

I-66 サブバンド EHMM を用いた低解像度文字画像の高解像度化

Resolution Enhancement for Low Resolution Character Images Using Subband EHMM

中里 智樹† Tomoki Nakazato 長井 隆行† Takayuki Nagai 嶺 竜治‡ Ryuji Mine 酒匂 裕‡ Hiroshi Sako 樽松 明† Akira Kurematsu

1. はじめに

近年、デジタル画像データを扱う機会の増加に伴って、デジタル画像中からテキストデータを取り出すために文字認識を行う必要性が大きくなっている。文字認識はスキャナ等で取り込んだ文書画像など高解像度の画像では十分な認識率を誇るが、デジタルカメラの画像など低解像度の画像ではその認識率は低下する。この問題点を解決するために低解像度画像を高解像度化して、認識率の向上を図ることが考えられる。

本論文では、サブバンド HMM (サブバンド隠れマルコフモデル) [1]を用いた高解像度化法を提案する。サブバンド HMM はサブバンドごとの信号を独立にモデル化するものであるが、提案手法ではサブバンド間 (低解像度と高解像度) の対応関係を同時に学習することで、低解像度の画像からは観測できない高周波成分を復元可能とする。また、2次元信号である画像を HMM でモデル化するために、HMM を擬似的に2次元に拡張した Embedded HMM (EHMM) を用いる[2]。

本論文では、提案手法により高解像度化した文字画像の2乗平均誤差 (MSE) および OCR による文字認識を通して、提案手法が有効であることを示す。

2. 高解像度化の問題

画像の高解像度化とは標本化間隔を狭めることであり、一般的に画素間を補間することで実現できる。しかし、単純な補間では画像の高周波成分を復元できないため、高解像度化した画像はぼけた画像となる。これを克服するため、近年、超解像法が提案されている[3]。超解像法とは、入力画像に制約を設けることにより高周波成分を推定可能とし、その高周波成分を用いて高解像度化することでぼけの少ない高解像度画像を復元できるものである。文献[3]では顔画像を対象とした超解像法を提案しているが、計算量やメモリの使用量が膨大となる問題がある。そこで本論文では、サブバンド HMM を用いた超解像法を提案する。提案手法では、[3]に比べ少ない計算量とメモリ使用量で効果的な高解像度化が可能である。

3. 提案手法

3.1 概要

提案する高解像度化手法は、学習フェーズと復元フェーズに分けられる。高解像度化の概要を図1に示す。学習フェーズでは、組となる高解像度と低解像度の画像から特徴ベクトルを抽出し、HMMによりモデル化した後、各解像

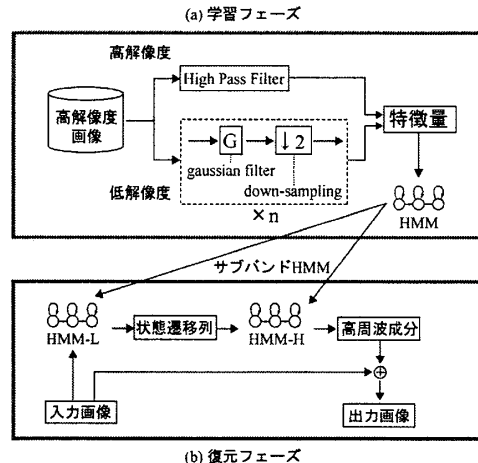


図1 高解像度化概要

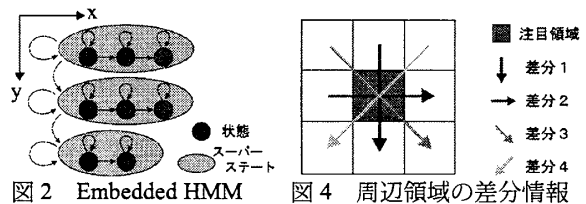


図2 Embedded HMM

図4 周辺領域の差分情報

度の HMM に分割する。復元フェーズでは低解像度の入力画像と低解像度の HMM (HMM-L) から最も尤度の高い状態遷移列を計算する。この状態遷移を高解像度の HMM (HMM-H) で行うことで入力画像に対応した高周波成分が得られ、高解像度画像を復元することができる。

3.2 サブバンド EHMM

画像をモデル化するために HMM を擬似的に2次元に拡張した Embedded HMM (EHMM) を用いる。EHMM では1次元 HMM を1つの状態 (スーパーステート) と考え、図2のように並べることにより2次元に拡張している。また、サブバンド HMM は信号を各サブバンドで独立にモデル化するものであり、これを EHMM に適用したものがサブバンド EHMM である。提案手法ではサブバンド EHMM を遷移確率を共有するように学習することで、低解像度と高解像度のモデルの対応関係を学ばせるようにしている。

遷移確率を共有して EHMM を学習するためには、まず各サブバンドで独立に特徴ベクトルを抽出し、結合して全体としての EHMM を学習する。その後、出力確率を各サブバンドに分割することで、遷移確率を共有した2つのサブバンド EHMM (低解像度と高解像度) が作成される。

3.3 特徴ベクトル

提案手法では低解像度画像の画素に対応した領域で画像の特徴量を抽出する。低解像度画像では注目画素と周辺8

† 電気通信大学 電子工学専攻
The University of Electro-Communications

‡ (株)日立製作所 中央研究所
Hitachi, Ltd., Central Research Laboratory

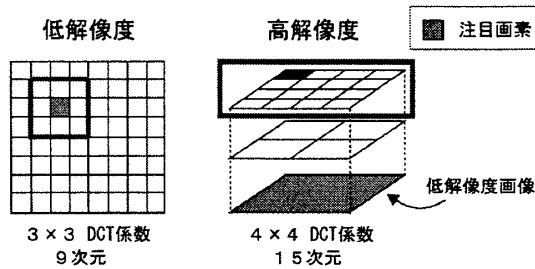


図3 特徴ベクトル (高解像度化倍率4倍)

