

# ワイヤレスセマンティックコミュニケーション によるデジタルツインの構築

2024年3月19日

情報工学・人間科学系 梅原大祐  
電気電子工学系 高井伸和

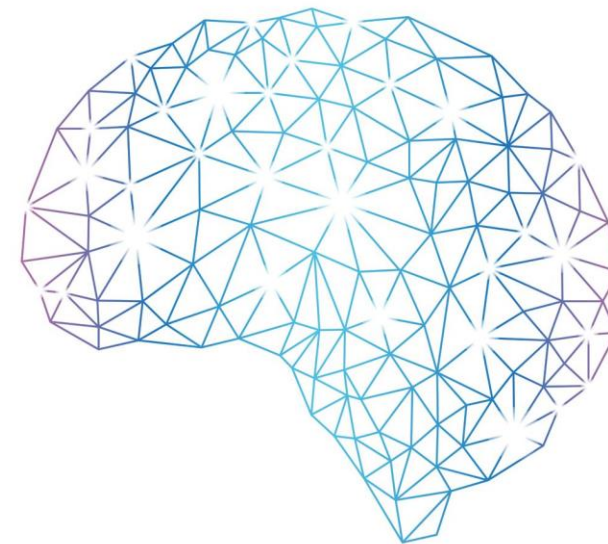
# デジタルツインの課題

- 現実空間（フィジカル空間）を無線センシング/モニタリング
  - 果たして現実空間の情報を常時送るだけの無線回線容量はあるのか？

フィジカル空間



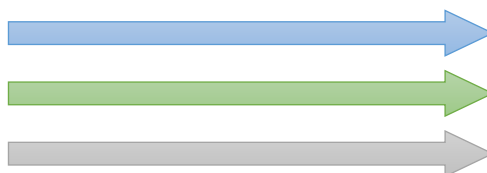
サイバー空間  
(デジタルツイン)



医療DX/ヘルスケアICT  
マテリアルDX  
産業/農業ICT  
エンターテインメント

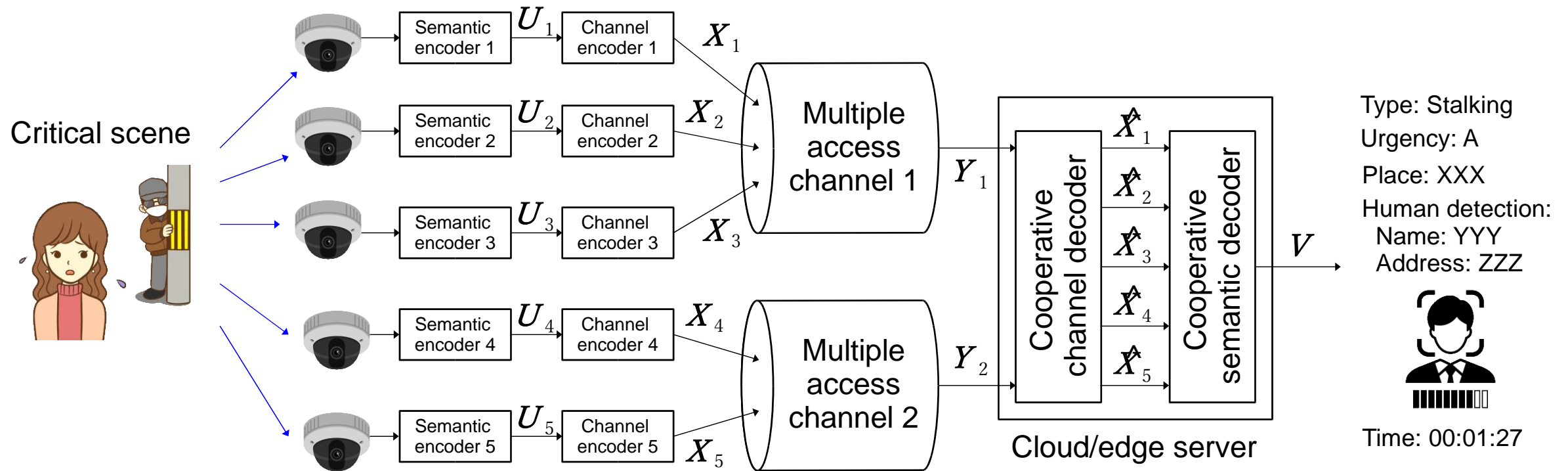
6G/IoT  
プライバシー保護  
センシング/モニタリング

機械学習/AI  
クラウド/エッジコンピューティング  
サイバーセキュリティ  
量子ICT



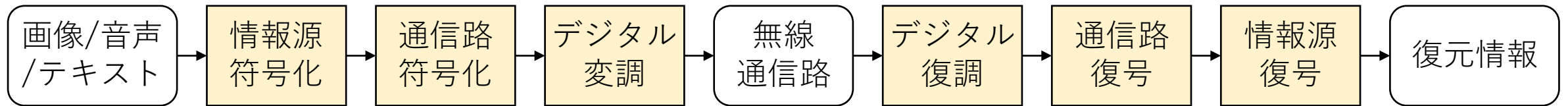
# セマンティックコミュニケーションの適用

- インシデントを検知するセマンティックコミュニケーションの例
  - 複数のセンサ：インシデントを検知したときに情報をサーバに伝送
  - クラウド/エッジサーバ：複数のセンサ情報からインシデントを読み取る

