

# 100年以上未来の人たちに使ってもらえる 歪みの生じないロスレス符号化

携帯電話やVoIP、音楽配信に不可欠な音声・音響信号の圧縮符号化技術。100年以上未来のひとたちにも使ってもらえる歪みの生じないロスレス符号化技術が近く国際標準化される予定だ。この標準化に深く関わっている音声・音響圧縮符号化の世界的権威である守谷健弘NTT R&Dフェロー・NTTコミュニケーション科学基礎研究所人間情報研究部長に、お話をうかがった。

## 電気通信研究所は在学中からの憧れ

守谷フェローは、学生時代から当時の電電公社電気通信研究所に憧れていたということですが...

**守谷** 私は在学中から電気通信研究所に対して、おおいなる憧れを抱いていました。磁束量子を用いた論理回路を卒論のテーマにしましたが、その際に電気通信研究所の方々が書かれた多くの論文を参考にさせ

ていただき、とにかく“すごい”という印象を持ちました。さらに修士課程に進んで生体信号処理を勉強するうちに、音声音響信号の分野で世界の一流の研究者が何人もいるということがわかり、電気通信研究所への就職を希望しました。すでに推薦枠で一人決まっていたのですが、その人が心変わりしたために、就職することができました。電電公社は、音声を伝えることを生業にしており、音声は情報をやりとりする一番根元

となるものです。しかも音声は、認識・合成・伝達といろいろあり、学問の領域としても面白く、非常にやりがいのある魅力的な研究分野だと思っていました。入所してすぐの訓練期間中に、現在のNTTコミュニケーション科学基礎研究所（CS研）の管村昇所長に指導していただき、基礎研究部第四研究室には、板倉文忠・現名城大学教授や嵯峨山茂樹・現東京大学教授、東倉洋一・現国立情報学研究所教授をはじめ、錚々たる方々がおられることがわかり、第四研究室への配属を希望しましたが、競争率は非常に高かったようですが、運良く希望通り配属されました。毎日が勉強でしたが、第四研究室の研究活動のいずれもが世界的なレベルで戦っているということを実感しました。

## 胸ポケットに入る電話を作りたい

第四研究室に入られ、どのような基礎研究をされたのですか。

**守谷** 当時の高原靖研究開発本部長が行った新入社員に向けた研究開発テーマの紹介の中に、「ワイシャツの胸ポケットに入る電話を10年以内に作りたい」というのがありま



人間情報研究部長  
守谷 健弘 氏  
NTT R&Dフェロー  
NTTコミュニケーション科学基礎研究所  
もりや たけひろ

1980年電電公社電気通信研究所入社。音声・音響信号の高圧縮符号化技術の研究開発に従事。1995年電子情報通信学会（論文賞/業績賞/小林記念特別賞）、1998年科学技術庁（現文部科学省）長官表彰（注目発明）、2003年全国発明表彰特許庁長官賞を受賞、2003年IEEE Fellow、2004年NTT R&Dフェロー。現在、NTTコミュニケーション科学基礎研究所人間情報研究部長。工学博士（東京大学）、MPEG-4の拡張機能に関する標準化にも従事。

した。それを実現するためには、電池や回路、無線といった技術に加え、音声をデジタル化し圧縮して符号化する技術がキーになるということから、その基礎研究を私のテーマとしました。当時、音声を圧縮する技術では、板倉先生が世界のトップを走っていました。しかし、5年、6年と経つにつれ世の中の大勢としては、情報を圧縮する意義がうすれ、携帯電話のデジタル化の実現可能性も遠のいてきました。私自身、8年間の基礎研究をベースにしたベクトル量子化に基づく中帯域音声符号化の研究で学位を取得しましたが、違う研究を行うことを考えていました。ちょうどその時、ベル研究所に行けることになり、情報の圧縮・符号化と違う分野の研究を行いたいと思っていました。ベル研に行ってみると音声符号化の研究は非常に活発であり、数カ月経ったところで、米国とヨーロッパで、政府主導によるデジタル携帯電話の導入に向けた取組みが加速されていることがわかりました。そこで、私も改めてベル研で低遅延音声符号化の研究をやり始めました。ところが、その最中に日本でもデジタル携帯電話を行うことになり、海外メーカーも含めたコンテストを実施し、一番良いものを採用することになりました。当時、NTTには音声符号化の研究者がきわめて手薄になっていたため、急遽日本に帰り3カ月で作ってコンテストに応募しましたが、10社以上の応募中、惜しくも次点でした。

