深層学習に基づく一括符号化変調を用いた水中音響画像伝送

*井上 文彰[1]，久野 大介[1]，丸田 一輝[2]，原 祐子[3]，中山 悠[4]

大阪大学大学院工学研究科[1]
東京工業大学 超スマート社会卓越教育院[2]
東京工業大学 工学院情報通信系[3]
東京農工大学 大学院工学研究科[4]

一括符号化 (JSCC)

- 一般的な画像伝送システム

  - 2 段階で符号化 (情報圧縮，冗長性付与)

- 一括符号化 (**JSCC**) **Joint Source-Channel Coding**

  - 送信画像を直接，伝送ビット列に変換

一括符号化変調 (JSCCM)

一括符号化変調 (**JSCCM**) **Joint Source-Channel Coding and Modulation**

- **エンコーダ** $f_{\mathrm{enc}} : \mathbb{R}^{H \times W \times 3} \to \mathbb{C}^{K \times 1}$
  - 入力された**画像データ** $\boldsymbol{x}$ に対し**複素シンボル列** $\boldsymbol{s}$ を出力
- **デコーダ** $f_{\mathrm{dec}} : \mathbb{C}^{K \times 1} \to \mathbb{R}^{H \times W \times 3}$
  - 入力された**複素シンボル列** $\tilde{\boldsymbol{s}}$ に対し**画像データ** $\tilde{\boldsymbol{x}}$ を出力

Deep JSCCM

深層学習に基づく一括符号化変調 (**Deep JSCCM**)[1]

- 多層ニューラルネットワークによるパラメータ化
  - エンコーダ関数 $f_{\mathrm{enc},\boldsymbol{\theta}}$，デコーダ関数 $f_{\mathrm{dec},\boldsymbol{\theta}}$
- 誤差逆伝播法により，パラメータ群 $\boldsymbol{\theta}$ を学習
  - エンコーダ入力画像とデコーダ出力画像の誤差を最小化
  - 確率的勾配降下法 (SGD: Stochastic gradient descent) を利用
- シミュレーション評価により，従来法に対する優位性を確認[1]

[1] E. Bourtsoulatze et al., in *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 3, pp. 567–579, 2019.

Deep JSCCM

深層学習に基づく一括符号化変調 (**Deep JSCCM**)[1]

- 多層ニューラルネットワークによるパラメータ化
  - エンコーダ関数 $f_{\mathrm{enc},\boldsymbol{\theta}}$，デコーダ関数 $f_{\mathrm{dec},\boldsymbol{\theta}}$
- 誤差逆伝播法により，パラメータ群 $\boldsymbol{\theta}$ を学習
  - エンコーダ入力画像とデコーダ出力画像の誤差を最小化
  - 確率的勾配降下法 (SGD: Stochastic gradient descent) を利用
- シミュレーション評価により，従来法に対する優位性を確認[1]
- ただし，既存研究では単純な理論的通信路のみを想定
  - 加法的白色ガウス雑音 (AWGN)，緩慢レイリーフェージング

本研究の概要

- **音響通信による水中画像伝送**への Deep JSCCM の応用
  - 伝送速度ならびに雑音耐性の向上を期待
- 以下の要素を新たに考慮した Deep JSCCM 設計法を提案
  - 長遅延マルチパス通信路
  - 送受信器における帯域制限フィルタ
  - マルチパス干渉を補償する (時間領域) 動的等化器