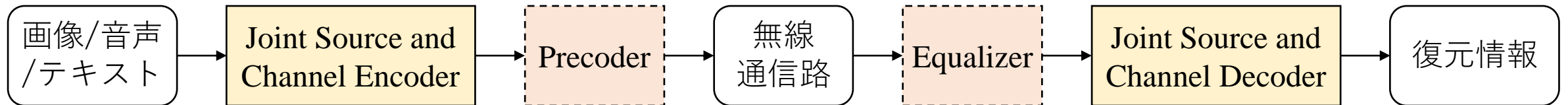


# セマンティックコミュニケーションとは

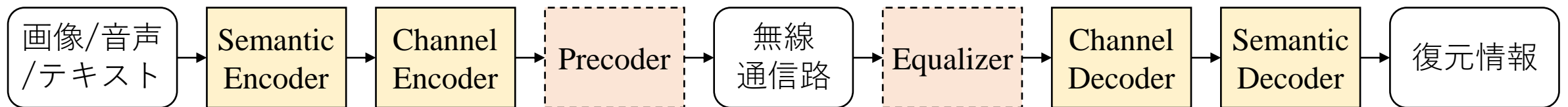
- 現在の無線通信システムのブロックダイアグラム



- Joint Source and Channel Coding (JSCC) のブロックダイアグラム

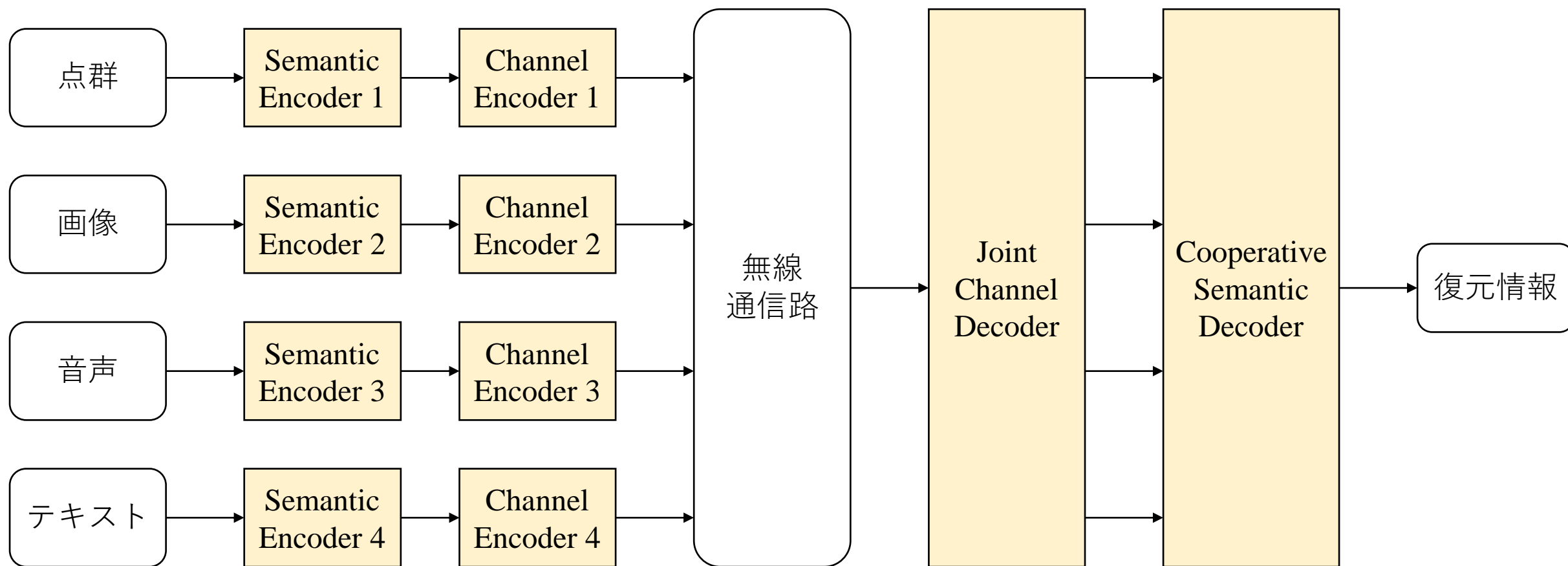


- セマンティックコミュニケーションのブロックダイアグラム



# 協調セマンティックコミュニケーション

- 協調セマンティックのブロックダイアグラム



# 2023年度の取り組み

- 既存のセマンティックコミュニケーションを実装，その特性を検証
  1. 無線音声伝送におけるセマンティックコミュニケーション
    - ✓ 梅原研究室のB4が担当，DeepSC-S [1] の性能を評価
    - ✓ 異言語音声間の汎化能力について検証
    - ✓ コミュニケーションシステム研究会（3/14）にて研究発表
  2. 無線画像伝送におけるセマンティックコミュニケーション
    - ✓ 梅原研究室のM1が担当，WITT [2] の性能を評価
    - ✓ 情報工学専攻の中間発表（1/19）にて研究の進捗状況を発表
  3. PAPRを低減する無線セマンティックコミュニケーション
    - ✓ 高井研究室のM1が担当，DeepJSCC-Q [3] を検証中

[1] Z. Weng *et al.*, “Semantic communication systems for speech transmission,” IEEE JSAC, vol. 39, no. 8, 2021.

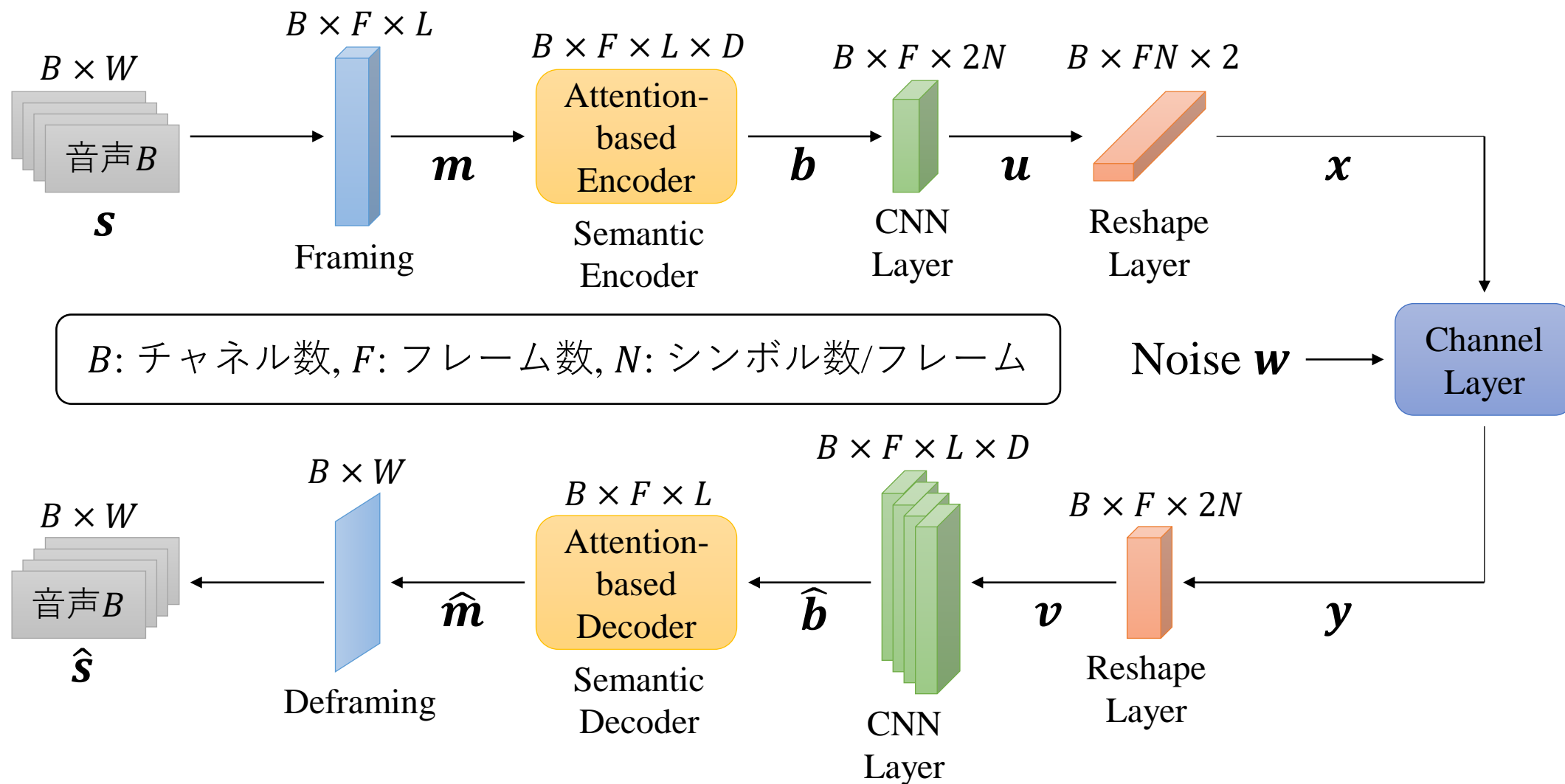
[2] K. Yang *et al.*, “WITT: A wireless image transmission transformer for semantic communications,” Proc. IEEE ICASSP 2023, 2023.

[3] T.-Y. Tung *et al.*, “DeepJSCC-Q: Constellation constrained deep joint source-channel coding,” IEEE JSAIT, vol. 3, no. 4, 2022.

