URL https://journal.ntt.co.jp/article/41199
DOI https://doi.org/10.60249/26111101

NTTコミュニケーション科学基礎研究所
フェロー

守谷健弘 Takehiro Moriya

声を圧縮し，世界をつなぐ——「知の泉」から生まれた符号化技術と国際標準

音声や音楽の信号をデジタル化し，携帯電話やアプリなどを通じて，高い品質を保ちながら効率良く届けるのに不可欠な「音声・音響符号化技術」．40年以上にわたり，この研究と国際標準化を牽引してきた守谷健弘フェローは，現在，立体的な音場まで伝える音声・音響符号化技術「IVAS」の標準化に取り組んでいます．研究という「知の泉」から，いかに技術を育て，世界で使われる「恵み」へつなげるのか．これまでの歩みと，次世代の研究者への思いを伺いました．

携帯電話の音声を支える基礎技術に携わって

守谷さんは長年，音声・音響符号化の研究に携わってこられました．どのような変化を感じておられますか．

まず，携帯電話の音質は非常に良くなりました．携帯電話では使える電波が限られていますから，同時に利用できる人数を増やすためには，極限まで情報量を圧縮する必要があります．方式や通信条件に応じて，従来のデジタル固定電話の数分の1，場合によっては10分の1程度の情報量で音声を伝送しなければなりませんでした．初期の携帯電話では，符号化技術も発展途上であり，通信網にも制約があったため，音声が途切れたり，遅れたり，誰が話しているのか分からないほど音質も悪かったのです．その後，さまざまな技術開発により，今では遅延をほとんど意識することなく，良い音質でスムーズに会話ができるようになりました．また，通信技術全体の進歩やサービスの普及に伴い，電話料金も大幅に下がりました．

携帯電話が普及したことによる変化もあります．その1つが交通事故による死亡者数が1990年代をピークに減少したことです．もちろん，その背景にはさまざまな要因がありますが，1995年から携帯電話の加入者数が一気に増えたことで，事故直後の救急への通報が容易になり，医療機関への搬送までの時間が短縮され，多くの命が救われたのではないかと私は推察しています．そうしたことから私は，運転中に限らず，寝室，トイレ，浴室にも常にスマホを持つことをお勧めしています（寝室では，枕元は安眠の妨げになるので足元をお勧めします）．もちろんスマホを楽しむ目的で使うのではなく，もし怪我や病気で倒れて動けなくなったときに救急通報の命綱として使うためです．このように自分の研究が間接的であれ人命に役立っているのは，研究を続けるうえで大きなモチベーションになっています．

さて，**図1**は音声・音響符号化と国際標準の50年の歴史を簡単に示したものです（ここでは電話音声用途に限定しています）．1975年ごろから基本技術が考案されてきましたが，それ以前には，音声・音響信号処理の礎を築いた「予測符号化」の故ビシュヌ・アタルさん（ベル研究所）と，「最尤推定」の板倉文忠さん（NTT）という偉大なお2人がいらっしゃいます．偶然にもお2人は，同時に別の立場から「全極型線形予測モデル」[*1]を提唱され，これがその後のLSP（Line Spectrum Pair：線スペクトル対）[*2]やCELP