**1位は、どこでしたか。**

**守谷** 米国でも採用されたモトローラ社の方式が、一番音声品質が良かったので、これを基に日本で最初のデジタル携帯電話サービスが開始されました。デジタル方式（PDC）は、広範囲で雑音が少なく、秘話性などの利点がありますが、当初のシステム（フルレート）では情報量が比較的多く、アナログ方式と比べ電波利用効率の点ではメリットが少ない状態でした。そこで3年後に改めてハーフレート（情報量半分）という条件でコンテストを行うことになりました。その3年間は当時無線研（現在NTTドコモ）の研究者も含めた若手研究者たちとプロジェクト体制で研究開発を進めました。その成果が「PSI-CELP」（Pitch Synchronous Innovation-Code Excited Linear Prediction）です。情報量が半分でありながら、最初スタートした際のフルレートと同等以上の音声品質を達成し、標準方式として採用されました。これが現在のMovaに用いられていますし、基本技術はFOMAへも継承されています。

**ハーフレート方式とは、どのようなものですか。**

**守谷** 音声のデジタル符号化は、入力信号である音声を10ms、20msといった短い時間で区切り、その間の音声の特徴を統計分析してパラメータやテーブル番号といったデジタル情報に置き換えて送信信号とし、受けた側で個別のパラメータ等から元の音声に近いように再構成するという仕組みです。これを携帯電話に

実装できるようにするために、信号処理にいろいろな工夫を凝らしたアルゴリズムを採用して、品質を改善しつつ、処理量を極力減らすようにしています。最初のフルレート方式は64kbit/sの音声データを11.2kbit/s（音声6.7kbit/s）に圧縮符号化していますが、ハーフレート方式はさらにその半分の5.6kbit/s（音声3.45kbit/s）にまで圧縮しています。

**MSが採用、国際標準規格にもなったTwinVQ**

**守谷フェローが開発されTwinVQ（Transform-domain Weighted Interleave Quantization）は、ストリームサービスを支える技術と聞いていますが、どのようなところに特徴がありますか。**

**守谷** TwinVQは、音楽データを原音の約1/10から1/100まで圧縮できる技術で、圧縮されたビット列の一部からでも音楽を再生できる機能（スケーラブル符号化）への対応、高圧縮時の優れた音質、伝送路上で発生する符号誤りに対する耐性など、数多くの優れた特徴を備えた音楽圧縮技術です。マイクロソフト社には、Windows Mediaをリリースする前に2年間採用していただいたほか、動画及び音楽の国際標準規格であるMPEG-4オーディオに正式採用されました。前述のハーフレート方式やITU-T標準のG.729など、一貫して音声の情報圧縮技術の研究開発に従事してきましたが、この技術で音楽を表現するのに使ったらどうだろうというのが、TwinVQ開発