# DeepSC-S の異言語間汎化能力の検証

# 学習システムアーキテクチャ

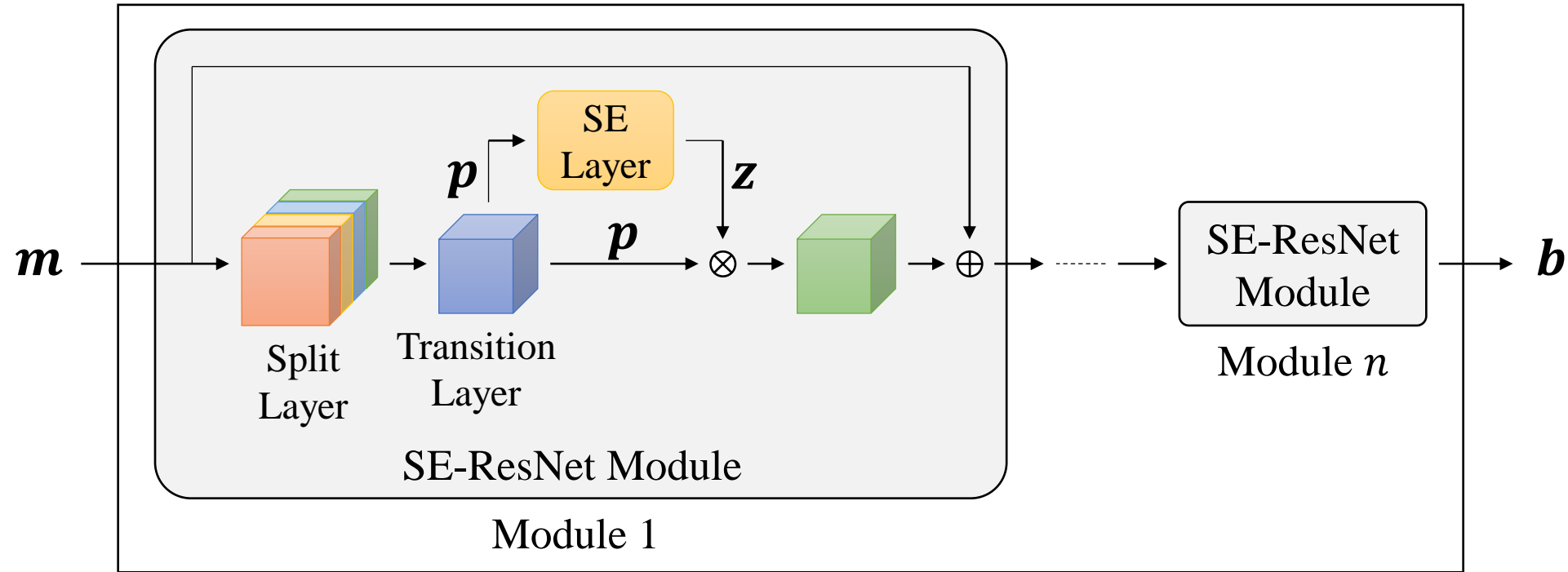
- DeepSC-S 学習システムアーキテクチャ



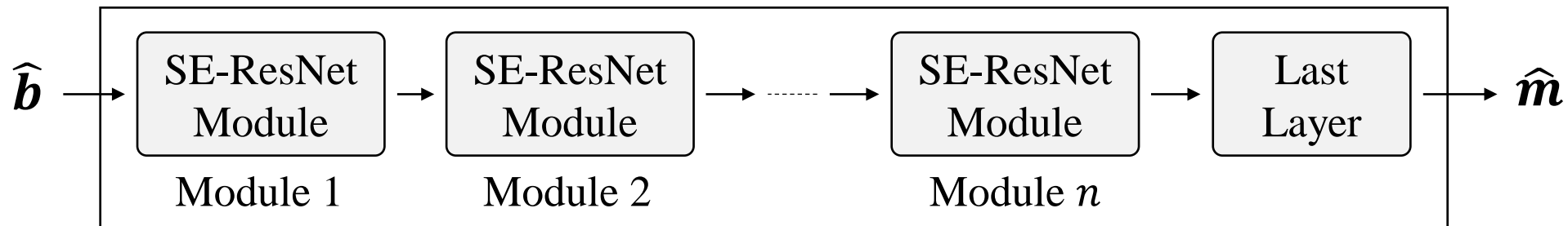


# Attention-based Semantic Encoder/Decoder

- Attention-based semantic encoder

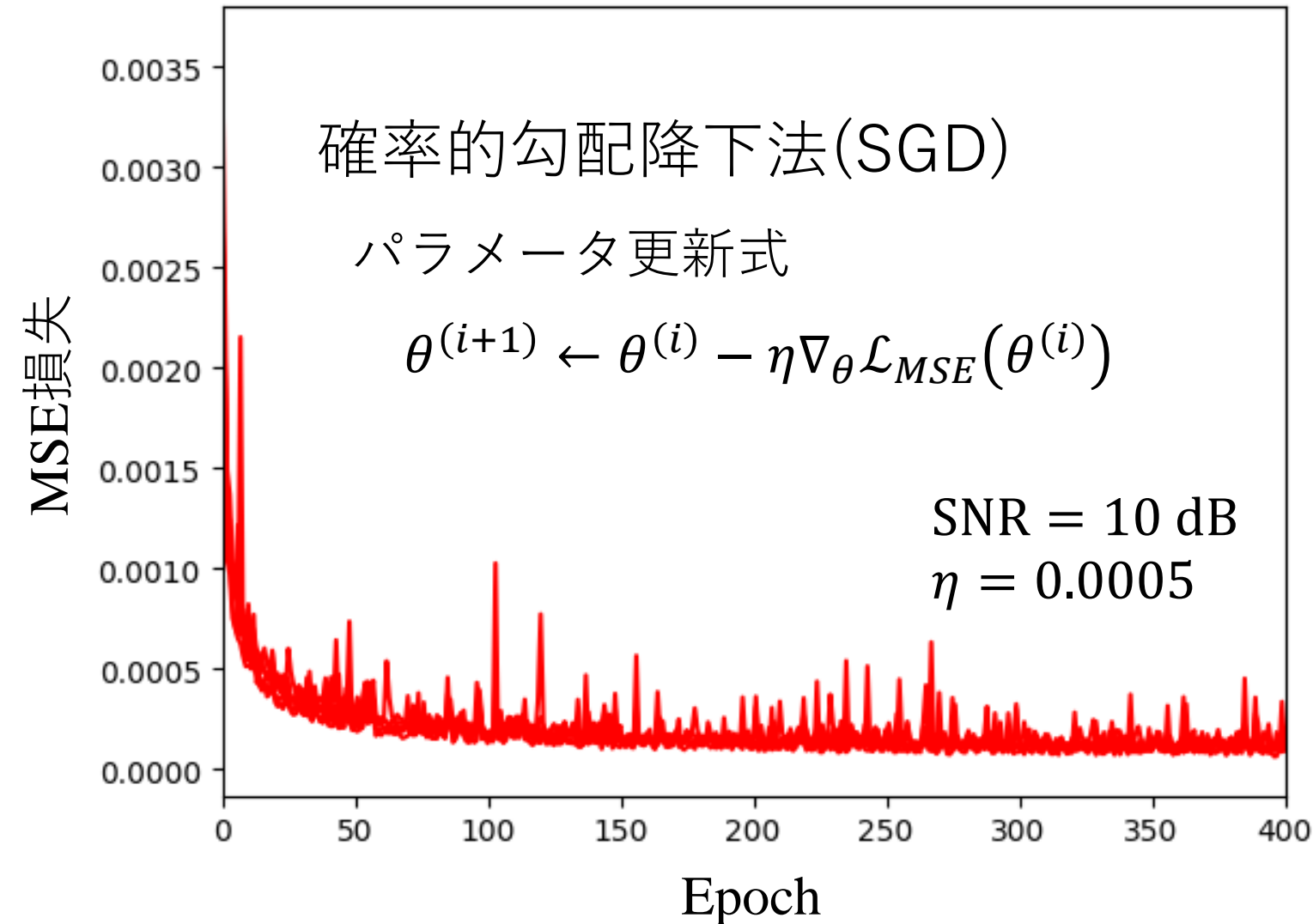


- Attention-based semantic decoder



# Encoder/Decoder の学習

- MSE損失関数による確率的勾配降下法 (SGD)