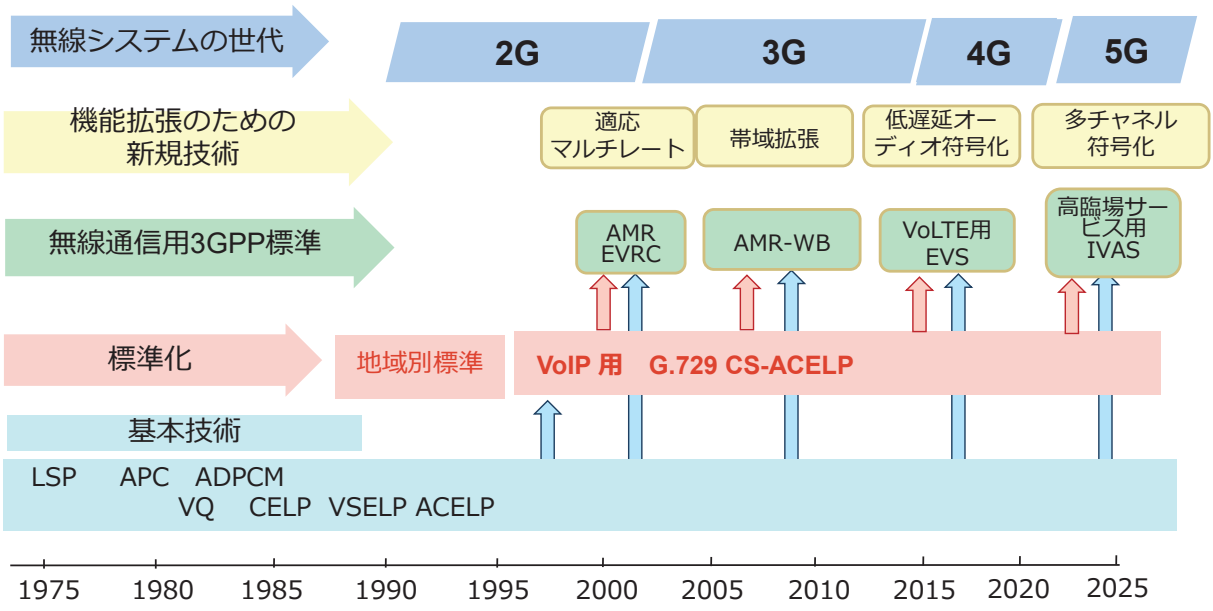

図 1　電話音声用の音声・音響符号化と国際標準の50年の歩み

(Code-Excited Linear Prediction：符号励振線形予測)[*3]といった基本技術へと発展していくことになりました.

しかしながら，私が入社した1980年当時は，「情報を圧縮して何になるのか」「アナログの携帯電話で十分ではないか」という意見もあったのです．周囲の研究者は次第に別テーマへ移っていき，一時は私 1 人で研究していたほどでした.

そうした中，1990年代になると世界で携帯電話のデジタル化の流れが一気に進み，NTTグループはPSI-CELPを開発するに至りました．この技術は，ARIB（電波産業会）主催の圧縮技術のコンテストで優勝し，1993年に第 2 世代の携帯電話のハーフレート国内標準化方式に採用されました.

ただし，限られた電波帯域で利用希望者の急速な拡大にこたえるために3.45 kbit/sという低レートにせざるを得なかったこともあり，品質は十分とはいえず，国内利用にとどまりました.

その後，国際統一規格制定の機運が高まり，NTTが欧米の有力機関と協調して開発したのがCS-ACELPです．この方式は 1995 年 に ITU-T（International Telecommunication Union-Telecommunication Standardization Sector）の国際標準G.729に制定されました．通常の音声符号化の 8 分の 1 の 8 kbit/sの情報量で公衆通信規格の品質を実現した最初の方式となり，インターネットを介して世界共通で使われるようになりました．また無線通信においても，私たちが貢献したCS-ACELPと同様の基本技術を使いながら，世代ごとに新たな付加価値を加えた符号化技術が開発・標準化されて，大きく進展していきました.

＊1 全極型線形予測モデル：音声のスペクトル包絡を，過去の音声信号から予測し，極だけを用いたフィルタで表現するモデル.
＊2 LSP：声道の響き方を表す複数の数値を効率良く表現したパラメータのこと．圧縮や伝送に適しています．1975年に板倉文忠が発表し，携帯電話の音声符号化に欠かせない重要な基礎技術として，2014年にIEEEマイルストーン賞に認定されました.
＊3 CELP：声道の特性を線形予測で表し，声帯などの音源に相当する「励振信号」をコードブックから選んで音声を再現する方式のこと．人間の声を低いビットレートで高品質に伝送できます.

世界中のスマホを支える「EVS」の革新

ここ10年ほどで，携帯電話の音質が格段に良くなった印象があります．

それはまさに，第 4 世代の移動体通信ネットワーク，VoLTE（Voice over Long-Term Evolution）向けに，私たちが世界の有力機関と共同で開発・標準化に大きく貢献したEVS（Enhanced Voice Services）によるものです．音質が向上した要因の 1 つは，扱える音声帯域が広がったことによります．2000年ごろまでは，伝えられる音声帯域は3.5 kHzと昔の黒電話と同程度でしたが，EVSではFM放送並みの14 kHzまで帯域を広げることができました．これは，周波数帯域の拡大に比例して単純にビットレートを増やすのではなく，限られた情報量の中で，より広い帯域の音を効率良く表現できるようにしたというものです．これにより子音や息づかい，背景雑音までも伝えられるようになり，より自然な音声通信が可能になりました．

またEVSでは，CELPに加えて新たに開発された音楽向け符号化技術も採用されています．電話は双方向の会話のため，EVSでは音を20 ms単位の短いフレームに区切り，低遅延で処理しなければなりません．これを音楽に適用すると非常に音が悪くなってしまうのです．そこで，私たちは音楽が持つ周波数成分や時間的な構造を理解し，短い処理単位でも高品質に低遅延で音楽を伝送できるよう工夫を重ねました．さらには，従来の回線交換では符号誤りに弱いため可変長符号化[*4]を使うことは困難でしたが，パケット通信では誤りのあるパケットは破棄できることから，これを取り入れ，圧縮効率を高めることができるようになりました．そして，パケットが失われた場合にも，音が不自然に途切れないよう，パケット消失を補償する高度な信号処理も組み込みました．