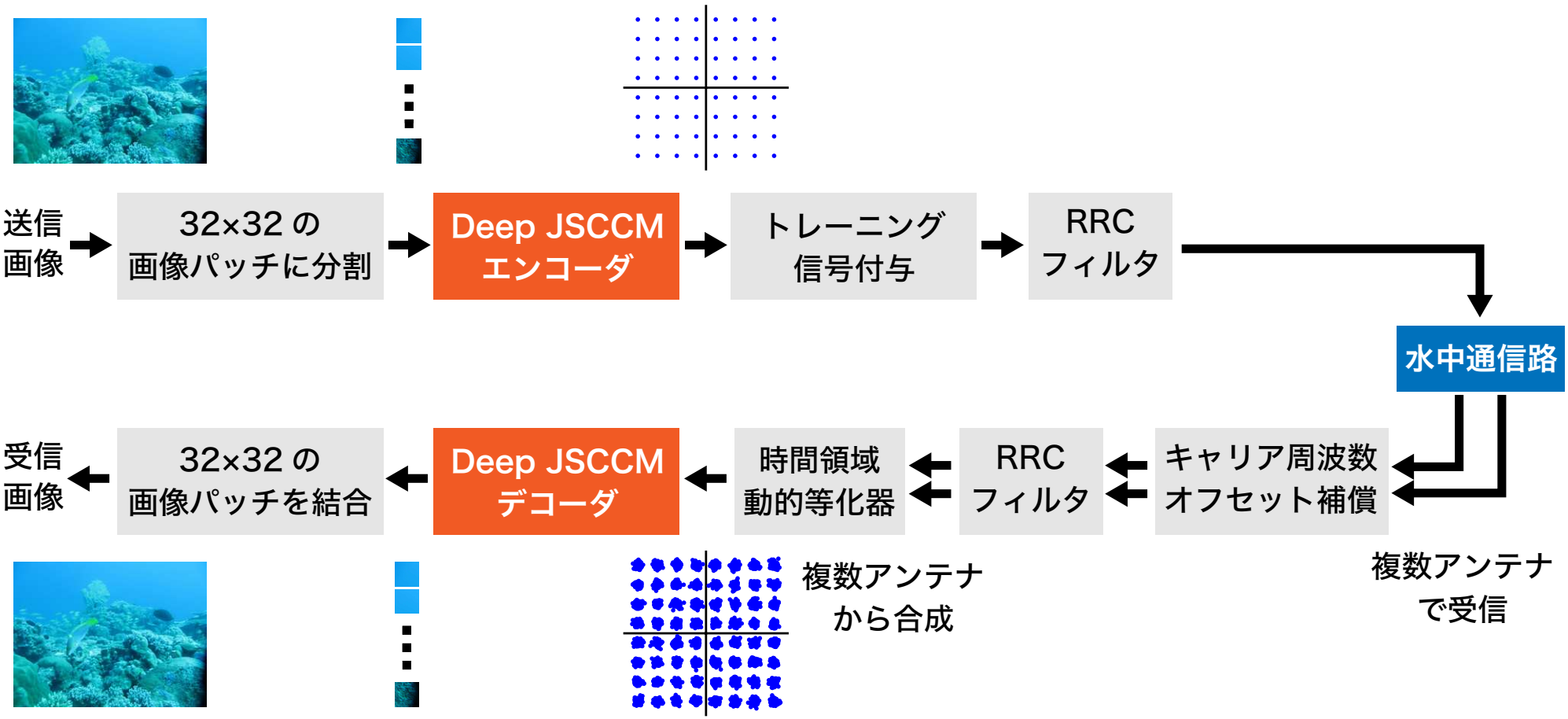
エンコーダとデコーダの層構造や学習過程に独自の工夫 (後述)

- シミュレーション実験により，本手法の有効性を評価

システムモデルと提案手法

システムの全体像

- 下図に，本研究で考察するシステムの概略を示す
  - SIMO (Single-Input-to-Multi-Output) 通信路を考える



通信路モデル

$\mathbf{x} \in \mathbb{C}^{1 \times N_\mathrm{s}}$: 送信信号ベクトル　(N_s: 送信サンプル数)

$\mathbf{y}_i \in \mathbb{C}^{1 \times N_\mathrm{s}}$: i 番目のアンテナの受信信号ベクトル，$\mathbf{Y}$: $\mathbf{y}_i$ を並べた行列

- 本研究では，次の線形 AWGN 通信路モデルを考える

$$\mathbf{Y} = \mathbf{H}\mathbf{x} + \mathbf{Z}$$

$\mathbf{H}$: 通信路行列，$\mathbf{Z}$: AWGN 行列

通信路モデル

$\mathbf{x} \in \mathbb{C}^{1\times N_\mathrm{s}}$：送信信号ベクトル　($N_\mathrm{s}$: 送信サンプル数)

$\mathbf{y}_i \in \mathbb{C}^{1\times N_\mathrm{s}}$: i 番目のアンテナの受信信号ベクトル，$\mathbf{Y}$: $\mathbf{y}_i$ を並べた行列