平均二乗誤差(MSE)

$$\mathcal{L}_{MSE}(\theta) = \frac{1}{W} \|\mathbf{s} - \hat{\mathbf{s}}\|^2$$

$\mathbf{s}$ : 元の音声信号系列

$\hat{\mathbf{s}}$ : 復号した音声信号系列

$W$ : 音声信号系列の長さ

$\eta > 0$ : 学習率

$\theta = (\theta^T, \theta^R)$ : ネットワークパラメータ

$\theta^{(i)}$ :  $i$ -th epoch のパラメータ値

# データセットと評価指標

- データセット

- 英語音声：男女2名による1-3秒の音声，824ファイル，48 kHzサンプリング
- 日本語音声：女性3名による5-10秒の音声，300ファイル，48 kHzサンプリング

- 性能評価指標

- SDR (Signal-to-Distortion Ratio, 信号対歪み比)
- PESQ (Perceptual Evaluation of Speech Quality) [4]
  - ✓ 知覚・認知モデリングにより音声品質の劣化量を計算してMOS値を推定

$$\text{SDR} = 10 \log_{10} \left( \frac{\|\mathbf{s}\|^2}{\|\mathbf{s} - \hat{\mathbf{s}}\|^2} \right)$$

- 評価項目

	実験1	実験2	実験3	実験4
学習	英語音声	日本語音声	英語音声	日本語音声
テスト	英語音声	英語音声	日本語音声	日本語音声

[4] ITU-T P.862, *Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-band Telephone Networks and Speech Codecs*, Feb. 2001.

# 学習モデル諸元

- エンコーダ/デコーダの学習モデル

	Layer name	Filters	Activation
Tx	6 × SE-ResNet	6 × 64	ReLU
	CNN layer	8	None
Rx	CNN layer	8	ReLU
	6 × SE-ResNet	6 × 64	ReLU
	Last layer (CNN)	1	None

- 学習パラメータ

Parameter	Value
サンプルレート	48 kHz
サンプル長 $W$	16,384
バッチサイズ $B$	32
フレーム数 $F$	128
フレーム長 $L$	128
学習率 $\eta$	0.0005
学習エポック数	400
学習SNR	0, 5, 10, 15, 20 dB

# 復元した音声信号 (SNR = 10 dB)

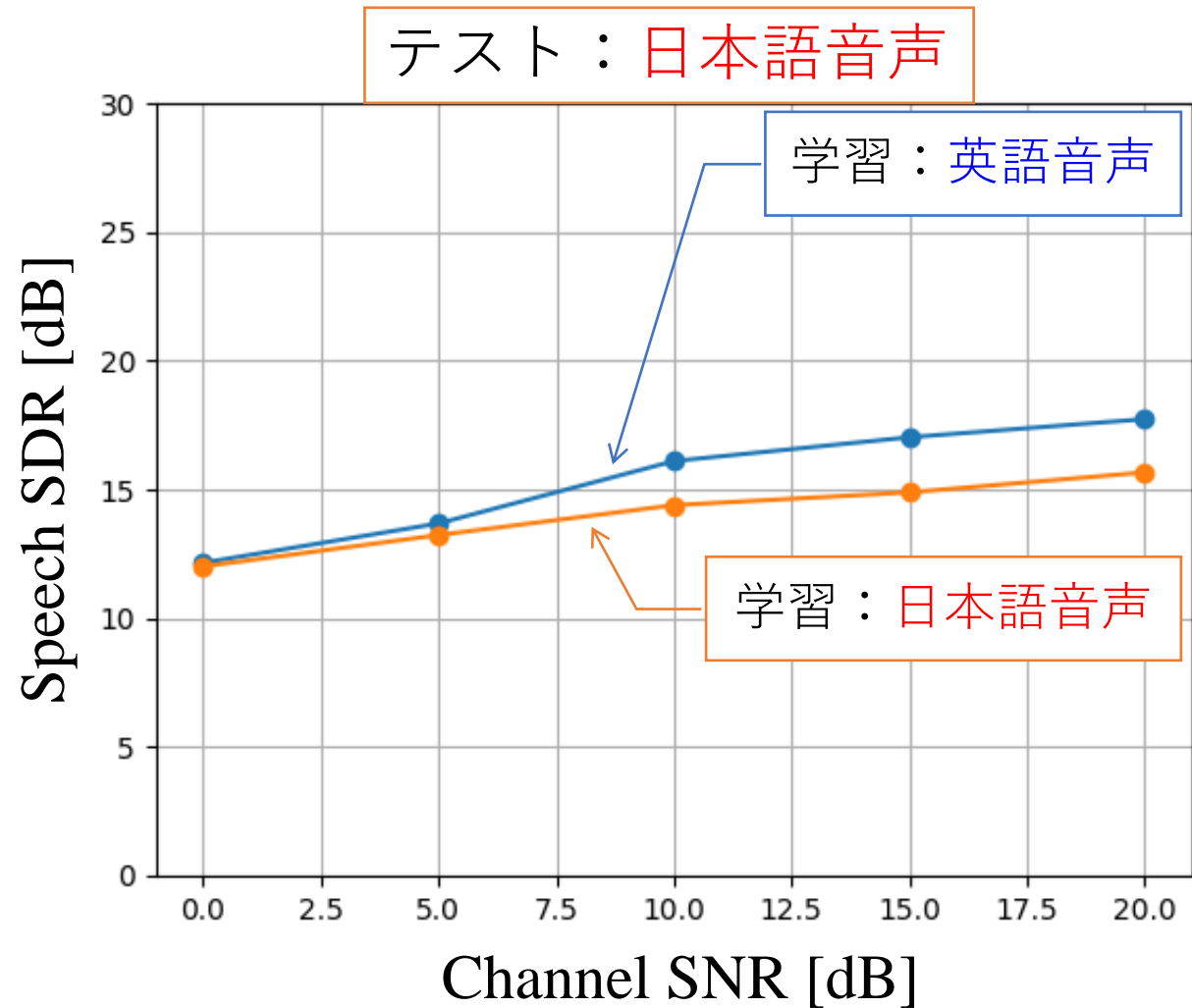
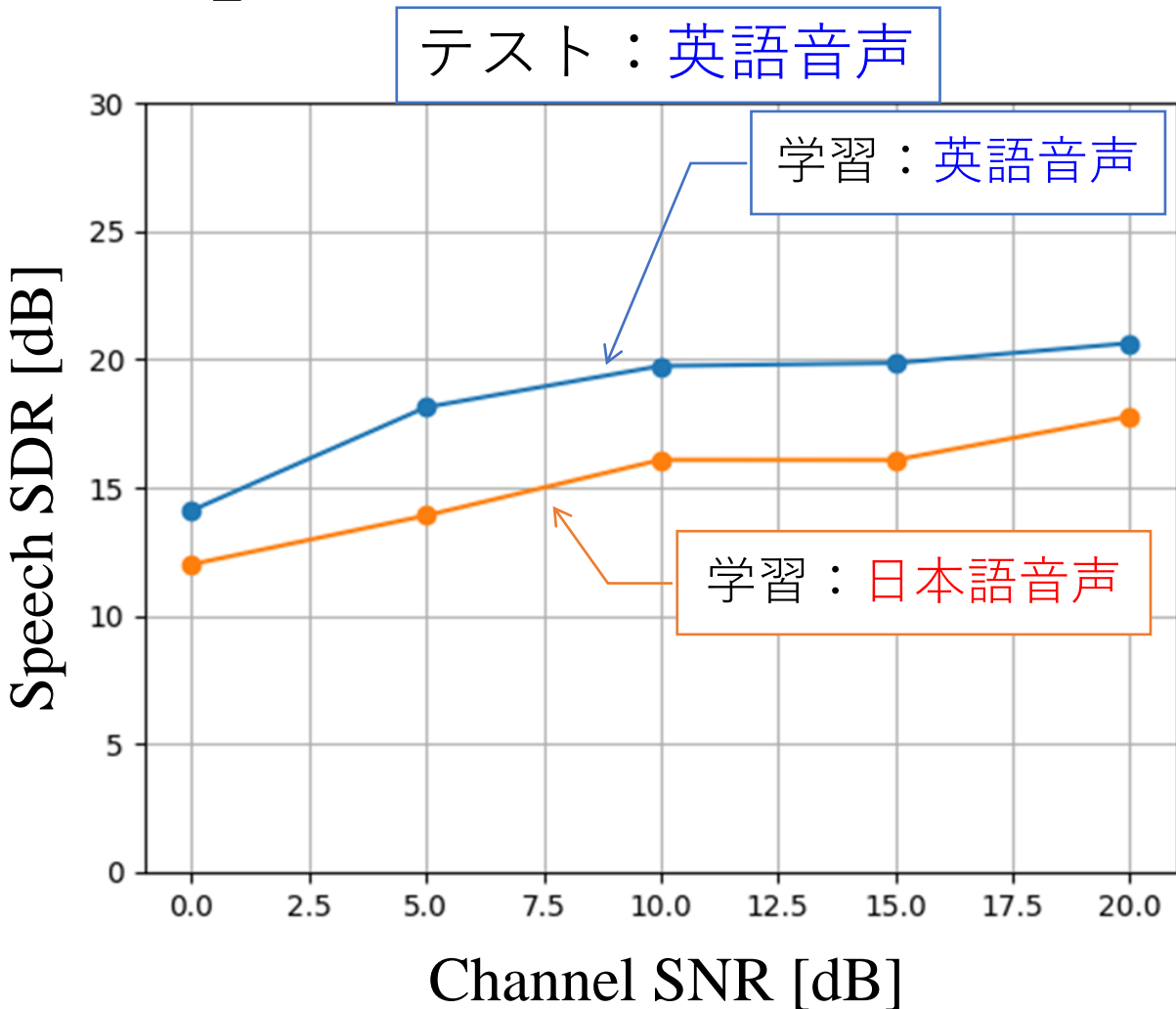
- 英語音声

