

ニューラル陰関数表現を用いた無線画像伝送手法に関する一検討

桑原 明大[†]大迫 勇太郎[‡]藤橋 卓也[‡]秋濃 (小池) 俊昭^{‡†}渡辺 尚[‡][†]大阪大学工学部[‡]大阪大学大学院情報科学研究科^{‡†}Mitsubishi Electric Research Laboratories

1 はじめに

無線ネットワークを介して画像情報を効率的に伝送する方法として Joint Photographic Experts Group (JPEG) を用いた伝送手法が検討されている。本稿では、より効率的な画像情報の無線伝送手法としてニューラル陰関数表現を用いた伝送手法を新たに提案するとともに、その効果について議論する。

2 関連研究

無線伝送路を介して画像情報を伝送する場合、JPEG に代表されるデジタル符号化技術を各画像をビット列に圧縮する。得られたビット列に対してはチャネル符号化を適用して誤り耐性を高めるとともに、デジタル変調技術を用いて無線伝送する。デジタル符号化を用いた伝送技術は伝送路品質に変動がない場合、そのデータレートに合わせた符号化パラメータを設定することで画像品質を向上できる。近年では、ニューラル陰関数表現を用いることで画像情報をさらに圧縮するための方法が検討されている。代表的な手法として COmpression with Implicit Neural representations (COIN) [1] が提案されている。COIN はニューラル陰関数表現の1つである Sinusoidal representation networks (SIREN) [2] を用いて画素値の座標情報を入力、対応する画素値を出力とするネットワーク構造を画像ごとに学習する。学習済のネットワーク構造をネットワークを通じてユーザへ伝送するとともに、ユーザは受信したネットワーク構造に対して座標情報を入力することで画像を取得することが可能となる。

上述の手法を用いて画像情報を無線伝送した場合、無線伝送路品質の変動に起因する画像品質の低下が問題となる。無線通信における伝送路品質は受信端末の位置や周辺環境に応じて頻繁に変動する。例えば、画像伝送中に無線伝送路品質が改善された場合、JPEG に基づく伝送手法は量子化に起因する品質劣化を復元することができず、ニューラル陰関数表現に基づく伝送手法はネットワーク構造の規模に起因して画像品質を向上することができない。

本稿では、無線伝送路における画像情報の効率的な伝送手法を設計するために、ニューラル陰関数表現を用いた画像伝送手法を新たに提案する。より具体的には、ニューラル陰関数表現 SIREN とアナログ伝送を組み合わせることで画像伝送中に生じる伝送路品質の変化に応じて画像品質を向上できる伝送手法の実現を目指す。

3 提案手法

3.1 全体像

図1に提案手法の全体像を示す。送信画像ごとにニューラル陰関数表現を用いて画素の座標情報と色情報との対応付けを学習する。学習済のニューラル陰関数表現は送信側と受信側とで共有する。その後、送信側では元画像とニューラル陰関数表現から得られた画像との間の残差情報を取得するとともに、得られた残差情報をアナログ伝送する。最後に、ニューラル陰関数表現から得られた画像と伝送した残差信号を受信側で加算することで画像を復元する。

3.2 ニューラル陰関数表現

図2に提案手法で用いたニューラル陰関数表現のネットワーク構造を示す。ネットワーク構造は SIREN に基づいて定めた。入力層は対象画像の座標成分 (x, y) 、出力層は座標成分に対応する画素値 (r, g, b) と定めた。入力層と出力層の間は M 個のノードを保有する L 個の隠れ層で接続されている。各隠れ層においては活性化関数 \sin を用いる。提案手法におけるニューラル陰関数表現は次式のとおりに定義できる。

$$\Phi(\mathbf{x}) = \mathbf{W}_{L+1}(\phi_L \circ \phi_{L-1} \circ \dots \circ \phi_0)(\mathbf{x}) + \mathbf{b}_{L+1}, \quad (1)$$

$$\mathbf{x}_i \mapsto \phi_i(\mathbf{x}_i) = \sin(\mathbf{W}_i \mathbf{x}_i + \mathbf{b}_i). \quad (2)$$

ここで、 \mathbf{W}_{L+1} はすべての層に対する重み、 \mathbf{b}_{L+1} はすべての層に対するバイアス、 \mathbf{x} はニューラル陰関数表現に対する入力情報を意味している。第 i 層のネットワークでは $\phi_i: \mathbb{R}^{M_i} \mapsto \mathbb{R}^{N_i}$ 、 $\mathbf{W}_i \in \mathbb{R}^{N_i \times M_i}$ 、 $\mathbf{b}_i \in \mathbb{R}^{N_i}$ 、 $\mathbf{x}_i \in \mathbb{R}^{M_i}$ を満たす。活性化関数として \sin 関数を用いた理由はその導関数が元の \sin 関数から位相がずれた関数となるためである。SIREN は任意の次数の導関数を使用することで、微分操作を介した逆伝播において表現能力が劣化せず、勾配消失や爆発問題を回避しながら、複雑な入力信号を表現することができる。