- 本研究では，次の線形 AWGN 通信路モデルを考える

$$\mathbf{Y} = \mathbf{H}\mathbf{x} + \mathbf{Z}$$

$\mathbf{H}$: 通信路行列，$\mathbf{Z}$: AWGN 行列

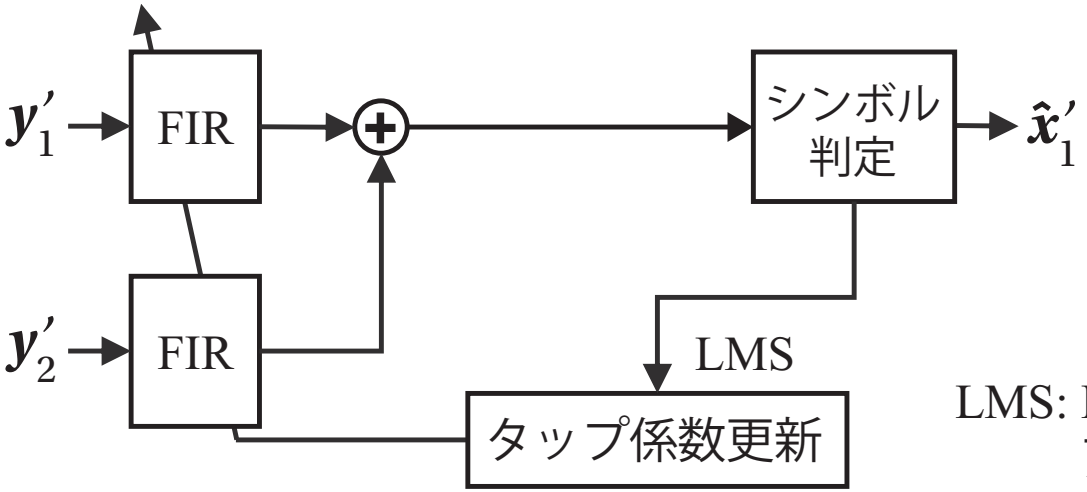
- ただし $\mathbf{H}$ は，各アンテナに対応するブロック $\mathbf{H}_i$ を並べた行列
  - $\mathbf{H}_i$ の対角要素は直接波，非対角要素は遅延波に対応[2]

$$h_i^{(k,l)} = \sqrt{L_i^{(k,l)}} \exp\left[j(2\pi f_i^{(k,l)} k T_s + \phi_i^{(k,l)}) \right]$$

$L_i^{(k,l)}$: 遅延波と直接波の電力比，$f_i^{(k,l)} := f_d \cos\alpha_i^{(k,l)}$，$f_d$: ドップラー周波数
T_s :サンプリング間隔，$\phi_i^{(k,l)}$: 初期位相，$\alpha_i^{(k,l)}$: 受信アンテナ入射角

[2] K. Shima et al., in *IEEE Access*, vol. 9, pp. 18361–18372, 2021.

