

システム情報工学研究科修士論文概要

年 度	平成 24 年度	学位名		修士(工学)
専 攻	知能機能システム	専攻	著者氏名	澤田 真吾
指導教員氏名 星野 准一				
論文題目				
Web 中の文章とリスト構造を用いたバンドメンバー自動収集手法に関する研究				
論文概要				
<p>本論文では Web 中にある文章とリスト構造から音楽のバンドやグループに関連するメンバーを自動的に収集する手法を述べる。音楽のバンドやグループには、ボーカルやギタリスト、ベーシストなど、多くのミュージシャンがメンバーとして参加している。そして、どのようなボーカルが歌っているか、どのようなギタリストが演奏しているかという情報は、エンドユーザーが購入する楽曲を選ぶ手がかりとなる。しかし、バンドの現メンバーや過去のメンバーなどのすべてのメンバーを集約したデータが記載されたデータベースや Web サイトは少ないため、人の手で情報を集めるのは難しい。そこで我々は、Web マイニングを用いて、音楽のバンドやグループに関係するミュージシャンを、複数の Web ページから自動で収集する手法を提案する。我々の手法では Web ページ上の文章とリスト構造体を利用する。Web ページの文章中に出現するミュージシャン名は、'guitarist', 'singer' などの特徴的な単語と共に出現しやすいという性質がある。そこで周囲の単語列から次に出現する単語を予測する N-gram モデルを用いて文章からメンバー抽出を行う。抽出後、取得したメンバーの中で、抽出頻度が一定以上のものを正解のメンバー、一定未満のものをリザーブメンバーとする。次に、Web 中のリストを用いてメンバー収集を行う。Web 中には人名が箇条書きの形で記載されたリストが存在する。このリストに一人でも正解のメンバーが含まれていれば、そのリストはバンドメンバーを表すリストであり、他の人名もバンドのメンバーであると思われる。そこで N-gram によって得た正解メンバーとリスト中の人名とマッチングを行い、マッチした場合、リスト中の人名をすべて抽出する。最後にリストから抽出した人名と、N-gram で抽出したリザーブメンバーとのマッチングを行い、マッチした人名を正解のメンバーとする。我々はこの手法を用いて、バンド数 50 組に対しメンバー収集を行った。その結果、提案手法は、従来のメンバー収集手法より正確で、より多くのメンバーを収集できることを示した。</p>				
審査日	平成 25 年 1 月 31 日			
審査員	(大学名 職名)	(学位)	(氏名)	
主査	筑波大学 准教授	博士(情報科学)	星野 准一	
副査	筑波大学 講師	博士(工学)	浜中 雅俊	
副査	筑波大学 教授	博士(工学)	宇津呂 武仁	