

大輪 勤
OWA Tsutomu

後藤 哲也
GOTOU Tetsuya

川端 基夫
KAWABATA Motoo

当社は、「駅前探険倶楽部」をはじめとするWeb上のサービスをパソコン・携帯電話向けに提供してきた。その経験を基に、従来のテキスト情報に加え、複数キャリア・端末向けに画像の配信を行う画像配信システムを開発した。これは、各端末の表示能力に応じたフォーマット、サイズ、色数の画像を、配信システム側で自動的に判別して配信するものである。また、音声合成技術と駅前探険倶楽部で用いられている地図エンジン、時刻表案内サービスとを組み合わせ、最寄駅の時刻表を読み上げる音声時刻表案内システムを開発した。更に、音声合成技術を基盤としたコンテンツビジネス展開の第一歩として、CG(Computer Graphics)タレントが合成音声とリップシンクしながらテキスト読上げを行うバーチャルアナウンサーシステムを開発した。放送業界、インターネットや携帯電話、ゲームなど、様々な分野での事業展開を図っている。

Toshiba has been providing various services over the Internet to PC and mobile Web-enabled phone users including the "Ekimae-Tanken Club" Internet-based information and guidance service (ekitan.com). Toshiba has now developed an image delivery system based on our long experience with ekitan.com. The system distributes still pictures as well as text information to multiple carriers and multiple devices. It adapts original pictures to each device according to its display capabilities such as screen size, color depth, and acceptable image file format.

Toshiba has also developed a voice-based railway timetable guidance system, which reads out the timetable for the nearest station using Toshiba's text-to-speech(TTS) technology, a map engine, and the railway timetable service. Moreover, as the first step in contents business deployment based on TTS technology, Toshiba has produced a virtual announcer system that performs lip-synched animation. Toshiba is aiming at business deployment in various fields such as the digital broadcasting industry, the Internet, the cellular phone industry, and the game industry.

1 まえがき

携帯電話、パソコンからのインターネット利用者が急増している。同時に、通信速度の向上も著しい。また、位置情報が取得できる携帯電話の普及も進んでいる。この環境の変化に伴い、新しいサービスが次々と提案されている。

当社は、駅前探険倶楽部をはじめとするインターネットサービス・コンテンツビジネスに早くから注力してきた。ここでは、当社が事業化を目指す三つの新しいマルチメディアサービス・コンテンツビジネスについて紹介する。

2 画像配信システム

当社の持つ技術の一つにMMS(Mobile Multi-protocol Service)がある。これは、記述言語の異なるキャリア、そして表示能力の異なる携帯端末に対して、適切な形で情報を配信するものである。この技術の自然な拡張として、従来のテキスト情報に加え、複数キャリア・端末へ適切な形で画像を配信することが考えられる。

携帯端末にて表示可能な静止画画像フォーマットはキャリ

ア・端末ごとに異なる。ある端末はBMP(Bit MaP)画像、別の端末ではPNG(Portable Network Graphics format)画像が表示できるといった具合である。表示可能な画像サイズも端末ごとに異なる。このため、複数キャリア・端末向けのコンテンツの作成には、様々なフォーマット、サイズの画像を用意するとともに、それぞれへのリンクを用意する必要があり、煩雑な手作業が発生していた。

動画においても同様な状況が予想される。現時点では各キャリアの動画ファイルフォーマット、端末仕様は公開されていないが、他と異なる、特色あるサービスを提供したいキャリアの立場からは、同一仕様とはなりにくい。

このような問題を解決するために画像配信システムを開発した。システム概要を図1に示す。このシステムは、変換・登録サーバ、配信サーバ、端末情報データベース(DB)、画像DBに大別される。変換・登録サーバは入力された画像を端末ごとに適切なフォーマット、サイズに変換し画像DBに登録する。配信サーバは、適切な画像データを画像DBから検索し、端末へ配信する。端末情報DBは、端末ごとの表示可能画像フォーマット、画像サイズなどを保持しており、変換・登録サーバ及び配信サーバの両者から参照される。

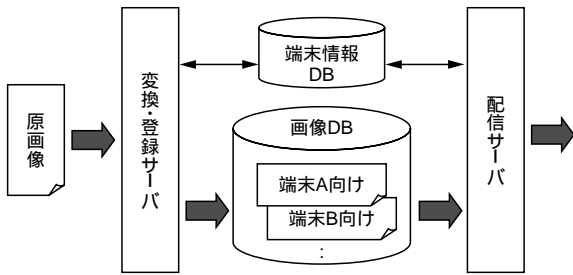


図1. 画像配信システム 原画像は端末別に適切なフォーマット, サイズに変換, 登録, 配信される。
Image delivery system

この画像配信システムの特長は, 端末情報をシステム側で管理し, 一つの原画像から異なる端末向けに適切な形式の画像を生成, 蓄積する点と, アクセスしてきた端末向けの画像を配信側にて適切に選択, 配信する点にある。

