



# 音声通信の歴史と未来

## —音声音響符号化技術の視点から—

2022年9月28日

日本電信電話株式会社

NTTコミュニケーション科学基礎研究所

守谷健弘

# ICT技術の進展と電話



50年のICTの進展は日常生活にも大きな貢献  
電話はその象徴で、電話をみれば時代がわかる  
音声通信は電話から多様で高機能サービスへ



総務省情報通信白書令和元年「携帯電話の登場・普及とコミュニケーションの変化」

# 内容



## 1. 音声通信の歴史

便利になったか？

## 2. 音声音響符号化の貢献

音質はよくなったか？

## 3. 現在の音声音響通信の問題

音質は満足か？

## 4. 将来の音声音響通信

使う人たちのためか？

# 内容



## 1. 音声通信の歴史

便利になったか？

## 2. 音声音響符号化の貢献

音質はよくなったか？

## 3. 現在の音声音響通信の問題

音質は満足か？

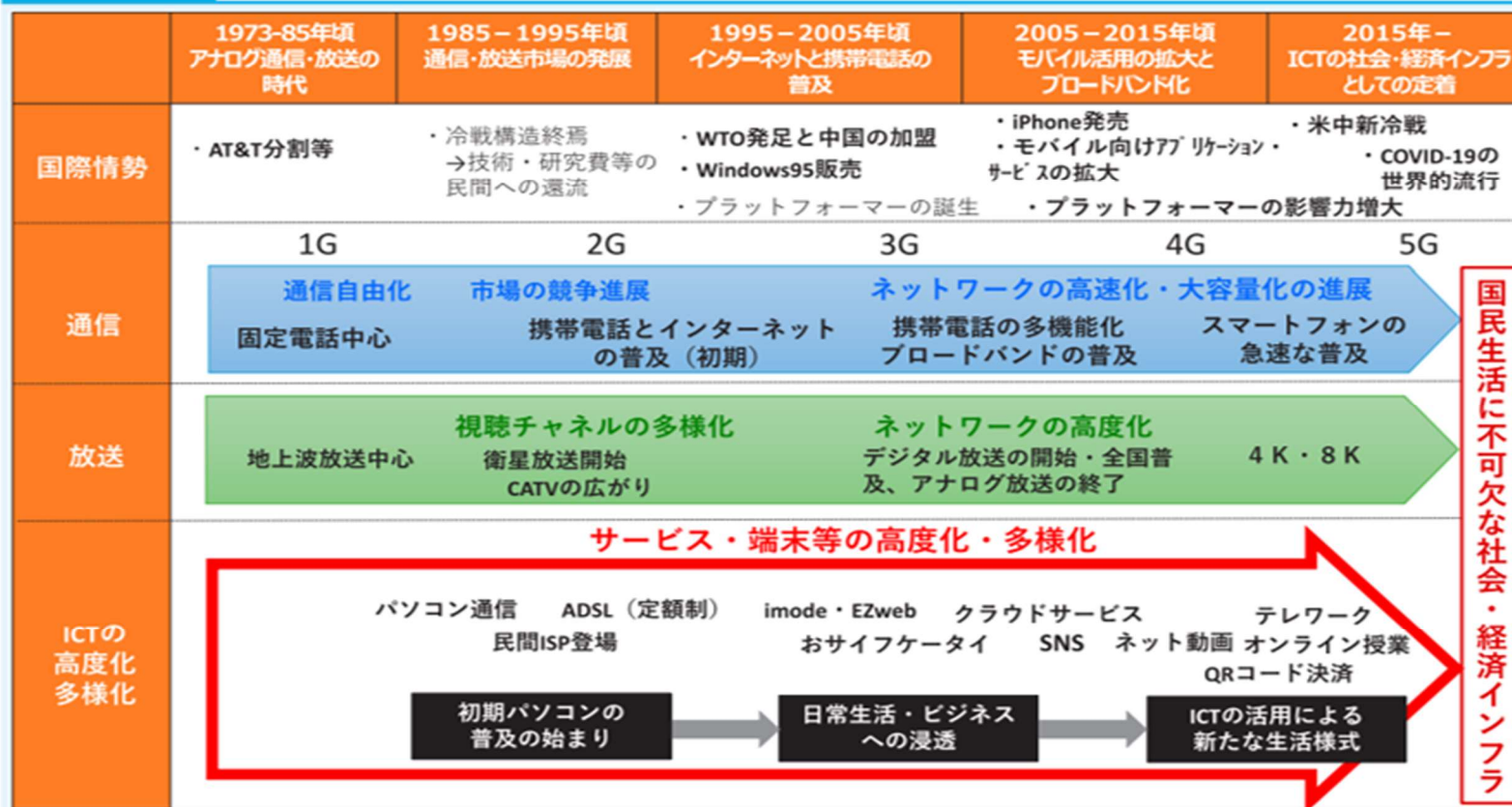
## 4. 将来の音声音響通信

使う人たちのためか？

# 過去50年間の情報通信分野の動向



図表 1-0-1-1 過去50年間の情報通信分野の動向