画素値の座標情報を $\mathbf{x}_i = (x_i, y_i)$ 、座標情報に対応する画素値 rgb を $f(\mathbf{x}_i)$ 、座標情報と画素値が対となったデータセットを $D = \{\mathbf{x}_i, f(\mathbf{x}_i)\}$ とするとき、ニューラル陰関数表現を用いて座標情報から画素値を求める関数は $\Phi: \mathbb{R}^2 \mapsto \mathbb{R}^3$ 、 $\mathbf{x} \mapsto \Phi(\mathbf{x})$ と定義できる。提案手法のニューラル陰関数表現は次式の損失関数を最小化するように各層の重みとバイアスを学習することで座標情報から正確な画素情報を生成することができる。

$$\mathcal{L} = \sum_{i \in D} \|\Phi(\mathbf{x}_i) - f(\mathbf{x}_i)\|^2 \quad (3)$$

3.3 残差伝送部

送信端末は原信号となる画像情報とニューラル陰関数表現から生成した画像情報を用いて残差情報を生成する。ここで、座標 (i, j) における原信号の画素値 $p_{i,j}$ 、ニューラル陰関数表現から生成された画素値 $\hat{p}_{i,j}$ とすると、対応する残差情報 $r_{i,j}$ は次式のとおりに求まる。

$$r_{i,j} = p_{i,j} - \hat{p}_{i,j}. \quad (4)$$

得られた残差情報に対しては2次元離散コサイン変換 (two-dimensional discrete cosine transform: 2D-DCT) を用いることで対応する DCT 係数を取得する。その後、得られた係数の大きさに応じて送信電力を割り当てるとともに、送信電力を割り当てた DCT 係数を2つずつ取り出して送信信号の同相成分・直交成分に割り当てて送信する。

ここで i 番目の送信信号を x_i とする。各送信信号は $2i$ 番目および $2i+1$ 番目の残差情報の DCT 係数 s_i, s_{i+1} とそれぞれに対するスケール係数 g_{2i} および g_{2i+1} を用いて次式のとおりに表現できる。

$$x_i = g_{2i}r_{2i} + jg_{2i+1}r_{2i+1}. \quad (5)$$

ここで j は虚数単位を表す。このとき、 i 番目の係数に対するスケール係数 g_i は、平均送信電力 P 下において送信前の残差情報の DCT 係数 r_i と受信後の残差情報の DCT 係数 \hat{r}_i との間の平均二乗誤差 (Mean Square Error: MSE) を最小化するように割り当てられる。すなわち、以下の式のように表現される。

$$\min_{\{g_i\}} \text{MSE} = \mathbb{E}[(r_i - \hat{r}_i)^2] = \frac{1}{N} \sum_i \frac{\sigma^2 \lambda_i}{g_i^2 \lambda_i + \sigma^2}, \quad (6)$$

$$\text{s.t.} \quad \frac{1}{N} \sum_i g_i^2 \lambda_i = P. \quad (7)$$

ここで $\lambda_i = r_i^2$ 、 σ^2 は無線伝送路における雑音電力、 N は残差情報の総数である。このとき、MSE を最小化可能な g_i は以下の式にしたがうことが知られている。

$$g_i = m \lambda_i^{-1/4}, \quad m = \sqrt{\frac{NP}{\sum_j \lambda_j^{1/2}}}. \quad (8)$$