これらにより, 配信者はワンソース, マルチユースなサービスを容易に構築でき, ユーザーは自分が使っている端末の種類や表示能力を気にすることなく, このURL(Uniform Resource Locator)へアクセスすれば自分の端末にあった画像を得られるという利点がある。

もう一つの特長は, 異なる端末向けにデータを生成するモジュールを独立させた点にある。こうして拡張性を確保している。現在, 静止画間の変換モジュールの開発は終了し, 動画間の変換モジュールを開発中である。技術的には難しい面もあるが, 着信メロディ系データの変換を行うことで, 画像配信システムからマルチメディア配信システムへの発展が見込まれる。

2.1 変換・登録サーバ

変換・登録サーバはWebサーバ上のアプリケーションである。Web管理者は, 配信する原画像ファイル名, 画像番号などの表1に示す変換・登録のためのパラメータを指定し, ファイルのアップロードを行う。

表1. 変換・登録パラメータの例(抜粋)
Example of registration and conversion parameters

項目名	値
原画像ファイル名	/homes/foo/images/butterfly.jpeg
画像番号	00000000/00000001
著作権者	TOSHIBA
公開開始日時	2002/12/24 00:00:00
公開終了日時	2002/12/31 11:59:59
アスペクト比保存	

これらのパラメータ指定と端末情報に基づきフォーマット, サイズ変換がなされる。変換後のファイルをブラウザにて確認後, 問題がなければ配信サーバへ登録する。ここで, 画像

番号はWeb管理者が画像ごとに付加するユニークな番号であり, この画像にアクセスする際のURLの一部に使用される。

2.2 配信サーバ

一方, 配信サーバは, 前述のURLを含むHTTP(Hyper Text Transfer Protocol)GETリクエストを受けると, 端末固有の情報を含むUserAgentフィールドから端末種別を判別する。この情報を用いてフォーマット, サイズなどを端末情報DBから取得し, 適当な画像データを特定, 配信する。

配信サーバの性能は数百同時アクセス/1台を見込んでおり, 同時アクセス数の伸びに応じて, 負荷分散機の下に配置される配信サーバの数を増やすことで, スケーラビリティを確保することができる。

今後, MMSと画像配信システムとを組み合わせることで, マルチメディアサービスを当社にて提供するとともに, マルチメディアコンテンツを配信したいコンテンツプロバイダーへASP(Application Service Provider)展開を図る。

3 音声時刻表案内デモシステム

近年, 位置情報を取得可能な携帯電話があいついで発売されている。この位置情報と当社の駅前探険倶楽部で提供している時刻表案内とを組み合わせた“ 音声時刻表案内 ”のデモシステムを開発した。目の不自由な方の支援をコンセプトとして, 携帯電話から送付される位置情報から最寄駅を検索し, その駅の時刻表を音声で案内するものである。音声時刻表案内の例を図2に示す。

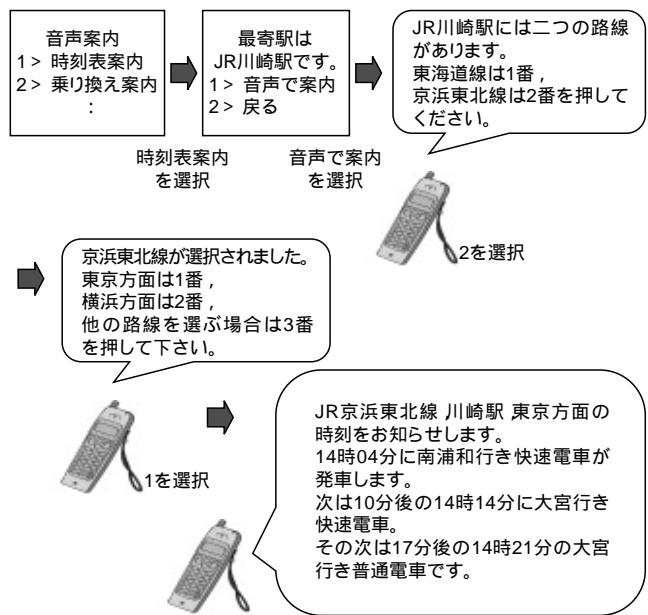


図2. 音声時刻表案内例 駅の時刻表を音声で案内する例を示す。
Example of voice-based railway timetable guidance