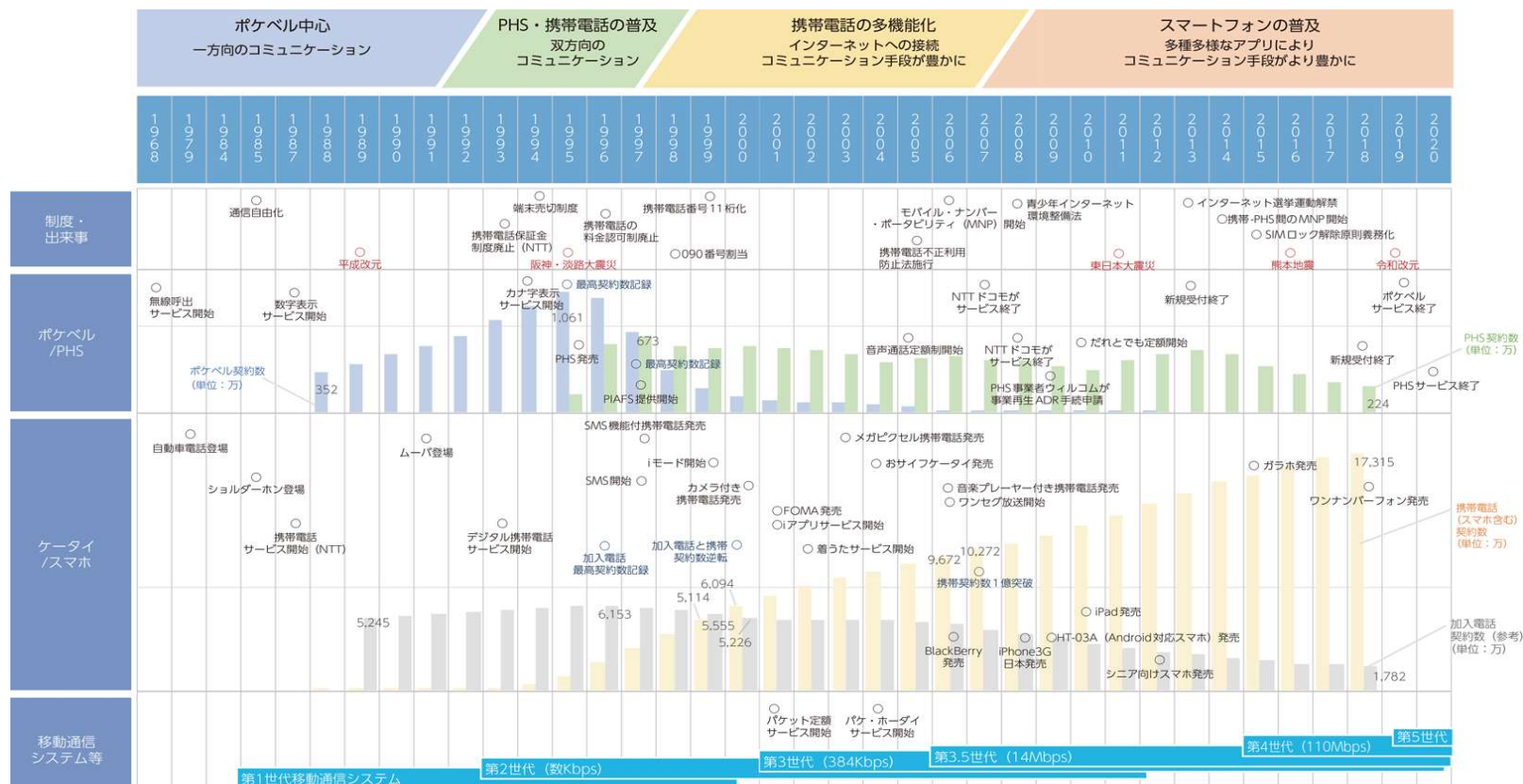


国民生活に不可欠な社会・経済インフラ

総務省(2022)  
「デジタル社会における  
経済安全保障に関する  
調査研究」

(出典) 総務省 (2022) 「デジタル社会における経済安全保障に関する調査研究」

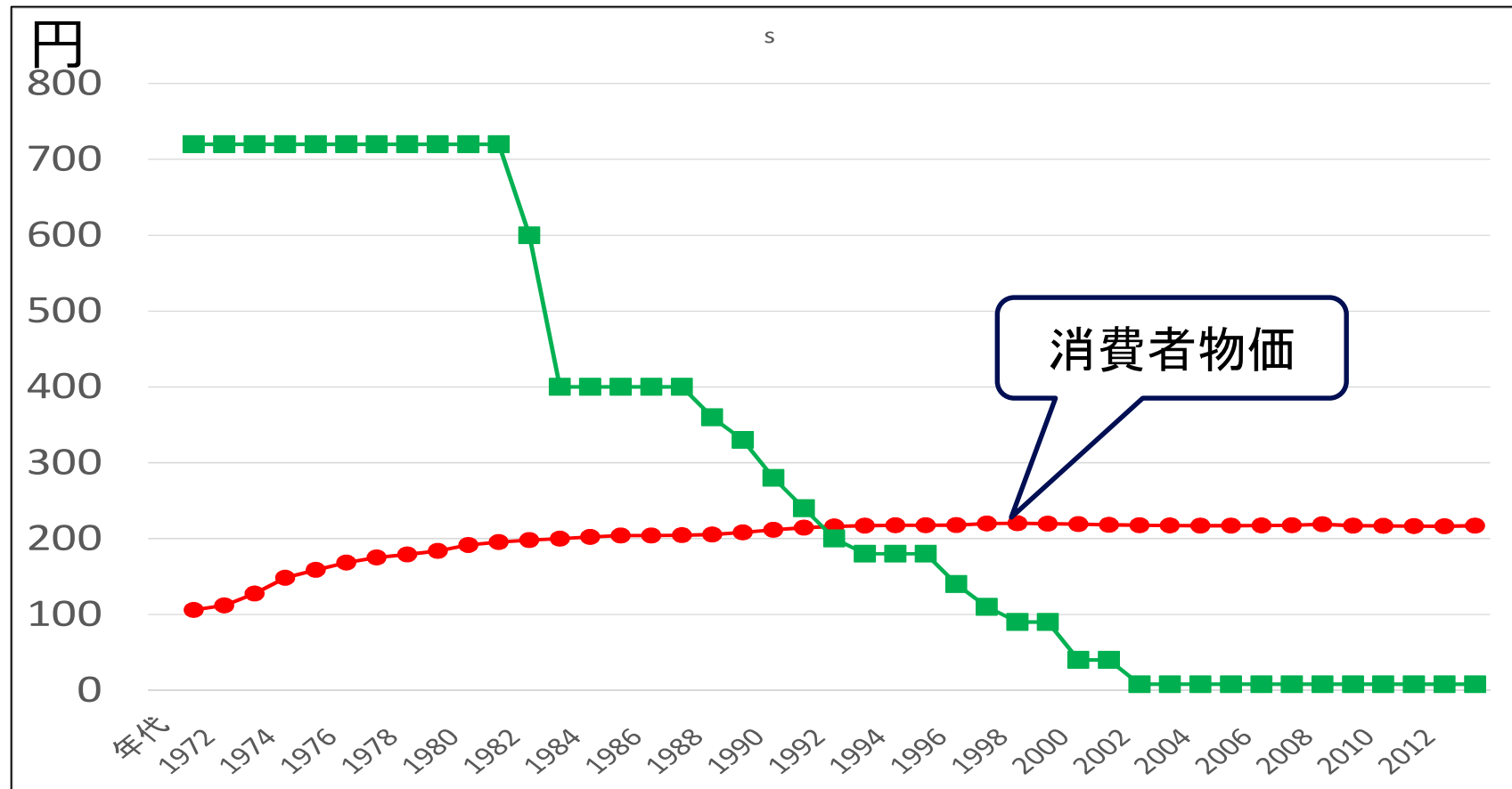
# 音声通信の歴史



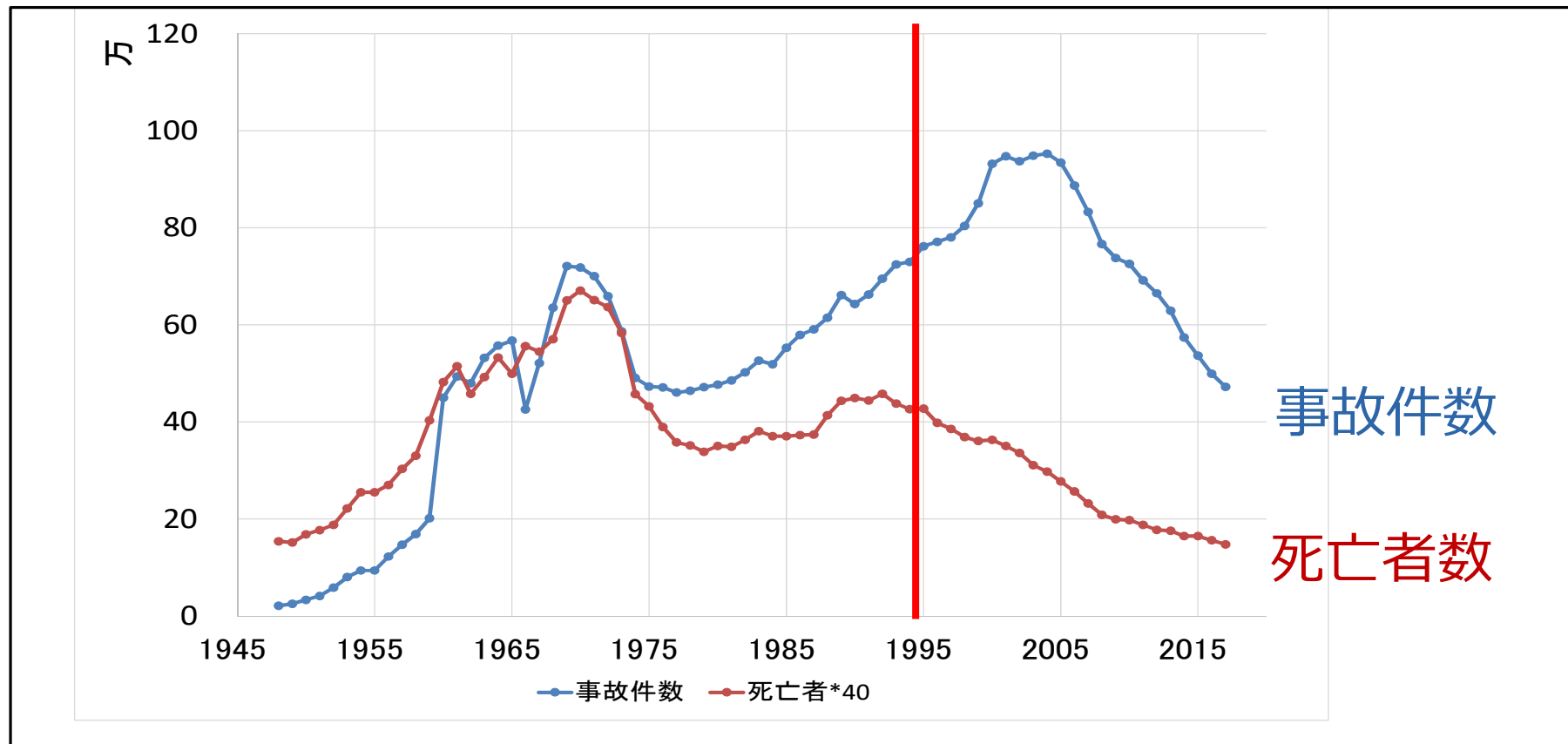
総務省情報通信白書令和元年「携帯電話の登場・普及とコミュニケーションの変化」

Copyright 2022 NHT Corporation

# 長距離通話料金（3分）

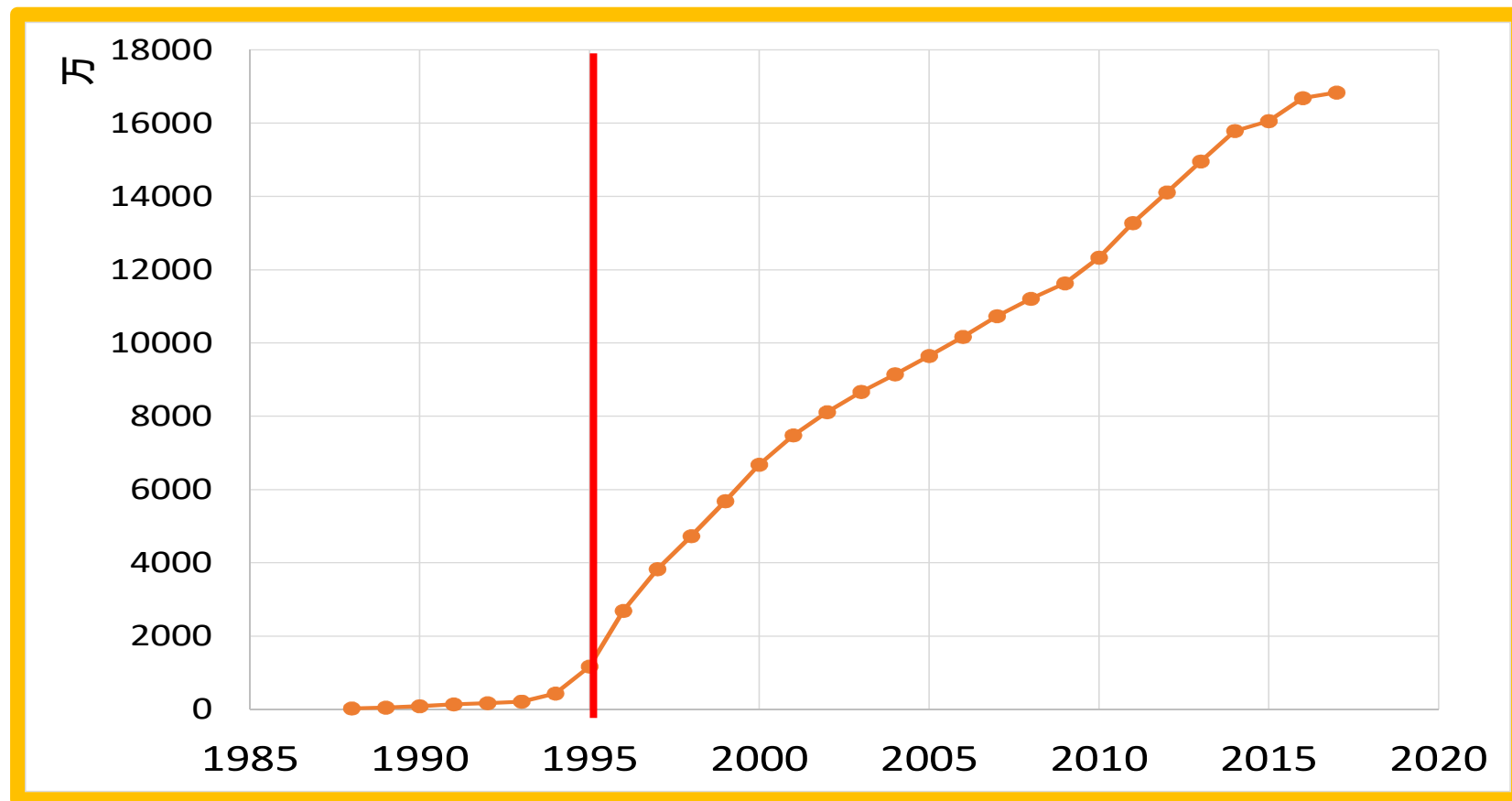


# 交通事故件数と死亡者数（日本、年間）

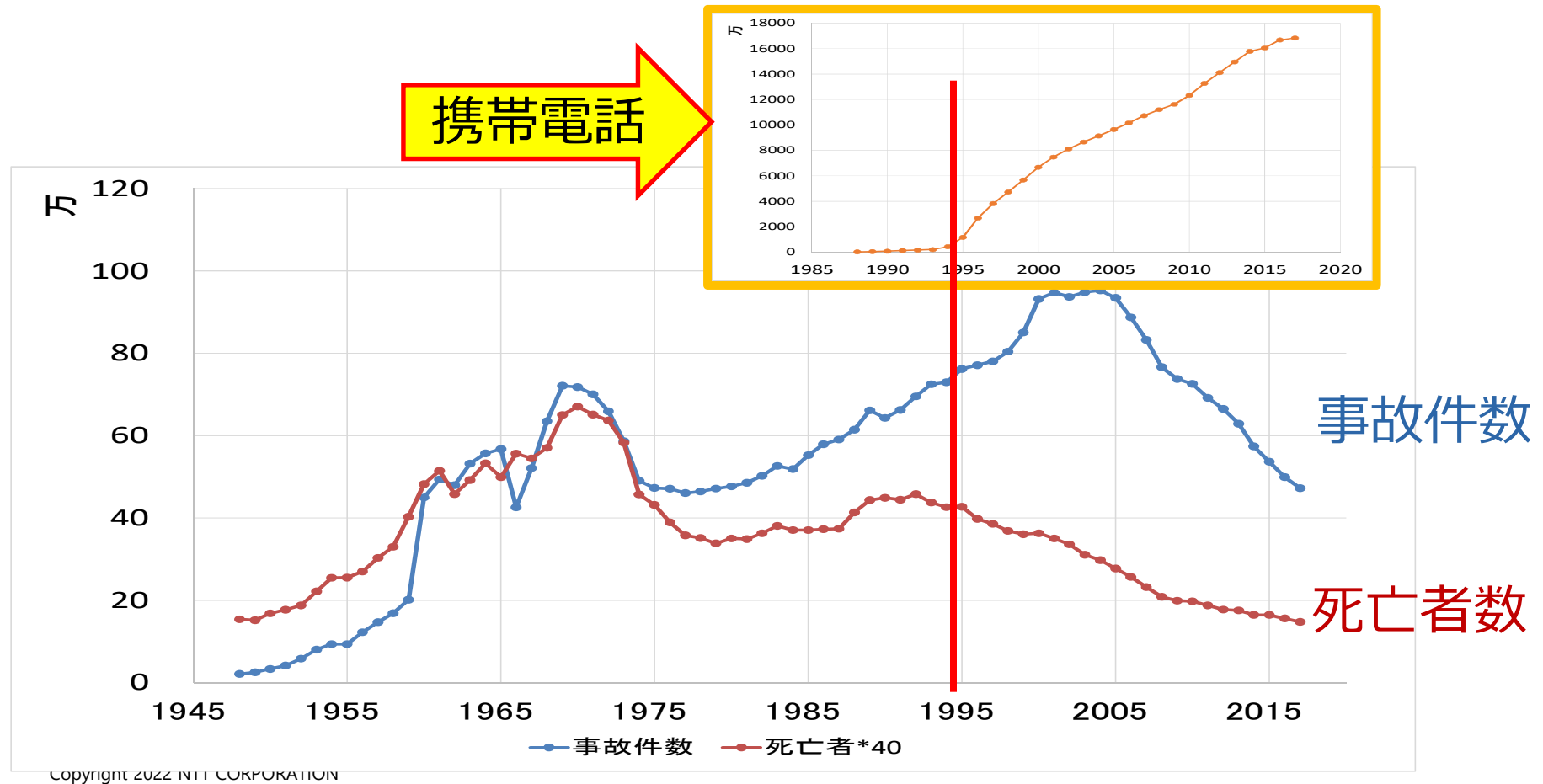




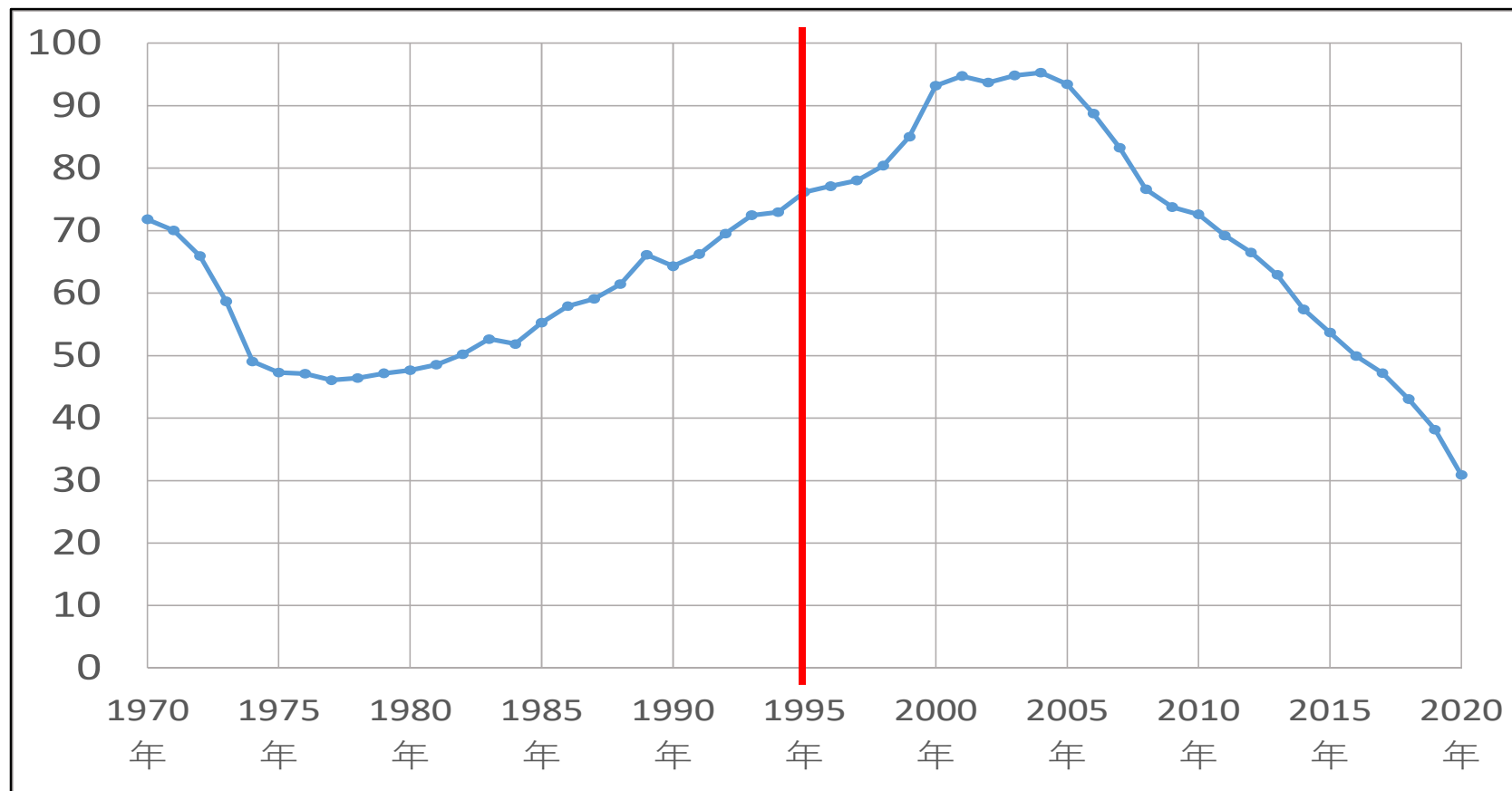
# 携帯電話の加入者数（日本）



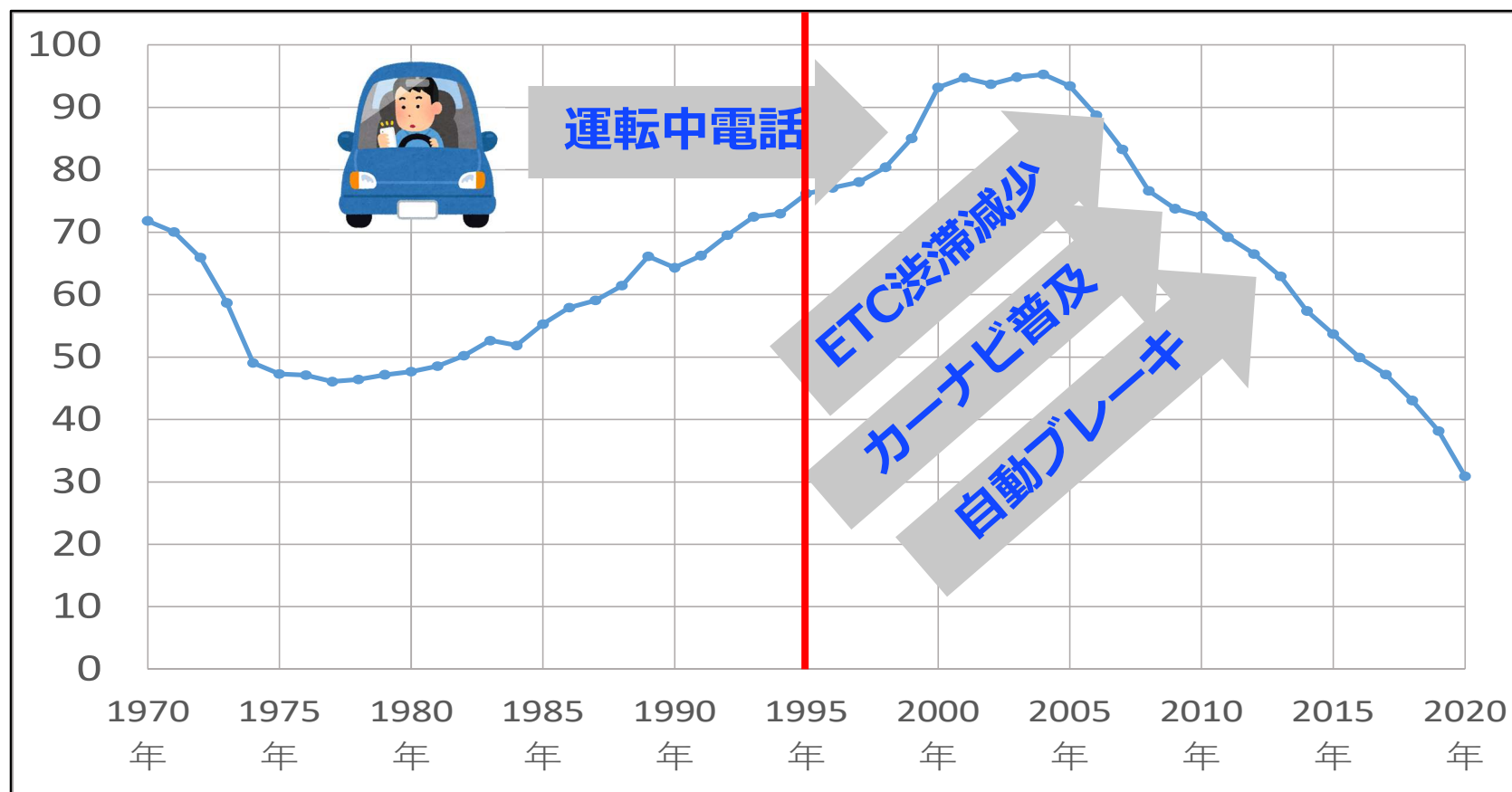
