特別企画
広帯域音声符号化技術とNGN端末向け
プロトコル試験機の開発に取り組むNTT研究所

広帯域音声符号化技術「G.711.1」への取組み
～よりクリアな電話音声を実現～

NTT研究所は長年にわたり、音声符号化技術による音声・楽音通信への研究を進めてきた。最近では、広帯域音声符号化の技術を活かし、従来の電話音声「G.711」の発展系として、肉声に近い電話音声を実現した「G.711.1」への取組みがある。

NTTサイバースペース研究所
第一推進プロジェクト
主幹研究員
大室 仲氏

NTT研究所の音声符号化研究開発の歩み

NTT研究所は、音声符号化技術の研究を基礎に、最先端のネットワーク環境やアプリケーション環境の要請に応える形で、音声・楽音通信技術の研究開発を進めてきた（図1参照）。1990年代始め頃までの音声符号化技術は低速なネットワークでの活用が主体だったため、高い圧縮率が求められていた。やがて研究が進む中、低遅延・低演算量符号化技術や高品質音楽用ベクトル符号化技術などが開発され、圧縮率とともに音声の品質が求められるようになってきた。主な方式として、CS-ACELP、TwinVQなどがある。2000年以降に入ると、圧縮率よりは品質の高度化および利便性に重点が置かれるようになってきた。この背景には、ブロードバンド回線などネットワークの広帯域化が進み、大容量の音声データを送受信することが可能になったことが大きい。

NGN時代に向けて、NTT研究所が現在、最も注力しているのが、G.711の発展系である「G.711.1」である。

NTTが独自開発した技術で高品質・広帯域な音声通話を実現

G.711.1は、NTT研究所が独自に開発した広帯域音声符号化技術「UEMCLIP（ユーエムクリップ：mU-law EMbedded Coder for Low-delay IP communication）」をベースとしている。UEMCLIPは、ITU-T G.711として広く利用されている電話音声符号化の標準方式を拡張し、高品質な広帯域音声での通信を可能にする符号化技術である。VoIP等の音声通信アプリケーションや、多地点での音声会議用途を目的として、NTT研究所で開発された。

UEMCLIPの技術を基に、ETRI（韓国）、FranceTelecom（フランス）、Huawei（中国）、VoiceAge（カナダ）との共同提案により、ITU-Tで標準化されたのがG.711.1である。

NTTサイバースペース研究所第一推進プロジェクト　主幹研究員の大室 仲氏は、G.711.1について次のように語っている。

「現在の公衆電話回線の音声であるG.711は、肉声とは違って聞こえるのが現状です。何を言っているのかはわかりますが、耳には『電話を通した声』として伝わってきます。それを補うのが広帯域化の技術にな