のきっかけです。オーディオと音声は別の研究分野として、相互交流もほとんどありませんでした。聴覚を専門とする新入社員だった岩上直樹君を中心にして音声の符号化技術を活用したオーディオ符号化を研究しました。その結果、MPEG (MP3) やMDに匹敵する音質で、しかも20kbit/s、16kbit/s、8kbit/sというそれまでのオーディオの領域では考えられなかった低ビットレート(高圧縮)でも、非常に劣化が少ないものが開発できたということです。オーディオ業界の常識にはあまりとらわれずに、自分たちの技術を元にオリジナルな工夫を凝らしたからこそ実現できたと思います。しかし、MPEG-4オーディオの国際標準規格にはなっていますが、現在は光ファイバーやADSLなどの広帯域サービスが中心で、オーディオ信号を8kbit/sに圧縮するという新しい需要がほとんどなくなりました。私自身、苦勞して国際標準規格にまでしたものの、世の中に十分な貢献ができていないという点で、非常に悔しい思いが強いですね。G.729は、携帯電話やVoIPで、世界中の方々に使っていただいています。TwinVQは一時は沢山使っていただきましたが、お客さまへのアピールの方法なども含めもう少し上手くできなかったのかなと反省しています。

**MSは、TwinVQを使って、どのようなことをやったのですか。**

**守谷** 当時は、ISDN (64kbit/s) しかありませんでしたから、そこに音楽とビデオを流すストリーミング

サービスです。実は、MS以外のいろんなところからも声をかけていただき、使っていただきましたが、最大手のリアルネットワークスには契約上の問題から使われませんでした。MS社も自分たちでいろいろ開発していましたが、オーディオを圧縮する部分だけは、NTTが開発したものを採用せざるを得なかったということは、随分悔しかったと思います。ですから2年後に、圧縮処理の軽い独自の方式を開発し、現在のWindows Mediaとしてリリースしました。

### 歪みのないロスレス符号化

**音声放送などストリーミングサービスがインターネット上で容易に楽しめるようになりましたが、現在、音声・音響信号の圧縮符号化については、どのようなことに取り組まれていますか。**

**守谷** 歪みのまったくないロスレス符号化に取り組んでいます。前述のTwinVQやインターネットで普及しているMP3は、圧縮率が高いものの、符号化・復号化の処理過程で歪みが生じたり音質が劣化します。これに対して、ロスレス符号化は、原音を完全に再現でき、しかも圧縮ができるという符号化方式です。

**圧縮率はどのくらいですか。**

**守谷** TwinVQの場合は、CD1枚の情報を1/30とか1/50くらいに圧縮できますが、ロスレスの場合は1ビットも変わらずに再構成できる代わりに、1/2とか1/3くらいの圧縮率です。

このロスレス符号化については、ブロードバンド化やユーザーの高品質志向に対応すべく、大学やNTTも含む各企業でいろいろな技術をバラバラに開発していました。このままでは利用者には使ってもらえないと考え、私は世界の専門家に国際標準化の必要性を呼びかけました。最初は、標準化の手間や意義、過去の標準化との整合などが議論になりましたが、音楽業界も含めて次第に多くの賛同が得られるようになりました。そして2002年にISO/IECのMPEGオーディオグループがロスレス符号化の国際標準化作業を開始しました。コンテストも終了し、複数の技術を組み合わせる最高性能を達成する形で2005年秋頃には国際標準化が完了する予定です。もちろんNTTも複数の要素技術を提案しています。

**ロスレス符号化の用途として、どのようなものがあげられますか。**

**守谷** 高品質オーディオ信号の蓄積、ライブ演奏などの高音質のストリーミングなどが想定されています。また、音楽だけではなく、地震の波形や生体信号など各種の多チャンネルの信号が歪みなく圧縮できることから、幅広い分野での応用が期待できます。国際標準化により世界中の人が安心して利用できますし、100年以上未来の人たちにも使ってもらうことが可能になります。例えばレコード業界の方が、音楽データを永久保存したいという場合、仮にフリーソフトを利用して圧縮しておくと、10年経っても解凍できると

いう保証は何もないわけです。国際標準規格として仕様を決め、ソースコードも管理しておくことにより、100年経っても解凍することが可能です。また、特許など知財の権利関係も明確化されていますので、安心です。このロスレス符号化は、ある種実用的な研究ですが、基礎研究の立場でいうと、非常に普遍性があり、時代を越え国境を越えて利用できることを狙ったものといえます。

