

NTT持株会社ニュースリリース

(ニュースリリース)

2011年4月7日

双方向通信向け高臨場感音声符号化技術「G.711.1 Annex D」を国際標準化 ～テレプレゼンス等の臨場感をさらに高める音声符号化技術～

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:三浦惺)は、ETRI(韓国)、France Telecom(フランス)、華為技術(中国)、VoiceAge(カナダ)と共同で、テレビやFMラジオなどの放送と同程度である14kHz帯域の音声
を符号化※1できる方式をITU-T※2に提案し、G.711.1 Annex Dとして国際標準化されました。

本方式をテレプレゼンスやハイビジョン会議装置等に搭載することで、双方向での臨場感にあふれた音声通信が可能になります。さらに、本方式は、電話に広く用いられている音声符号化技術であるG.711※3やG.711.1※4とも相互接続できることから、幅広い用途での利用が期待できます。

1. 背景

NTTのサイバースペース研究所(以下、「NTTの研究所」)では、従来より音声符号化技術の国際標準化に取り組み、1996年に国際標準として採択されたVoIP電話向けのG.729※5を始め、NGN向けのG.711.1などを成立させてきました。

従来のデジタル電話やVoIP電話では、電話帯域音声(300Hz～3.4kHz)を符号化するのに、音声コーデック※6のITU-T標準G.711が広く用いられてきましたが、電話帯域音声は、会話を行うには十分なものの、符号化できる帯域が狭いため、人間の声を忠実に再現することは困難でした。

また、MPEG AAC※7など、放送やVODなどのコンテンツ向けに、より広い帯域の音声を再生可能なコーデック技術も標準化されていますが、テレビ会議などの双方向通信に用いるには伝送遅延が大きいため、これからますます普及が見込まれる高臨場感通信サービスに適用可能な、帯域が広く、リアルタイム性の高い音声符号化技術が求められていました。

2. 今回の成果

NTTの研究所では、一般の電話で使われる音声帯域の4倍である50Hz～14kHzの帯域の符号化を可能にし、高音質な電話サービスへの適用が期待される、音声符号化の新方式G.711.1 Annex Dを標準化しました。

3. 14kHz帯域音声符号化方式G.711.1 Annex Dの特徴

(1) 14kHz帯域音声を再生可能 [別紙1](#)

従来の電話音声(300Hz～3.4kHz)や7kHz広帯域音声(50Hz～7kHz)よりもはるかに広い帯域の音声により、人の声のみならず、環境音や音楽まで忠実に再生することができ、相手に話し手の雰囲気まで伝えることが可能になります。

(2) 双方向通信に適した少ない音声遅延

携帯電話の符号化方式3GPP AMR※8や放送用途の方式MPEG AACでは、音声信号を20ms(ミリ秒)以上のブロック単位で処理するため、音声の遅れが避けられません。本方式は、G.711.1と同様に、最小5ms単位での音声処理が可能のため、音声の遅延を抑えられます。つまり、遅れの少ない自然な通話が可能です。

(3) 従来電話との相互接続が容易 [別紙2](#)

従来、複数の地点を接続するような会議では、全ての端末が同じ音声コーデックを搭載している必要がありました。従来電話が一地点でも混在した場合、全ての端末のコーデックをG.711に合わせるか、装置間で符号データを変換する装置を別途

設ける必要があります。本方式は、G.711、G.711.1で再生可能な符号データに、音声帯域を広げるための符号データ(16～32kbit/s)を追加して伝送するため、従来電話が混在しても、G.711の符号データを取り出すだけで符号変換が可能ですので、コーデックをG.711に合わせる必要はありません。

(4) 多地点音声ミキシングが容易

通信会議用途の場合、複数地点が同時に通話できるよう、それぞれの地点からの音声を足し合わせる多地点音声ミキシングの機能が欠かせません。従来の音声ミキシングでは、各地点からの符号データを音声に戻し、足し合わせて、それぞれの地点向けの音声を作った後、再び符号データに符号化するという膨大な処理が必要でした。本方式と部分選択ミキシング※9と組み合わせることで、それらの処理を大幅に減らすことができ、一桁以上少ない処理量でのミキシングが可能になります。

