

＜令和6年度卒業研究論文＞

ウェーブレットとL1畳み込み

スパース表現を使用した

単一画像超解像

指導教員: 黒木祥光 教授

制御情報工学科5年 60242 アジュマル

# 目次

<b>1</b>	<b>はじめに</b>	<b>2</b>
<b>2</b>	<b>技術背景</b>	<b>3</b>
2.1	CSR . . . . .	3
2.1.1	係数最適化 . . . . .	4
2.1.2	辞書最適化 . . . . .	6
2.2	SWT . . . . .	8
2.3	CSR および SWT による SISR . . . . .	10
<b>3</b>	<b>提案手法</b>	<b>13</b>
3.1	学習過程 . . . . .	13
3.1.1	LR および HR フィルタの学習 . . . . .	13
3.1.2	マッピング関数の学習 . . . . .	14
3.2	評価過程 . . . . .	16
<b>4</b>	<b>実験</b>	<b>18</b>
<b>5</b>	<b>むすび</b>	<b>22</b>
	<b>参考文献</b>	<b>24</b>

# 第1章

## はじめに

本研究では, 定常ウェーブレット変換 (SWT : Stationary Wavelet Transform) と畳み込みスパース表現 (CSR : Convolutional Sparse Representation) を使用して, 単一画像超解像 (SISR : Single Image Super-Resolution) のモデルを設計することを目的とする. SISR は, 低解像度 (LR : Low Resolution) 画像から高解像度 (HR : High Resolution) 画像を推定するタスクである. 従来の SISR モデルでは, スパースコーディング手法を用い, 入力 LR 画像を重複するパッチに分割した後, それらを基に HR 画像を推定する. しかし, この手法では, 重複するパッチ間の一貫性が考慮されないという問題がある. この課題に対応するため, 近年では CSR を活用した手法 [1] が提案され, 画像全体をスパースコーディングするアルゴリズムが開発されている. さらに, 従来のバイキュービック補間による劣化したモデルを用いるが, 新たなアプローチとして SWT を適用し, HR および LR のサブ画像に分解する手法 [2] も提案されている. SWT は, アップサンプリングを伴う分解が可能である点に加え, スパース性, コンパクト性, 方向性, 冗長性といった特性を有するため, SISR タスクに適している. 本研究では, 従来の CSR における L2 ノルムの誤差項を L1 ノルムの誤差項に置き換えることで, 新たな SISR モデルを設計し, 超解像の精度向上を図る.

## 第2章

## 技術背景

### 2.1 CSR

CSR[3] とは, 画像  $\mathbf{s}_k \in \mathbb{R}^U$  をゼロ要素の多いスパースな係数  $\mathbf{x}_{k,m} \in \mathbb{R}^U$  と辞書フィルタ  $\mathbf{d}_m \in \mathbb{R}^V$  の畳み込みの和で近似する方法である. 画像は2次元信号なので, 実際の実装においては画像を行列として処理されるが, 計算の簡略化のため画像をベクトル化し, 以下のように定式化する.

$$\mathbf{s}_k \approx \sum_{m=1}^M \mathbf{d}_m * \mathbf{x}_{k,m} \quad (k = 1, 2, \dots, K) \quad (2.1)$$

ここで  $K$  は画像の枚数,  $M$  は辞書の枚数を表し,  $*$  は巡回畳み込みの操作を示す. 辞書フィルタは画像における特徴を, 係数は対応する辞書フィルタが画像内のどの位置に現れるかを示している. すなわち, スパース係数は辞書を通じて画像の特徴を表現するものであり, スパースな特徴マップと呼ばれることもある.

CSR の係数と辞書を導くために以下のような最適化問題を設定する。

$$\arg \min_{\mathbf{d}_m, \mathbf{x}_{k,m}} \frac{1}{2} \left\| \sum_{m=1}^M \mathbf{d}_m * \mathbf{x}_{k,m} - \mathbf{s}_k \right\|_2^2 + \lambda \sum_{m=1}^M \|\mathbf{x}_{k,m}\|_1 \quad s.t. \quad \mathbf{d}_m \in C_{PN} \forall m, \quad (2.2)$$

$$C_{PN} = \{\mathbf{y} \in \mathbb{R}^N \mid (\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{y} = \mathbf{0}, \|\mathbf{y}\|_2 \leq 1\} \quad (2.3)$$

式 (2.2) の第1項は, CSR による近似画像と元画像の誤差を評価する誤差項である. 式 (2.2) の第2項は係数をスパースに制約する項であり,  $\lambda$  はスパース性を制御するため

のパラメータである。また, 式 (2.3) の  $C_{PN}$ , ゼロパディング行列  $\mathbf{P}$  で  $\mathbf{d}_m$  を  $\mathbf{x}_{k,m}$  と同次元にし, 正則化するための制御集合である。ここから, 画像と畳み込み和の誤差を最小化するために, 係数と辞書を交互に最適化する処理を行う。

### 2.1.1 係数最適化

係数の最適化では以下の最小化問題を解く。

$$\arg \min_{\mathbf{x}_{k,m}} \frac{1}{2} \left\| \sum_{m=1}^M \mathbf{d}_m * \mathbf{x}_{k,m} - \mathbf{s}_k \right\|_2^2 + \lambda \sum_{m=1}^M \|\mathbf{x}_{k,m}\|_1 \quad (2.4)$$

この目的関数は凸関数であるため, ADMM を適用することで解く。ADMM に当てはまることができるように, 式 (2.4) を以下のように書き換える。

$$\arg \min_{\mathbf{x}_{k,m}} \frac{1}{2} \|\mathbf{g}_0\|_2^2 + \lambda \|\mathbf{g}_1\|_1 \quad \text{s.t.} \quad \begin{pmatrix} \mathbf{D} \\ \mathbf{I} \end{pmatrix} \mathbf{x}_k - \begin{pmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{s}_k \\ \mathbf{0} \end{pmatrix} \quad (2.5)$$

ここで, 巡回畳み込みで生成する辞書行列  $\mathbf{D}_m$  を  $\mathbf{D}_m \mathbf{x}_{k,m} = \mathbf{d}_m * \mathbf{x}_{k,m}$  が成り立つように定義し,  $\mathbf{D}$  および  $\mathbf{x}_k$  を以下のように定義する。

$$\mathbf{D} = (\mathbf{D}_1 \quad \mathbf{D}_2 \quad \dots \quad \mathbf{D}_M) \quad \mathbf{x}_k = (\mathbf{x}_{k,1} \quad \mathbf{x}_{k,2} \quad \dots \quad \mathbf{x}_{k,M})^T \quad (2.6)$$