オリジナル	学習：英語音声	学習：日本語音声
		

- 日本語音声

オリジナル	学習：英語音声	学習：日本語音声
		

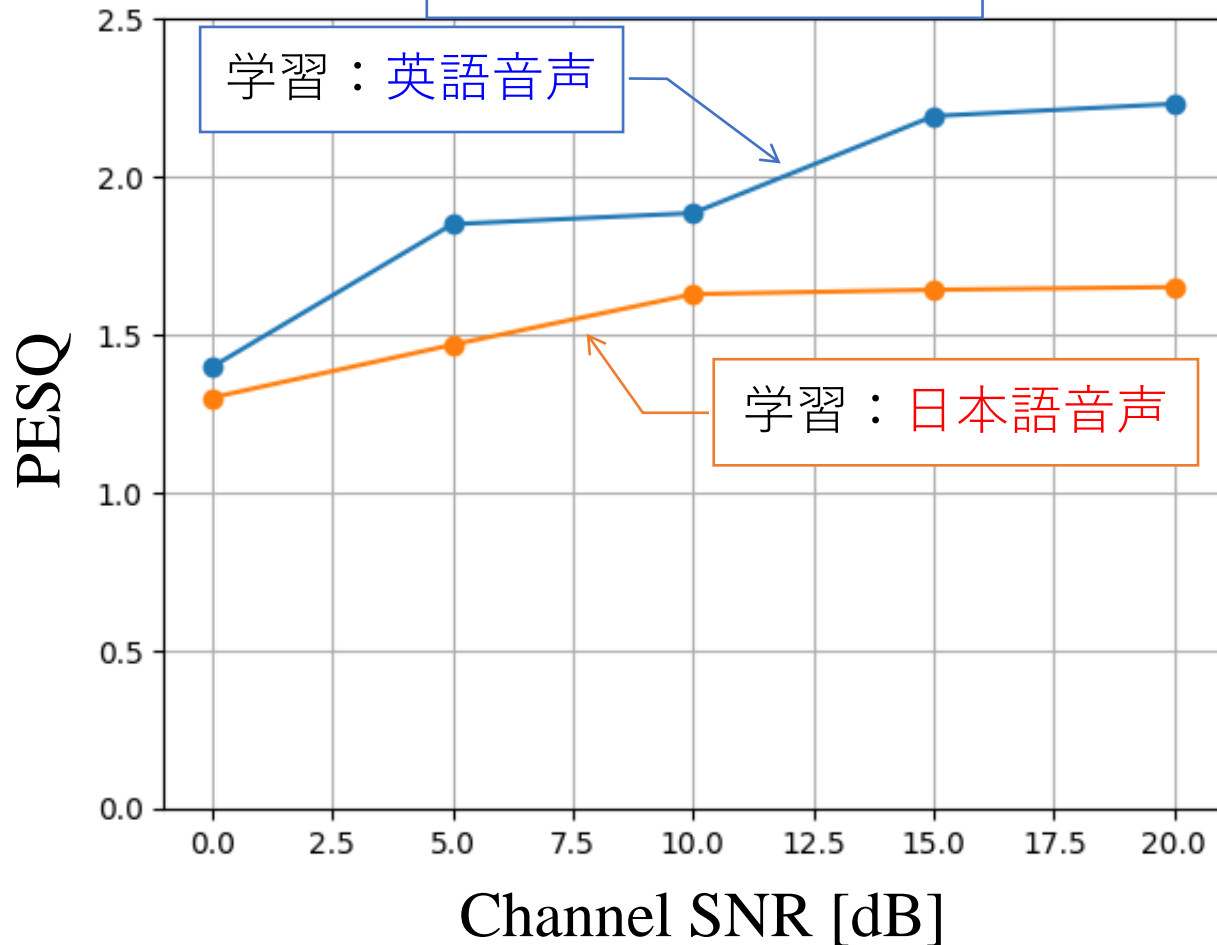
# Speech SDR vs Channel SNR



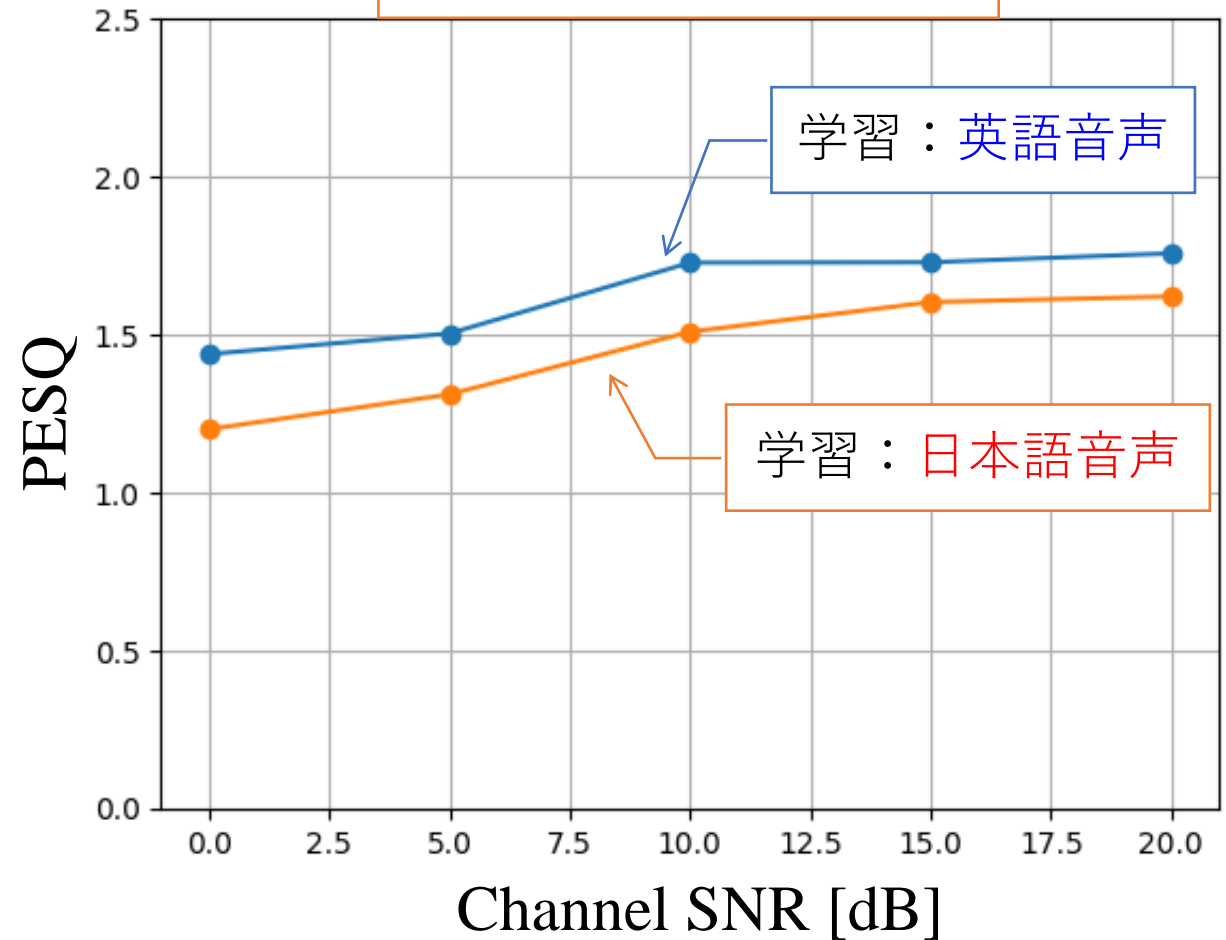
- どちらの言語音声で学習しても日本語音声の復元忠実度が低い
- 日本語音声で学習したときに英語・日本語音声の復元忠実度が低い

# PESQ vs Channel SNR

テスト：英語音声



テスト：日本語音声



- どちらの言語音声で学習しても日本語音声の復元品質が低い