### 国際標準を先導することの重要性

**標準化や知的財産の重要性についてお聞かせください。**

**守谷** 市場を創造するという点で、国際標準を先導することは非常に重要です。国際標準化のメリットとしては、国際的な相互接続性・継続性が担保されるとか、前述したように知的保有者が明確化される、ソフトウェアの普及やハードウェアの低価格が促進されるといった点があげられます。特に、大規模サービスを展開する企業にとっては、サブマリン特許などを回避できることは非常に重要です。国際標準化された技術については、特許保有者が集まった任意の団体が技術提供/ライセンス供与を行うのが一般的です。例えば、MPEG LAという任意団体がありますが、この団体はMPEG-4技術の特許保有者で構成されるライセンス機関で、MPEG-4を使う際の料金条件を定め、MPEG-4のライセンスに関する一種の精算機関としての役割を果たしています。標準化と特

許というのはリンクしており、特に10年くらい前から、標準化技術についても、無償ではなくて権利を主張する傾向にあります。したがって、できるだけ良い技術を持ち寄って、標準化しようというようになってきました。ライセンス収入のみという会社もたくさんあります。基礎研究の分野でも、特に標準化ということになると、他社との罅迫り合いもあり、戦略的に特許を出すことが必須です。

### 35年前から世界に先駆けて音声信号の符号化研究を実施したNTT

**音声信号の符号化でNTTが果たした役割と、今後の取組みについてお聞かせください。**

**守谷** NTTは35年程前から、世界に先駆けて計算機で音声进行分析・合成する研究を行っています。PARCOR (PARTIAL auto-CORrelation) 方式は、信号処理一般にいろいろなところで使われており、ロスレス符号化でも使われています。

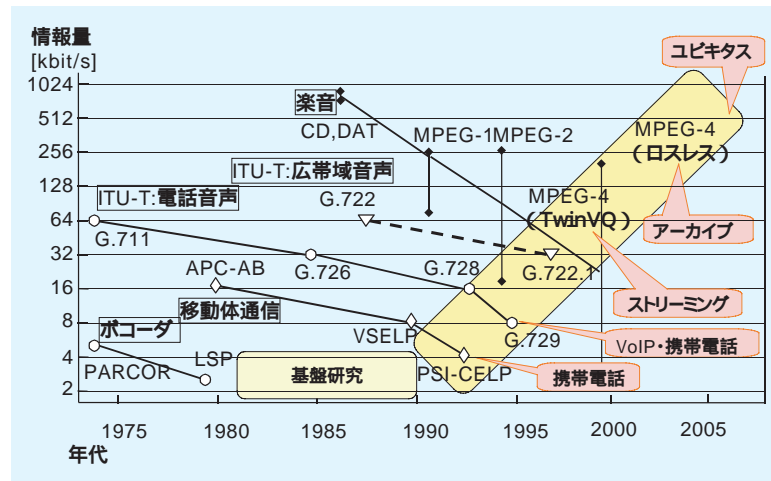


図1 音声・音響信号符号化の研究と標準化

その後のLSP (Line Spectrum Pair) 方式は、基礎技術として世界中の携帯電話に使われています。研究に着手したのも早いし、レベル的にも間違いなく世界のトップを走っていました。私自身は、そういう非常にアクティブな研究室に新米として入り、自然にいろいろ教えていただきました。音声信号の符号化は、自分の専門を決めてそれを延々やるというタイプの研究分野ではありません。標準化が完了すると、積極的に新たな異なる条件分野にでていって戦わないと、遅れをとってしまいます。視野を広げて、本当に世の中が必要としているものは何かを考え、かつ自分たちの実績・強みを踏まえたうえで、次に何を狙って研究するかが重要になります。先ほどのロスレス符号化も半年後には国際標準が決まる予定ですので、普及のための開発を続けつつ、次の戦いの場を考えているという状況です。

**本日は有り難うございました。**

(聞き手・構成：編集長 河西義人)