スケール形式の ADMM[4] を用いるため, この最小化問題を以下のような拡張ラグランジュ関数に書き換える。ここでは,  $\rho$  はラグランジュ関数のペナルティパラメータを示す。

$$\begin{aligned} L\{\mathbf{x}_k, \mathbf{g}_0, \mathbf{g}_1, \mathbf{h}_0, \mathbf{h}_1\} &= \frac{1}{2} \|\mathbf{g}_0\|_2^2 + \lambda \|\mathbf{g}_1\|_1 + \frac{\rho}{2} \|\mathbf{D}\mathbf{x}_k - \mathbf{g}_0 - \mathbf{s}_k + \mathbf{h}_0\|_2^2 \\ &\quad + \frac{\rho}{2} \|\mathbf{h}_0\|_2^2 + \frac{\rho}{2} \|\mathbf{x}_k - \mathbf{g}_1 + \mathbf{h}_1\|_2^2 + \frac{\rho}{2} \|\mathbf{h}_1\|_2^2 \end{aligned} \quad (2.7)$$

この問題の更新式は以下のようになる.

$$\mathbf{x}_k^{(i+1)} = \arg \min_{\mathbf{x}_k} \frac{\rho}{2} \left\| \mathbf{D}\mathbf{x}_k - \mathbf{g}_0^{(i)} - \mathbf{s}_k + \mathbf{h}_0^{(i)} \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{x}_k - \mathbf{g}_1^{(i)} + \mathbf{h}_1^{(i)} \right\|_2^2 \quad (2.8)$$

$$\mathbf{g}_0^{(i+1)} = \arg \min_{\mathbf{g}_0} \frac{1}{2} \left\| \mathbf{g}_0 \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{D}\mathbf{x}_k^{(i+1)} - \mathbf{g}_0 - \mathbf{s}_k + \mathbf{h}_0^{(i)} \right\|_2^2 \quad (2.9)$$

$$\mathbf{g}_1^{(i+1)} = \arg \min_{\mathbf{g}_1} \lambda \left\| \mathbf{g}_1 \right\|_1 + \frac{\rho}{2} \left\| \mathbf{x}_k^{(i+1)} - \mathbf{g}_1 + \mathbf{h}_1^{(i)} \right\|_2^2 \quad (2.10)$$

$$\mathbf{h}_0^{(i+1)} = \mathbf{h}_0^{(i)} + \mathbf{D}\mathbf{x}_k^{(i+1)} - \mathbf{g}_0^{(i+1)} - \mathbf{s}_k \quad (2.11)$$

$$\mathbf{h}_1^{(i+1)} = \mathbf{h}_1^{(i)} + \mathbf{x}_k^{(i+1)} - \mathbf{g}_1^{(i+1)} \quad (2.12)$$

実領域において畳み込み演算を行っている  $\mathbf{D}\mathbf{x}_k$  を解くことは容易ではないため, フーリエ領域で解くことを考える.  $\hat{\mathbf{s}}_k$  を  $\mathbf{s}_k$  のフーリエ変換,  $\hat{\mathbf{x}}_k$  を  $\mathbf{x}_k$  のフーリエ変換,  $\hat{\mathbf{d}}$  を  $\mathbf{d}$  のフーリエ変換とし,  $\hat{\mathbf{D}}$  を以下のように定義する.

$$\hat{\mathbf{D}} = \begin{pmatrix} \text{diag}(\hat{\mathbf{d}}_1) & \text{diag}(\hat{\mathbf{d}}_2) & \dots & \text{diag}(\hat{\mathbf{d}}_M) \end{pmatrix} \quad (2.13)$$

フーリエ領域では畳み込み演算が要素積になることと, L2 ノルムが保存されることを利用すると, 式 (2.8) は以下のように書き換えることができる.

$$\hat{\mathbf{x}}_k^{(i+1)} = \arg \min_{\hat{\mathbf{x}}_k} \frac{\rho}{2} \left\| \hat{\mathbf{D}}\hat{\mathbf{x}}_k - \hat{\mathbf{g}}_0^{(i)} - \hat{\mathbf{s}}_k + \hat{\mathbf{h}}_0^{(i)} \right\|_2^2 + \frac{\rho}{2} \left\| \hat{\mathbf{x}}_k - \hat{\mathbf{g}}_1^{(i)} + \hat{\mathbf{h}}_1^{(i)} \right\|_2^2 \quad (2.14)$$

これを最小化する  $\hat{\mathbf{x}}_k$  は以下を満たす.

$$\hat{\mathbf{D}}^H \left( \hat{\mathbf{D}}\hat{\mathbf{x}}_k - \hat{\mathbf{g}}_0^{(i)} - \hat{\mathbf{s}}_k + \hat{\mathbf{h}}_0^{(i)} \right) + \left( \hat{\mathbf{x}}_k - \hat{\mathbf{g}}_1^{(i)} + \hat{\mathbf{h}}_1^{(i)} \right) = \mathbf{0} \quad (2.15)$$

$\hat{\mathbf{x}}_k$  の更新式は以下のようになる.

$$\hat{\mathbf{x}}_k^{(i+1)} = \left( \hat{\mathbf{D}}^H \hat{\mathbf{D}} + \mathbf{I} \right)^{-1} \left\{ \hat{\mathbf{D}}^H \left( \hat{\mathbf{g}}_0^{(i)} + \hat{\mathbf{s}}_k - \hat{\mathbf{h}}_0^{(i)} \right) + \left( \hat{\mathbf{g}}_1^{(i)} - \hat{\mathbf{h}}_1^{(i)} \right) \right\} \quad (2.16)$$

式 (2.16) の  $\hat{\mathbf{D}}^H \hat{\mathbf{D}}$  はサイズが大きく計算コストが増加するため, Sherman-Morison-Woodbury の逆行列の補助定理を用いて変形する.

$$\hat{\mathbf{x}}^{(i+1)} = \left( \hat{\mathbf{g}}_1^{(i)} - \hat{\mathbf{h}}_1^{(i)} \right) + \hat{\mathbf{D}}^H \left( \hat{\mathbf{D}} \hat{\mathbf{D}}^H + \mathbf{I} \right)^{-1} \left\{ \hat{\mathbf{g}}_0^{(i)} + \hat{\mathbf{s}} - \hat{\mathbf{h}}_0^{(i)} - \hat{\mathbf{D}} \left( \hat{\mathbf{g}}_1^{(i)} - \hat{\mathbf{h}}_1^{(i)} \right) \right\} \quad (2.17)$$

式 (2.9) は,  $\mathbf{g}_0$  についての勾配を得ることで求めることができる.