音声時刻表案内システムの概要を図3に示す。まず、ユーザーがこのデモサイトにアクセスすると位置情報を取得するための画面が音声時刻表案内サーバから携帯電話に送付される。ここで“時刻表案内”を選択すると、位置情報を含むHTTPリクエストが音声時刻表案内サーバに送付される。この位置情報を地図エンジンに渡し最寄駅を得る。次に、音声案内に用いる電話番号を含む画面を携帯電話に送付する。この画面上で“音声で案内”を選択すると、携帯電話からCTI(Computer Telephony Integration)エンジンに電話が掛かる。CTIでは、掛かってきた電話番号の取得、電話の数字キー押し下げの検知がなされる。音声時刻表案内サーバにはこの電話番号と最寄駅との組を保存してあるため、CTIが取得した電話番号から最寄駅を判別でき、この最寄駅に発着する路線、時刻表を時刻表エンジンに問い合わせることができる。路線、上り、下りの選択は数字キーを通じて行う。最後に、時刻表を自然な日本語とした後、音声合成エンジンにてテキストから音声データに変換し、携帯電話に対し読上げを行う。

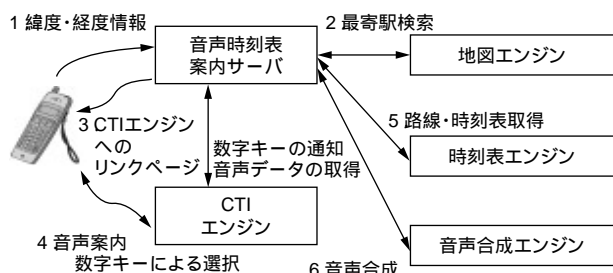


図3. 音声時刻表案内システム 音声時刻表案内サーバが他のエンジンをコントロールして時刻表読上げを行う。
Voice-based railway timetable guidance system

音声合成には、当社の日本語音声認識 / 合成ソフトウェアであるLaLaVoice™2001⁽¹⁾を用いた。また、時刻表を自然な日本語に変換する部分にも、当社の自動案内テキスト生成技術⁽²⁾を用いている。

パケット課金であるブラウザ上でのサービスに比べ、音声通話での読上げサービスは通話料が負担となるという課題もあるが、携帯電話に表示される文字は小さく読みにくいので、音声で案内が聞けるのは便利との声もある。当社にて提供中の道案内サービスでの道案内文の読上げ、あるいはWebサイトの読上げサービスにも応用を図り、従来のテキストを“目で見える方法”に加え、“耳で聞く”というもう一つの方法をユーザーに提供していきたい。

4 音声合成技術を基盤としたコンテンツビジネス展開

当社が1998年に開発した音声合成技術 TOTally Speaker

Driven Text To Speech : TOS Drive TTS⁽³⁾は現在、当社パソコンDynaBookにプリインストールされている“東芝音声システム”やパッケージソフトウェアのLaLaVoice™2001などのテキスト読上げソフトウェアとして製品化されている。また、電話機や最近ではカーナビゲーションの音声出力のためのミドルウェアとして、東芝の音声合成が広く使われるようになってきた。

一方、各メディア及びコンテンツ業界では、デジタル化対応並びにブロードバンド化対応が進むに伴い、有力なコンテンツが不足していくことが懸念されている。特に、放送業界では、既存の地上波キー局に加え、BS(放送衛星)放送、CS(通信衛星)放送、ケーブルテレビ、デジタル放送など、メディアの多角化が進んでおり、上記懸念が現実となる可能性も大きい。

当社では、より一般ユーザーに触れると考えられるエンタテインメント系メディア及びコンテンツ分野での音声合成の利用が、表2に示すように多岐に展開できると考え、以下を重点的に研究してきた。

- (1) 感情表現、口語表現の実現への基礎研究
- (2) 音声合成辞書の作成期間短縮
- (3) 有名人の声をを用いた模倣精度の検討

特に、本格的なビジネス展開をする際、迅速な顧客対応をするために重要な(2)項に関しては、従来の1/4の時間で音声合成辞書作成を実現可能とした。

表2. エンタテインメント系コンテンツ分野での音声合成の利用分野
Application of speech synthesis

分野	想定利用例	メリット	課題
映画	音声吹替え	MA費用削減	感情表現、口語表現の実現
ゲームソフト	RPGテキスト読上げ	付加価値の提供	
テレビ番組	アニメキャラクター音声 CM音声 バーチャルタレント音声	MA費用削減 即時制作対応	
玩具(がんく)	ロボット・人形音声	制作費の低減	
ラジオ	DJ音声		
音楽	バーチャル歌手		

RPG : ロールプレイングゲーム
MA : ミキシング&オーディオ
DJ : ディスクジョッキー

また、感情表現や口語表現が自在となり、声に様々な表情が付加可能となれば、音声合成技術により作り出された個々の声は個性を持つことができる。したがって、表2に示す各分野に音声合成技術を随時供給していくビジネススキームとは別に、俳優や声優が、映画、テレビ番組、CM (CoMmercial) など様々なコンテンツに出演するように、音声合成技術で創出された個性ある声も一つの声で様々なコンテンツに“出演”でき、いわば“音声合成技術による声のプロデュース業”が可能となると考える。