- 日本語音声で学習したときに英語・日本語音声の復元品質が低い

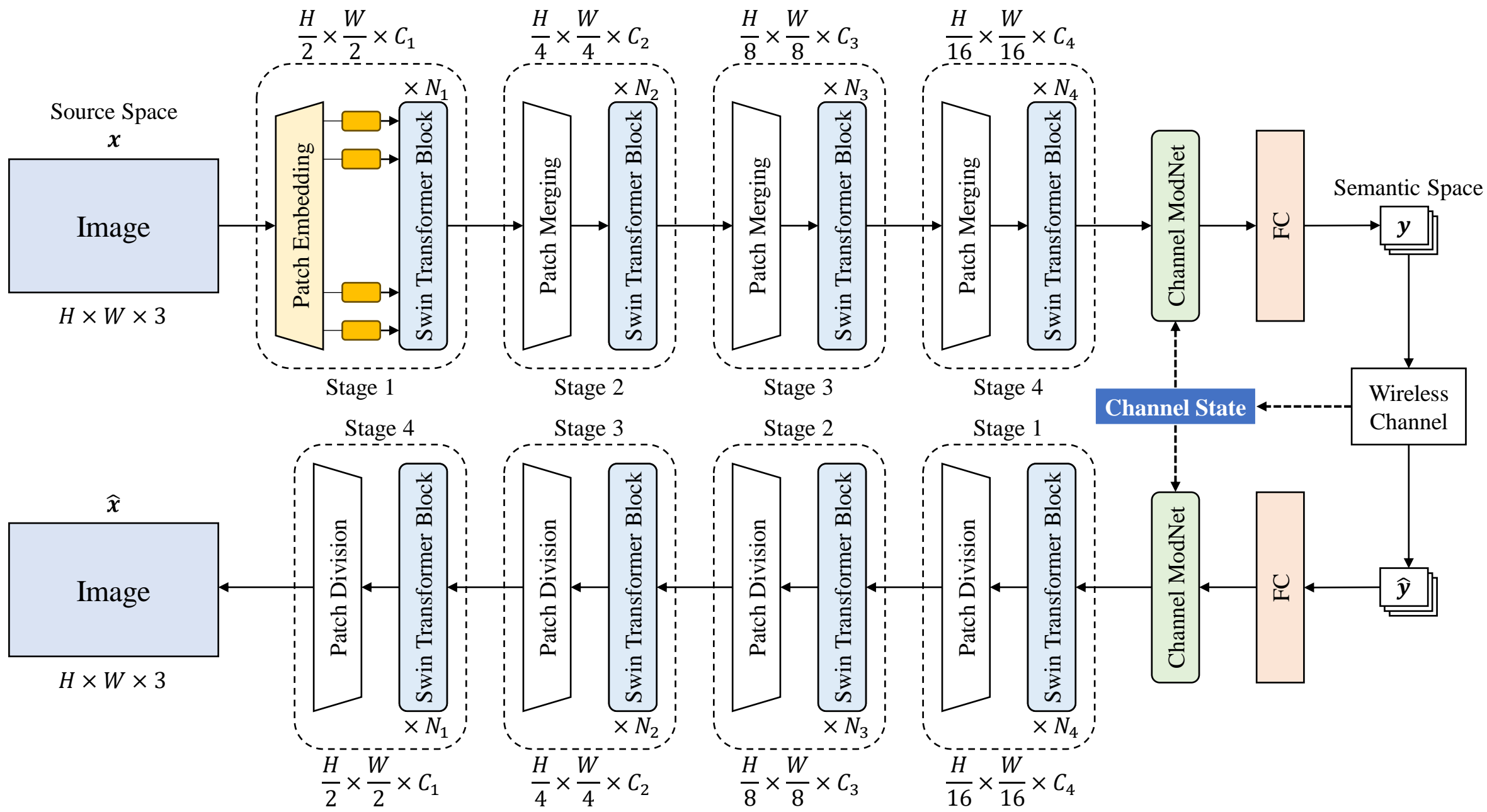
# DeepSC-S の検証のまとめ

- 今回の検証では，異言語音声間の汎化能力が十分ではなかった
  - 適切な学習データセットについての検証
  - 適切なハイパーパラメータチューニングの必要性
  - SE-ResNetのAttention機能の解明
- 今後に向けての取り組み
  - 音声信号伝送以外の異なるタスクの検証
    - ✓ 音声から抽出したテキストを無線伝送してサイバー空間でテキストを復元
    - ✓ 音声から声紋鑑定に必要となる情報を抽出してサイバー空間で声紋鑑定