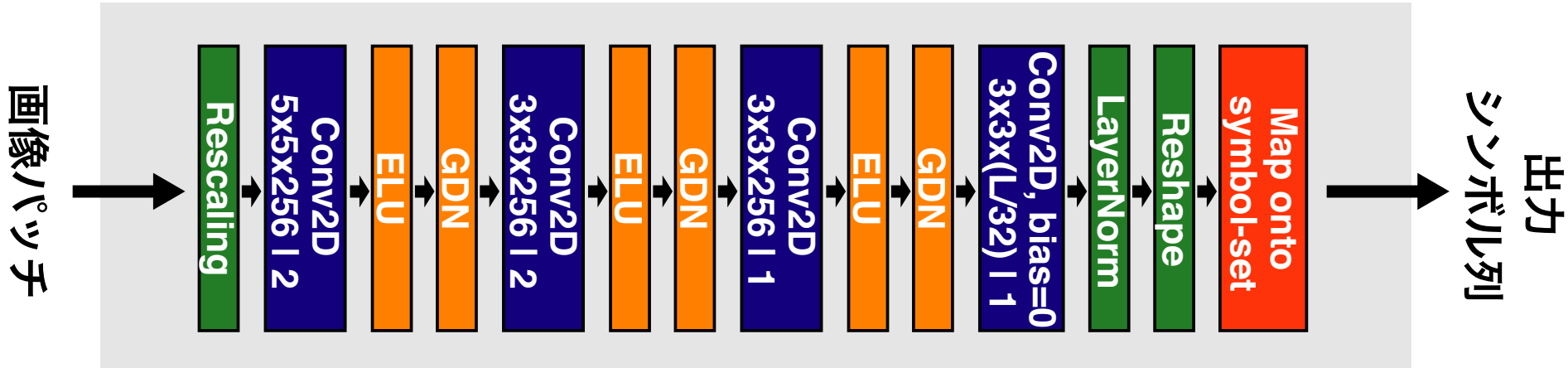
時間領域動的等化器



LMS: Least Mean Square
アルゴリズム

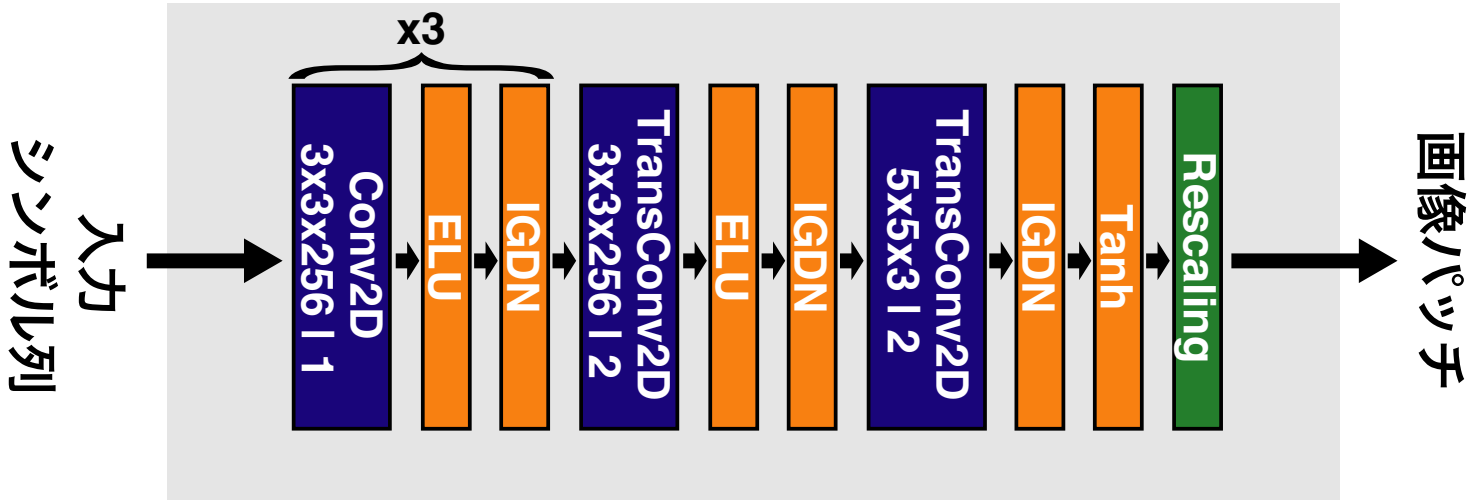
- アンテナの数だけ FIR (finite impulse response) フィルタを用意
  - 出力を加算して合成
- シンボルごとに FIR フィルタの係数を更新
  - シンボル判定結果をフィードバック

JSCCM エンコーダ



- Conv2D: 畳み込み層　(パラメータ記法: kernel_shape | strides)
- ELU: Exponential Linear Unit
- GDN: Generalized Divisive Normalization
- Map onto symbol-set　★動的等化器との結合に必須
  - 入力に対し，候補シンボル集合 $S \subset \mathbb{C}^M$ 上の最近傍点を出力
    (M: 候補シンボル点の総数)

JSCCM デコーダ



- TransConv2D: 転置畳み込み層
  (パラメータ記法: kernel_shape | strides)
- IGDN: Inverse Generalized Divisive Normalization

JSCCM エンコーダとデコーダの学習 (1)

