選奨を、MiruSingerでは情報処理学会インタラクシオン2007でインタラクティブ発表賞を受賞している。それらを見ても、内外の学会でも高い評価と反響を受けており、すでに研究者としての地位が確立されていると言えよう。

もちろん扱っているテーマに対し、研究内容が十全というわけではない。例えば歌唱力評価の実験はうまい／へたについて、かなり極端な差があるデータによる分析のため、相対音高とヴィブラートという特徴量が顕著になった。しかし一定以上の水準の歌唱者のデータを対象とした場合には、より精密な判定を行うために別の特徴量が浮上してくると考えられる。また口ドラム実験にしても、2種類のドラム音だけしか対象としないのは限定性が強い。

ただしこれらは本研究に対する否定的見解として述べているのではなく、むしろ本研究の成果を踏まえて、今後様々に発展・精密化させていく余地があることを指摘したものである。著者自身によるものも含め、今後の研究の進展に期待したい。本研究で得られた結果はそれ自身として、新しい研究領域を切り開く端緒として高く評価すべきである。新規の取組みにおいてそのような成果を達成しえたことは、課題設定などがバランスよく行われていることの反映でもあり、研究上のセンスや地に足のついた研究態度を感じさせる。

このように研究成果の面では極めて高く評価できる反面、それを述べる論文の構成や記述内容には課題が感じられる。研究内容を具体的に記した2～7章の記述に比して、序章や8・9章の記述には相対的に「弱さ」が感じられる。これらの部分は研究及び論文の本筋には直接関係しないとはいえ、せつかくの高い成果が十分効果的にアピールされていないきらいがある。

最終試験においても、質問や指摘が行われたのは、研究内容や成果というよりは、このような点についてであった。研究の構想や全体的輪郭を的確かつ効果的に表現すること、広い見地から研究分野全体を概観し、自分の研究を正当に位置づけ、今後の展望を示していくことなどは、直接研究を遂行していく上での能力とは別の面の洞察力や表現力を要求することではあるが、著者の今後の大成を願う上では、そういった面での精進も大いに期待したい。

よって、著者は博士（情報学）の学位を受けるに十分な資格を有するものと認める。