I-010

文字を構成する線画の組み合わせの機械学習
Machine learning of the combination of the drawing which constitutes a letter

草野周太 † 酒井智弥 †
Shuta Kusano Tomoya Sakai

1. はじめに

文字認識技術は今日の情報化社会において非常に重要であり，例えば光学文字認識は郵便物の宛先読み取りなどのため古くから研究開発されてきた [1][2]．また，携帯端末による文字の取得や，抽出した文字を頼りに関連する情報をユーザに提供するなど様々な応用が開発されている．

情景画像の文字認識の典型的な既存手法は，文字を検出するための文字らしさについての学習と，学習した文字らしさを探索する手続きから構成されている．例えば長井ら [5] の手法では，文字らしさの学習データとして情景画像の中から文字が存在する可能性が高い看板領域を抽出し，その抽出した画像領域のウェーブレット変換，独立成分分析によって文字の特徴を学習している．國重ら [6] の手法では，画像中の任意の位置からブロック状に取り出した小領域の離散コサイン変換を特徴量に用いている．

情景画像中の文字は，低解像度によるつぶれや，動きによるブレやボケ，他の物体に部分的に隠されることによる欠損，明暗や影による光学的な変化，観測ノイズ等による汚損が考えられる．文字検出はこれらに影響され難い頑健性が求められる．文字には，文字に似た特徴をもつ図形や物体にはない文字独特の線画の組合せが存在するはずである．すなわち，これが文字らしさである．この文字らしさを文字の線画の組合せと考え，基底として学習する辞書を作成することができれば，文字一文字を学習して文字検出を行う場合と異なり，文字の線画の組み合わせを学習して文字検出を行うため，上に述べた他の物体に部分的に隠されることによる欠損や，明暗や影による光学的な劣化等が存在する画像に対しても，文字検出を行うことができるのではないかと考える。

本研究では，文字らしさを基底にもつ辞書を作成することを目指す．特に少数の基底で文字を合成するスパース表現のための辞書 [7][8] に着目し，漢字のように比較的複雑な構造を持つ文字の文字らしさを線画等の構成要素の組合せとして学習する可能性を検討する．文字はいくつかの線画が組み合わさって構成されている．例えば，図 1 の「海」という漢字は「さんずい」と「毎」という字の組合せで構成されている．さらに「さんずい」は 3 つの曲線状の線画の組合せにより構成されている．同様に「毎」も線画の組合せにより構成されている．

様々な文字に共通の典型的な線画や構成要素を基底に持つ辞書をスパースコーディングによって学習できれば，これを用いて画像が簡潔に表現されるか否かに基づき文字・非文字を判別できる可能性がある．特に

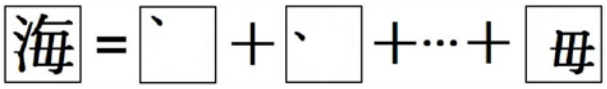

図 1: 漢字は偏や旁で構成されており，それらは線画の組合せで構成される．ゆえに，文字を構成する線画の線形和で文字を合成できると仮定する．

漢字は単に基本的な構成要素を持つだけでなく，それらの構成要素には典型的な組合せがある．基本的な構成要素から成り立っていること，かつ，それらが漢字に見られる典型的な組合せになっていることが漢字らしさであろう．そのような文字らしさを文字画像の集合から学習するため，本研究では，文字の構成要素の学習と，構成要素の典型的な組合せの学習の 2 段階で文字らしさの学習を試みる．

2. 文字の構成要素とその組み合わせの学習

2.1. 文字画像のスパース表現

ひとつの文字画像を複数の画像の線形結合で表現することを考える．文字画像はすべて正規化済みであるものとする．正規化では文字のサイズや濃淡がをあらかじめ定められた値になるように加工されている．本研究では，簡単のため図 1 のように，それぞれの文字がサイズ $h \times w$ 画素の正方形の 2 値化画像で与えられるものとする．