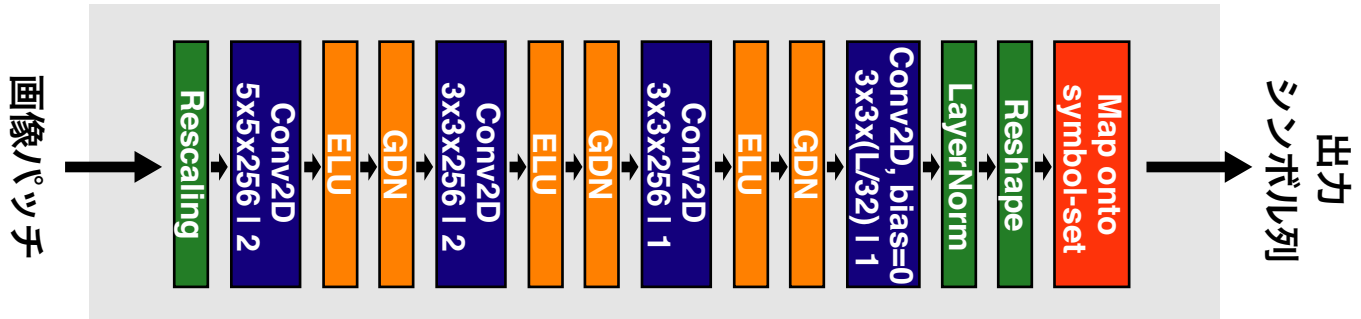
- 以下の要素は 微分不可能な関数 とみなす
  - 送受信器の帯域制限 (RRC) フィルタ
  - 線形 AWGN 通信路
  - 受信器の時間領域動的等化器
- 微分不可能な関数 ψ に対し，以下の手法で誤差逆伝播を実行
  - 入出力関係 $\boldsymbol{y} = \psi(\boldsymbol{x})$ を考える
  - 以下の式で，上の関係式を置き換える

$$\boldsymbol{y} = \boldsymbol{x} + \mathrm{sg}(\psi(\boldsymbol{x}) - \boldsymbol{x})$$

sg: 勾配計算時に定数とみなすという演算子 (Stop gradient の意)

JSCCM エンコーダとデコーダの学習 (2)

- パラメータ学習には，Adamax オプティマイザ[3] を使用
- 損失関数はエンコーダ入力画像とデコーダ出力画像の自乗誤差
  - ただし，エンコーダ最終層の候補シンボル集合 S は除く



- 候補シンボル集合 S の学習
  - 他の層と同時に，Adamax オプティマイザで学習
  - 前段の出力と S 上の最近傍点との自乗誤差を損失関数とする

[3] D. Kingma and J. Ba, in *Proc. ICLR 2015*, 2015.

JSCCM エンコーダとデコーダの学習 (3)

- 本モデルは中間層に複雑な微分不可能関数を含む
  - 長遅延マルチパス通信路
  - 送受信器における帯域制限フィルタ
  - マルチパス干渉を補償する (時間領域) 動的等化器

  ➡ そのままでは 初期段階のパラメータ学習 がうまく進まない

- 本研究では, 2 段階 に分けて学習を実行する手法を考案
  - **事前学習**: 単純な AWGN 通信路を用いてパラメータを学習
    - 画像圧縮・復元に適したパラメータ初期値を獲得
  - **本学習**: 本来の通信路モデルを用いてパラメータを学習

性能評価

シミュレーションパラメータ

- 2 パスモデルを使用
  - 直接波と単一の遅延波

比較対象には以下を用いる

- 画像圧縮:

  JPEG, JPEG2000
- 変調方式:
  - QPSK, 16QAM
- 前方誤り訂正:
  - Turbo 符号 (1/2, 3/4)

(パラメータは文献[4] を参考に決定)

項目	値
アンテナ数	2
キャリア周波数	300 kHz
システム帯域幅	200 kHz
訓練シンボル数	4000 シンボル
RRC フィルタのロールオフ率	0.2
等化 FIR フィルタのタップ数	21
SNR	18–26 dB
遅延経路数	1
送信器の移動速度	1 m/s
遅延波の相対遅延量	520 サンプル
遅延波と直接波の電力比 $L_i^{(k,l)}$	0.4
音速	1500 m/s
入射角 $\alpha_i^{(k,l)}$	一様ランダム
初期位相 $\phi_i^{(k,l)}$	一様ランダム

[4] H. Fukumoto et al., in *Proc. Global Oceans 2020*, 2020.

使用データセット

学習時は，性能評価 (テスト) 用とは別の画像データセットを使用

- 学習用データセット : Downsampled Imagenet[5]
  - 多数の 32×32 カラー画像を含むデータセット
  - 訓練データ数 1,281,167，検証データ数 50,000

[5] P. Chrabaszcz et al., arXiv preprint, arXiv:1707.08819, 2017.

使用データセット

学習時は，性能評価 (テスト) 用とは別の画像データセットを使用

- 学習用データセット : Downsampled Imagenet[5]
  - 多数の 32×32 カラー画像を含むデータセット
  - 訓練データ数 1,281,167，検証データ数 50,000
- テスト用データセット : 文献[6] の水中画像データセット (890 枚)

[5] P. Chrabaszcz et al., arXiv preprint, arXiv:1707.08819, 2017.
[6] C. Li et al., in *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389, 2020.