$$\mathbf{g}_0^{(i+1)} = \frac{\rho}{1+\rho} \left( \mathbf{D} \mathbf{x}_k^{(i+1)} - \mathbf{s}_k + \mathbf{h}_0^{(i)} \right) \quad (2.18)$$

式 (2.10) は, 指示関数近接作用素 [4] を用いて解くことができる. ここで,  $\odot$  は要素ごとの積を表す.

$$\mathbf{g}_1^{(i+1)} = \text{sign} \left( \mathbf{x}_k^{(i+1)} + \mathbf{h}_1^{(i)} \right) \odot \max \left( 0, |\mathbf{x}_k^{(i+1)} + \mathbf{h}_1^{(i)}| - \frac{\lambda}{\rho} \right) \quad (2.19)$$

### 2.1.2 辞書最適化

辞書の最適化では以下の最小化問題を解く.

$$\arg \min_{\mathbf{d}_m} \frac{1}{2} \left\| \sum_{m=1}^M \mathbf{x}_{k,m} * \mathbf{d}_m - \mathbf{s}_k \right\|_2^2 + \sum_{m=1}^M \iota_{C_{PN}}(\mathbf{d}_m), \quad (2.20)$$

$$\iota_{C_{PN}}(\mathbf{y}) = \begin{cases} 0 & \mathbf{y} \in C_{PN} \\ \infty & \text{otherwise} \end{cases} \quad (2.21)$$

この目的関数も係数の場合と同様に凸関数であるため, ADMM を適用することで解く. ADMM に当てはまるように, 式 (2.20) を以下のように書き換える.

$$\arg \min_{\mathbf{d}} \frac{1}{2} \|\mathbf{g}_2\|_2^2 + \iota_{C_{PN}}(\mathbf{g}_3) \quad \text{s.t.} \quad \begin{pmatrix} \mathbf{X}_k \\ \mathbf{I} \end{pmatrix} \mathbf{d} - \begin{pmatrix} \mathbf{g}_2 \\ \mathbf{g}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{s}_k \\ \mathbf{0} \end{pmatrix} \quad (2.22)$$

ここで、巡回畳み込みで生成する辞書行列  $\mathbf{X}_{k,m}$  を  $\mathbf{X}_{k,m}\mathbf{d}_m = \mathbf{x}_{k,m} * \mathbf{d}_m$  が成り立つように定義し、 $\mathbf{X}_k$  および  $\mathbf{d}$  を以下のように定義する.

$$\mathbf{X}_k = (\mathbf{X}_{k,1} \quad \mathbf{X}_{k,2} \quad \dots \quad \mathbf{X}_{k,M}) \quad \mathbf{d} = (\mathbf{d}_1 \quad \mathbf{d}_2 \quad \dots \quad \mathbf{d}_M)^T \quad (2.23)$$

式 (2.7) のように拡張ラグランジュ関数を導出し、この問題の更新式は以下のようになる.

$$\mathbf{d}^{(i+1)} = \arg \min_{\mathbf{d}} \frac{\rho}{2} \left\| \mathbf{X}_k \mathbf{d} - \mathbf{g}_2^{(i)} - \mathbf{s}_k + \mathbf{h}_2^{(i)} \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{d} - \mathbf{g}_3^{(i)} + \mathbf{h}_3^{(i)} \right\|_2^2 \quad (2.24)$$

$$\mathbf{g}_2^{(i+1)} = \arg \min_{\mathbf{g}_2} \frac{1}{2} \left\| \mathbf{g}_2 \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{X}_k \mathbf{d} - \mathbf{g}_2^{(i)} - \mathbf{s}_k + \mathbf{h}_2^{(i)} \right\|_2^2 \quad (2.25)$$

$$\mathbf{g}_3^{(i+1)} = \arg \min_{\mathbf{g}_3} \iota_{CPN}(\mathbf{g}_3) + \frac{\rho}{2} \left\| \mathbf{d} - \mathbf{g}_3^{(i)} + \mathbf{h}_3^{(i)} \right\|_2^2 \quad (2.26)$$

$$\mathbf{h}_2^{(i+1)} = \mathbf{h}_2^{(i)} + \mathbf{X}_k \mathbf{d} - \mathbf{g}_2^{(i)} - \mathbf{s}_k \quad (2.27)$$

$$\mathbf{h}_3^{(i+1)} = \mathbf{h}_3^{(i)} + \mathbf{d} - \mathbf{g}_3^{(i)} \quad (2.28)$$

実領域において畳み込み演算を行っている  $\mathbf{X}_k \mathbf{d}$  を解くことは容易ではないため、フーリエ領域で解く解くことを考える.  $\hat{\mathbf{s}}_k$  を  $\mathbf{s}_k$  のフーリエ変換,  $\hat{\mathbf{d}}$  を  $\mathbf{d}$  のフーリエ変換,  $\hat{\mathbf{x}}_k$  を  $\mathbf{x}_k$  のフーリエ変換とし、 $\hat{\mathbf{X}}_k$  を以下のように定義する.

$$\hat{\mathbf{X}}_k = \begin{pmatrix} \text{diag}(\hat{\mathbf{X}}_{k,1}) & \text{diag}(\hat{\mathbf{X}}_{k,2}) & \dots & \text{diag}(\hat{\mathbf{X}}_{k,M}) \end{pmatrix} \quad (2.29)$$

係数最適化と同様に、式 (2.24) は以下のように書き換えることができる.

$$\hat{\mathbf{d}}^{(i+1)} = \arg \min_{\hat{\mathbf{d}}} \frac{\rho}{2} \left\| \hat{\mathbf{X}}_k \hat{\mathbf{d}} - \hat{\mathbf{g}}_2^{(i)} - \hat{\mathbf{s}}_k + \hat{\mathbf{h}}_2^{(i)} \right\|_2^2 + \frac{\rho}{2} \left\| \hat{\mathbf{d}} - \hat{\mathbf{g}}_3^{(i)} + \hat{\mathbf{h}}_3^{(i)} \right\|_2^2 \quad (2.30)$$

これを最小化する  $\hat{\mathbf{d}}$  は以下を満たす.

$$\hat{\mathbf{X}}_k^H \left( \hat{\mathbf{X}}_k \hat{\mathbf{d}} - \hat{\mathbf{g}}_2^{(i)} - \hat{\mathbf{s}}_k + \hat{\mathbf{h}}_2^{(i)} \right) + \left( \hat{\mathbf{d}} - \hat{\mathbf{g}}_3^{(i)} + \hat{\mathbf{h}}_3^{(i)} \right) = \mathbf{0} \quad (2.31)$$