決められた順に並べた $p = hw$ 個の画素値を成分に持つベクトル $\boldsymbol{y} \in \mathbb{R}^p$ で 1 枚の画像を表す．n 枚の画像は $\boldsymbol{y}^{(1)}, \cdots, \boldsymbol{y}^{(n)}$ であり，これらを並べた行列を $\boldsymbol{Y} = [\boldsymbol{y}^{(1)}, \cdots, \boldsymbol{y}^{(n)}] \in \mathbb{R}^{p \times n}$ とする．同様に，文字を構成する典型的な線画の画像 (サイズ $h \times w$) も p 次元ベクトル $\boldsymbol{d} \in \mathbb{R}^p$ で表す．本論文では，これを文字画像の基底ベクトルと呼び，k 本の基底ベクトルを並べた行列 $\boldsymbol{D} = [\boldsymbol{d}^{(1)}, \cdots, \boldsymbol{d}^{(k)}] \in \mathbb{R}^{p \times k}$ を文字画像の辞書と呼ぶ．第 j 番目の文字画像 $y^{(j)}$ の線形合成は，

$$\boldsymbol{y}^{(j)} = \alpha_1^{(j)} \boldsymbol{d}^{(1)} + \cdots + \alpha_k^{(j)} \boldsymbol{d}^{(k)} = \boldsymbol{D}\boldsymbol{\alpha}^{(j)} \quad (1)$$

と表される．ここで，$\boldsymbol{\alpha}^{(j)} = [\alpha_1^{(j)}, \cdots, \alpha_k^{(j)}]^T \in \mathbb{R}^k$ は，j 番目の画像を辞書 $\boldsymbol{D}$ で表したときの線形結合係数を成分にもつベクトルである．n 枚についてまとめて表すと

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{A} \quad (2)$$

となる．ここで $\boldsymbol{A} = [\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(n)}] \in \mathbb{R}^{k \times n}$ である．

$\boldsymbol{d}^{(1)}, \cdots, \boldsymbol{d}^{(k)}$ が，文字画像を合成するための典型的な線画の画像であれば，どの文字画像もこれらのうち

†長崎大学大学院工学研究科

の何枚かで合成できるはずである．このとき，文字画像の合成に使われた基底ベクトルの結合係数のみが非ゼロで，それ以外はゼロとなる．つまり，どの $\boldsymbol{\alpha}^{(j)}$ もスパースなベクトルとなる．逆に，どの $\boldsymbol{\alpha}^{(j)}$ もスパースならば，基底ベクトルは多くの文字に共通の典型的な線画等の文字の構成要素を表す画像であるといえる．そのような辞書 $\boldsymbol{D}$ と結合係数 $\boldsymbol{A}$ による $\boldsymbol{Y}$ の表現を，文字画像のスパース表現と呼ぶ．

2.2. 文字の構成要素の抽出

与えられた文字画像の集合から，それらの文字を形づくる典型的な構成要素を抽出するには式 (2) がスパース表現となるような辞書 $\boldsymbol{D}$ を求めることである．スパース表現とは，データを数個の基底の線形結合で表現する手法で，基底の線形結合係数の非ゼロ成分の個数が少ない．文字画像の行列 $\boldsymbol{Y}$ の各列ベクトル $\boldsymbol{y}^{(j)}$ が共通の辞書 $\boldsymbol{D}$ でスパース表現されるとき，結合係数の行列 $\boldsymbol{A}$ はスパースである．与えられた文字画像集合の行列 $\boldsymbol{Y}$ に対して係数行列 $\boldsymbol{A}$ が最もスパースになるような辞書 $\boldsymbol{D}$ を作成する問題を定式化すると

$$\min_{\boldsymbol{D}\in\mathcal{D}} ||\boldsymbol{Y}-\boldsymbol{D}\boldsymbol{A}||_F \quad \text{subjec to } \forall j,\ \ ||\boldsymbol{\alpha}^{(j)}|| \le m \tag{3}$$

$$\mathcal{D} = \{\boldsymbol{D}\in\mathbb{R}^{p\times k} \mid \forall j,\cdots,k,\ \ ||\boldsymbol{d}^{(j)}||_2 \le 1\}$$