# 交通事故死亡者数と携帯電話の普及



# 交通事故件数



# 交通事故件数



# 交通システムへのICTの貢献



- **携帯電話の普及**

病院までの時間短縮      => 交通事故死者の減少

運転中の通話が増加      => 交通事故件数の増加

- **ETCの普及**

- **カーナビの普及**

- **自動ブレーキの普及**

## 1. 音声通信の歴史

便利になったか？

- 音声通信の価格は大幅に下がった
- 携帯電話で音声通信は便利になった
- 間接的に人の命を救うこともできる



# 目次



## 1. 音声通信の歴史

便利になったか？

## 2. 音声音響符号化の貢献

音質はよくなったか？

## 3. 現在の音声音響通信の問題

音質は満足か？

## 4. 将来の音声音響通信

使う人たちのためか？

# 音声音響符号化技術



歪を許す  
高圧縮  
符号化

- 歪を許さない符号化
  - 長期保存、ハイレゾ配信

- 片方向通信用符号化
  - デジタル放送、音楽プレーヤ、光ディスク

- 双方向通信用符号化
  - 携帯電話、IP電話

- 信号分析、処理、情報理論
- 人間の感覚、生成
- ハードウェア

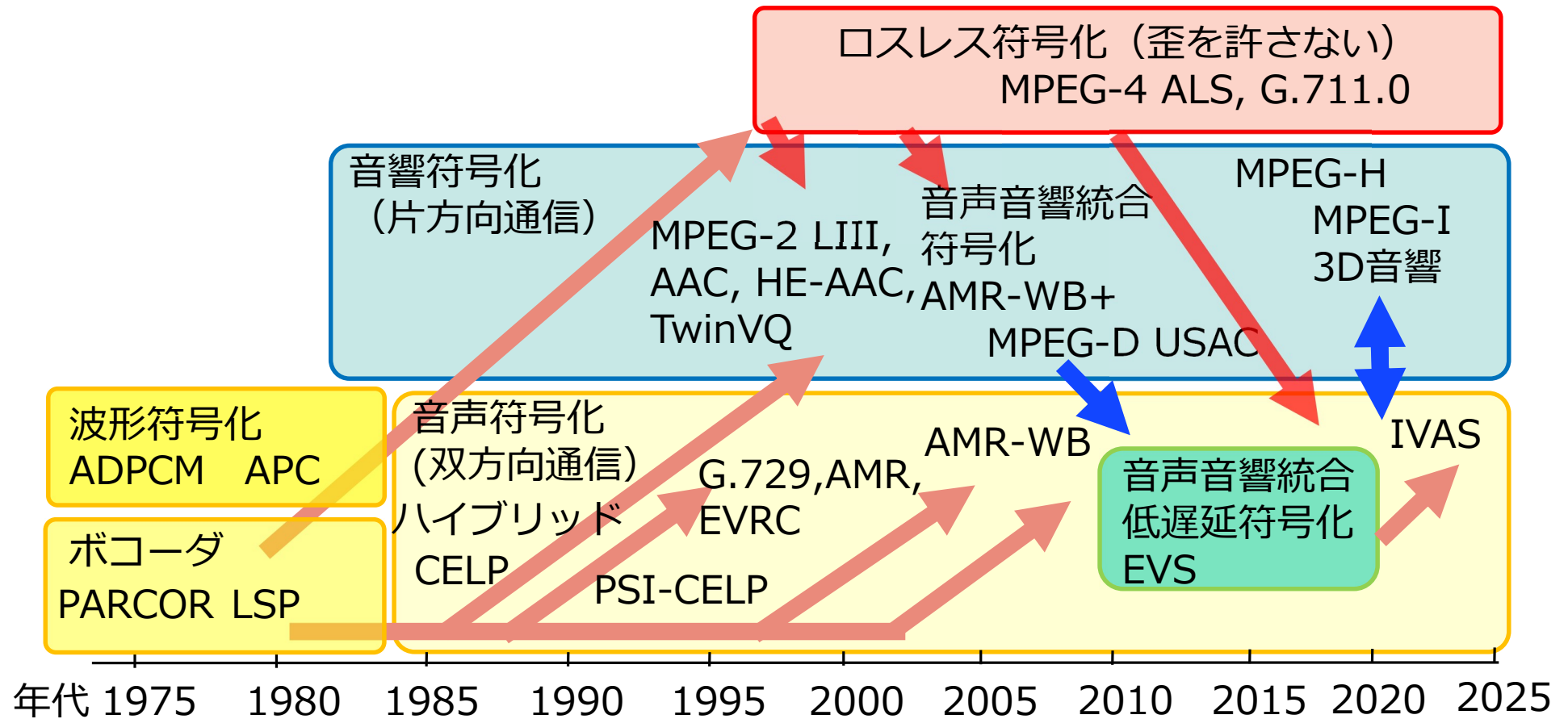


# 音声音響符号化技術の分類

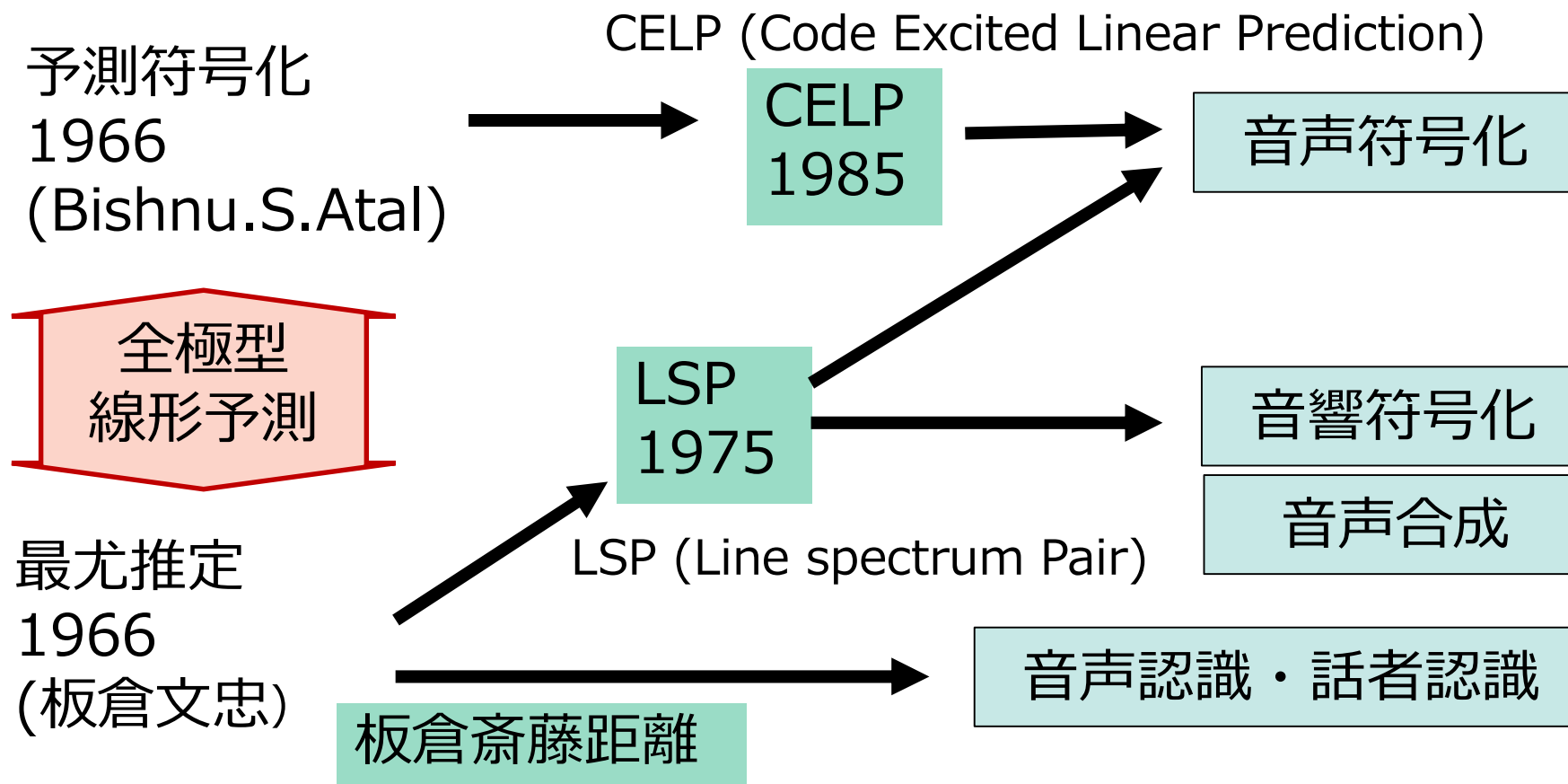


	符号化歪	可変長 符号化	聴覚特性 利用	フレーム 長
歪を許さない符号化	なし	◎	×	任意
片方向通信 符号化	あり	◎	◎	80 ms
双方向通信 符号化	あり	×	◎	20 ms

# 音声音響符号化技術の歴史

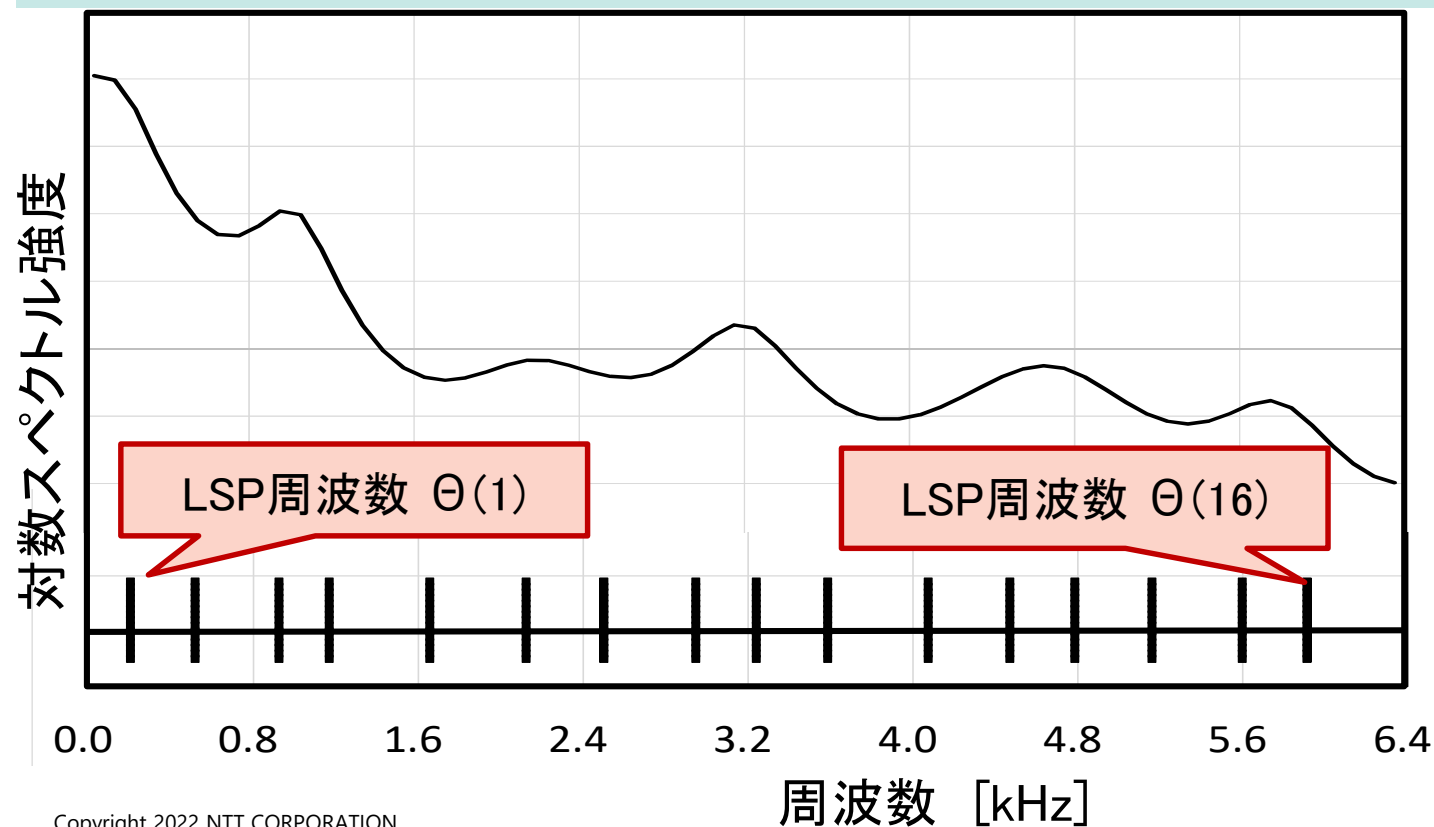


# 基礎技術：線形予測



# 基礎技術: LSPとスペクトル包絡

◆ LSPはスペクトル包絡を効率よく表現するパラメータ



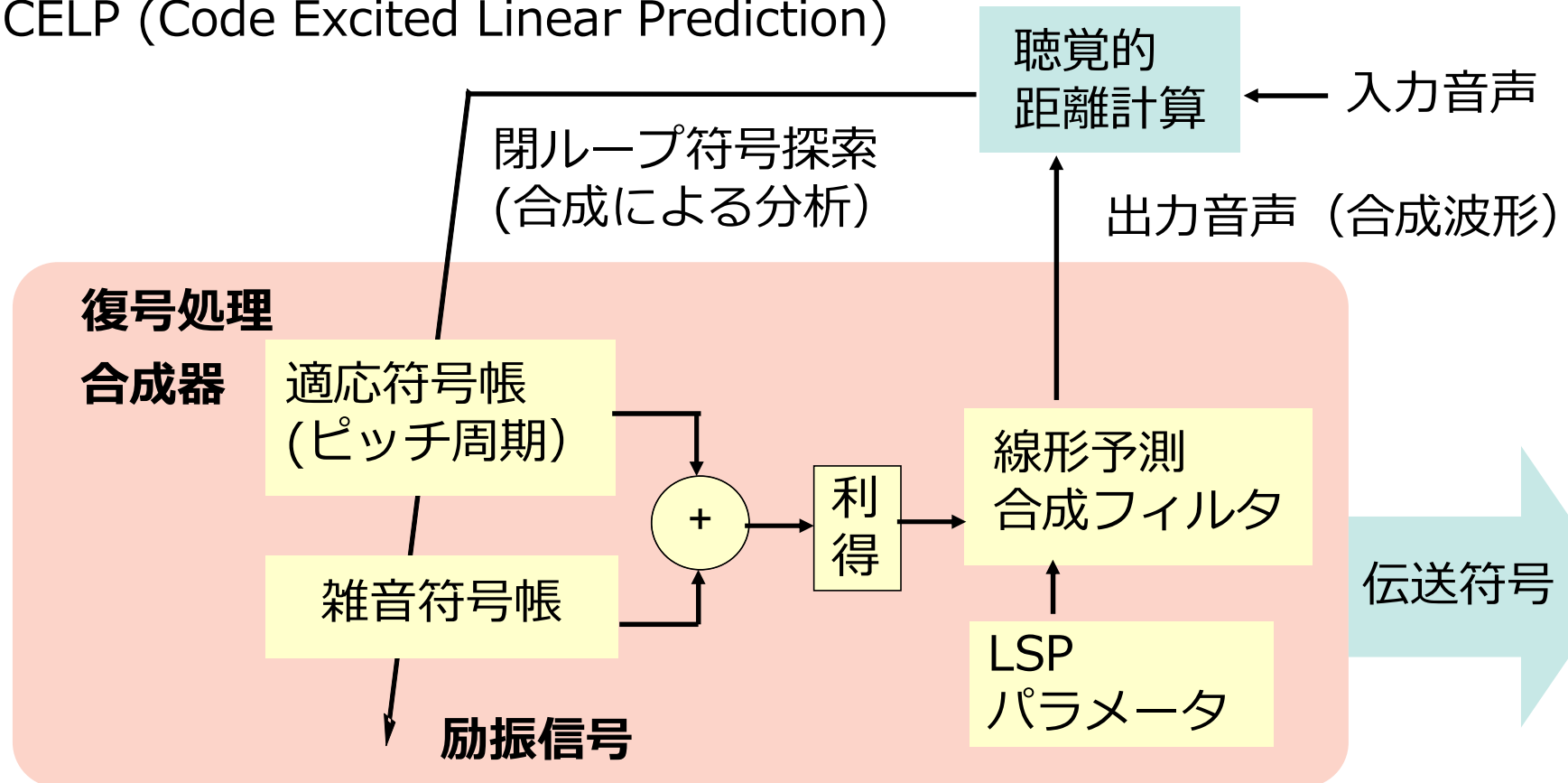
スペクトル包絡：  
• 口の形を表現  
• 音韻特徴を表現

山谷の深さ：  
• サンプル間相関  
• 予測の程度  
• 情報圧縮の程度

# 基礎技術: CELPの符号器の基本構成

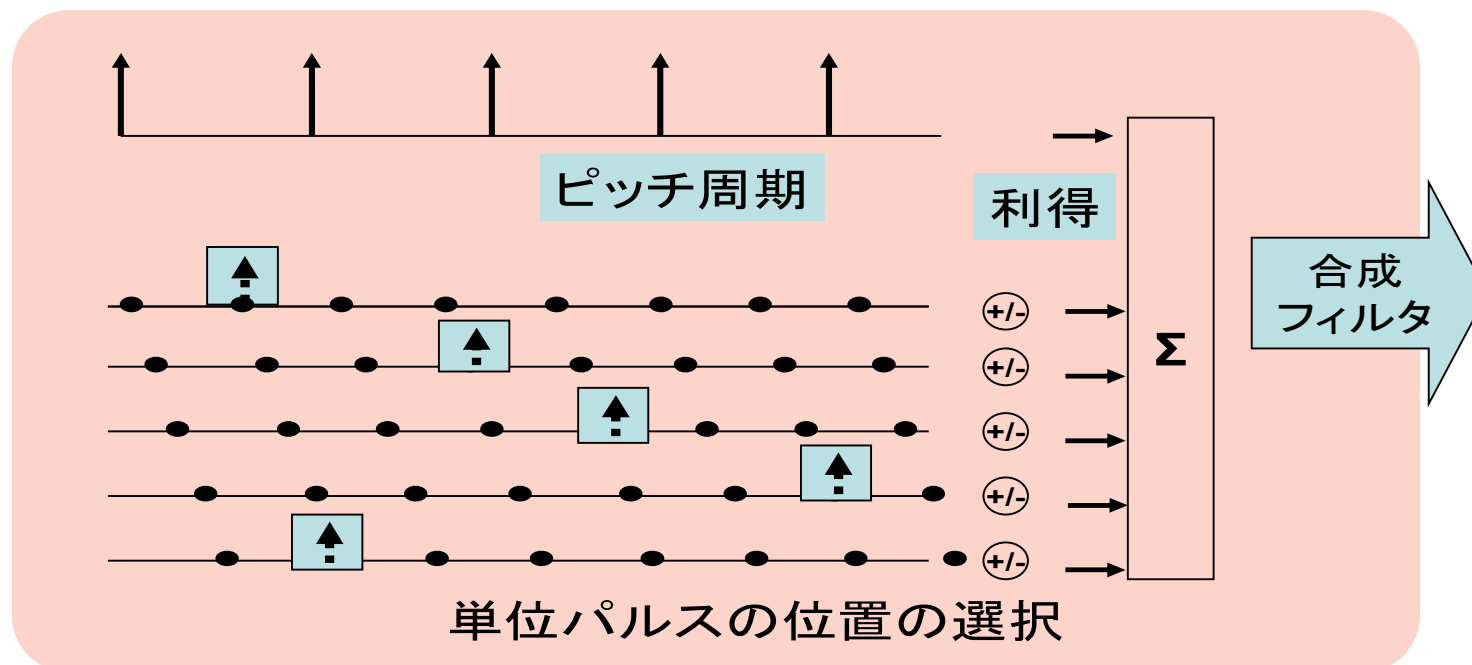


CELP (Code Excited Linear Prediction)