使用データセット

学習時は，性能評価 (テスト) 用とは別の画像データセットを使用

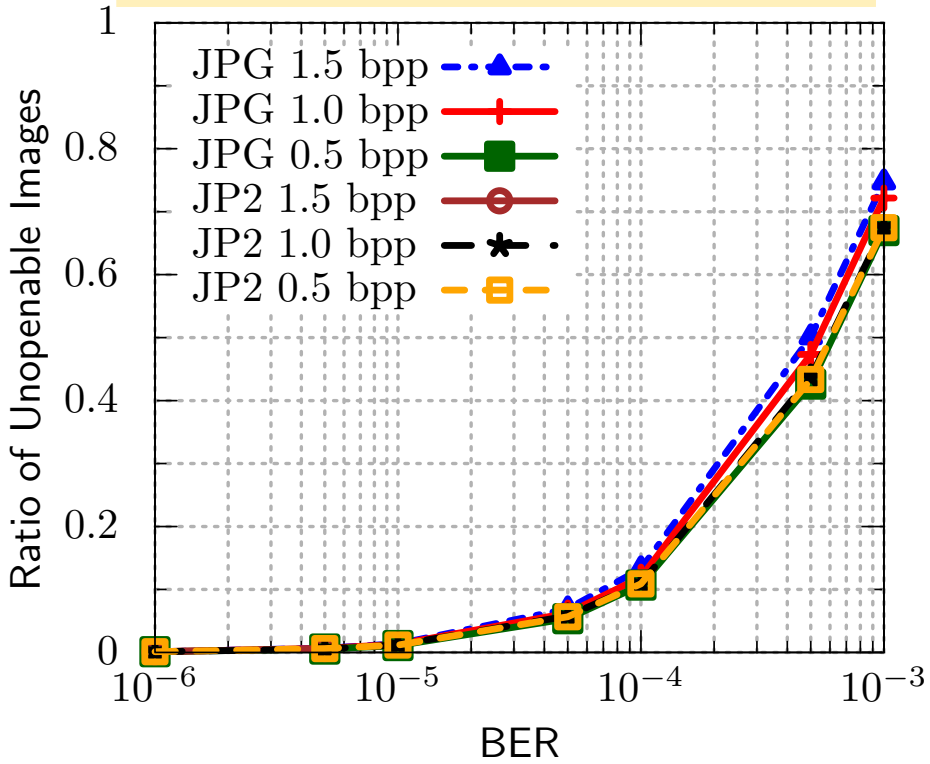
- 学習用データセット : Downsampled Imagenet[5]
  - 多数の 32×32 カラー画像を含むデータセット
  - 訓練データ数 1,281,167，検証データ数 50,000
- テスト用データセット : 文献[6] の水中画像データセット (890 枚)
  - 画像サイズを揃える前処理を実行
    - アスペクト比が $4/3$ 以上の画像のみ抽出 (825 枚)
    - リサイズおよび切り抜きによりサイズを 256×192 に統一

[5] P. Chrabaszcz et al., arXiv preprint, arXiv:1707.08819, 2017.
[6] C. Li et al., in *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389, 2020.

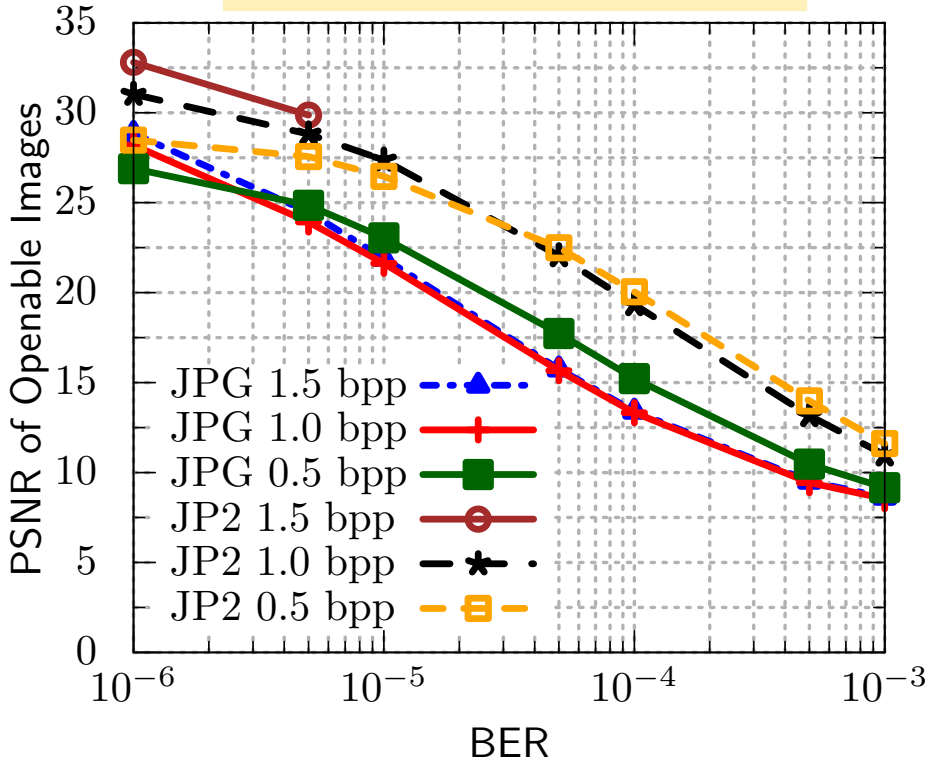
比較対象の選定 (1)

