

# 歌声合成ソフトウェア VOCALOIDのメカニズム

剣持 秀紀

最近、歌声合成技術が大きな注目を集めています。本稿では、歌声合成技術VOCALOIDのしくみを説明します。

(編集部)



## VOCALOIDとは

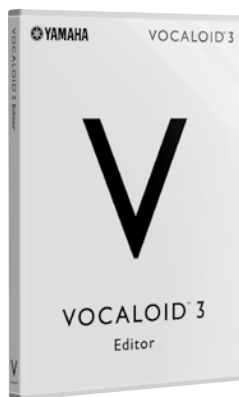
### ● さまざまな歌声で歌ってくれるソフトウェア

VOCALOID<sup>注1</sup>は、ヤマハが開発したパソコン用歌声合成ソフトウェアです。音符とテキストを入力すると楽譜に忠実に歌ってくれます。歌声ライブラリを差し替えると、さまざまな声種のボーカルをプロデューサー気分で楽しめます。

この技術は2003年に発表されて以来バージョンアップ<sup>注2</sup>を重ね、2007年には写真1(a)のVOCALOID2、2011年には写真1(b)のVOCALOID3を発表しています。図1にVOCALOID全体の構成を示します。歌詞や音符、表現を入力し、歌声合成部で歌声ライブラリを結合します。



(a) VOCALOID2の「初音ミク」(クリプトンフューチャーメディア)はライブラリとツールのセット



(b) VOCALOID3 Editor (ヤマハ)は外付けライブラリが必要だが、VOCALOID2より機能が増えている

写真1 歌声合成ソフトウェアVOCALOIDは、家電量販店などで販売されている

### ● 合成方法はシンプル

VOCALOIDでは、実際の人間の歌声から取り出した音声素片を接続することで歌声を合成します。技術的なロマンを追求するのであれば、全てを物理モデルで表現し、何もないところから歌声を作り出すようなものが理想的かもしれませんが、「使える」歌声合成技術への早道として、音声素片を接続する方法を採用しています。

素片を接続するといっても、現在のテキスト音声合成技術で主流となっている朗読などの自然言語を収集したデータベースである大規模コーパスを使う方法ではなく、音と音のつながり(diphone: 子音+母音ならCV, 子音+母音ならVCと表す)と母音の伸ばし音という単純な素片を用いています。これは、大規模コーパス方式により歌声を合成することは、入力となる楽譜の音符の高さ/長さで歌詞の組み合わせが非常に多くなるので非現実的であるためです。

もちろん、特定のジャンルの歌、旋律やリズムという制約を加えれば、大規模コーパスによる歌声合成も現実的だとは思いますが。

VOCALOIDでは、テキスト音声合成方式としては枯れ

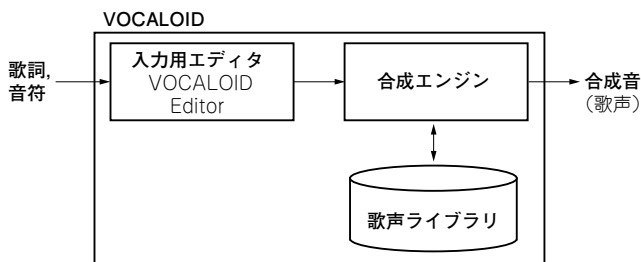


図1 歌声合成ソフトウェアVOCALOIDは歌詞と音符を入力すると歌声を生成する

注1: “VOCALOID”とは、Vocalとoidを組み合わせた造語。“oid”とは、ギリシャ語のeidōsに由来し、「～のような」という意味を作る接尾辞。限りなく人間の声に近づけたいという憧れがこめられた名前である。

注2: 合成アルゴリズムの抜本的改良をもって「バージョンアップ」としている。