# ACELP (Algebraic CELP) の合成モデル NTT

- ◆ 基本原理と品質は同じ、大規模な符号帳を大胆にモデル化
- ◆ 大幅な演算量削減、大幅なメモリ量削減



# 携帯電話の課題（2010年）とEVS



- 従来の携帯電話の音声帯域は狭い
- 従来の携帯電話では音楽の品質が低い
- 回線交換からパケット (LTE)通信になると決定

2010年 3 GPP EVSの標準化開始

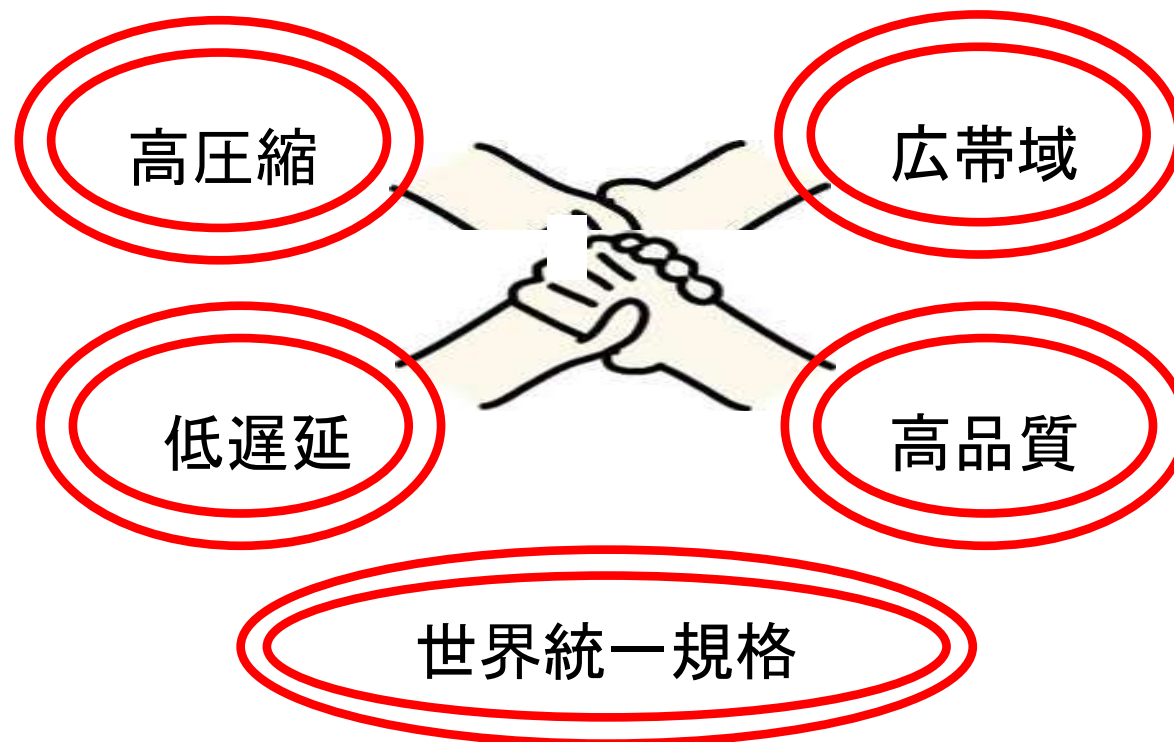
- 開発エンジニア：50人以上＊3年
- 主観評価試験：のべ5万人、300万回評点
- 規格書：文書800ページ以上＋参照ソフト

2014年 3GPP EVSの標準化完成

# EVSの到達点



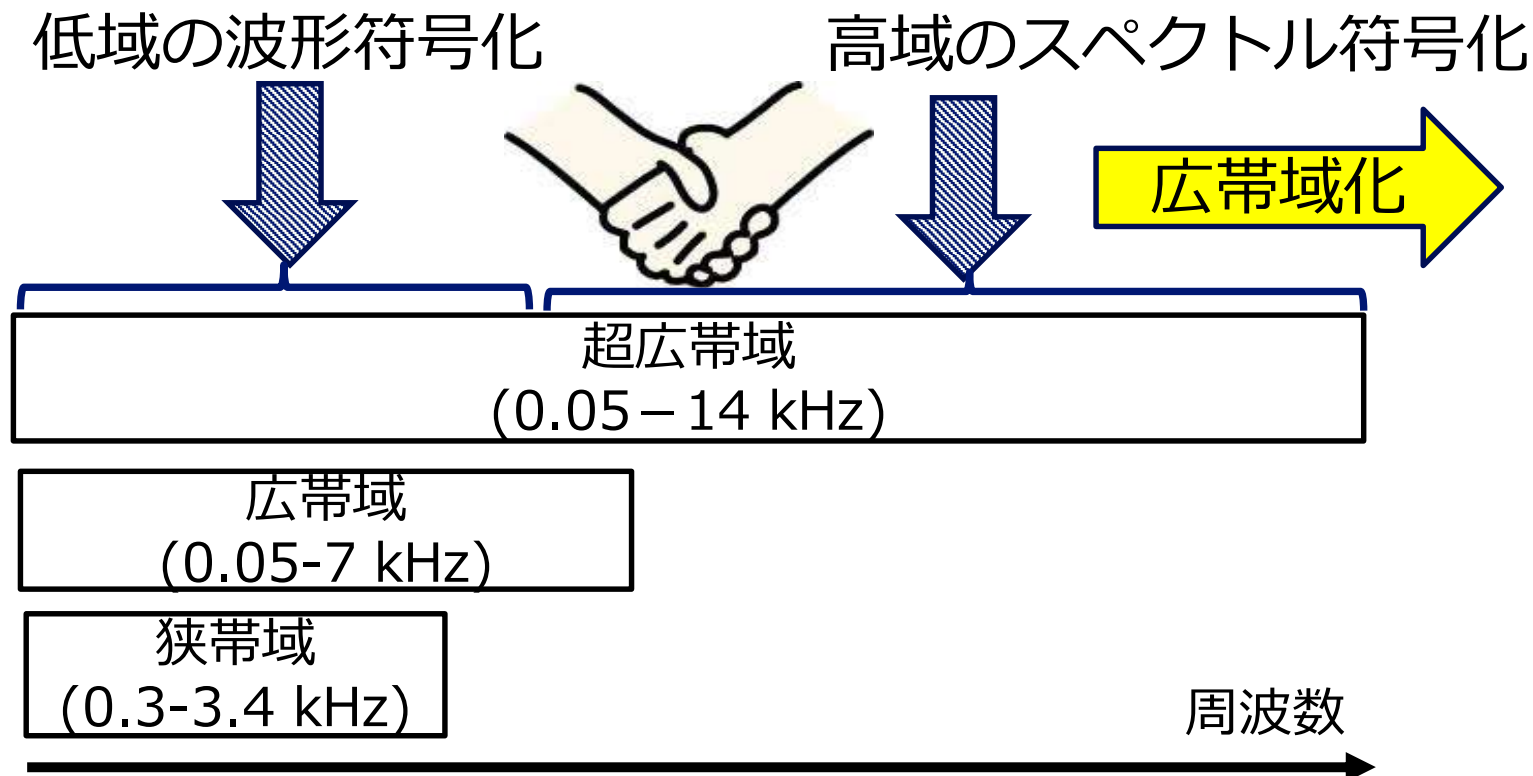
EVS (**E**nhanced **V**oice and **S**ervices)





# EVRSの技術の統合（広帯域化）

## ◆ 低域のみ波形再現（全ての波形再現には4倍の情報が必要）



## EV Sの技術の統合（音響符号化の低遅延化） NTT

### ◆ 20msフレームでも音響信号に対する符号化歪を大幅に軽減

音声符号化

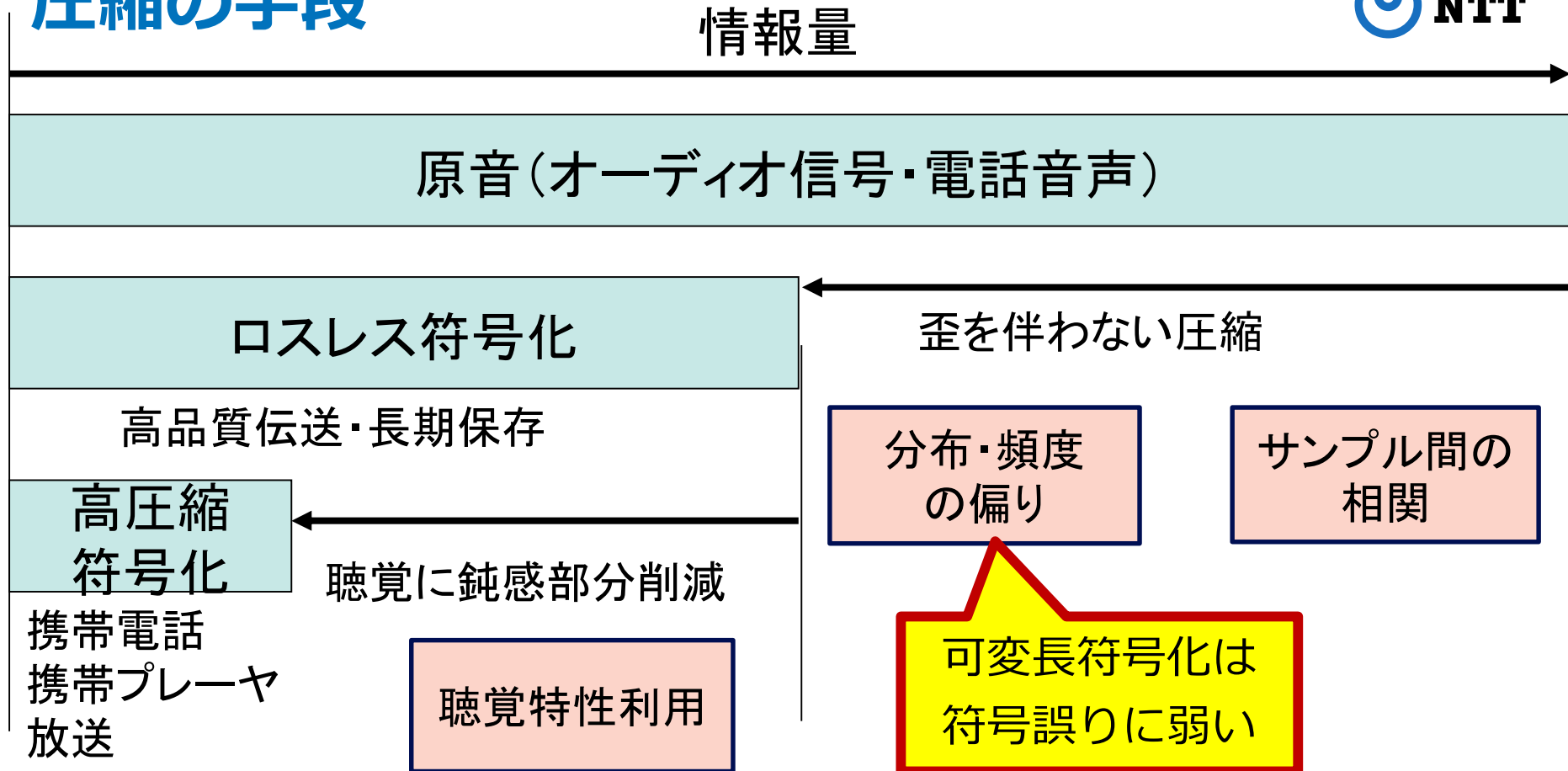
音響符号化



- 従来の音響符号化のフレーム長： 80ms
- 音声符号化に必須のフレーム長： 20ms
- **20msフレームで音楽符号化の品質劣化は致命的だった**
- **周波数領域の周期性モデルによる符号化歪の大幅軽減**