画素の DCT 係数を特徴量とし、高解像度画像では低解像度画像の注目画素に相当する $M \times M$ 領域 (高解像度化倍率 M 倍) において、DCT 係数を特徴ベクトルとする (図3)。また、3.4 節で述べる尤度最大化を行うために高解像度画像において注目領域の周辺の差分情報を特徴ベクトルとして保存する (図4)。

3.4 周辺情報を用いた尤度最大化

サブバンド EHMM によって得られる高周波成分は、状態遷移の各状態における平均値ベクトルを並べたものであり、復元する画像は領域ごとに不連続な画像となる。これを解決するために、周辺の差分ベクトルを用いて高周波成分 p_h の尤度を最大化することを考える。

差分情報を含まない特徴ベクトルを p_h 、含んだ特徴ベクトルを O_h とする。いま、

$$O_h = W p_h \quad (1)$$

と表すと、EHMM のパラメータ λ_h 、状態遷移列 q のとき、 $P(O_h | q, \lambda_h)$ の対数尤度は次のように書ける。

$$\log P(O_h | q, \lambda_h) = \alpha(q, c_{jm}^i) - \frac{1}{2} (W p_h - \mu_h)^T \Sigma_h^{-1} (W p_h - \mu_h) + \frac{1}{2} \log |\Sigma_h^{-1}| - \beta \quad (2)$$

ただし、 μ_h は平均ベクトル、 Σ_h は共分散行列であり、 α は状態遷移列および混合重みの関数であり、 β は定数である。式(2)は p_h に関する2次形式であり、

$$\frac{\partial \log P(O_h | q, \lambda_h)}{\partial p_h} = 0 \quad (3)$$

のとき最大となる。したがって、次の線型方程式を解けば尤度を最大化することができる。

$$\Gamma p_h = v \quad (4)$$

ただし、

$$\Gamma = W^T \Sigma_h^{-1} W, \quad v = W^T \Sigma_h^{-1} \mu_h$$

である。

4. 実験

PC フォントデータを用いて 8×8 画素から 32×32 画素への4倍の高解像度化を行った。モデル化に使用したEHMMはスーパー状態数8個、スーパー状態内の状態数は各8個、各状態につき混合数は20個とした。

高解像度化した画像の例を図5に示す。図より、提案手法では3次補間法に比べてぼけの少ない画像が得られていることが分かる。また、高解像度化した画像の評価として、オリジナルの画像とのMSEを表1に、OCRによる認識実

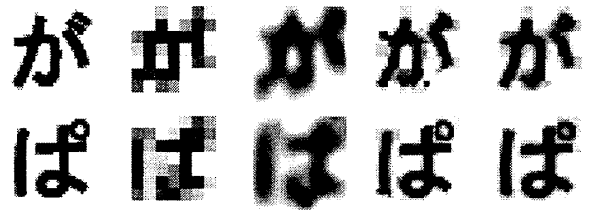


図5 高解像度化の例

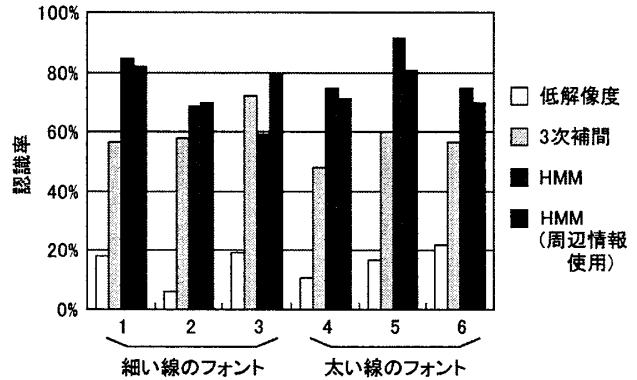


図6 フォント種別の文字認識結果

表1 オリジナルの画像とのMSE

	3次補間	文献[3]	HMM	HMM (周辺情報使用)
MSE	3150	2281	2402	2000

験結果を図6に示す。テストデータには、HMMのトレーニングに用いていないオープンデータの平仮名およびアルファベットを用いた。表1から提案手法は、3次補間法、文献[3]よりも良い結果が得られているのが分かる。また、図6からは提案手法が文字認識にも有効であることが確認できる。

5. まとめ

本論文では、サブバンドHMMをベースとした画像の高解像度化手法を提案した。提案手法を用いることでぼけの少ない高解像度画像が得られ、文字認識においても有効であることを示した。

今後の課題としては、画像の位置ずれ、EHMMの状態数による違いの評価などが挙げられる。

参考文献

- [1] 細木光宏, 長井隆行, 樽松明, “サブバンドHMMによる音声の広帯域化とノイズ除去”, 第16回デジタル信号処理シンポジウム講演論文集, pp.1-6, November, 2001.
- [2] Ara V. Nefian and monson H. Hayes III, “Face Recognition Using an Embedded HMM”, IEEE International Conference Audio Video Biometric based Person Authentication, pp.19-24, March, 1999.
- [3] Simon Baker and Takeo Kanade, “Hallucinating faces”, Fourth International Conference on Automatic Face and Gesture Recognition, March, 2000.