I_028

多様体学習に基づく単フレーム超解像

Manifold Learning Based Single-Frame Super-Resolution

溝内 義章 †　　末松 伸朗 †　　林 朗 †
Yoshiaki Mizouchi　　Nobuo Suematsu　　Akira Hayashi

1　はじめに

近年デジタルカメラが普及しデジタル画像の取り扱いが一般的になってきた．しかしそれらの解像度には限界があり，また高解像度画像を扱うものは高いコストを必要とする．これに対し低解像度画像から高解像度画像を構築する超解像という画像処理技術が注目されている．

超解像では多視点の画像集合を用いたもの，特にカメラの外部パラメータ推定をともなう手法が研究されてきた．本稿で扱う単フレーム超解像はこれと異なり，対応する高精細画像が既知の教師データを利用することで，一枚のみの入力画像から高精細画像を構築する．したがって，教師データの持つ規則性をいかに見出し，新規画像の超解像へ利用するかがこの手法の鍵となる．

近年急速に発展している多様体学習は，主に非線形次元削減の文脈で利用される技術である．単フレーム超解像では，教師データから規則性を抽出する方法として応用可能であり，Locally Linear Embedding (LLE)[2] と呼ばれる多様体学習を応用した手法が提案されている [1]．

本稿では，データ集合に多様体が潜在するという考え方をより積極的に用いる単フレーム超解像法を提案する．ただし，この手法はデータ空間と多様体上の空間の間の相互変換を必要とする．そこで，そのような相互変換の得られる多様体学習のひとつである Charting[3] をベースとした単フレーム画像に対する超解像法を検討する．

2　提案手法

単フレーム超解像では，画像圧縮法と同様に，画像をたとえば 8 ピクセルの小領域（パッチ）へ分割し，画像パッチを単位に処理が行われる．画像パッチをデータ点としてとらえ，その集合を考えると，それは画像の持つ規則性に従っているためにある多様体上近くに偏在していると予想される．超解像の対象となる画像と同様の画像的特徴を持つ学習データが手に入るものとする．つまり，低解像度の画像パッチと，それに対応する高解像度画像パッチのペアの集合を教師データとして使用する．ここでは，低解像度画像パッチ集合には，高解像度画像パッチ集合と同じ構造の多様体が隠れていることを仮定する．この多様体を学習することで，処理対象の低解像度画像から高解像度画像を構築する．

† 広島市立大学大学院情報科学研究科
〒731-3194 広島県広島市安佐南区大塚東 3–4–1
Email: mizo@robotics.im.hiroshima-cu.ac.jp

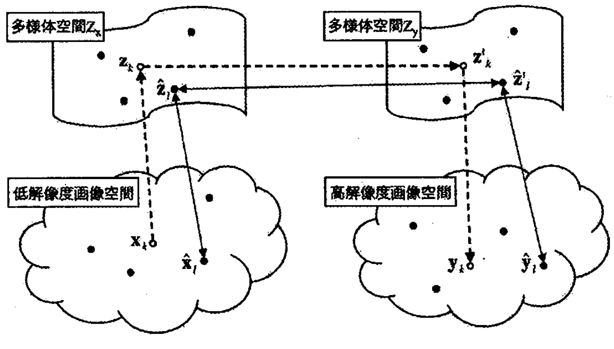

図 1　画像空間と多様体空間の関係

ある高解像度パッチを $\mathbf{y}_k \in \mathbf{R}^{D_y}$，それに対応する低解像度パッチを $\mathbf{x}_k \in \mathbf{R}^{D_x}$ とする．本稿では，高解像度パッチの集合と低解像度パッチの集合は内在的次元 $d(D_y > D_x > d)$ の多様体を持ち，その多様体は同じ構造であると仮定する．したがって，$\mathbf{y}_k$ と $\mathbf{x}_k$ は多様体上の同じ点に対応するものとする．ただし，多様体学習の結果は座標の取り方に自由度を持っていることや二つの多様体が異なる次元のデータから学習されたものであるので，それらの座標は直接比較できることはないだろう．しかし，二つの多様体表現の間の変換は，少なくとも局所的には比較的簡単に対応がつくものであろう．