現状、上記ビジネススキームを展開するためには、現技術レベルの音声（文語表現重視の音声）に個性を持たせることが必要である。そこで、文語表現に合うコンテンツとして、ニュース、天気予報、解説、非人間キャラクター分野を想定、更に視覚的効果の付加をねらいCGキャラクターと組み合わせることにした。そこで誕生したのが、バーチャルアナウンサー“川本愛”である（http://www.toshiba.co.jp/about/press/2001_06/pr_j0501.htm）

CGキャラクター及びモーションプログラム部分は、CGタレント“飛飛（フェイ フェイ）”などで実績のあるイーフロンティア社（<http://www.e-frontier.co.jp/>）と提携し供給された。当社からは、CGキャラクターの雰囲気にあった音声を制作し、音声ファイルとともに、口の動きを指定するリップシンクデータを供給することにより、リップシンクも可能にした（図4）。

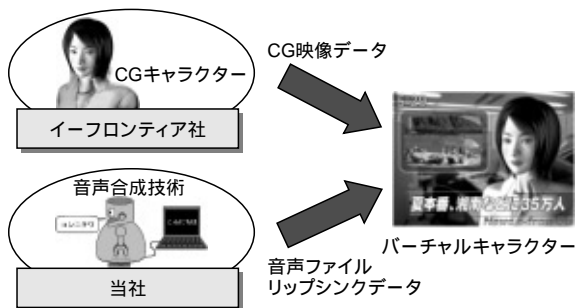


図4．バーチャルアナウンサー“川本愛” CGキャラクターの雰囲気にあった音声を制作，リップシンクを可能にした。
Virtual announcer, Kawamoto Ai

バーチャルアナウンサー“川本愛”は、ニュースや天気予報などのテキスト原稿さえあれば、テキスト原稿を自動認識し読み上げていくことができ、高い即時性と制作費削減効果を持つ。また、リップシンクの精度、表情や動作バリエーションを広げるなど課題は多いが、改良を重ね、放送業界や番組制作会社向けの有力コンテンツとして提供していくことで、メディアの多角化によるマーケットニーズにこたえていきたいと考えている。また、非人間CGキャラクターとの組合せにより、新たなニーズの模索も実施したいと考えている。

5 あとがき

当社の開発した画像配信システム、音声時刻表案内システム、バーチャルアナウンサーシステムについて述べた。

画像配信システムは、画像を携帯端末向けに適切なフォ

ーマット、サイズで配信するBtoC（Business to Consumer）サービスのほかに、ユーザー間での画像交換といったCtoC（Consumer to Consumer）サービスへの応用も考えられる。これにより、ユーザーは自分の持つ端末の種類や相手の端末の種類を気にすることなく画像をやり取りすることができるようになり、利便性の向上が期待できる。

音声時刻表案内は、近年、携帯電話から取得できるようになった位置情報の応用の一つであり、当社道案内サービスにも適用が期待できる。従来の案内文を“目で見える”方法に加え、“耳で聞く”というもう一つの方法をユーザーは選択することができるようになった。

バーチャルアナウンサーシステムは、CGタレントが合成音声とリップシンクしながらテキスト読上げを行うものである。CGキャラクターと組み合わせているのは個性を持たせるためである。これにより、テキスト原稿を自動認識し読み上げていくことで、高い即時性と制作費削減効果を持つ。今後は、音声合成技術の向上及びバーチャルキャラクターの改良はもちろんのこと、放送業界のほか、インターネットや携帯電話、ゲームなど様々な分野での事業展開を図るとともに、他国語での展開も検討する。

文献

- (1) www2.toshiba.co.jp/pc/lalavoice/product.htm
- (2) 長谷川 保,ほか：“歩行者向け対話型道案内サービスのための案内テキスト生成”. 電子情報通信学会技術研究報告101,227,HIP2001-13,2001,p.43-48.
- (3) 籠嶋岳彦.高品質音声合成(口調・声質を学習する音声合成システムTOS Drive TTS). 東芝レビュー.56,5,2001,p.70-71.



大輪 勤 OWA Tsutomu
iバリュー クリエーション社 技術部主務。
Webアプリケーションの開発に従事。
iValue Creation Co.



後藤 哲也 GOTOU Tetsuya
iバリュー クリエーション社 技術部。
Webアプリケーションの開発に従事。
iValue Creation Co.



川端 基夫 KAWABATA Motoo
e-ソリューション社 メディア&コンテンツ事業部 コンテンツ事業開発部主務。音声合成関連ビジネスの開発に従事。
e-Solutions Co.