ユーザは無線伝送路を介して送信信号を受信する。このとき、受信信号は次式にしたがう。

$$y_i = x_i + n_i. \quad (9)$$

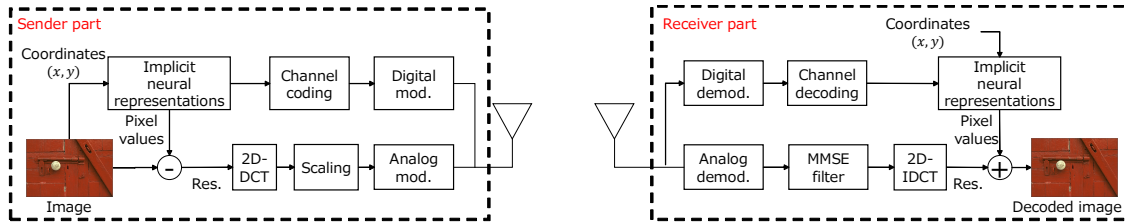


図 1: 提案手法の全体像

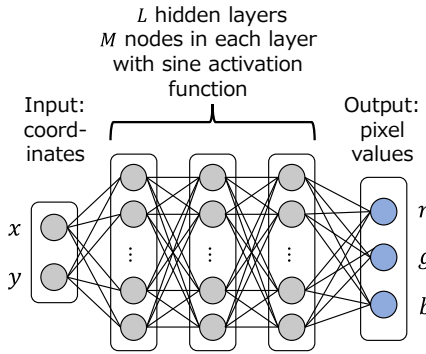


図 2: ニューラル陰関数表現のネットワーク構造

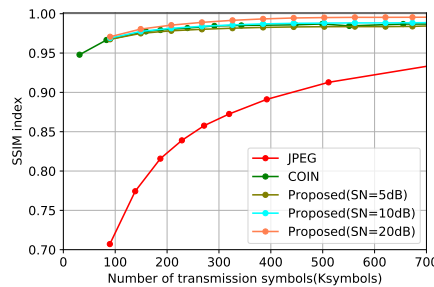


図 3: データサイズに対する映像品質

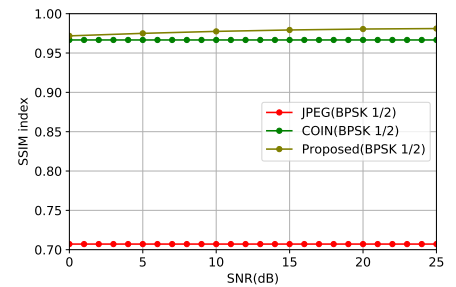


図 4: 無線伝送路品質に対する映像品質

ここで y_i は i 番目の受信信号, n_i は送信信号 x_i に対してかかる加算性白色ガウス雑音 (Additive white Gaussian noise) である. 雑音の分散は σ^2 にしたがるものとした. 受信信号に含まれる雑音を除去するために提案手法では Minimum MSE (MMSE) フィルタを用いる.

$$\hat{r}_i = \frac{g_i \lambda_i}{g_i^2 \lambda_i + \sigma^2} \cdot y_i \quad (10)$$

雑音の影響を除去した残差情報の DCT 係数に対して 2 次元逆離散コサイン変換 (2D-IDCT) を用いて残差情報を取得する. 最後に, ニューラル陰関数表現から得られた画像情報と残差情報とを加算して復元後の画像情報を得る.

4 性能評価

4.1 評価環境

提案手法の有効性を確認するために, Python によるシミュレーションを用いてデータサイズに対する映像品質, 無線伝送路品質に対する映像品質を評価した.

データセット:

評価に用いたデータセットとして kaggle が提供する kodak データセットに含まれる kodim01 から kodim10 までを使用した. 評価結果においては kodim02 を用いた結果を示している. 解像度は 768×512 画素である.

比較手法:

比較手法として JPEG を用いた伝送手法, COIN を用いた伝送手法を用意した. JPEG を用いた伝送手法においては Python ライブラリ Pillow を用いて送信画像をビット列へ圧縮するとともに圧縮したビット列に対して 1/2 レートの畳込み符号と binary phase shift keying (BPSK) を用いて無線伝送するものとした. COIN を用いた伝送手法においては学習済のニューラル陰関数表現を 1/2 レートの畳込み符号と BPSK を用いて無線伝送するものとした. 受信側では受信したニューラル陰関数表現に対して座標情報を入力することで対応する画像を生成する.

ネットワーク構造:

提案手法および COIN を用いた伝送手法において, ニューラル陰関数表現の隠れ層で用いたノード数 M は 28 と定め, 隠れ層数 L は 1 から 10 に変化させ, ネットワーク構造の規模がもたらす影響を評価した.

評価指標:

性能評価指標として Structural Similarity (SSIM) を使用した. SSIM はウィンドウ単位で計算する. 画像内の全ウィンドウの平均値を SSIM として求めた. SSIM は 0 から 1 までの範囲で値を取る. 値が 1 に近いほど受信画像が元画像に近いことを表している.

4.2 評価結果

図 3 にデータサイズに対する映像品質を示す. 提案手法においては無線伝送路品質 (Signal-to-Noise Ratio: SNR) が 5 dB, 10 dB, 20 dB 時における性能を示している. 図 4 に無線伝送路品質に対する映像品質を示している. このときの送信シンボル数は 150 Ksymbols を想定している. 評価結果から, 以下のことが分かる.

- 特に, 帯域が限られている環境において, 提案手法が JPEG を用いた伝送手法および COIN を用いた伝送手法より高い品質を達成できること
- 提案手法の SSIM は無線伝送路品質に比例して映像品質が改善していること

5 おわりに

本研究ではより効率的な画像情報の無線伝送手法としてニューラル陰関数表現を用いた伝送手法を提案した. 評価結果から提案手法は既存手法と比較して限られた帯域下において高い SSIM を実現することが分かった. 今後は動画への拡張を考えている.

謝辞

本研究は JSPS 科研費 (JP19H01101, JP22H03582) の支援の下で行った.

参考文献

- [1] Dupont, Emilien, Hrushikesh Loya, Milad Alizadeh, Adam Goliński, Yee Whye Teh, and Arnaud Doucet. "COIN++: Data Agnostic Neural Compression," 2022."
- [2] Sitzmann, Vincent, et al. "Implicit neural representations with periodic activation functions." Advances in Neural Information Processing Systems 33 (2020): 7462-7473.