超解像を行いたい低解像度パッチを $\mathbf{x}_k$，それに対応する真の高解像度パッチを $\mathbf{y}_k$ とする．また，それぞれの多様体空間へ写像された点を $\mathbf{z}_k$ と $\mathbf{z}'_k$ とすると，$\mathbf{x}_k$ から直接 $\mathbf{y}_k$ を推測するという困難な問題は，$\mathbf{z}_k$ から $\mathbf{z}'_k$ を推測するという易しい問題へ変換される．この間の変換は，教師データ中のその近傍の点から容易に推定できる簡単な変換となるはずである．

本手法は学習ステップと超解像ステップの 2 ステップで構成される．学習ステップにおいて学習用画像パッチの多様体を求める．超解像ステップにおいて入力画像パッチの多様体上の点と学習データから高解像度画像を構築する．

学習ステップ

1. 低解像度画像パッチデータ $\{\hat{\mathbf{x}}_l\}_{l=1}^{N}$ を多様体空間 Z_x に写像する
2. 高解像度画像パッチデータ $\{\hat{\mathbf{y}}_l\}_{l=1}^{N}$ を多様体空間 Z_y に写像する．次元は Z_x とそろえる

超解像ステップ

1. 超解像を行う低解像度画像パッチデータ $\{\mathbf{x}_k\}_{k=1}^{n}$ を得る
2. 各 $\mathbf{x}_k$ に対応する多様体空間 Z_x 上の点 $\{\mathbf{z}_k\}_{k=1}^{n}$ を得る
3. 各 $\mathbf{z}_k$ の近傍のデータを使用して多様体空間 Z_y との間の

最適な対応付けを行う変換 T を得る

4. その変換を使用して求めたいデータ点に対応する多様体空間 Z_y 上の点 $\{\mathbf{z}'_k\}_{k=1}^n$ を得る
5. 逆変換により各 $\mathbf{z}'_k$ に対応する高解像度画像パッチデータ $\{\mathbf{y}_k\}_{k=1}^n$ を得る
6. 得られた高解像度パッチデータを組み合わせて高解像度画像を構築する

2.1 Charting a manifold

本稿で使用する多様体学習法 Charting について説明する．Charting はある次元空間に存在する多様体を，その内在的次元空間へ非線形写像を行う．多様体上のデータ点に対し局所線形性を保つ近傍へソフトクラスタリングを行う charting ステップと，それにより得られた各 chart を同一空間へと写像し接続する connection ステップで構成される．

2.1.1 Charting ステップ

多様体上のデータ点を局所線形性が保たれる近傍で分割する．データ点を Gaussian mixture model(GMM) でモデル化することでソフトクラスタリングを行う．決めなければならないのは，GMM の持つ平均 μ_j と共分散行列 Σ_j である．これに対して，事前分布を $p(\mu,\Sigma)=\exp[-\sum_{i\neq j} m_i(\mu_j)D(\mathcal{N}_i\|\mathcal{N}_j)]$ とし，MAP 推定を行うのが Charing ステップである．ここで，$\mathcal{N}_j$ は平均 μ_j，共分散行列 Σ_j の正規分布であり，$D(\mathcal{N}_i\|\mathcal{N}_j)$ は $\mathcal{N}_i$ と $\mathcal{N}_j$ の Kullback Leibler ダイバージェンス．また，$m_i(\mu_j)$ は，μ_i と μ_j の近さをはかる尺度で，問題に応じて適宜選ばれる．

[3] では，各コンポーネントの平均値をデータ点自体に固定した場合に，残るパラメータである共分散行列が解析的に解けると主張し，その利用を前提としている．

2.1.2 Connecting ステップ

各コンポーネントの平均 μ_k と共分散行列 Σ_k に基づき，主成分分析と同様に次元削減を行うと局所線形部分多様体が定義される．これをチャートと呼ぶ．

Chart を適切に配置し（並進と回転を定め），重み付き平均によりそれらを滑らかにつなぐことで多様体の低次元パラメータ化が完成する．$d\times(d+1)$ 行列 G_k で表される適切な並進と回転は，データ点のチャートによる表現が各チャートの重みも考慮した上でなるべく一致するように定める．[3] では，この最適化問題も解析的に解けることを示している．