となる．ここでの $||\cdot||_F$ は，フロベニウスノルムである．

係数行列 $\boldsymbol{A}$ のどの列ベクトル $\boldsymbol{\alpha}^{(j)}$ の非ゼロ成分の個数も m 以下となる条件の下で，辞書 $\boldsymbol{D}$ による復元 $\boldsymbol{D}\boldsymbol{A}$ が $\boldsymbol{Y}$ になるべく近くなるように，$\boldsymbol{D}$ と $\boldsymbol{A}$ を求める [7]．ただし，辞書 $\boldsymbol{D}$ の列（基底ベクトル）の l^2 ノルムはどれも 1 以下とする．$||\boldsymbol{\alpha}^{(j)})||_0$ は $\boldsymbol{\alpha}^{(j)}$ の非ゼロ成分の個数を与える l^0 ノルムである．

l^0 ノルムを l^1 ノルムに緩和した最適化問題

$$\min_{\boldsymbol{D}\in\mathcal{D},\boldsymbol{\alpha}^{(j)}\in\mathbb{R}^k} \sum_{j=1}^{n}\left(\frac{1}{2}||\boldsymbol{y}^{(j)}-\boldsymbol{D}\boldsymbol{\alpha}^{(j)}||_2^2+\lambda||\boldsymbol{\alpha}^{(j)}||_1\right) \tag{4}$$

も式 (3) と同じ解を持つ．式 (4) は辞書 $\boldsymbol{D}$ と結合係数 $\boldsymbol{\alpha}^{(j)}$ による復元 $\boldsymbol{D}\boldsymbol{\alpha}^{(j)}$ と $\boldsymbol{y}^{(j)}$ を比べた 2 乗誤差と，$\boldsymbol{\alpha}^{(j)}$ の l^1 ノルム（絶対値ノルム）を同時に最小化する問題である．l^1 ノルムの最小化は，ベクトルをスパースにする効果がある．これは l^0 ノルム最小化と同様の効果である．

式 (4) について，$\boldsymbol{D}$ を固定して ${}^\forall\boldsymbol{\alpha}^{(j)}$ に関する最小化問題を解き，${}^\forall\boldsymbol{\alpha}^{(j)}$ を固定して $\boldsymbol{D}$ に関する最小化問題を交互に反復することで解 $\boldsymbol{D}$ と $\boldsymbol{A}$ が得られる．両最小化問題は共に凸最適化問題となっており，解が一意となる．式 (3) が一意の解を持つとき，式 (3) の解と式 (4) の解が一致するような λ が存在する．K-SVD アルゴリズム [7] は式 (3) または式 (4) の解法のひとつである．辞書 $\boldsymbol{D}$ を固定して係数行列 $\boldsymbol{A}$ を求めた後，$\boldsymbol{A}$ を固定せずに $\boldsymbol{D}$ の各列 $\boldsymbol{d}^{(i)}$ とそれに対応する係数 ($\boldsymbol{A}$ の第 i 行) を同時に更新することで，交互に固定する解法よりも早い収束を達成する．$\boldsymbol{d}^{(i)}$ と対応する係数の更新では，$\boldsymbol{d}^{(i)}$ を除外した辞書による復元の残差の合計が最小になるように基底 $\boldsymbol{d}^{(i)}$ と係数 ($\boldsymbol{A}$ の第 i 行) を特異値分解で求めている．

式 (4) はオンライン辞書学習のアルゴリズム (online dictionary learning) が実用的である [8]．係数 $\boldsymbol{A}$ と辞書 $\boldsymbol{D}$ を交互に更新する算法であるが，基底 $\boldsymbol{d}^{(i)}$ の更新に確率勾配降下法を応用することで $\boldsymbol{Y}$ の各列ベクトルを逐次処理し，早い収束と記憶領域の削減を達成している．新たな学習データを $\boldsymbol{Y}$ に追加して $\boldsymbol{D}$ と $\boldsymbol{A}$ を更新することも可能である．本研究では文献 [8] の著者が提供するパッケージ SPAMS を使用する．

2.3. 文字の構成要素の組合せの学習

式 (4) の解 $\boldsymbol{D}$ は，文字画像集合から文字を形づくる典型的な構成要素を表す k 枚の画像を基底ベクトルとして得たものになっていると考えられる．$\boldsymbol{Y}$ を文字画像から作成した場合，基底ベクトルは線画等の画像に相当するであろう．漢字画像の場合は 偏や旁またはそれらを構成する線画などであると予想される．式 (1) において，スパースなベクトル $\boldsymbol{\alpha}^{(j)}$ の非ゼロ成分は $\boldsymbol{y}^{(j)}$ の画像を構成する線画（基底ベクトル）等の組合せを表している．その組合せが，与えられた文字の集合に特有の典型的な「文字らしさ」を表す図形といえるであろう．