# 圧縮の手段

情報量



# EVISの技術の統合（高圧縮化）



## ◆ パケット化で“可変長符号化”による高圧縮が利用可能に

高圧縮符号化

歪のない符号化



01111010

1110

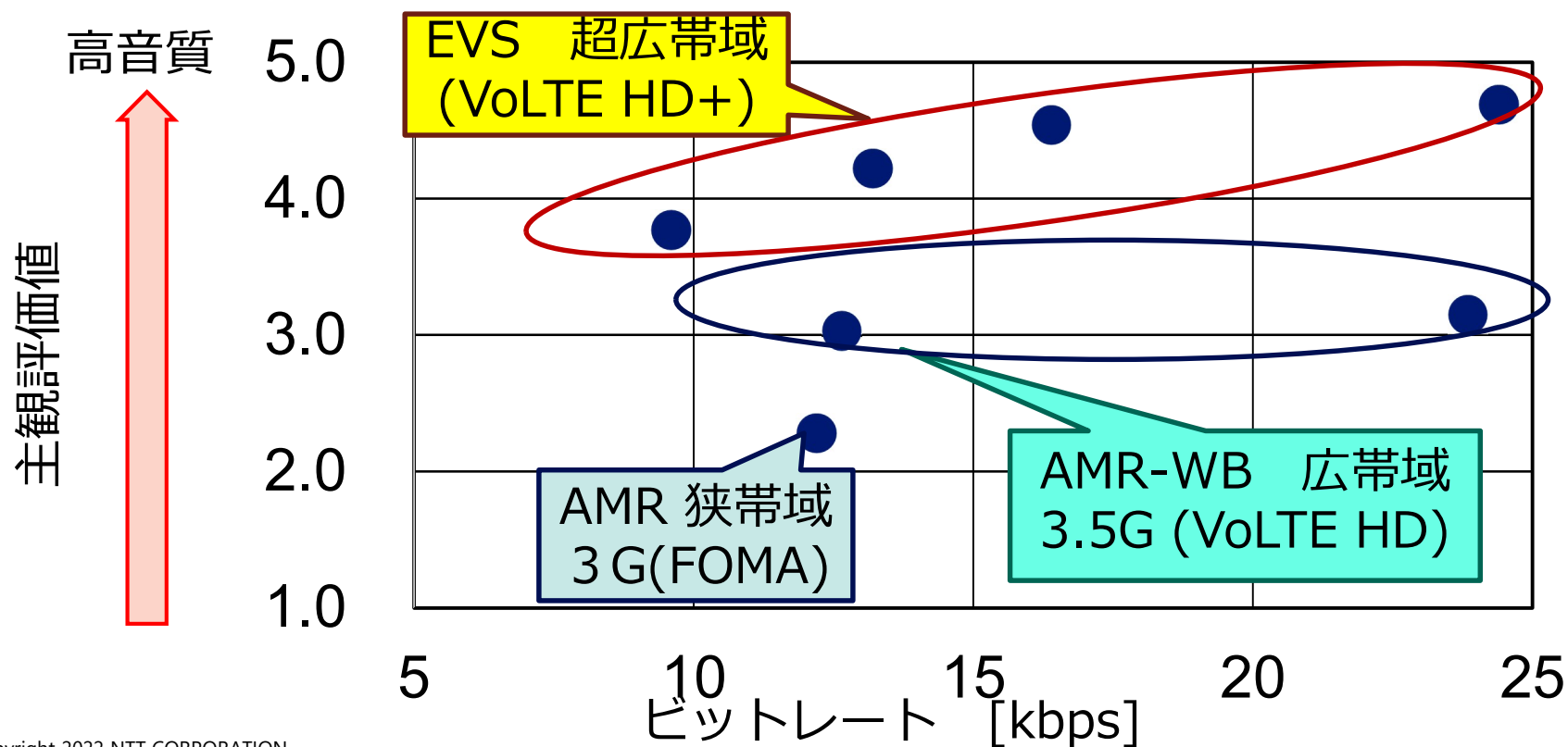
10111101110

（可変長符号化）

- 従来の回線交換では符号誤りに弱い“可変長符号化”は不可
- パケット消失に対応する高度な信号処理の組み込み

# EVNSの主観評価（高品質化）

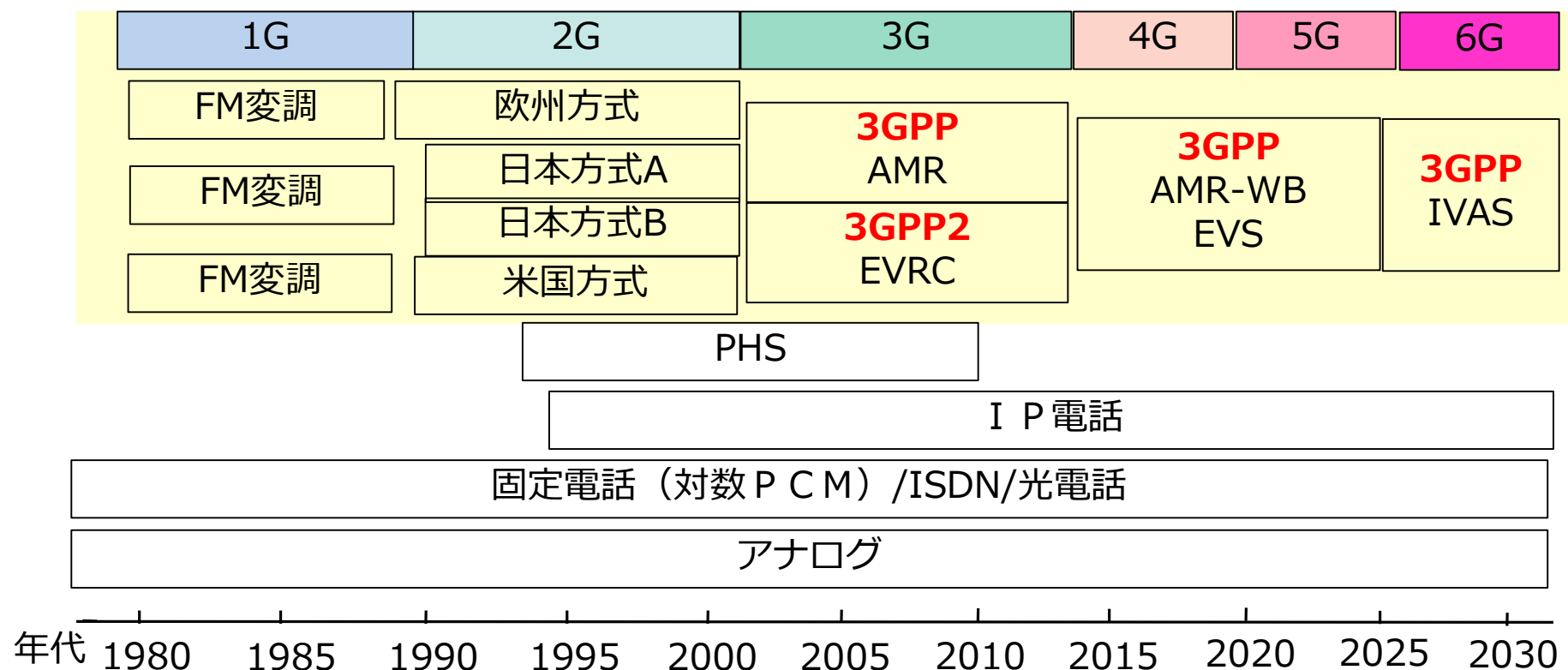
## ◆ 従来の電話と同時に聞き比べると大幅な品質改善



# 移動電話と符号化方式



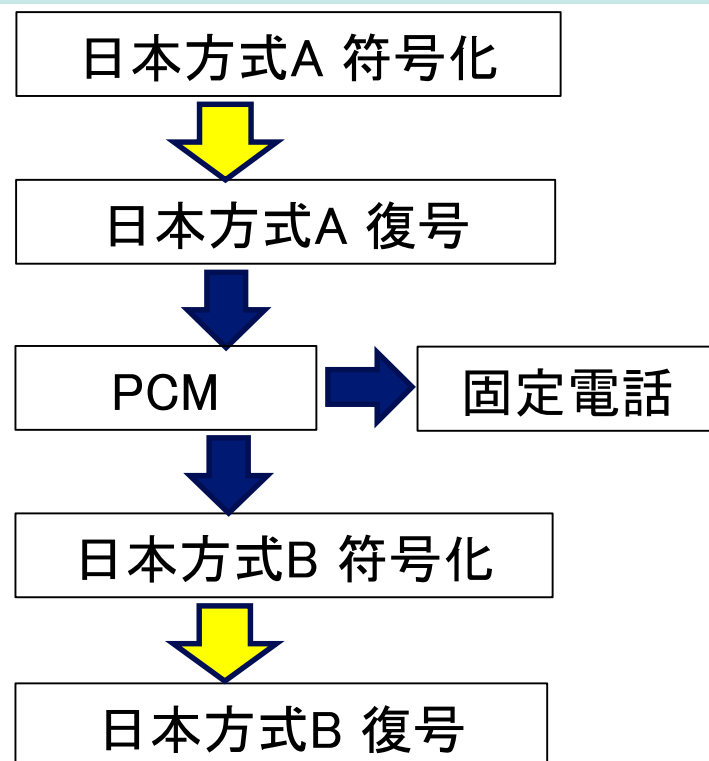
◆ 符号化方式は無線システムに依存しつつ、世界的に統一



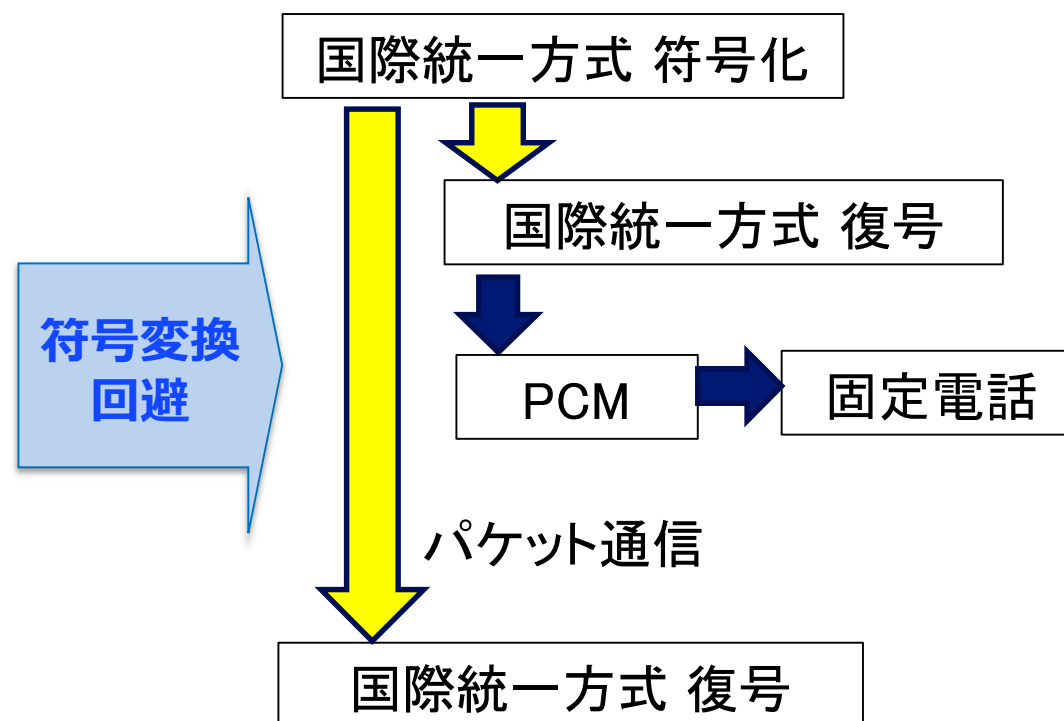
# 移動電話の相互接続（符号変換回避）



2018年以前（符号変換多数）



2018年以降（符号変換無）



## まとめ

## 2. 音声音響符号化の貢献

音質はよくなったか？

- 無線システムの更新のたびに新符号化で性能改善
- EVS (VoLTE) で究極の**電話品質**に到達
  - これまでの符号化技術の集大成
  - 低ビットを維持したまま、音声と音楽に高品質
  - 音声帯域（周波数）は3.5 kHzから 14 kHzに拡大
- 日本国内の電話会社間で符号変換回避