# WITT (Wireless Image Transmission Transformer) の性能検証

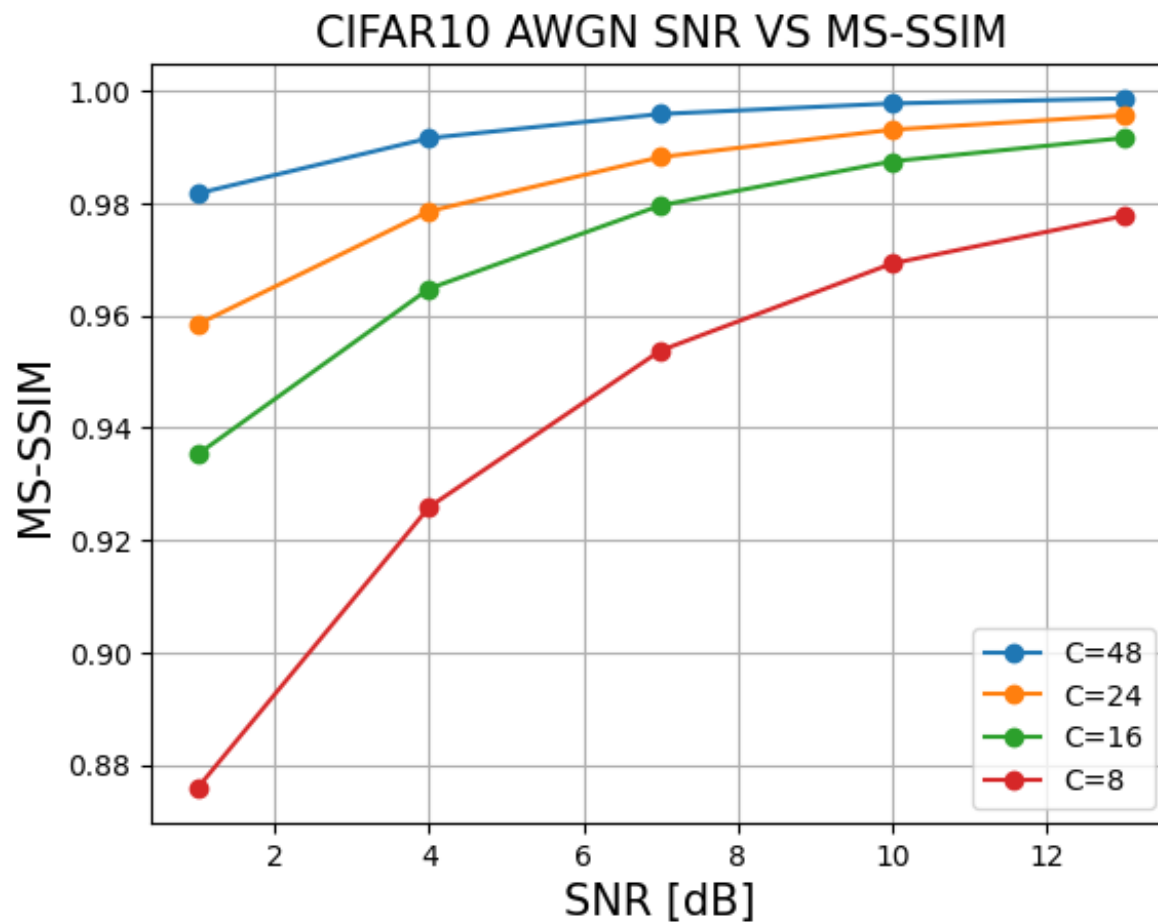
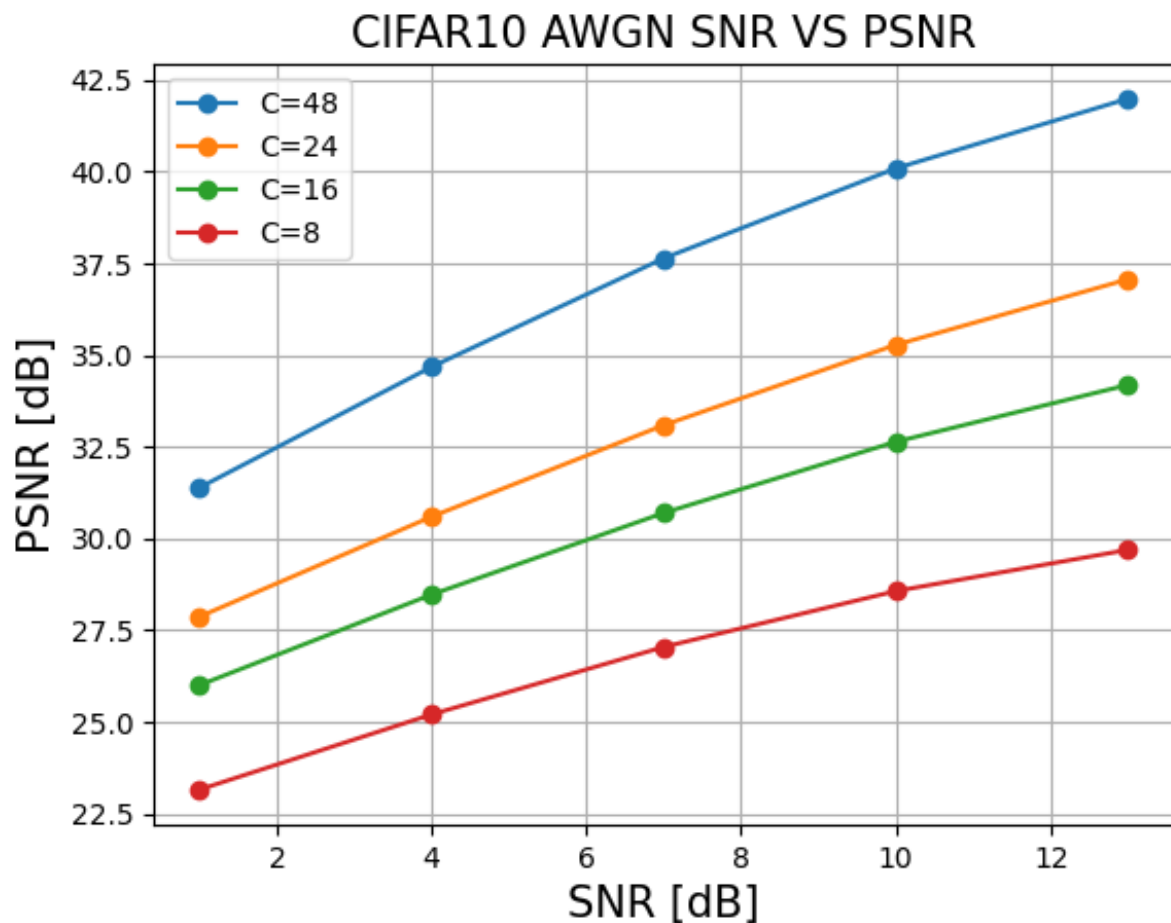
# WITTシステムアーキテクチャ