$\hat{\mathbf{d}}$  の更新式は以下のようになる.

$$\hat{\mathbf{d}}^{(i+1)} = \left( \hat{\mathbf{X}}_k^H \hat{\mathbf{X}}_k + \mathbf{I} \right)^{-1} \left\{ \hat{\mathbf{X}}_k^H \left( \hat{\mathbf{g}}_2^{(i)} + \hat{\mathbf{s}}_k - \hat{\mathbf{h}}_2^{(i)} \right) + \left( \hat{\mathbf{g}}_3^{(i)} - \hat{\mathbf{h}}_3^{(i)} \right) \right\} \quad (2.32)$$

式 (2.32) の  $\hat{\mathbf{X}}_k^H \hat{\mathbf{X}}_k$  はサイズが大きくなり計算コストが増加するため, Sherman-Morison-Woodbury の逆行列の補助定理を用いて変形する.

$$\hat{\mathbf{d}}^{(i+1)} = \left( \hat{\mathbf{g}}_3^{(i)} - \hat{\mathbf{h}}_3^{(i)} \right) + \hat{\mathbf{X}}_k^H \left( \hat{\mathbf{X}}_k \hat{\mathbf{X}}_k^H + \mathbf{I} \right)^{-1} \left\{ \hat{\mathbf{g}}_2^{(i)} + \hat{\mathbf{s}} - \hat{\mathbf{h}}_2^{(i)} - \hat{\mathbf{X}}_k \left( \hat{\mathbf{g}}_3^{(i)} - \hat{\mathbf{h}}_3^{(i)} \right) \right\} \quad (2.33)$$

式 (2.25) は,  $\mathbf{g}_2$  についての勾配を得ることで求めることができる.

$$\mathbf{g}_2^{(i+1)} = \frac{\rho}{1 + \rho} \left( \mathbf{X}_k \mathbf{d}^{(i+1)} - \mathbf{s}_k + \mathbf{h}_2^{(i)} \right) \quad (2.34)$$

式 (2.26) は, 指示関数近接作用素を用いて解くことができる.

$$\mathbf{g}_3^{(i+1)} = \text{prox}_{1/\rho \iota_{C_{PN}}(\cdot)} \left( \mathbf{d}^{(i+1)} + \mathbf{h}_3^{(i)} \right), \quad (2.35)$$

$$\text{prox}_{\gamma \iota_{C_{PN}}(\cdot)}(\mathbf{y}) = \begin{cases} \mathbf{P} \mathbf{P}^T \mathbf{y} & \|\mathbf{P} \mathbf{P}^T \mathbf{y}\|_2 \leq 1 \\ \frac{\mathbf{P} \mathbf{P}^T \mathbf{y}}{\|\mathbf{P} \mathbf{P}^T \mathbf{y}\|_2} & \|\mathbf{P} \mathbf{P}^T \mathbf{y}\|_2 > 1 \end{cases} \quad (2.36)$$

## 2.2 SWT

2次元 SWT は, 画像信号  $S$  を複数レベルのサブ画像に分解し, それぞれのレベルにおいて近似係数  $Ca$  と詳細係数  $Cd$  を生成する方法である. 各レベルのサブ画像は前レベルの近似係数に依存して決定される. SWT の構造は図 2.1 に示す.

ウェーブレット分解は, 信号を各レベル  $j$  においてローパスフィルタ  $h_j$  とハイパスフィルタ  $g_j$  で畳み込むことによって実行される. ローパスフィルタ (すなわちスケーリング関数) での畳み込みによって近似係数  $Ca$  が得られ, ハイパスフィルタ (すなわち

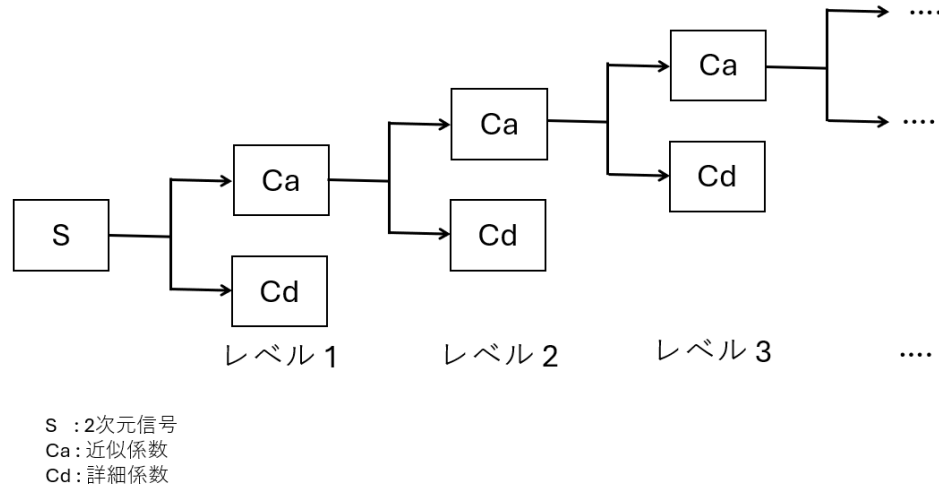


図 2.1: SWT の構造

ウェーブレット関数)での畳み込によって詳細係数  $Cd$  が得られる。詳細係数には、水平詳細、垂直詳細、および対角詳細の3つのサブバンド画像が含まれて、信号の高周波成分(エッジ、特徴、ノイズなど)を表示する。

$$Ca_j[n] = \sum_k h_j[k] Ca_{j-1}[k-n] \quad (j = 0, 1, \dots) \quad (2.37)$$

$$Cd_j[n] = \sum_k g_j[k] Ca_{j-1}[k-n] \quad (j = 0, 1, \dots) \quad (2.38)$$

SISR モデルにおいては、画像情報の損失と計算コストの増大が主要な課題とされている。この問題を解決するために定常ウェーブレット変換 (SWT) を導入する。SWT は、従来の離散ウェーブレット変換 (DWT : Discrete Wavelet Transform) とは異なり、ダウンサンプリングを行わず、より詳細な画像情報を保持できる点で優れている。また、SWT を適用すると、変換後も画像サイズが変化しないため、情報の損失を抑えることが可能となる。さらに、分解レベルが増加することで、異なる解像度の類似画像を生成できる点 SWT のも利点の一つである。

従来の SISR モデルではバイキュービック劣化を利用して LR 画像を取得するが, [2] では SWT を活用することで, スパース性と方向性を強調し, 精度向上と計算コストの削減を図る. ウェーブレット領域では, 画像の方向性特徴を高精度に捉えることが可能であり, 空間領域ベースの手法よりも有利である. また, ウェーブレットのスパース性により, フィルタサイズを小型化し, 計算コストを削減できる.

