

不確実性ガイド付き拡散モデルを用いた 無線画像伝送システムモデル

1 システムモデル

本稿では、高解像度画像データセット DIV2K を対象とした、拡散モデルベースの無線画像伝送システムを検討する。受信機において、拡散モデルのサンプリング過程でピクセル単位の不確実性を推定し、復元品質を向上させる。

1.1 送信機およびチャネルモデル

入力画像を $s \in \mathbb{R}^{H \times W \times 3}$ とする (DIV2K データセット)。送信機は、ソース画像をチャネル入力シンボル $x \in \mathbb{C}^k$ にマッピングするエンコーダ $f_\theta(\cdot)$ を有する。

$$x = f_\theta(s) \quad (1)$$

ここで、送信信号 x は平均電力制約 $\frac{1}{k}\mathbb{E}[\|x\|^2] \leq P$ を満たすように正規化される。

無線チャネルとして、加法性白色ガウス雑音 (AWGN) チャネルを仮定する。受信信号 y は次式で表される。

$$y = x + n \quad (2)$$

ここで、 $n \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ は独立同一分布に従う複素ガウス雑音ベクトルである。

1.2 拡散モデルによる受信と不確実性推定

受信機では、条件付き拡散モデルを用いて、受信信号 y から元の画像 s (拡散過程における x_0) を推定する。提案手法に基づき、サンプリングステップ t ごとにデノイジングスコアの分散を用いて不確実性を推定する。

1.2.1 不確実性推定アルゴリズム

タイムステップ t における不確実性マップ \mathbf{U}_t の推定は、モデル出力の感度 (Sensitivity) を代理指標として用いる。具体的な手順は以下の通りである。

1. **クリーン画像の近似:** 現在のノイズあり画像 \mathbf{x}_t とスコア推定値 $\epsilon_\theta(\mathbf{x}_t, t)$ を用いて、 \mathbf{x}_0 の近似値 $\hat{\mathbf{x}}_0$ を計算する。

$$\hat{\mathbf{x}}_0 = \frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}} \quad (3)$$

ここで、 $\bar{\alpha}_t$ は拡散プロセスのノイズスケジュールである。

2. **撮動サンプルの生成:** 近似した $\hat{\mathbf{x}}_0$ に対し、拡散プロセスの順方向 (Noising) 分布 $q(\tilde{\mathbf{x}}_t | \hat{\mathbf{x}}_0)$ に従って M 個の異なるノイズサンプル $\{\hat{\mathbf{x}}_t^i\}_{i=1}^M$ を生成する。

$$\hat{\mathbf{x}}_t^i = \sqrt{\bar{\alpha}_t} \hat{\mathbf{x}}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_i, \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(0, \mathbf{I}) \quad (4)$$

この撮動スキームは、拡散モデル専用に設計されたものである。

3. **分散による不確実性算出:** M 個のサンプルに対するスコア $\epsilon_\theta(\hat{\mathbf{x}}_t^i, t)$ を計算し、その分散を不確実性 \mathbf{U}_t とする。

$$\mathbf{U}_t = \text{Var}(\{\epsilon_\theta(\hat{\mathbf{x}}_t^i, t)\}_{i=1}^M) \quad (5)$$

この分散は、ノイズ分布の対数尤度の二階微分（曲率）と関連していることが示されている。

Algorithm 1 ピクセル単位の不確実性推定 (Pixel-wise Uncertainty Estimation)

Require: 画像 \mathbf{x}_t , タイムステップ t , サンプル数 M

Ensure: 不確実性マップ \mathbf{U}_t

- 1: 真のスコア $\epsilon_\theta(\mathbf{x}_t, t)$ を計算
- 2: \mathbf{x}_0 の近似値を計算: $\hat{\mathbf{x}}_0 \leftarrow \frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}}$
- 3: **for** $i = 1$ to M **do**
- 4: ノイズ再注入: $\hat{\mathbf{x}}_t^i \leftarrow \sqrt{\bar{\alpha}_t} \hat{\mathbf{x}}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_i$
- 5: スコア計算: $\boldsymbol{\epsilon}'_i \leftarrow \epsilon_\theta(\hat{\mathbf{x}}_t^i, t)$
- 6: **end for**
- 7: 分散を計算: $\mathbf{U}_t \leftarrow \text{Var}(\{\boldsymbol{\epsilon}'_i\}_{i=1}^M)$
- 8: **return** \mathbf{U}_t

1.2.2 再送マスク生成と信号合成

不確実性マップ \mathbf{U} の各要素 u_j が所定の閾値 τ を超える場合、そのシンボルインデックス j を再送対象とするバイナリマスク \mathbf{M} を生成する。

$$m_j = \begin{cases} 1, & \text{if } u_j > \tau \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

受信機はこのマスク \mathbf{M} をフィードバックし、送信機はマスクされた部分信号

$$\mathbf{x}_{\text{sub}} = \mathbf{x} \odot \mathbf{M} \quad (7)$$

のみを再送する。

受信機では、再送された信号 \mathbf{y}_{re} と、バッファされている過去の受信信号 \mathbf{y}_{old} を最大比合成 (AWGN 下では単純平均) し、信号対雑音比を改善する。

$$\mathbf{y}_{\text{new}}[j] = \frac{\mathbf{y}_{\text{old}}[j] \cdot c[j] + \mathbf{y}_{\text{re}}[j]}{c[j] + 1} \quad (8)$$

ここで、 $c[j]$ はシンボル j の累積受信回数である。更新された受信信号 \mathbf{y}_{new} を新たな条件 (Condition) として CDDM プロセスを継続することで、不確実性の高い領域を選択的に「浄化」し、最終的な画質を向上させる。