# 目次



## 1. 音声通信の歴史

便利になったか？

## 2. 音声音響符号化の貢献

音質はよくなったか？

## 3. 現在の音声音響通信の問題

音質は満足か？


## 4. 将来の音声音響通信

使う人たちのためか？

## 現在の課題



- 個人用音声通信は**電話から無料電話, 常時接続**に移行
- 業務用音声通信は**電話からweb会議システム**に移行

 **NTT Data**  
Trusted Global Innovator

コンサルティング  
サービス

コンサルタン  
ト紹介

ニュースリ  
リース

イベ  
ント

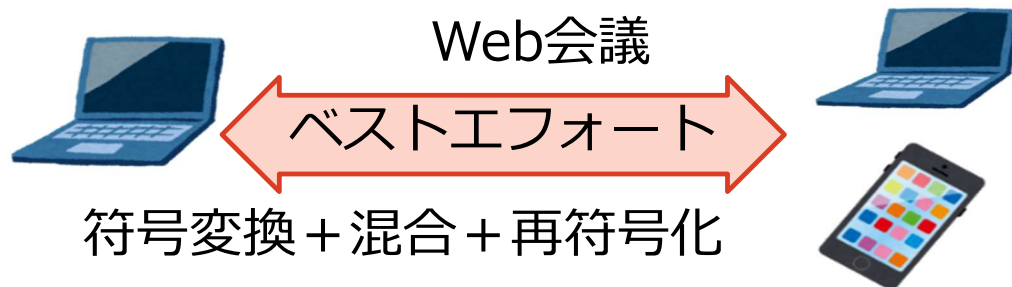
ナ  
レ  
ッ  
ジ

企業  
情報

採用  
情報

📧 お問い合わせ

# 現状の音声通信の問題点



無料電話



高臨場高機能



音声品質が不安定



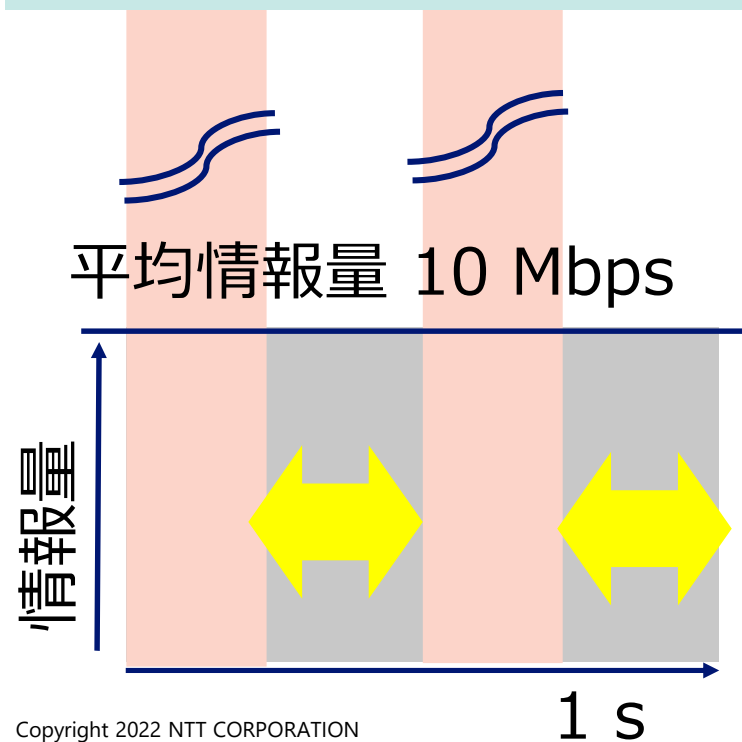
拡張性が乏しい



安定な音声品質

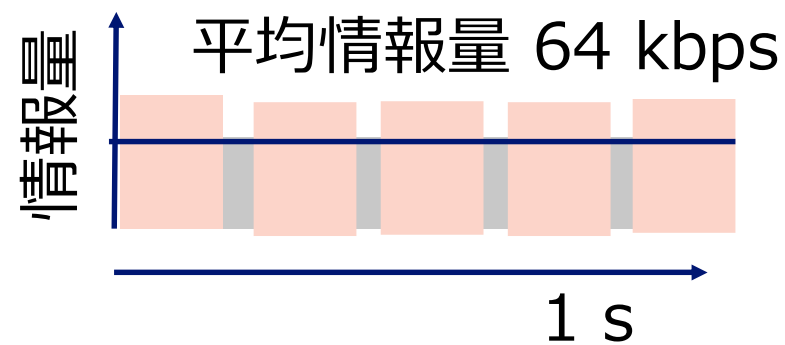
# パケットの隙間の問題

- ◆ パケットの隙間は双方向通信のネック
- ◆ 高ビットレートでも音声品質低を維持できない

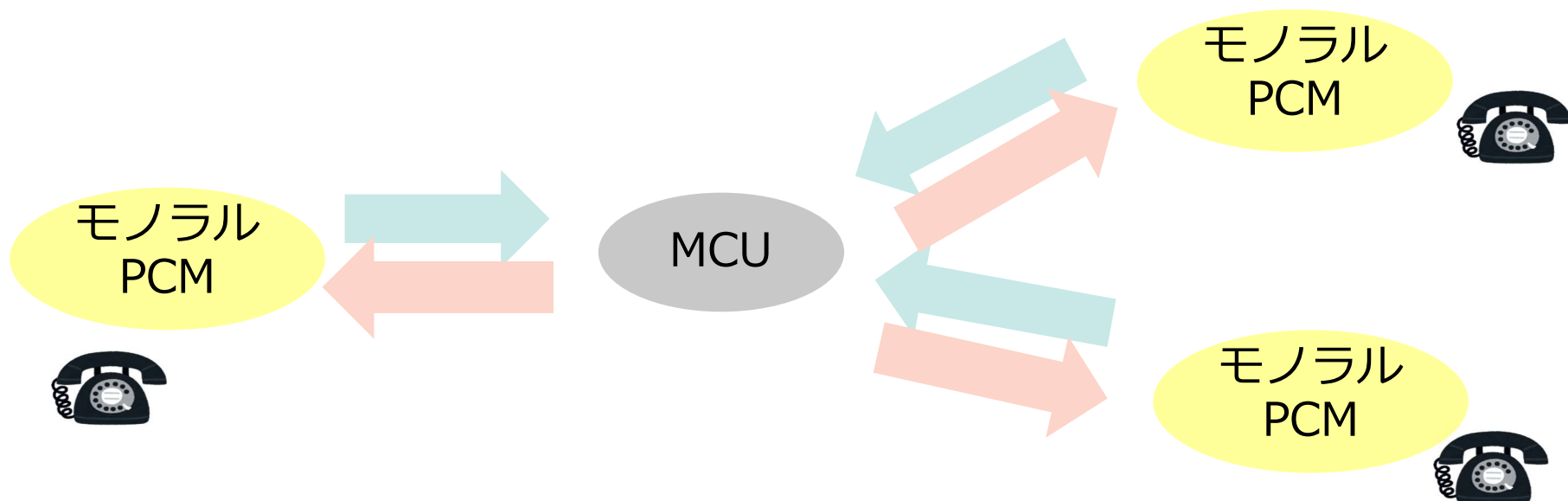


## 音声通信に適したパケット

- 優先制御
- 短いパケット



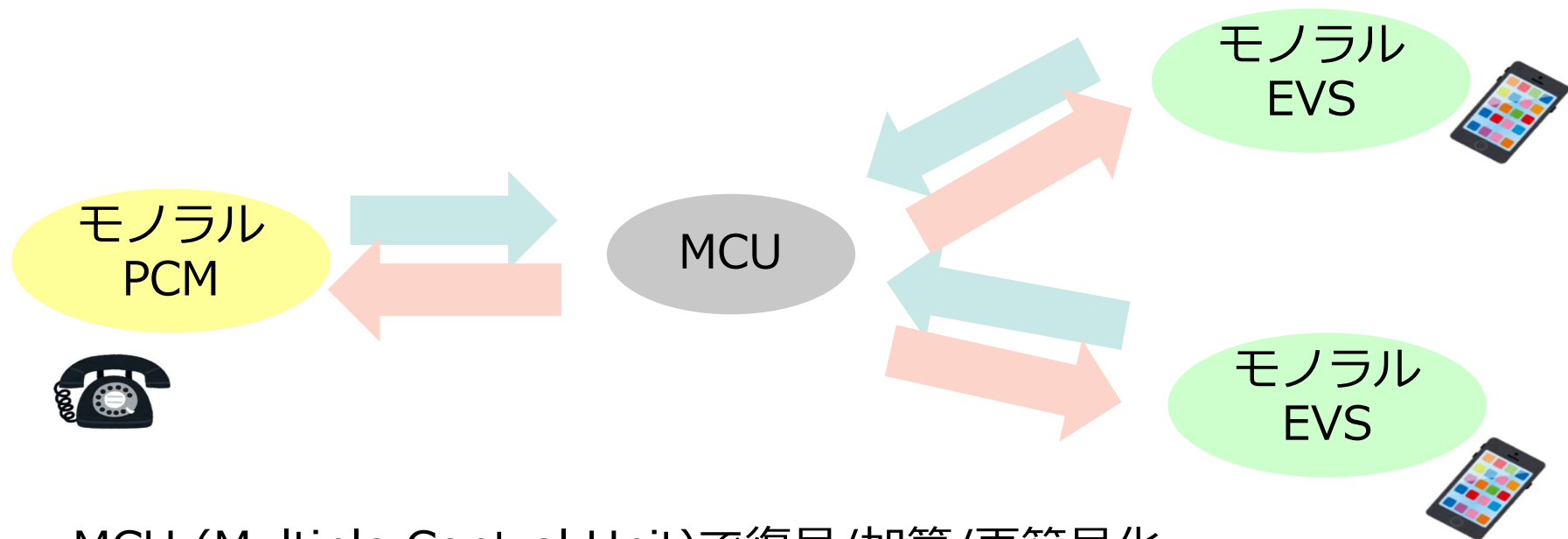
# 多地点会議電話（固定電話時代）



- MCU (Multiple Control Unit)で復号/加算/再符号化
- 復号/加算/再符号化による劣化は小さく、処理は軽い

## ◆ 音声通話のみ

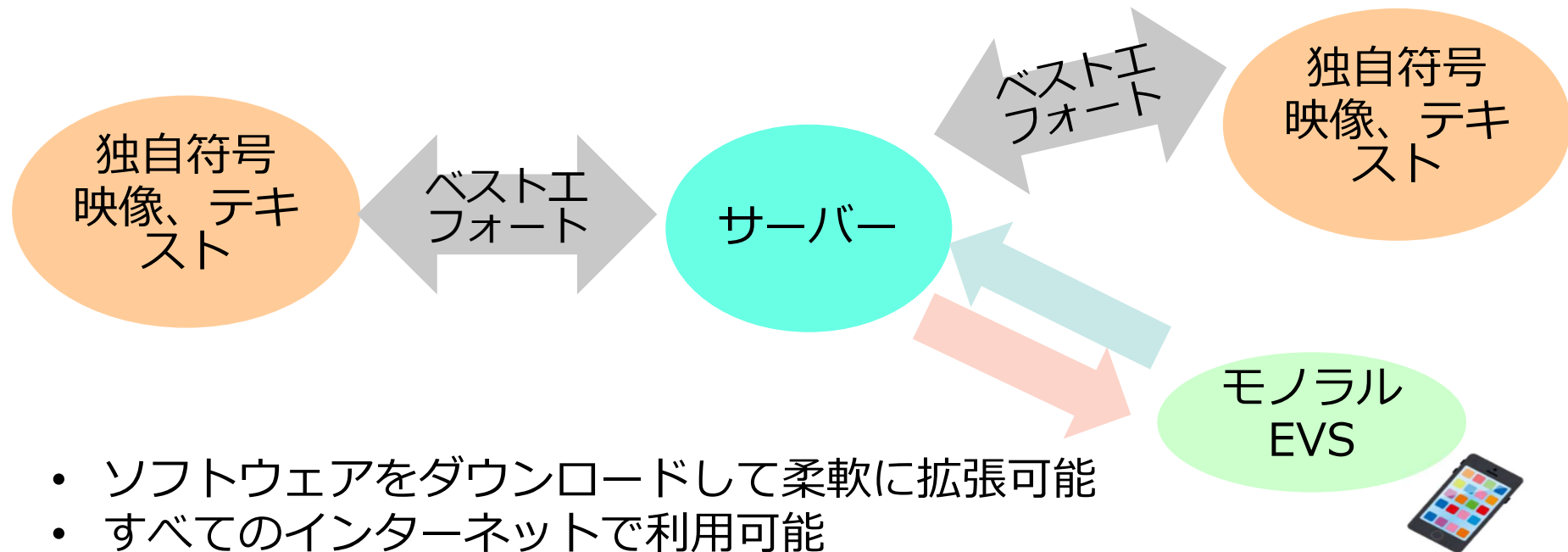
# 多地点会議電話（混在時代）



- MCU (Multiple Control Unit)で復号/加算/再符号化
- 各回線はモノラルであるが優先制御（4G/NGN/回線）

◆ EVS復号、加算、再符号化による品質劣化や遅延が大きい

# 多地点会議電話（ソフトウェア）



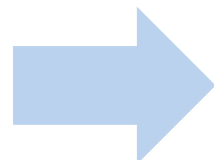
- ソフトウェアをダウンロードして柔軟に拡張可能
- すべてのインターネットで利用可能
- 電話とは符号変換