- ビット誤り率 (BER: Bit error rate) が受信画像品質に与える影響

破損して開けない画像の割合



受信画像品質 (PSNR)

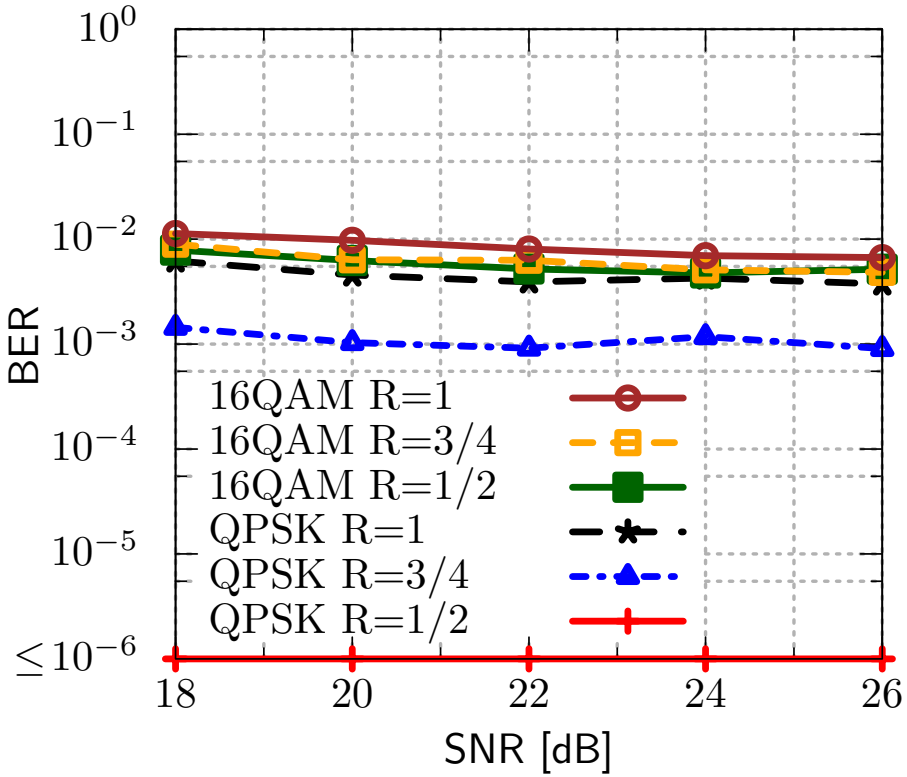


bpp: bit per pixel (圧縮度の指標)　　$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\mathrm{MSE}}$ (大きいほど高品質)

JPEG・JPEG2000 画像の安定受信には 10^{-6} 以下の BER が必要

比較対象の選定 (2)

- QPSK ならびに 16QAM における BER (R: 符号化率)



- QPSK ($R = 1/2$):
  - 左図で唯一 $\mathrm{BER} \leq 10^{-6}$ を満たす
  - ➡ 提案手法との比較に使用

以下，提案手法との比較では
$\mathrm{SNR} \geq 18\mathrm{dB}$ で $\mathrm{BER} = 0$ とみなす

性能評価 (1)