この広帯域，高音質，高圧縮，高品質という要件を満たすEVSは，2014年に移動通信システムの国際標準化団体である3GPP（3rd Generation Partnership Project）の音声・音響符号化方式として国際標準化されました．第三者機関によって大規模な主観品質評価試験が行われた結果，従来方式よりも格段に高品質であることが確認されました．2018年以降は，これまでバラバラだった国内の方式もEVSに統一され，さらには世界の通信事業者や通信機器メーカーなどが一斉に採用したことで，いまや世界中の何十億台ものスマホで使われています．日本から海外へ携帯電話をかけても，国内と同じような品質でつながりますし，周囲を見渡して，ほとんどの人がスマホを手に持っている姿を見ると，まるで夢のようだと感じます．

どのようなご苦労がありましたか．

標準化活動というのは，企業や国の利害が絡むため，学会とは違って，非常に厳しい指摘が飛び交うことがあります．どの方式を世界標準にするのかの争いになれば，相手の方式の欠点を喧嘩腰であげつらう場面も見られます．私は職人肌で，裏方で技術を磨くほうが向いていて，ときには言い負かされてしまうこともあり，「弱気の守谷」といわれていました．

また，あれほど苦労して標準化したにもかかわらず，全く使われないままの技術もあります．それでも，いったん同じ方向を向くことができれば，世界中のエンジニアが協力し合い，ときには寝る間を惜しんでより良い技術をつくろうと皆で力を合わせる姿は素晴らしいものがあります．そして最終的には何十億人もの人々に自分が貢献した技術が使われ，ときには命を救うこともあると考えれば，エンジニア冥利に尽きるというものでしょう．

＊4 可変長符号化：頻繁に現れる情報には短い符号を，あまり現れない情報には長い符号を割り当てることで，データ全体の量を減らす技術．

立体的な音の場を届ける次世代規格「IVAS」

現在の活動について教えてください.

今もこうしてWeb会議システムを通じてお話ししていますが，ときどき音声が途切れたり，遅延があってしゃべりづらかったりして，音声品質が不安定ですよね．実際に，NTTデータ経営研究所と音響機器メーカーの「Shure」が，オンライン会議の音声コミュニケーションがもたらす生体ストレス反応について調べたところ，85％もの利用者が音質に不満を抱えていることが分かりました．そうしたこともあり，ここ数年は3GPPにおいて，EVSを拡張するIVAS（EVS extension for Immersive Voice and Audio Services）の国際標準化に取り組んできました（**図2**）．2025年に主要な仕様が固まり，現在も規格の改善や普及に向けて，週2回ほどWeb会議で議論を重ねています．

EVSは基本的にモノラルの音声通信ですが，IVASではステレオで立体的な音場を双方向に伝えます．リアルタイムで送受信される複数の音声ストリームを扱うことで，多地点の会話にも対応します．オンライン会議に加えて，専用のゴーグルやメガネを使ったXR（Extended Reality：現実世界と仮想世界を融合する技術）やメタバースなどにおいても，あたかもその場にいるかのような臨場感を味わえることを想定しています．

なおこの規格は，Web会議システムなど，ネットワーク上のさまざまな通信アプリケーションの通信技術WebRTCとの連携も視野に入れています．そうしたことから多様で高機能な実時間通信のための共通基本ツールなどの開発が行われています．一方で，私たち電話会社としては，移動通信（VoLTE）や固定網（NGN）などの制御技術を連携させた音声品質の保証をねらっています．また，すでに世界中に普及しているEVSの端末との互換性を考慮して，符号変換を行わずにIVASの端末と接続できるような相互互換性を備えた仕様も盛り込みました．