◆ 通話品質(音質、遅延)が保証されない  
◆ ネットワークの高速化では解決しない

# 情報伝達手段（電話から手段の多様化）



- 固定電話
  - 黒電話
  - 公衆電話
- 移動電話
  - 自動車
  - 肩掛け
  - ポケット携帯
  - 腕時計型
  - スマートホン
  - コードレス



**多様なデバイス  
マルチモーダル、臨場感**

- 行動中
  - 眼鏡型
  - HMD
- 会議
  - ヘッドセット
- エンターテインメント
  - 大画面没入
  - ゲーム

**基本的音質が重要**

**高臨場感  
高機能  
柔軟拡張性**

**双方向の音声通信  
の品質を改善する  
手段がない**

**音声品質保証**



## まとめ



### 3. 現在の音声音響通信の問題

音質は満足か？

- 固定網と移動網の電話が独立
  - 機能拡張には膨大なコストがかかり、維持コストも必要
- 業務用通信はWeb会議にシフト
  - 多様なデバイス、多様な環境での無限の進化
- 誰も安定した音声通信を使えない、提供できない
  - 永年の先人の努力が失われる



# 目次



## 1. 音声通信の歴史

便利になったか？

## 2. 音声音響符号化の貢献

音質はよくなったか？

## 3. 現在の音声音響通信の問題

音質は満足か？

## 4. 将来の音声音響通信

使う人たちのためか？

# 音声品質を保証する高機能電話



## Beyond 5G課題 (3GPP Rel.18)

- WebRTCと3GPP 優先制御の  
**連携プロトコル**
- 多地点、高臨場  
**IVAS符号化**

高臨場感  
高機能  
柔軟拡張性

WebRTC

連携プロトコル  
IVAS符号化

3GPP優先制御

音声品質保証

OTT  
ベンダの  
仕事

電話会社  
の仕事

# 連携プロトコルの位置づけ



## WebRTCの規格

- 実時間通信のための共通基本ツール
- 多様なサービスに拡張性が高い



## eiRTCW (extension of immersive Real-time Comm. for WebRTC)

- アプリケーションソフトから優先制御を使う
- 制御プロトコルの共通化、インタフェースの公開

## 3GPP IMS (IP Multimedia Subsystem) の規格

- 優先制御などのプロトコル群
- 移動網(VoLTE)、固定網(NGN) の電話の基盤

# IVASの目標



(EVS Extension for **I**mmersive **V**oice and **A**udio **S**ervices)

\* 2023年中の仕様凍結めざし、オープン開発中

- **立体音場の収音、圧縮伝送、再生による高臨場感双方向通信**

\* EVSはモノラル通信

- **複数ストリームと再生合成機能による多地点双方向通信**

\* EVSは1対1の地点の通信

- **相互接続機能によるEVSとの符号変換なし双方向通信**

\* EVSは世界中に普及

# IVASの公開開発



- 開発中のソースコードはGitlab上での公開開発に移行  
<https://forge.3gpp.org/rep/ivas-codec-pc/ivas-codec>

IVAS Gitlab



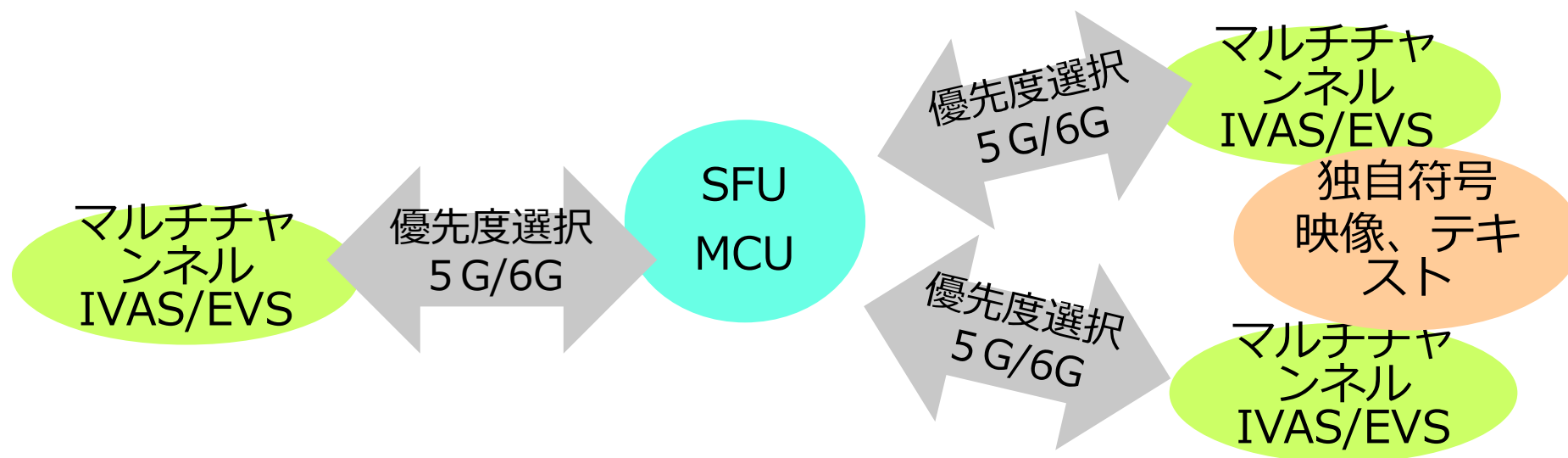
- 誰でも最新のソースコードをダウンロードできる
- 仕様の議論、ソースの改訂は共同開発契約会社限定  
(主観評価費用支払いなど規約に合意した11社のみ)
- 従来 of 標準化は、競争選定、閉鎖的共同開発



Copyright 2022 NTT CORPORATION

2年半ぶりの開発技術会議 2022/9/13-15

# 将来の多地点会議（音声優先とOTT）



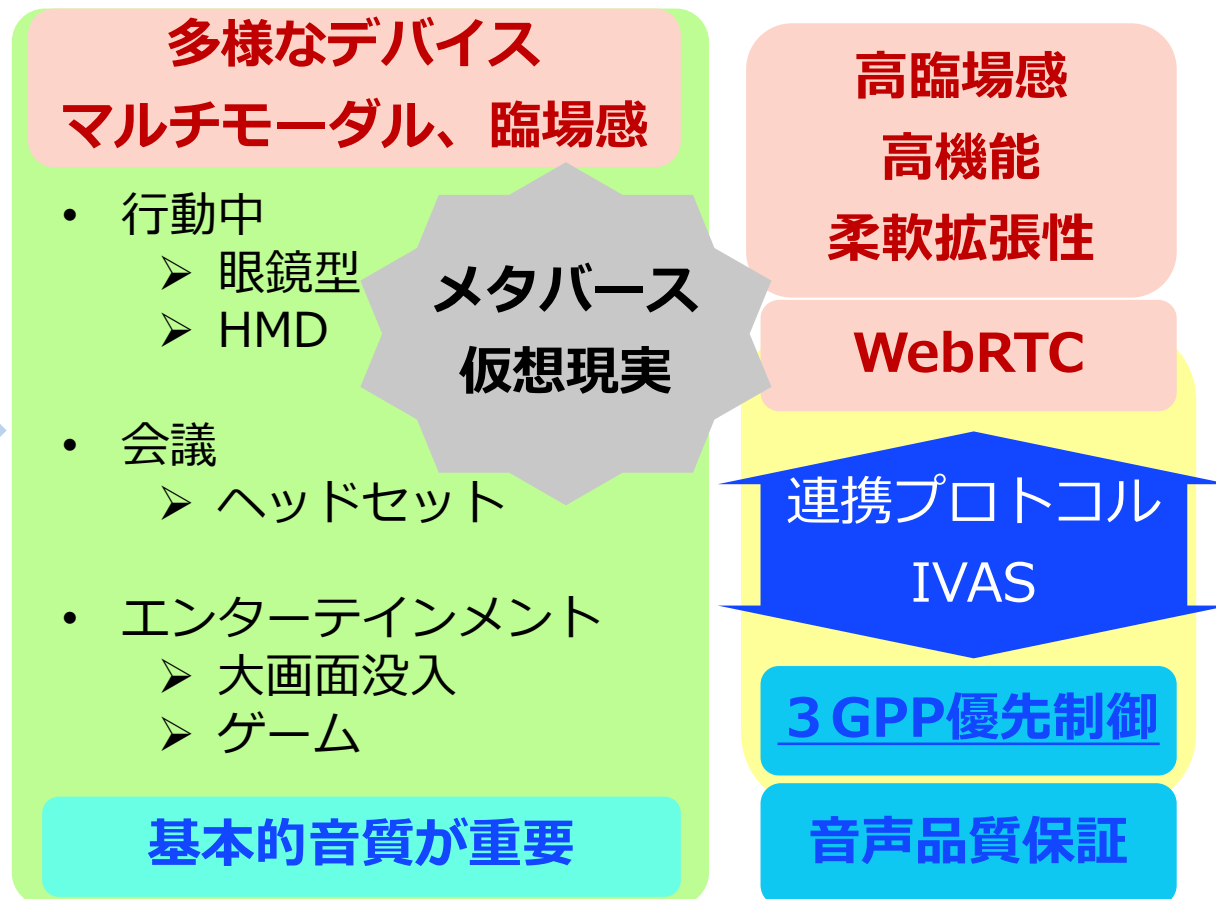
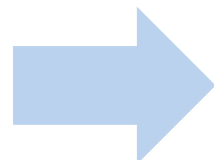
## **\*SFU (Select and Forward Unit)**

- 各地点からの高圧縮ストリームの中から許容本数を選択
- 選択したストリームを他地点に転送
- 複数ストリームから自分に合わせた音声を合成（IVASの機能）

# 情報伝達手段（電話から手段の多様化）



- 固定電話
  - 黒電話
  - 公衆電話
- 移動電話
  - 自動車
  - 肩掛け
  - ポケット携帯
  - 腕時計型
  - スマートホン
  - コードレス





## まとめ

# 4. 将来の音声音響通信

使う人のためか？

### 使う人

- 多様なサービスや製品を使うことができる
- 基本音声品質を維持と無料電話を選択できる

### 作る人

- 多様なサービスや製品を作成できる
- 基本音声品質を維持できる

### 電話会社

- 優先制御を有料で提供し、基本音声品質を保証



## まとめ



1. 音声通信は便利になった
2. 電話の音質はよくなった
3. 現在の多様な音声通信の音質はよくない
4. 将来の音声通信は使う人たちのために

### 使う人の立場に立った音声通信の未来

- 通信の原点は音声通信で、その音質は最優先であるべき
- 高速ネットワークだけでは高音質にはならない
- 音声通信の品質に責任をもつ会社が必要

# 質疑応答、補足



# お客さまの声



## 音声品質が悪い

- 送り側で波形と照合して、番号を送り、受け側で再構成する
- 自分の息子の声さえ区別がつかない => 詐欺の原因
- 電話音声認識サービスの認識率低下 => サービス低下

## 圧縮する必要はない

- 非圧縮なら品質は最高

## 音声品質は十分

- 新たな研究や高臨場より、料金を安く

# 音声符号化と音響符号化の対比



	音声符号化	音響符号化
サンプリング	8 kHz	48 kHz
処理	時間領域	周波数領域
圧縮	予測符号化	変換符号化
用途	双方向通信	片方向通信
チャンネル	モノラル 単音	ステレオ 複合音
フレーム長	20 ms	80 ms
学会	speech 情報システム/音声B	audio and acoustics 基礎境界/電気音響

# LSPの発明からの世界への普及



# EVSの位置づけ



	電話用 符号化	放送用 符号化	音声音響 符号化	低遅延音 響符号化	VoLTE用 符号化	
レート [kbps]	10	48	16	24	5.9-128	高圧縮
遅延 [ms]	30	80	80	30	32	低遅延
帯域 [kHz]	4	24	24	24	24	広帯域
音楽品質	劣化	良好	良好	良好	良好	高音質
標準例	ITU-T G.729 3GPP AMR	MPEG MP3 MPEG AAC	AMR- WB+ MPEG USAC	AAC-ELD	3GPP EVS	