- 1 ピクセル当たりのシンボル数 (spp) を 0.5 に固定
  - 256×192 画像は $24,576$ シンボル
- エンコーダのシンボル候補点数 $|S|$ を 256 に設定
- 学習は 70,000 ステップ実行 (うち 40,000 ステップは事前学習)

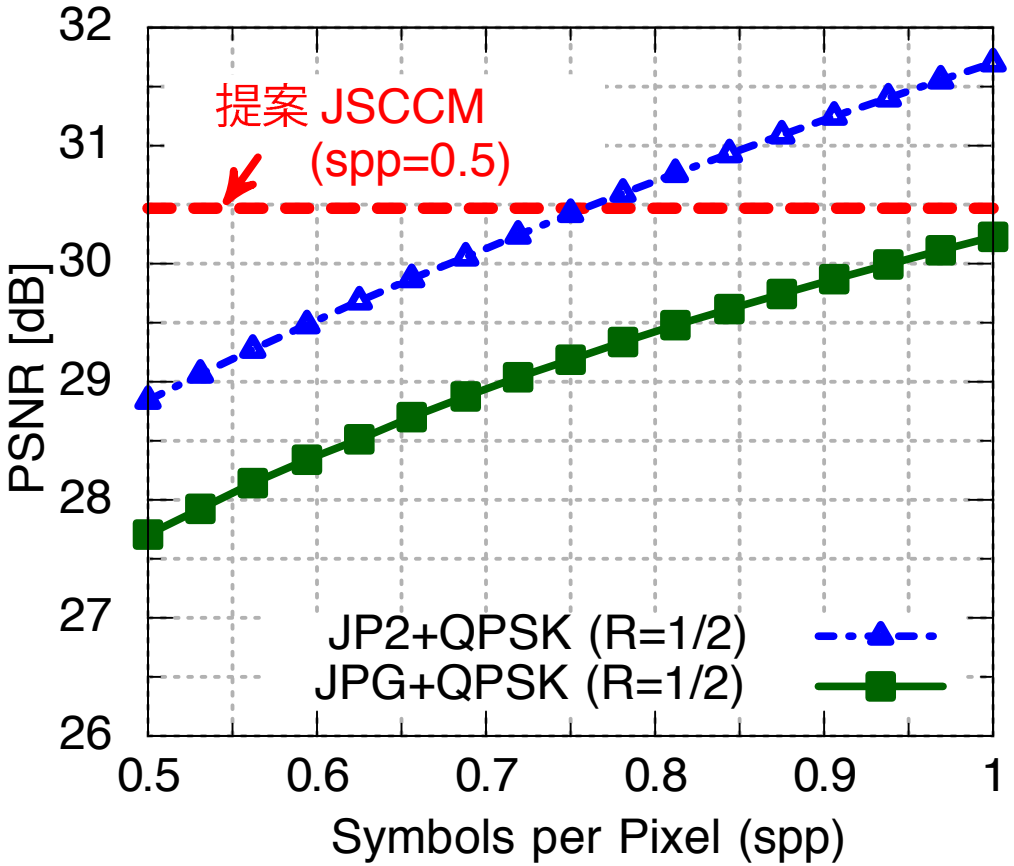
提案手法 (JSCCM) と従来手法の比較結果

- SNR を 18 dB に固定したときの PSNR

提案 JSCCM	QPSK + JPEG2000	QPSK + JPEG
30.470	28.842	27.706

性能評価 (2)

提案手法 (JSCCM) と従来手法の比較結果



(左図) 従来手法において
送信シンボル数が
画像品質に与える影響

- 提案手法と同等の画像品質の達成に必要なシンボル数
  - JPEG2000:
    提案手法の約 1.5 倍
  - JPEG :
    提案手法の約 2 倍

➡ **提案手法により,**
伝送速度が大幅に向上

まとめ

- 音響通信による水中画像伝送への Deep JSCCM の応用を考察
- 以下の要素を新たに考慮した Deep JSCCM 設計法を提案
  - 長遅延マルチパス通信路
  - 送受信器における帯域制限フィルタ
  - マルチパス干渉を補償する (時間領域) 動的等化器
- シミュレーション実験により，本手法の有効性を評価
  - JPEG2000+QPSK と比べて約 1.5 倍の伝送速度
  - JPEG+QPSK と比べて約 2 倍の伝送速度
- 今後の課題: 実機実験に基づく評価，ならびに層構造の改良

受信画像の例

元画像

提案 JSCCM

JPEG2000 + QPSK

JPEG + QPSK

学習されたシンボル配置