ここでは更に，基底ベクトルの典型的な組合せを $\boldsymbol{\alpha}^{(1)},\cdots,\boldsymbol{\alpha}^{(n)}$ から求める．そのような組合せを非ゼロ成分で表すベクトルを $\boldsymbol{c}\in\mathbb{R}^{k\times l}$ とする．さらに，典型的な組合せがいくつか存在すると仮定する．$\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}$ を並べた行列を $\boldsymbol{C}=[\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}]\in\mathbb{R}^{k\times l}$ とする．この $\boldsymbol{C}$ は線画の組合せ行列といえるであろう．典型的な組合せを十分な数だけ用意できれば，線画の組合せを表す $\boldsymbol{\alpha}^{(j)}$ は $\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}$ を基底に使って合成できるであろう．

$$\boldsymbol{\alpha}^{(j)} = \beta_1^{(j)}\boldsymbol{c}^{(1)}+\cdots+\beta_l^{(j)}\boldsymbol{c}^{(l)} = \boldsymbol{C}\boldsymbol{\beta}^{(j)}, \tag{5}$$

$$\boldsymbol{\beta}^{(j)} = [\beta_1^{(j)},\cdots,\beta_l^{(j)}]^T\in\mathbb{R}^l \tag{6}$$

$j=1,\cdots,n$ の n 枚の画像についてまとめて書くと

$$\boldsymbol{A}=\boldsymbol{C}\boldsymbol{B} \tag{7}$$

$$\boldsymbol{B}=[\boldsymbol{\beta}^{(1)},\cdots,\boldsymbol{\beta}^{(n)}]\in\mathbb{R}^{l\times n} \tag{8}$$

となる．

基底ベクトル $\boldsymbol{d}^{(1)},\cdots,\boldsymbol{d}^{(k)}$ の組合せを表す $\boldsymbol{\alpha}^{(j)}$ を合成するために適した典型的な構成要素の組合せを $\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}$ が表しているならば，どの組合せ $\boldsymbol{\alpha}^{(j)}$ も，これら $\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}$ のうちの何個かで合成できることになる．このとき，組合せ $\boldsymbol{\alpha}^{(j)}$ の合成に使われたいくつかの $\boldsymbol{c}^{(i)}$ の結合係数 $\beta_i^{(j)}$ のみが非ゼロで，それ以外はゼロとなる．つまり，どの $\boldsymbol{\beta}^{(j)}$ もスパースなベクトルとなる．逆に，どの $\boldsymbol{\beta}^{(j)}$ もスパースならば，$\boldsymbol{c}^{(1)},\cdots,\boldsymbol{c}^{(l)}$ は，構成要素の組合せの多くに共通の典型的な組合せ

を表すベクトルであるといえる．そのような $\boldsymbol{C}$ と $\boldsymbol{B}$ による $\boldsymbol{A}$ の表現は，スパース表現に他ならない．

$\boldsymbol{Y}$ をスパース表現する $\boldsymbol{D}$ と $\boldsymbol{A}$ を求めたときと同様に，$\boldsymbol{A}$ をスパース表現する $\boldsymbol{C}$ と $\boldsymbol{B}$ はスパースコーディングによって得ることができる．式 (4) と同様に定式化すると，

$$\min_{\boldsymbol{C}\in\mathcal{C},\boldsymbol{\beta}^{(j)}\in\mathbb{R}^l} \sum_{j=1}^{n} \left(\frac{1}{2}||\boldsymbol{\alpha}^{(j)} - \boldsymbol{C}\boldsymbol{\beta}^{(j)}||_2^2 + \lambda||\boldsymbol{\beta}^{(j)}||_1\right) \quad (9)$$

$$\mathcal{C} = \{\boldsymbol{C} \in \mathbb{R}^{k\times l} |\ \forall j, \cdots, l,\ \ ||\boldsymbol{c}^{(j)}||_2 \leq 1\} \quad (10)$$