2.1.3 逆射影

提案する単フレーム超解像法は，多様体上からデータ空間への逆射影も必要とする．これにはいろいろ選択肢があるが，以下が比較的素直なものである．

座標空間における補間的密度

$$p_{\mathbf{z}|k}(\mathbf{z})=\mathcal{N}(\mathbf{z};\mathbf{G}_k\begin{bmatrix}\mathbf{0}\\1\end{bmatrix},\mathbf{G}_k\begin{bmatrix}[\mathbf{I}_d,\mathbf{0}]\Lambda_k[\mathbf{I}_d,\mathbf{0}]^\top & \mathbf{0}\\ \mathbf{0} & 0\end{bmatrix}\mathbf{G}_k^\top)$$

を定義する．$p(\mathbf{y}|\mathbf{z},k)$ は，平均が $\mathbf{z}$ の線形関数である Dirac のデルタ関数であるとする．このとき，$\mathbf{z}$ を生成するチャートを周辺化することで事後平均は得られる．

$$\hat{\mathbf{y}|\mathbf{z}}=\sum_k p_{k|\mathbf{z}}\left(\mu_k+\mathbf{W}_k^\top\left(\mathbf{G}_k\begin{bmatrix}\mathbf{I}\\0\end{bmatrix}\right)^+\left(\mathbf{z}-\mathbf{G}_k\begin{bmatrix}\mathbf{0}\\1\end{bmatrix}\right)\right) \tag{1}$$

ここで，$(\cdot)^+$ は疑似逆行列である．

2.2 回転と並進による対応付け

高解像度画像パッチ集合と，低解像度画像パッチ集合から学習された同じ次元の二つの多様体座標系は，比較的簡単な変換でよく一致すると期待される．

超解像の手続きでは，対象となる低解像度パッチは低解像度多様体上へ射影される．その点の近傍にある学習データ中のデータを n 個選んでくると，教師データがあるのでそれらの高解像度多様体上の対応点は分かる．同じ d 次元空間の n 個の点同士をなるべく一致させるような並進と回転からなる変換は以下のようにして見つけることができる．

d 次元の点 $\mathbf{z}_i$ が $\mathbf{z}'_i$ に対応するとし，$i=1,\ldots,n$ の全てがなるべく一致するような回転と並進の変換を見つける問題を考える．ただし，各点には重み w_i が与えられており，その重み付きでの総自乗誤差で一致の程度は測るものとする．

点の回転と並進を $d\times(d+1)$ 行列により

$$\mathbf{z}'=T\begin{bmatrix}\mathbf{z}\\1\end{bmatrix}$$

と表すものとすれば，

$$\begin{aligned}E&=\frac{1}{2}\sum_{i=1}^n w_i\|\mathbf{z}'_i-T\tilde{\mathbf{z}}_i\|^2\\&=\frac{1}{2}\|(Z'-T\tilde{Z})W^{1/2}\|_F^2\end{aligned}$$

を最小化する T を見つける問題となる．ここで、

$$\begin{aligned}\tilde{\mathbf{z}}_i&=\begin{bmatrix}\mathbf{z}_i\\1\end{bmatrix}\\ \tilde{Z}&=[\tilde{\mathbf{z}}_1\cdots\tilde{\mathbf{z}}_n],\quad Z'=[\mathbf{z}'_1\cdots\mathbf{z}'_n]\\ W^{1/2}&=\mathrm{diag}(\sqrt{w_1},\cdots,\sqrt{w_n})\end{aligned}$$

である．$\frac{\partial E}{\partial T}=0$ を解けば，

$$T=Z'W\tilde{Z}^T(\tilde{Z}W\tilde{Z}^T)^{-1} \tag{2}$$

が得られる．

3 まとめ

教師データを利用した単フレーム画像に対する超解像を提案した．低解像度画像空間と高解像度画像空間に存在する多様体を学習し，その多様体間の対応付けを行うことで超解像を実現する．今後は提案した手法の実験を行い，結果を確認して行く．

参考文献

[1] Hong Chang, Dit-Yan Yeung and Yimin Xiong. Super-resolution through neighbor embedding. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, July 2004.

[2] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500):2323-2326, 2000.

[3] Matthew Brand. Charting a manifold. NIPS2002.