## 2.3 CSR および SWT による SISR

ウェーブレットドメインを利用した SISR モデルでは, 入力画像を SWT によって 2 つのレベルに分解する. レベル 1 は LR サブ画像, レベル 2 は HR サブ画像として扱われる. 各サブ画像は近似係数と詳細係数で構成されるが, 超解像のために, 重要な情報を持つのは詳細係数である. そのため, [2] では, CSR による学習の入力信号として詳細係数のみを使用する.

本モデルは学習段階と評価段階の 2 つのステージで構成される. 学習は, 以下の 3 つのプロセスに基づいて実施される.

- (1) LR フィルタの学習: CSR を用いて LR サブ画像から LR 特徴マップを抽出する.
- (2) HR フィルタの学習: 同様に, CSR を用いて HR サブ画像から HR 特徴マップを抽出する.
- (3) マッピング関数の学習: 抽出された特徴マップを基に, LR 特徴マップから HR 特徴マップを再構築するマッピング関数を学習する.

学習済み LR フィルタ, HR フィルタ, およびマッピング関数は, 後の評価過程においてモデルの性能評価に使用される.

評価段階では、まず入力 LR 画像 に対し、学習済みの LR フィルタを適用して LR 特徴マップを抽出する。次に、学習済みのマッピング関数を用いて、LR 特徴マップから HR 特徴マップを推定する。最後に、HR 特徴マップと学習済み HR フィルタを用いて、HR 画像を再構築する。学習および評価処理の流れを図 2.2 と図 2.3 に示す。