となる．

さらに，$\boldsymbol{Y}$ のスパースコーディングによって得られた $\boldsymbol{D}$ と，$\boldsymbol{A}$ のスパースコーディングによって得られた $\boldsymbol{C}$ の積を

$$\boldsymbol{E} = \boldsymbol{D}\boldsymbol{C} \quad (11)$$

とすると，与えられた画像 $\boldsymbol{Y}$ は式 (2)，(7)，(11) より

$$\boldsymbol{Y} = \boldsymbol{E}\boldsymbol{B} \quad (12)$$

と表される．すなわち，式 (4) と式 (9) の 2 段階のスパースコーディングによって，線画の基底の典型的な組合せで作られた文字らしさの基底ベクトルからなる行列 $\boldsymbol{E}$ が得られる．$\boldsymbol{E}$ の列ベクトルは $\boldsymbol{Y}$ の各文字画像を式 (12) のようにスパース表現する．

3. 漢字らしさを表す線画の組合せ

3.1. 漢字画像とパラメータの設定について

学習の例として，文字らしさを表す線画の組合せが特徴的であると考えられる漢字画像を学習用の画像に使用する．漢字画像を学習用の画像からなる行列 $\boldsymbol{Y}$ として式 (4) と式 (9) の 2 段階でスパースコーディングを行うと，漢字らしさである偏や旁またはその線画などが辞書 $\boldsymbol{D}$ や基底の組合せ $\boldsymbol{E}$ として得られると予想される．

漢字画像集合として，全書体共通第一水準漢字テキスト検索 [9] に掲載されている漢字 1945 文字を使用する．1 文字（1 枚の画像）を 32×32 ($p = h \times k = 1024$) 個の画素値を成分にもつベクトルで表す．線画の画素値を 1(白)，背景を 0(黒) とする．このベクトル 1945 本を $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n (n = 1945)$ とする．

第 1 段階の基底学習では，1945 本のベクトルを k 本の基底で表す．また，第 2 段階の学習では，既定の組合せを l 本に求める．k と l によって得られる辞書 $\boldsymbol{D}$ や $\boldsymbol{E}$ の違いを観察するため，表 1 に示すパラメータの組によって学習を実行した．λ を 0.15 に設定し，SPAMS[8] による式 (4) の最小化は，Intel Xeon X5680@3.33GHz で 4 スレッド使用すると約 10 分で収束した．

3.2. 漢字を構成する共通の線画

$k = 400$(param2) として，第 1 段階目の基底学習で得られた辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times400}$ を図 3 に示す．第 1 段階の基底学習は文字の構成要素となる線画を $\boldsymbol{D}$ として得

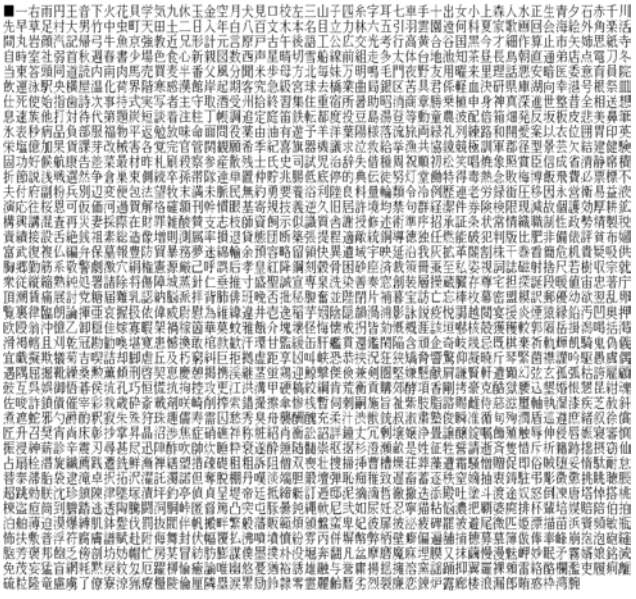

図 2: 漢字画像の一覧．全 1945 文字を並べて表示している．各文字は 32×32 画素で構成されている．

ることが狙いであるが，図 4 のような偏や旁などが得られた．多くの文字で重なりやすい線画が辞書の基底として選ばれていると考えられる．これに加えて，図 3 に示す基底には，図 5 のような線画が潰れたような文字が存在する．これは複数の線画が重なる交点で画素値が強め合う，または弱め合うことによる復元の誤差を低減するために使われる基底であると考えられる．