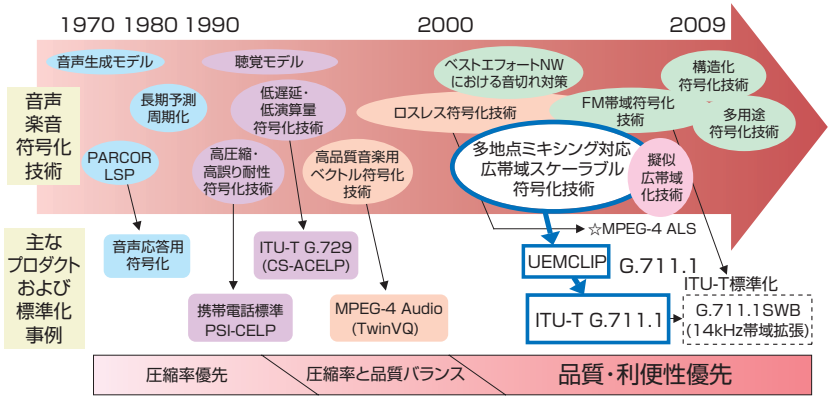

図1　NTTにおける音声符号化研究開発の歩み

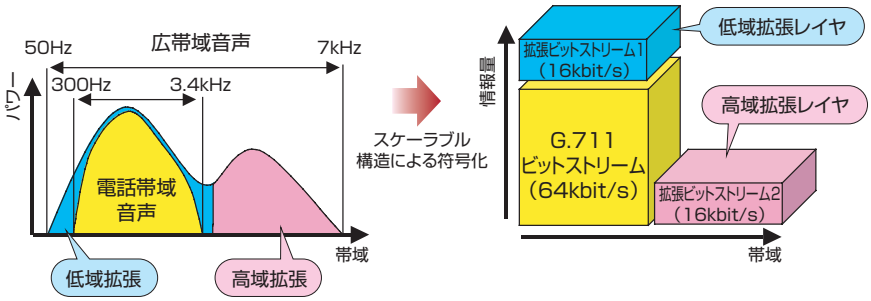

図２　G.711.1の音声広帯域化のイメージ

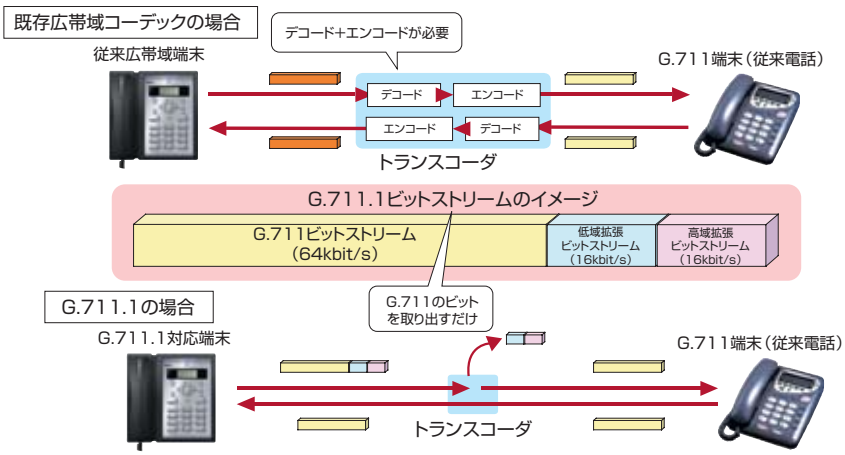

図３　G.711.1のトランスコーディングのイメージ

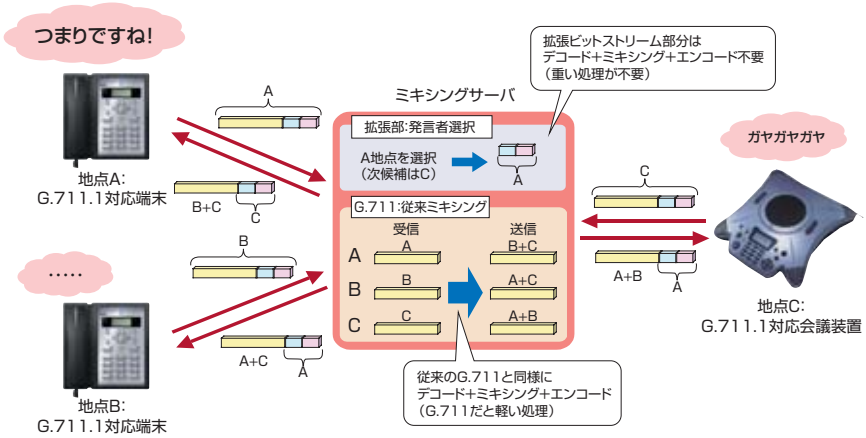

図４　G.711.1の低負荷多地点ミキシングのイメージ

ります。広帯域化を実現したG.711.1は、低域部分（50Hz～300Hz）も高域部分（3.4kHz～7kHz）も送れるようにすることで、肉声の音に近い鮮明な音声を電話で実現します（図２参照）。広帯域で通話できるようになれば、傍にいるかのようなリアルな肉声を届けることができ、聞き違いや電話相手の間違いを減らすこともできます。」

G.711.1の主な特長は次の通りである。

①７kHzの広帯域音声

電話帯域の倍の7kHzまで再生できる。肉声の音に近い、臨場感のある電話音声を実現できる。

②G.711とスケーラブルな構成

従来電話との接続をシームレスに行う処理である「トランスコーディング」が容易に行える。

G.711.1は、G.711のビットストリームに低域拡張ビットストリームおよび高域拡張ビットストリームを連結した構成になっており、G.711とスケーラブルな構成になっていることから、トランスコーダ上でG.711のビットを取り出すだけでトランスコーディングを実施する（図3参照）。このような技術により、ひかり電話とアナログ電話が相互接続する環境にも適しており、高音質IP電話機とアナログ電話機が混在する企業内電話システムの構築も容易となる。

③G.711と同等のミキシング負荷多地点通話を低コストで行う

広帯域音声と従来の音声が入り交じる環境で、従来の通話方式であるG.711と同等のミキシングを実現し、サーバにかかる負荷を大幅に下げている。従って、低コストによる多地点接続会議が可能となる。