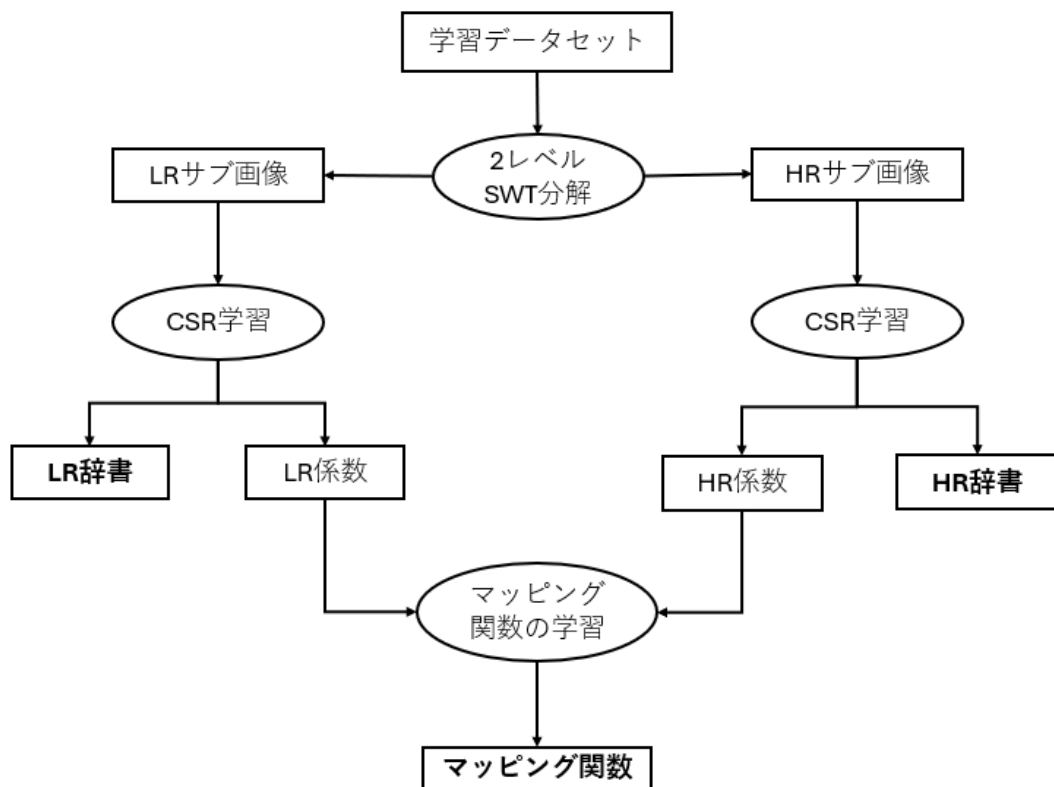


図 2.2: 訓練処理の流れ

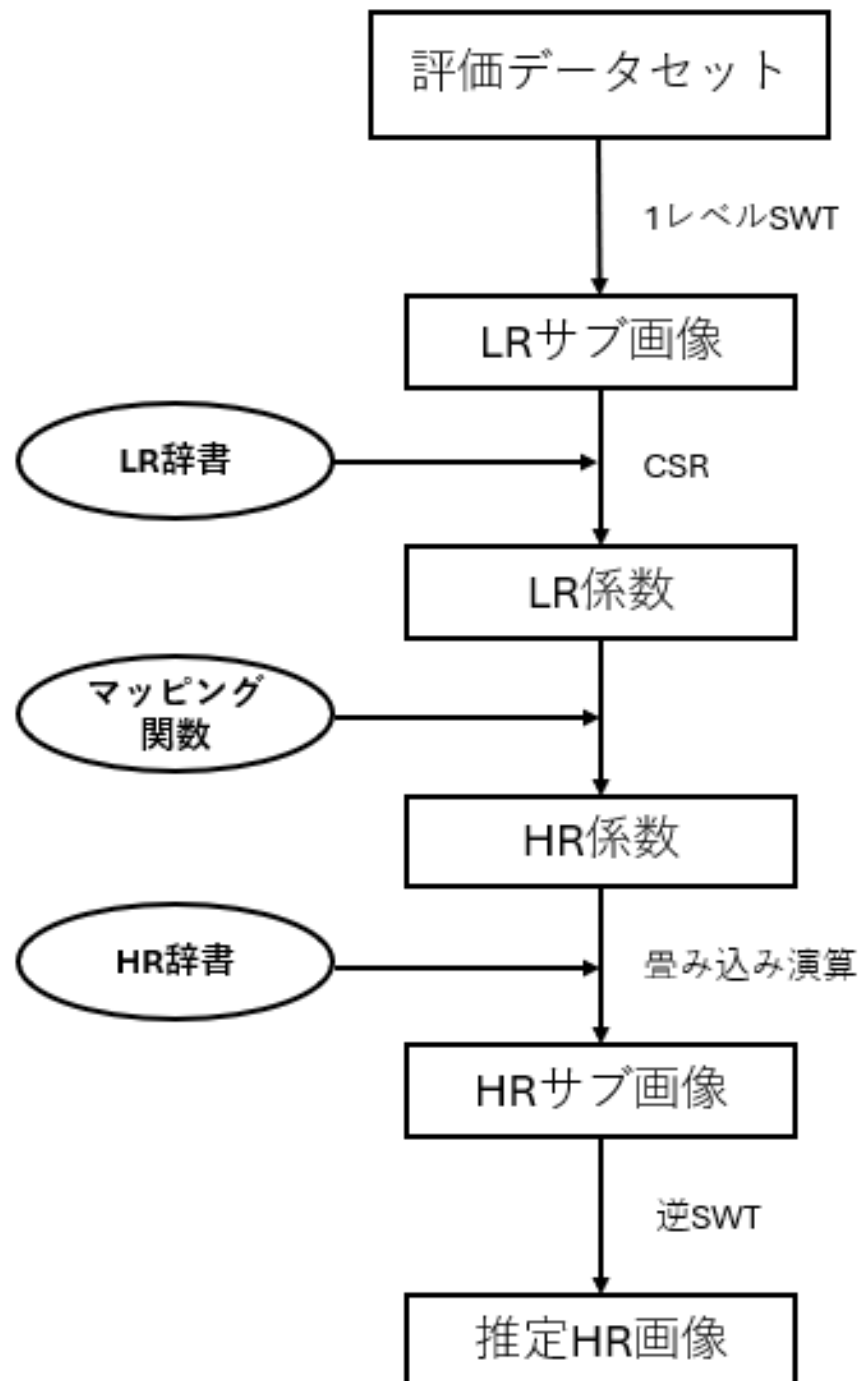


図 2.3: 評価処理の流れ

## 第3章

## 提案手法

本研究では, より高精度な超解像を実現するために, 従来の CSR における L2 ノルムの誤差項を L1 ノルムの誤差項に置き換えることを提案する. L1 ノルムは L2 ノルムに比べて外れ値に対する頑健性が高い [5] という特性が優れているため, モデルの近似誤差を評価することに適用する.

### 3.1 学習過程

#### 3.1.1 LR および HR フィルタの学習

図 2.2 に示すように, 入力画像  $\mathbf{s}_k$  はレベル 2 の SWT で LR サブ画像  $\mathbf{s}_k^l$  と HR サブ画像  $\mathbf{s}_k^h$  に分解される. LR サブ画像  $\mathbf{s}_k^l$  が与えられると, 式 2.2 に基づき, LR フィルタと特徴マップの学習を行う.

$$\arg \min_{\mathbf{d}_m^l, \mathbf{x}_{k,m}^l} \frac{1}{2} \left\| \sum_{m=1}^M \mathbf{d}_m^l * \mathbf{x}_{k,m}^l - \mathbf{s}_k^l \right\|_1 + \lambda \sum_{m=1}^M \|\mathbf{x}_{k,m}^l\|_1 \quad s.t. \quad \mathbf{d}_m^l \in C_{PN} \forall m \quad (3.1)$$

$$C_{PN} = \{\mathbf{y} \in \mathbb{R}^N \mid (\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{y} = \mathbf{0}, \|\mathbf{y}\|_2 \leq 1\} \quad (3.2)$$

ここで,  $\mathbf{x}_{k,m}^l$  は LR 特徴マップ,  $\mathbf{d}_m^l$  は LR フィルタを表し,  $M$  は LR フィルタの枚数を示す. これにより, 式 (2.4) と (2.20) のように係数と辞書の最適化問題を解いて, 式を

以下のように書き直す.

$$\arg \min_{\mathbf{x}_{k,m}^l} \left\| \sum_{m=1}^M \mathbf{d}_m^l * \mathbf{x}_{k,m}^l - \mathbf{s}_k^l \right\|_1 + \lambda \sum_{m=1}^M \|\mathbf{x}_{k,m}^l\|_1 \quad (3.3)$$

$$\arg \min_{\mathbf{d}_m^l} \left\| \sum_{m=1}^M \mathbf{x}_{k,m}^l * \mathbf{d}_m^l - \mathbf{s}_k^l \right\|_1 + \sum_{m=1}^M \iota_{C_{PN}}(\mathbf{d}_m^l) \quad (3.4)$$

$$\iota_{C_{PN}}(\mathbf{y}) = \begin{cases} 0 & \mathbf{y} \in C_{PN} \\ \infty & \text{otherwise} \end{cases} \quad (3.5)$$

スケール形式の ADMM を用いて最適化問題を解決するために, まず数式を変形し, 従来の L2 ノルムによる CSR と同様に L1 ノルムによる CSR を解ける. ただし, 式 (2.9) と式 (2.26) は以下のように書き直す必要がある.

$$\mathbf{g}_0^{(i+1)} = \arg \min_{\mathbf{g}_0} \|\mathbf{g}_0\|_1 + \frac{\rho}{2} \left\| \mathbf{D} \mathbf{x}_k^{(i+1)} - \mathbf{g}_0 - \mathbf{s}_k^l + \mathbf{h}_0^{(i)} \right\|_2^2 \quad (3.6)$$

$$\mathbf{g}_2^{(i+1)} = \arg \min_{\mathbf{g}_2} \|\mathbf{g}_2\|_1 + \frac{\rho}{2} \left\| \mathbf{X}_k \mathbf{d} - \mathbf{g}_2^{(i)} - \mathbf{s}_k^l + \mathbf{h}_2^{(i)} \right\|_2^2 \quad (3.7)$$

更新式は勾配法ではなく, 指示関数近接作用素を用いて解くことができる. そして, HR サブ画像  $\mathbf{s}_k^h$  も LR サブ画像  $\mathbf{s}_k^l$  と同様に,  $N$  枚の HR フィルタ  $\mathbf{d}_n^h$  と特徴マップ  $\mathbf{x}_{k,n}^h$  を学習する.

