



音声データ

「部品」と「設計図」に分解

音色・高さ

リズム・抑揚

NTT研 音楽検索の精度向上

NTTのコミュニケーション科学基礎研究所は、音声データに含まれる音の特徴を詳しく解析する技術を開発した。画像処理で使う数値解析を高次元化し、音声の構成単位となる基本の音を割り出し、それらの音がどんなリズムや抑揚で組み合わさっているかを分析する。大量の音楽から素早く検索したり、テレビ会議システムで目的の音だけを強調したりする技術として2年程度で実用化を目指す。

開発した技術を使えば、音声データのうち、「部品」にあたる音の高さや音色と、「設計図」にあたる音の組み合わせ方が同時に分かる。例えば、ピアノで演奏された楽曲データだと、使用頻度の高い鍵盤の音色と、楽譜の情報をそれぞれ取り出すことが可能という。膨大な音楽データを新技術で分析すれば、演奏

する楽器や歌手の歌声が違っていても、設計図を手がかりに同じ楽曲を効率よく探せる。従来は難しかった、似た曲をきちんと区別することもでき、検索精度が上がる。

従来の音声データ分析技術は、特定の音と完全に一致する音を探したり、音声を周波数成分にする楽器や歌手の歌声が違っていても、設計図を手がかりに同じ楽曲を効率よく探せる。従来は難しかった、似た曲をきちんと区別することもでき、検索精度が上がる。

ユーザーで推定する。同時に使う部品の数をできるだけ減らすように条件を設定し、最も適切な部品を選べるようになった。

分解して特徴を照らし合種類や発音の意味内容が微妙が違つと照合が難しかった。