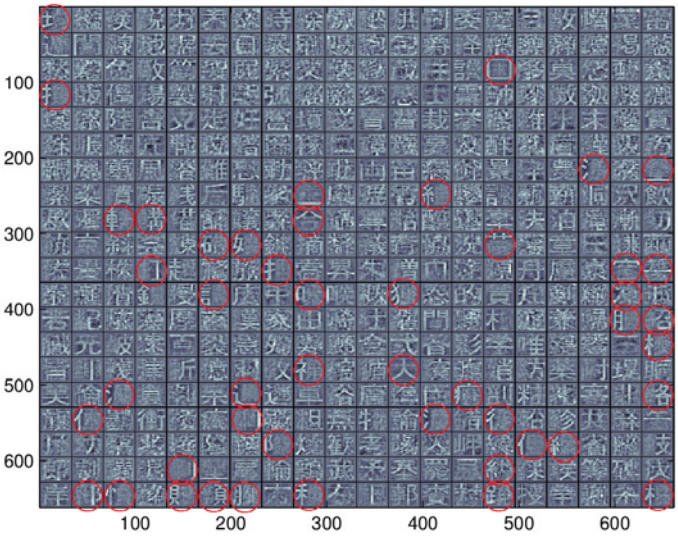

図 3: 辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times400}$(赤丸印は漢字の典型的な線画を表すと考えられる基底)．各基底を 32×32 の画素値の画像にして並べて図示している．

サイズ $k = 100$(param1) に設定したときに得られた辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times100}$ を図 6 に示す．図 3 の $k = 400$ よりも基底の数が少ないため，図 5 のような基底が多いように見える．しかし，漢字らしさである偏や旁なども得られている．

$k = 750$(param3) に設定したときに得られた辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times750}$ を図 7 に示す．図 5 のような線画を表

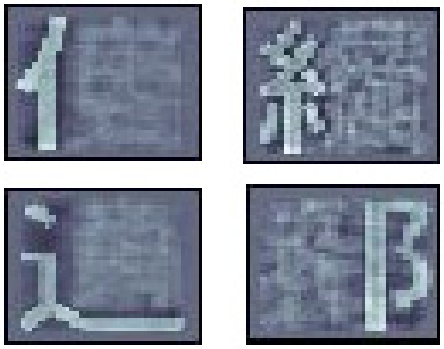

図 4: 辞書 $\boldsymbol{D}$ に含まれる偏や旁を表す基底の例.

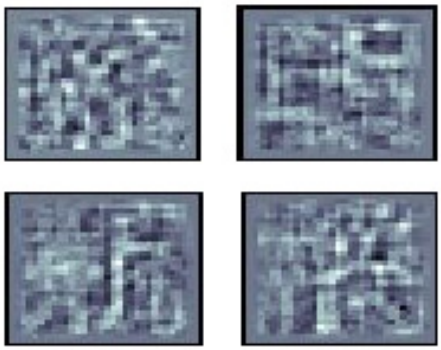

図 5: 辞書 $\boldsymbol{D}$ に含まれる基底のうち，線画を表していない基底の例.

表 1: 2 段階の学習における辞書のサイズ.

パラメータ名	辞書 D のサイズ k	辞書 E のサイズ l
param1	100	40
param2	400	40
param3	750	40

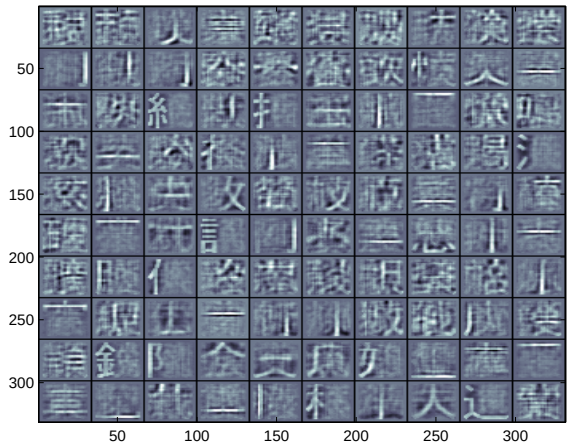

図 6: 辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times100}$．各基底を 32 × 32 の画素値の画像にして並べて図示している.