# CIFAR-10 PSNR, MS-SSIM vs AWGN SNR

- $C$ はチャンネル数, 学習時のSNRは10 dB

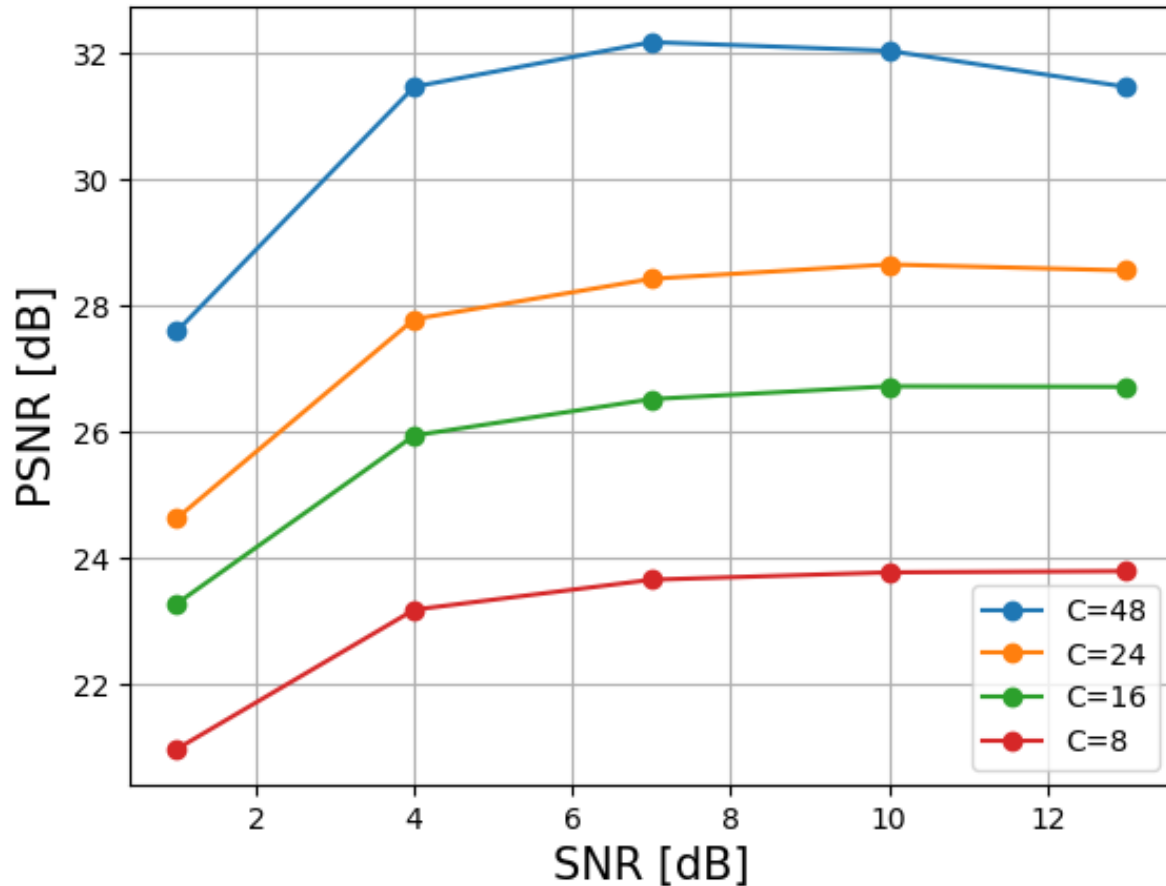
➤ Channel ModNet で異なるSNRでも追従できている? ⇒ 更なる検証が必要



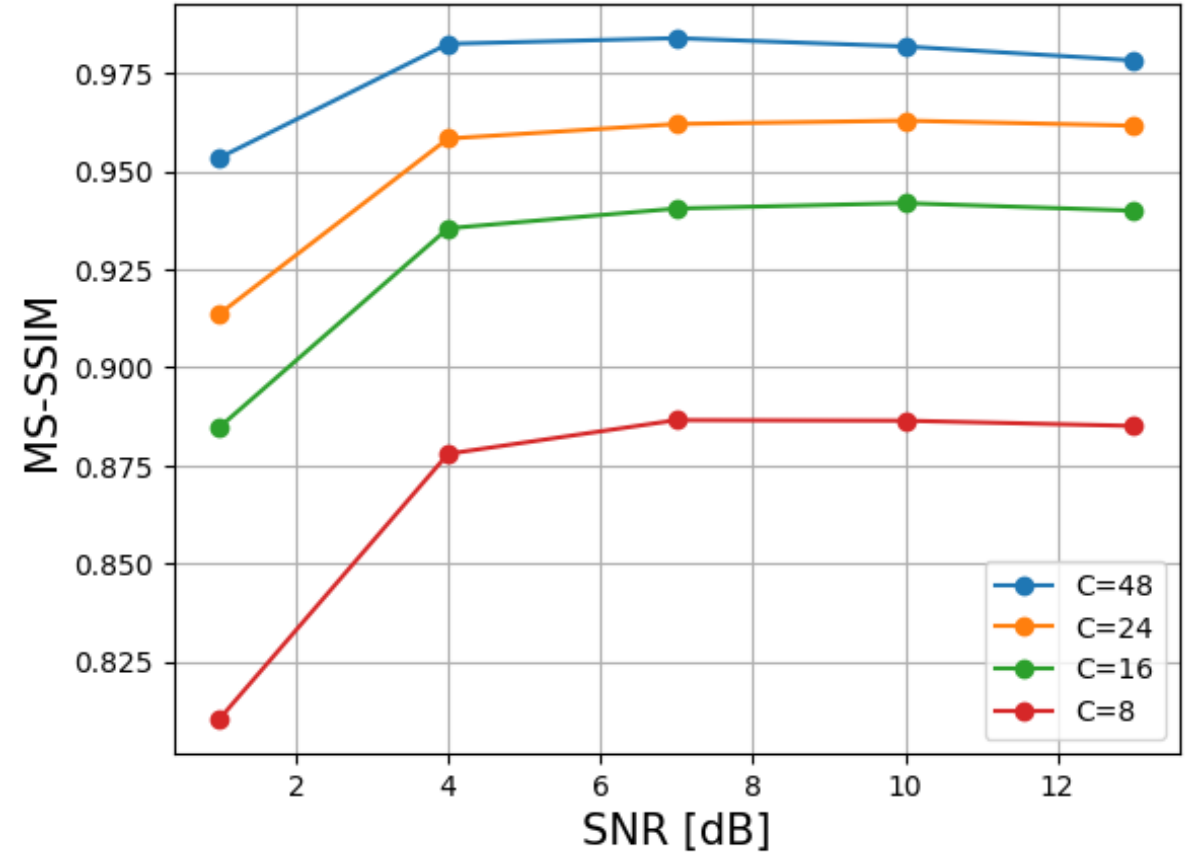
# CIFAR-10 PSNR, MS-SSIM vs Rayleigh SNR

- $C$ はチャンネル数, 学習時のSNRは3 dB
  - SNRの変化に追従できていない ⇒ 無線チャンネル補償技術の研究開発の必要性

CIFAR10 Rayleigh SNR VS PSNR



CIFAR10 Rayleigh SNR VS MS-SSIM



# WITT の検証のまとめ

- 今回の検証では，無線チャネルの CSI の変化に追従できていなかった
  - 新たな無線チャネル補償技術の研究開発を進める
- 今後に向けての取り組み
  - 画像伝送以外の異なるタスクの検証
    - ✓ 不審車両，不審人物の特定に必要な情報を抽出してサイバー空間に伝送
    - ✓ プライバシーを保護するための画像加工を施してサイバー空間に伝送

# 今後のプロジェクトの推進に向けて

- ワイヤレスセマンティックコミュニケーションについて
  - 無線デバイスと無線チャネルにおける要件を満足するための学習モデルの開発
    - ✓ 無線デバイスの電力効率の向上のためにはPAPRの低い信号が望まれる
    - ✓ RF回路の不完全性を補償する必要がある
    - ✓ 無線チャネルの利用にはそのスペクトラムについて制限がある
- デジタルツインの構築に向けて
  - サイバー空間にとって必要となる情報とは？
  - サイバー空間のどの情報をどの程度の頻度で更新するのか？
  - センシングした点群，画像，音声，テキストなどをどのように融合するのか？