具体的には、ミキシングサーバ内で、G.711ビットストリーム部分は従来のG.711と同様にデコード・ミキシング・エンコードを行い、連結している拡張ビットストリームは発言者を選択して各拠点に送信するという方式である（図４参照）。

以上の３つの特長の他にも、既存の広帯域音声符号化技術に比較して

特別企画

広帯域音声符号化技術とNGN端末向け
プロトコル試験機の開発に取り組むNTT研究所

低演算量や低遅延を実現し、パケット消失補償の高度化を行うといった特長がある。

従来の電話との通話でも高品質音声を実現した「擬似広帯域化」

G.711.1の普及のためには、専用の電話機間だけでなく、従来の電話機と通話しても高品質な音声であることが求められている。そこでNTT研究所は、受信した従来の電話音声からG.711.1の最高レベルである7kHzに近い、6kHz程度の広帯域までを推定して拡張する擬似広帯域化を実現した（図5参照）。この技術により、高音質電話機を持っていれば、携帯電話を含めてすべての相手との通話で、高音質を享受できるようになった。

前出の大室 仲氏は、擬似広帯域化について次のように語っている。

「NTT研究所では、音声の高度化の取組みをスタートしてから、従来の電話の音声通話と高音質の音声通話とを、同等に扱えるようにすることを考えていました。そのため、G.711との相互接続性を担保しながら十分な品質を確保することが求められます。擬似広帯域化の機能がG.711.1普及の後押しとなれば、G.711.1対応の電話端末の導入もスムーズに行われると考えています。」

G.711.1に対応した新しい電話端末を続々と提供

NTTグループは、G.711.1に対応した端末を続々と発売していく予定である。

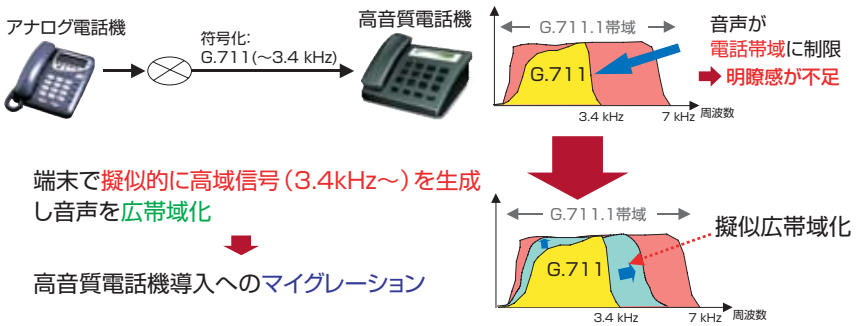

図5　擬似広帯域化

2008年3月にはIP電話会議装置「MEETING BOX」がフレッツ光ネクストの高音質電話に対応した。本製品は7kHz広帯域音声会議装置であり、4地点までの多地点通話が可能である。同年10月には、ビジネスニーズに特化し、7kHzを実現した「ビジネスホンαNXシリーズ」を発売した。2009年2月には、パソコンからひかり電話を利用できる「ひかりソフトフォン」の提供を開始した。ひかり電話加入者向けの通話料のみで利用できるソフトフォンであり、映像通信も可能である。

また2009年2月に発売した「ひかりクリアフォンHQ-100」は、個人・小規模企業向けの広帯域音声単体電話機であり、世界初のG.711.1コーデック搭載の商用端末である。なお、この端末は前述の擬似広帯域音声機能を利用できるので、従来電話との通話でも高品質音声の再生が可能である。

今後、NTT東日本・NTT西日本が中心になり、NTTブランド電話機やメーカーブランド電話などのNGN&G.711.1対応を推進していく予定である。

広帯域音声符号化技術のさらなる高度化に向けて

NTT研究所では今後もG.711.1について、さらなる音声の高品質化に向けた取組みを行っていく。具体的には、G.711.1SWB（スーパーワイドバンド）やG.711.1FB（フルバンド）の実現がある。

前出の大室 仲氏は、広帯域音声符号化の今後の展開について次のように語っている。

「NTT研究所では今後も試行錯誤を繰り返し、高品質と利便性を両立できるような音声符号化や通信の仕組みを考えていきます。我々は、100年続いた音声通話をコミュニケーションの基本として継続していく責任があり、そこで求められているのは、NGN時代に向けて高音質をスタンダードにする取組みです。アナログ電話がNGNひかり電話に切り替えられていくのを機に、さらなる高度な広帯域音声技術に取り組んでいきます。」

お問い合わせ先
NTTサイバーコミュニケーション総合研究所
企画部　広報担当
E-mail：randd@lab.ntt.co.jp