### 3.1.2 マッピング関数の学習

ほとんどの SISR アルゴリズムでは, LR 画像は HR 画像と同じサイズになるように補間される. しかし, 本研究では SWT を使用しているため, この補間処理は省略される. HR 特徴マップと LR 特徴マップのマッピングは次式のように表される.

$$\mathbf{x}_{k,n}^h \approx \sum_{m=1}^M \mathbf{W}_{n,m} \mathbf{x}_{k,m}^l \quad (3.8)$$

ここで、マッピング関数  $\mathbf{W}_{n,m} \in \mathbb{R}^{U \times U}$  は、LR 係数  $\mathbf{x}_{k,m}^l$  を HR 係数  $\mathbf{x}_{k,n}^h$  に写像するためのマッピング行列であり、LR 係数  $\mathbf{x}_{k,m}^l$  と同じ枚数を持っている。マッピング関数の最適化問題を解くために、次式のように定式化する必要がある。

$$\arg \min_{\mathbf{W}_{n,m}} \left\| \sum_{m=1}^M \mathbf{W}_{n,m} \mathbf{x}_{k,m}^l - \mathbf{x}_{k,n}^h \right\|_1 + \lambda \sum_{m=1}^M \|\mathbf{W}_{n,m}\|_1 \quad (3.9)$$

ここで、第2項の  $\mathbf{W}_{n,m}$  はスパースベクトルのマッピングに使用されるため、 $\mathbf{W}_{n,m}$  もスパースである必要がある。式 (3.9) の第2項目は  $\mathbf{W}_{n,m}$  のスパース性を制御して増加させるために含まれている。式 (3.9) は、計算簡略化のため、ベクトル表記に変形し、 $\mathbf{w}_n$  および  $\mathbf{x}_k^l$  を以下のように定義する。

$$\mathbf{w}_n = (\mathbf{W}_{n,1} \quad \mathbf{W}_{n,2} \quad \dots \quad \mathbf{W}_{n,M}) \quad \mathbf{x}_k^l = (\mathbf{x}_{k,1}^l \quad \mathbf{x}_{k,2}^l \quad \dots \quad \mathbf{x}_{k,M}^l)^T \quad (3.10)$$

これにより、式 (3.9) は以下のように書き換える。

$$\arg \min_{\mathbf{w}_n} \|\mathbf{w}_n \mathbf{x}_k^l - \mathbf{x}_{k,n}^h\|_1 + \lambda \|\mathbf{w}_n\|_1 \quad (3.11)$$

この目的関数も CSR の最適化問題と同様に凸関数であるため、ADMM を適用することで解く。ADMM に当てはまることができるように、式 (3.11) を以下のように書き換える。

$$\arg \min_{\mathbf{w}_n} \|\mathbf{g}_4\|_1 + \lambda \|\mathbf{g}_5\|_1 \quad s.t \quad \mathbf{w}_n \begin{pmatrix} \mathbf{x}_k^l \\ \mathbf{I} \end{pmatrix} - \begin{pmatrix} \mathbf{g}_4 \\ \mathbf{g}_5 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{k,n}^h \\ \mathbf{0} \end{pmatrix} \quad (3.12)$$



この問題の更新式は以下のようになる.

$$\mathbf{w}_n^{(i+1)} = \arg \min_{\mathbf{w}_n} \frac{\rho}{2} \left\| \mathbf{w}_n \mathbf{x}_k^l - \mathbf{g}_4^{(i)} - \mathbf{x}_{k,n}^h + \mathbf{h}_4^{(i)} \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{w}_n - \mathbf{g}_5^{(i)} + \mathbf{h}_5^{(i)} \right\|_2^2 \quad (3.13)$$

$$\mathbf{g}_4^{(i+1)} = \arg \min_{\mathbf{g}_4} \left\| \mathbf{g}_4 \right\|_1 + \frac{\rho}{2} \left\| \mathbf{w}_n^{(i+1)} \mathbf{x}_k^l - \mathbf{g}_4^{(i)} - \mathbf{x}_{k,n}^h + \mathbf{h}_4^{(i)} \right\|_2^2 \quad (3.14)$$

$$\mathbf{g}_5^{(i+1)} = \arg \min_{\mathbf{g}_5} \lambda \left\| \mathbf{g}_5 \right\|_1 + \frac{\rho}{2} \left\| \mathbf{w}_n^{(i+1)} - \mathbf{g}_5^{(i)} + \mathbf{h}_5^{(i)} \right\|_2^2 \quad (3.15)$$

$$\mathbf{h}_4^{(i+1)} = \mathbf{h}_4^{(i)} + \mathbf{w}_n^{(i+1)} \mathbf{x}_k^l - \mathbf{g}_4^{(i+1)} - \mathbf{x}_{k,n}^h \quad (3.16)$$

$$\mathbf{h}_5^{(i+1)} = \mathbf{h}_5^{(i)} + \mathbf{w}_n^{(i+1)} - \mathbf{g}_5^{(i+1)} \quad (3.17)$$

式 (3.13) を最小化する  $\mathbf{w}_n$  は以下を満たす.

$$\left( \mathbf{w}_n \mathbf{x}_k^l - \mathbf{g}_4^{(i)} - \mathbf{x}_{k,n}^h + \mathbf{h}_4^{(i)} \right) \mathbf{x}_k^{lT} + \left( \mathbf{w}_n - \mathbf{g}_5^{(i)} + \mathbf{h}_5^{(i)} \right) = \mathbf{0} \quad (3.18)$$

$\mathbf{w}_n$  の更新式は以下のようになる.

$$\mathbf{w}_n^{(i+1)} = \left\{ \left( \mathbf{g}_4^{(i)} + \mathbf{x}_{k,n}^h - \mathbf{h}_4^{(i)} \right) \mathbf{x}_k^{lT} + \left( \mathbf{g}_5^{(i)} - \mathbf{h}_5^{(i)} \right) \right\} \left( \mathbf{x}_k^l \mathbf{x}_k^{lT} + \mathbf{I} \right)^{-1} \quad (3.19)$$

式 (3.19) の  $\mathbf{x}_k^l \mathbf{x}_k^{lT}$  はサイズが大きく計算コストが増加するため, Sherman-Morison-Woodbury の逆行列の補助定理を用いて変形する.

$$\mathbf{w}_n^{(i+1)} = \left\{ \left( \mathbf{g}_4^{(i)} + \mathbf{x}_{k,n}^h - \mathbf{h}_4^{(i)} \right) \mathbf{x}_k^{lT} + \left( \mathbf{g}_5^{(i)} - \mathbf{h}_5^{(i)} \right) \right\} \left\{ \mathbf{I} - \mathbf{x}_k^l \left( \mathbf{I} + \mathbf{x}_k^l \mathbf{x}_k^{lT} \right)^{-1} \mathbf{x}_k^{lT} \right\} \quad (3.20)$$

式 (3.14) と 式 (3.15) は指示関数近接作用素を用いて解くことができる.