さらにIVASで特徴的なのは，開発中のソースコードをGitLab上に公開し，これまでの議事録もすべて誰でも閲覧できるようにしていることです．従来の標準化は競争的

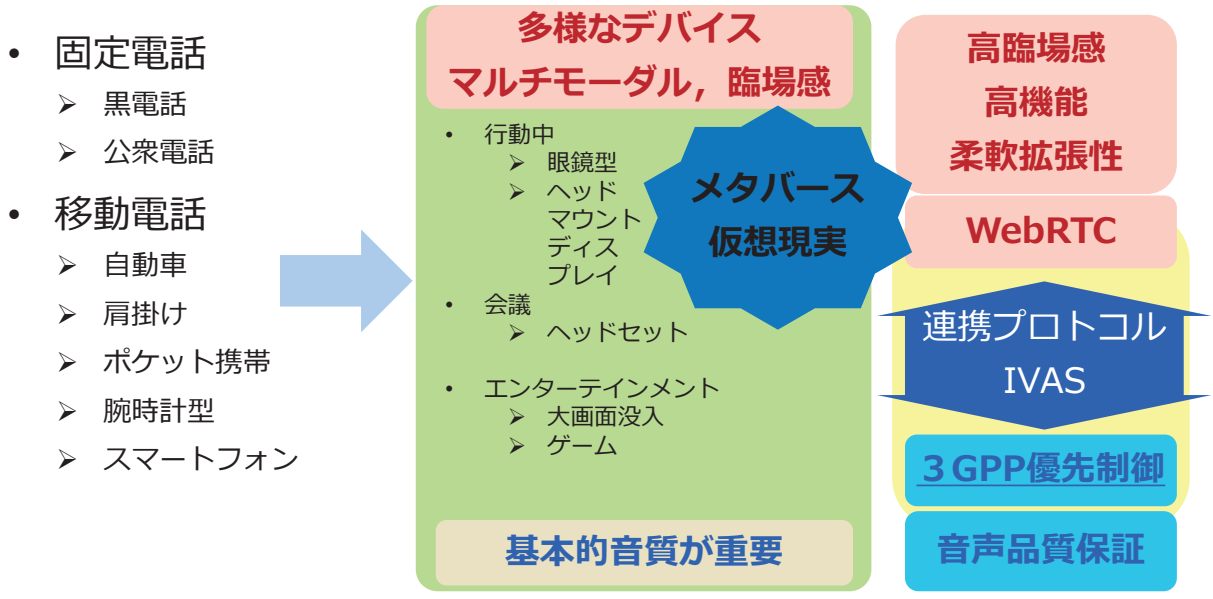

図2　情報伝達手段（電話から手段の多様化）

- **知の泉を汲んで研究し（必須）**
  - **新規性を基準に，発明，特許，論文などで学術貢献**
  - 厳しい競争，理解されない懸念もあるが，将来有益かもしれない．
- **実用化により（自信や機会があれば挑戦）**
  - **実用性に自信のある成果に限り，成果提供などで貢献**
  - 他技術との統合，コスト，競合技術などで製品化できない懸念．
- **世に恵みを具体的に提供しよう（協力者や利用者次第）**
  - **実際に世の中で使ってもらえるような商品のサポートで貢献**
  - 競合製品，利用者の嗜好などで，現実市場では受け入れられない懸念．

図３　知の泉の展開

かつ閉鎖的でしたが，いまや組織や国を超えて技術を磨く協力体制がとられています．開発過程を透明化し，参加者が相互に検証することで，より良い技術を採用していく．それこそが新しい国際標準のかたちといえます．今後，このIVASがゴーグル型デバイスや高品質のWeb会議などで広く活用されていくことになるのかどうか，期待を込めて注視していきたいと思います．

後進へのアドバイスをお願いします．

NTTのR&Dには，「知の泉を汲んで研究し実用化により世に恵みを具体的に提供しよう」という理念が息づいています（**図３**）．これは，NTT研究所の前身である電気通信研究所の初代所長・吉田五郎さんの言葉です．「知の泉を汲む」とは，新規性を基準に，発明や特許，論文などで学術に貢献するという，研究者として当たり前の責務です．ときに厳しい競争にさらされたり，新規性ゆえにすぐには理解されないこともあったりしますが，将来に有益な価値をもたらす可能性を秘めた重要な営みです．

しかし，そこから「実用化」につなげるのは，容易ではありません．他の技術と競合したり，コスト面などにより製品化に結びつかないことも多い．ただ私自身，40数年にわたり研究をしてきた中で感じるのは，５年あるいは10年に一度くらい，世の中のニーズと照らして「これはいける」と思えるような成果が出たなら，ぜひ，自信を持って周囲の人や事業会社などにアピールしてチャレンジしてほしいということです．うまくすれば，他の組織や企業の協力も得ながら，あるいは他の技術も使いながら，スマホの中で世界中で使われるアプリやサービスになるかもしれません．チャレンジすることで，まさに「世に恵みを具体的」に提供できることもあるのです．

もっとも私自身，何十億人もの利用者に役立った技術もあれば，全く使われなかった技術もあります．海外も含めれば私がかかわった特許は1700件を超えているようですが，それでも本当に世に恵みとして使われているものはその一部です．それだけ実用化は難しい．それでも，良い成果が出て，機会が巡ってきたなら，ぜひ自信を持ってチャレンジしてみてください．