4. 今後の予定

今後、ブロードバンドネットワークの更なる普及にともない、一層の発展が期待されるテレプレゼンスやハイビジョン通信会議などにおいて、より臨場感の高いサービスを提供するため、本技術のサービスへの実装を進めます。

用語解説

※1 符号化

音・画像などの情報を、伝送・記録できるよう、データに変換すること。

※2 ITU-T(国際電気通信連合電気通信標準化部門)

International Telecommunication Union Standardization Sector. 国際連合の専門機関である国際電気通信連合(ITU)の部門の一つで、情報通信技術に関する国際標準を策定している。

※3 G.711

ITU-Tで最初に標準化された音声符号化方式。電話帯域の音声信号を64kbit/sで伝送できる。全世界で最も普及している方式。

※4 G.711.1

NTT研究所主導の下、ITU-Tで標準化された7kHz広帯域音声符号化方式。G.711の符号データ(64kbit/s)に、音声帯域を広げるための符号データ(16～32kbit/s)を追加して伝送することにより、G.711との互換性を保ちつつ、高音質を実現できる。フレッツ光ネクストの高音質電話サービスに利用。

※5 G.729

ITU-Tで標準化された音声符号化方式で、電話音声をG.711の1/8の8kbit/sで伝送できる。伝送帯域が限られたネットワークでのVoIPに用いられている。

※6 コーデック

音・画像などの情報をデータに変換する符号化、そのデータを情報に戻す復号の処理を併せて行える装置、ソフト、方式を指す。

※7 MPEG AAC

ISOで標準化された音声符号化方式。デジタル放送、コンテンツ配信・蓄積などに広く用いられている。

※8 3GPP AMR

3GPPで標準化された携帯電話向けの音声符号化方式。ドコモのFOMAなどW-CDMA対応の携帯電話で広く用いられている。

※9 部分選択ミキシング

従来の音声ミキシングは、複数地点から受信した符号データを音声信号に復号し、それぞれの地点向けにミックスした後、再度符号データに符号化する。それに対し、部分選択ミキシングは、符号データの主要な部分のみ、それらミキシングの処理を行い、それ以外の部分は、話者地点を選択した後、その地点のデータをコピーし、その他全ての地点向けのデータとして代用する手法。G.711.1やG.711.1 Annex Dの符号データの内、処理が非常に軽いG.711の符号データのみ復号・ミキシング・再符号化し、処理の重いデータは話者地点からコピーすることで、広い音声帯域を保ったまま、G.711と同程度の少ない処理量でのミキシングが可能。

別紙・参考資料

▶ [別紙1 14kHz帯域音声を再生可能](#)

- ▶ [別紙2 従来電話との相互接続が容易](#)
- ▶ [別紙3 標準化成立までの経緯](#)
- ▶ [別紙4 関連して成立したその他の標準](#)

本件に関するお問い合わせ先

- 日本電信電話株式会社
サイバーコミュニケーション総合研究所
(サイバースペース研究所)

企画部広報担当

Tel: 046-859-2032

E-mail: randd@lab.ntt.co.jp

ニュースリリースに記載している情報は、発表日時点のものです。現時点では、発表日時点での情報と異なる場合がありますので、あらかじめご了承くださいとともに、ご注意をお願いいたします。

[NTT持株会社ニュースリリース インデックスへ](#)

NTT持株会社 ニュースリリース

▶ [最新ニュースリリース](#)

▶ [バックナンバー](#)

▶ [English is Here](#)

NTT持株会社 ニュースリリース内検索

1997 ▼ 年 04 ▼

月 ~

2021 ▼ 年 11 ▼ 月

検索

NTTグループの情報は
こちらからもご覧いただけます。



