

2004年2月6日

日本電信電話株式会社

NTTアドバンステクノロジー株式会社

演歌からポップスまで自在に歌声を作り出す 歌声合成技術“ワンダーホルン”のライセンス提供を開始

～“うたばら.com”で歌声合成のお試しができます～

日本電信電話株式会社（以下NTT、本社：東京都千代田区、代表取締役社長：和田紀夫）では、楽譜と歌詞から高音質の歌声を生成する歌声合成技術の開発に取り組んで参りました。このほど、ビブラートの入れ方・こぶし回しなどの「歌唱スタイル」表現といった新機能を開発するとともに、各種アプリケーションに実装するためのライブラリ群を整備し、「歌声合成技術“ワンダーホルン”」の名称で本日より、NTTアドバンステクノロジー株式会社（以下NTT-AT、本社：東京都新宿区、代表取締役社長：田嶋公郎）を通じて外部へ技術提供を開始いたします。

1. ワンダーホルンの技術開発成果

2000年9月にNTTから発表した歌声合成技術（開発コード“ホルン法”）は、あらかじめ個人の実声収録によって作製した個人歌声データベース<*1>と、その場で入力された楽譜と歌詞から、人間の歌声をコンピュータで合成する技術です（[図1：歌声合成技術“ワンダーホルン”の原理](#)）。本技術の特長は、歌声特有の倍音構造を忠実に再現するモデルの採用と、個人の歌声から実際に抽出した特殊ノイズをミックスすることです。今回、各種ライブラリ群を整備するとともに下記の改良を行い、エンターテインメントの実ビジネスへただちに提供できるまでレベルアップしました。

- (1) ビブラートやこぶしなどの「歌唱スタイル」の調整機能を組み込み、多様な歌唱表現を自在に付与することができるようになりました。これにより、これまでコンピュータ合成が困難であった演歌からポップスまで、さまざまな歌声を合成することができます。また、少人数の歌声の歌唱表現を変えて足し合わせることにより、大合唱を合成することも可能です。
- (2) 基本アルゴリズムの改良と個人歌声データベースの構造を変更し、個人の歌声の特徴をより精密に再現することが可能となりました。これにより特定の人の声での歌声メッセージを作成するなどの応用が可能になり

ました。([図2：自分だけの歌声メッセージの作り方](#))

- (3) アルゴリズムの軽量化と個人歌声データベースの圧縮により、高性能パソコンでなくても、家庭用ゲーム機やDSP(Digital Signal Processor) <*2> 搭載カラオケ機など幅広いプラットフォーム上で動作が可能となりました。
- (4) 楽譜・歌詞の入力インターフェースとしてMIDI規格<*3>に準拠しました。この結果、既存のカラオケ用MIDIデータからバックコーラス歌声を生成することができます。また、インターネットで流れているMIDIファイルや、市販MIDI編集ツールで作曲したメロディーに、簡単に歌詞をつけ、歌声として楽しむことができます。

2. NTT-ATのビジネス展開と「うたばらサイト」

“ワンダーホルン”はNTT-ATを通じて、アプリケーションごとの個別カスタマイズや装置への組み込み、各種SIなどの技術提供、さらには個人歌声データベース制作などの各種関連サービスの提供を本日より開始します。NTT-ATでは、技術提供に先立ち、皆様に気軽にワンダーホルンを試していただくサイト「うたばら.com (<http://www.utabara.com/>)」 <*4>を立ち上げました。本サイトでは、サイト上に用意されたサンプル歌声データベースを使って実際に歌声合成機能を試してみることができます。

NTT-ATでは、ゲーム業界（効果音やキャラクタ歌声用途など）、カラオケ業界（バックコーラスやデュエット曲用途など）、おもちゃ業界（歌って踊るロボット用途など）、教育業界（作曲ツールや合唱シミュレーション用途など）などへの技術提供を計画しています。技術提供の一環として、すでに株式会社セガ（本社：東京都大田区、代表取締役社長：小口久雄）が、ワンダーホルンを自社グループのゲームソフトなどの各種エンターテインメント事業向けに採用を検討しており、NTT-ATとセガとの間でワンダーホルンの包括ライセンス契約に向けて交渉を開始しました。