## 3.2 評価過程

本段階では, 前段階で構築された SISR モデルの性能を検証する. 評価には, 学習済み LR フィルタ, HR フィルタ, およびマッピング関数を用いる. まず, 入力 LR 画像

に対し, レベル 1 の SWT を適用して LR サブ画像を取得する. その後, 学習済み LR フィルタを使用して LR 特徴マップを抽出する. 次に, 学習済みマッピング関数と LR 特徴マップを使用して HR 特徴マップを推定する. 最後に, HR 特徴マップと対応する学習済み HR フィルタを畳み込むことで, 高解像度のサブ画像が作成される.

$$\mathbf{s}_k^h \approx \sum_{n=1}^N \mathbf{d}_n^h * \mathbf{x}_{k,n}^h \quad (3.21)$$

最終的に, HR サブ画像に対し, レベル 1 の逆定常ウェーブレット変換 (ISWT : Inverse Stationary Wavelet Transform) を適用して HR 画像を推定する.

## 第4章

# 実験

学習過程では, DIV2K データセット ( $256 \times 256$  ピクセル) から 25 枚の画像を使用して実験した. 辞書のブロックサイズは  $5 \times 5$  に設定され, LR 辞書には 20 枚, HR 辞書には 30 枚のフィルタが作成された. パラメータ  $\lambda$  は特徴マップとマッピング関数の学習で 0.000025 に設定され, パラメータ  $\rho$  はすべての学習で 0.00025 に設定された. 評価過程では, LR テスト画像と HR テスト画像を含む Set5 データセットの画像を使用した. 3 種類の異なる解像度 ( $256 \times 256$ ,  $170 \times 170$ ,  $128 \times 128$  ピクセル) の各 LR テスト画像を使用して実験し, 本モデルによって再構築された HR 画像を HR テスト画像と比較した.

L1 ノルムに基づく誤差項を用いたモデルと, L2 ノルムに基づく誤差項を用いたモデルの比較を行った. それぞれの方法で再構築された HR 画像の精度を比較するため, 元のテスト画像と再構成された画像の間の PSNR (Peak Signal-to-Noise Ratio) および SSIM (Structural Similarity Index Measure) を比較した. 提案手法, 比較手法で得られた PSNR 値および SSIM 値を表 4.1, 4.2, 4.3 に示す. また, 図 4.1, 4.2 は, image 1 であるテスト画像の再構築結果と比較画像を視覚的に比較する.

表 4.3 に示すように, 比較手法より提案手法の方が平均 PSNR および SSIM の値が高くなっている. 図 4.2 の画像比較では, 解像度が異なっても, L1 制約付き SISR モデルから再構成された画像の方が L2 制約付き SISR モデルよりも鮮明である. これらの比較により, 提案手法の SISR モデルの方が比較手法の SISR モデルよりも画像再

構成の精度が優れていることが分かった。

表 4.1: L1 の PSNR 値と SSIM 値

		image 1	image 2	image 3	image 4	image 5
X2	PSNR	32.951	32.980	32.965	32.981	32.966
	SSIM	0.9602	0.9786	0.9652	0.9792	0.9658
X3	PSNR	32.952	32.982	32.964	32.981	32.966
	SSIM	0.9573	0.9794	0.9621	0.9789	0.9634
X4	PSNR	32.958	32.983	32.964	32.982	32.965
	SSIM	0.9601	0.9802	0.9612	0.9795	0.9618

表 4.2: L2 の PSNR 値と SSIM 値

		image 1	image 2	image 3	image 4	image 5
X2	PSNR	29.970	29.985	29.977	29.985	29.978
	SSIM	0.9603	0.9786	0.9652	0.9792	0.9658
X3	PSNR	29.970	29.986	29.977	29.985	29.978
	SSIM	0.9572	0.9794	0.9621	0.9789	0.9635
X4	PSNR	29.974	29.986	29.977	29.985	29.977
	SSIM	0.9601	0.9802	0.9611	0.9795	0.9618

表 4.3: 平均 PSNR 値と SSIM 値の比較結果

image size [pixels]		256 × 256	170 × 170	128 × 128
PSNR [dB]	L1	<b>32.969</b>	<b>32.969</b>	<b>32.970</b>
	L2	29.979	29.979	29.980
SSIM	L1	<b>0.9700</b>	0.9682	0.9686
	L2	0.9698	0.9682	0.9686



図 4.1: HR テスト画像 (512 × 512)



(a) L1 の X2 (256×256)



(b) L1 の X3 (170×170)



(c) L1 の X4 (128×128)



(d) L2 の X2 (256×256)



(e) L2 の X3 (170×170)



(f) L2 の X4 (128×128)

図 4.2: L1 または L2 CSR で再構築された HR 画像の比較

## 第5章

### むすび

本研究では, L2 ノルム 制約付き CSR ではなく, L1 ノルム 制約付き CSR に基づく新しい SISR モデルを設計した. 提案もでると比較モデルの PSNR および SSIM 値を比較し, L1 ノルム誤差項のモデルが L2 ノルム誤差項のモデルよりも精度が高いことを証明することが成功した. 今後の課題として, グレースケール画像のみならず, カラー画像にも対応可能なモデルの設計を検討する. また, モデルの処理精度向上のために学習画像の枚数を増やすことや, メモリ使用量の最適化を図ることで計算負荷の軽減することなどが挙げられる.

## 謝辞

本研究及びその基礎の学習において，貴重な時間を割いてご指導して頂いた久留米工業高等専門学校制御情報工学科の黒木祥光教授，および同研究室の本科生，専攻科生に深く感謝致します．



## 参考文献

- [1] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng and L. Zhang, "Convolutional Sparse Coding for Image Super-Resolution," 2015 IEEE Int. Conf. on Comput. Vision (ICCV), Feb 2016, doi: 10.1109/ICCV.2015.212.
- [2] A. Ahmed, S. Kun, R. A. Memon, J. Ahmed and G. Tefera, "Convolutional Sparse Coding Using Wavelets for Single Image Super-Resolution," IEEE Access, vol. 7, pp. 121350-121359, Aug. 2019, doi: 10.1109/ACCESS.2019.2936455.
- [3] B. Wohlberg, "Efficient Algorithms for Convolutional Sparse Representations," IEEE Trans. Image Processing, vol. 25, no. 1, pp. 301-315, Jan. 2016, doi: 10.1109/TIP.2015.2495260.
- [4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Found. and Trends in Mach. Learn., vol. 3, no. 1, pp. 1-122, 2010, doi: 10.1561/22000000016.
- [5] K. Kumamoto, S. Matsuo and Y. Kuroki, "A Convolutional Dictionary Learning based l1 Norm Error with Smoothed l0 Norm Regression," 2019 Int. Symp. on Intell. Signal Process. and Commun. Syst. (ISPACS), Feb 2020, doi: 10.1109/ISPACS48206.2019.8986267.