3. 背景としてのNTTの総合プロデュース機能

本件は、研究開発と事業化との間にある、いわゆる“死の谷”<*5>を克服することを目的とし、2003年7月から開始している“総合プロデュース機能”に基づく事業化案件です。総合プロデュース機能とは、事業化の責任者として指名されたプロデューサーがNTTグループ内外の企業と協力しながら、NTT研究所の優れた研究成果の事業化を直接推進していく取り組みです。ワンダーホルンは肉声と遜色ない自然な歌声を再現できるため、通信分野に限らず、臨場感が求められるゲームソフトや業務用ゲーム機といった各種エンターテインメント分野などで幅広く採用されていくことが期待されます。NTTは今後も総合プロデュース機能により、さまざまな研究開発成果を幅広く事業化推進していく予定で

す。

<用語解説>

*1 個人歌声データベース

五十音すべての発声を収録し、そこから歌声素片を切り出した個人ごとのデータベース。このデータベースを切り替えることにより、異なった個人の歌声を合成させることができます。

*2 DSP(Digital Signal Processor)

音声や画像などの処理に特化した専用プロセッサ。カラオケなどの装置に組み込まれたり、パソコンに搭載されてCPUの処理を一部肩代わりしたりします。

*3 MIDI規格

Musical Instruments Digital Interfaceの略。シンセサイザや音源とパソコンとを接続して楽曲データをやりとりするための規格。音色、音程などのデータを送受信する手順が定められています。業界標準として広く普及しています。

*4 うたばら.com (<http://www.utabara.com/>)

NTT-ATの運営する、ワンダーホルン技術を自由にためすことのできるWebサイト。

*5 死の谷

もとは、米国商務省が発表した研究開発の展開容易性の曲線のこと、基礎研究と実用化の中間段階で投資が不足する状態を示しています。つまり、この両者の中間にプロジェクトが挫折するボトルネックがあり、俗に「死の谷」と呼ばれています。

注) ワンダーホルンの技術的特徴の詳細

通常のテキストを読上げる音声合成の多くでは、肉声の音声波形（1ピッチ程度）を重畳する「波形重畳方式」が利用されています。この方式は明瞭性も高く、人の話す音域（70Hz～300Hz程度）では自然性も高い高品質な音声合成を可能としますが、歌声のように音域が広くなると（100Hz～700Hz程度）、音声の高さ（ピッチ）の制御量が大きくなるため品質が劣化してしまう傾向があります。

また、歌声の、話す声に対する特徴として「倍音成分（Harmonic Overtones）」が非常に重要ですが、「波形重畳方式」は音声の局所的な波形特徴に重点を置いているため倍音をうまくモデル化できません。そこでワンダーホルンでは、歌声に特徴的な倍音を高品質で再現するため、純音である正弦波を多重に重畳する「正弦波重畳方式」を採用しています。

正弦波重畳方式だけで合成すると「楽器（トランペット）のような」不自然な合成音声になってしまいます。そこで、人間の肉声に含まれるノイズ成分を付加することによって単純な正弦波重畳方式と比べ飛躍的に肉声らしい歌声を実現しています。

- ・ [図1：歌声合成技術“ワンダーホルン”の原理](#)
- ・ [図2：自分だけの歌声メッセージの作り方](#)

＜本リリースに関する報道機関からの問い合わせ先＞

日本電信電話株式会社
（第一部門広報担当）

大道、奥泉
03-5205-5550

＜本リリースに関する報道機関以外からの問い合わせ先＞

日本電信電話株式会社
（第三部門プロデュース担当）

阪本、朝日
03-5205-5391

＜ライセンスに関する問い合わせ先＞

NTTアドバンステクノロジー株式会社
（メディア技術センタワンダーホルン担当）

加藤、板倉
0422-36-9303

NTT ニュースリリース 