さない基底があまり見られない．そのかわり，漢字画像をそのまま基底として学習したと考えられる基底が多く見られる.

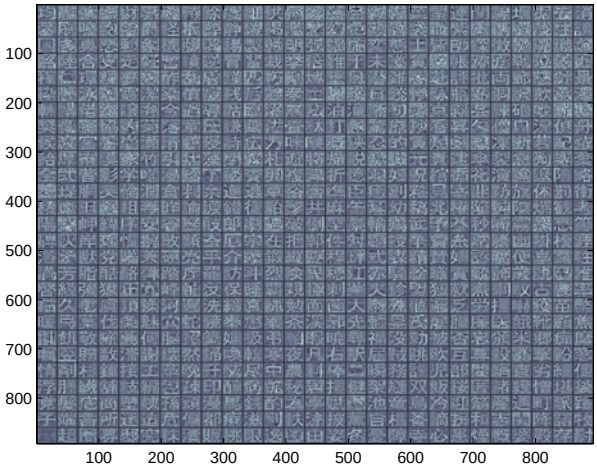

図 7: 辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times750}$．各基底を 32 × 32 の画素値の画像にして図示している.

3.3. 線画の組合せによる漢字らしさの表現

param2 では，第 2 段階の学習における基底の数を $l = 40$ としている．この理由は，$k = 400$ とした第 1 段階の学習で得られる辞書 $\boldsymbol{D} \in \mathbb{R}^{1024\times400}$ の中で，漢字を構成する典型的な線画を表すと考えられる基底 (図 3 の赤丸印) が約 40 本あり，これらの基底が表す線画の組合せによって漢字らしさが表現されるであろうと考えたためである.

表 1 の param1，param2，param3 の設定によって得られた $\boldsymbol{E} \in \mathbb{R}^{1024\times40}$ をそれぞれ図 8，図 9，図 10 に示す．図 3 や図 6，図 7 の辞書 $\boldsymbol{D}$ と比べると，多くの基底に偏や旁などが表れているように見える．図 5 のような基底は見られないが，各基底が持つ特徴的な線画が存在しない部分に同様の模様が見られる.

4. おわりに

本研究では，スパースコーディングによって文字画像から文字らしい共通の特徴を抽出する辞書の作成と評価を行った．今後の課題として，低解像度による文字のつぶれ，動きによるブレやボケ，他の物体に部分的に隠されていることによる欠損，明暗や影による光学的な変化，観測ノイズなどによる汚損に対する文字検出の性能評価が挙げられる．また，文字らしさは字種に依存しないため，字種に関係なく文字を検出することができ，かつ，文字以外の図形や物体を文字として誤検出しないことが望ましい．すなわち，文字の汎化性が求められる．汎化性をもつ辞書への改良も視野に入れて研究に取り組みたい.

5. 謝辞

本研究は JSPS 科研費 25330200 の助成を受けたものである.

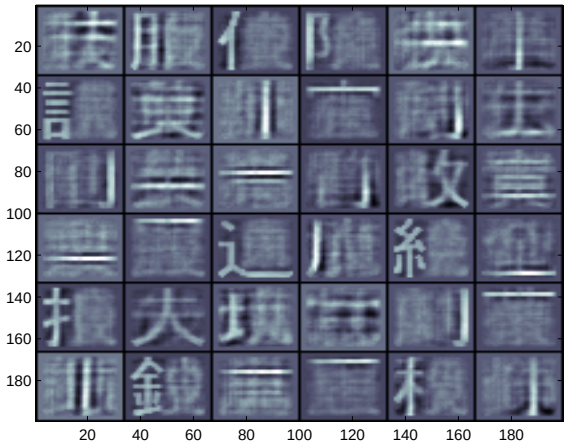

図 8: 1 文字を 32 × 32 の画素値の画像にして得られた辞書 $\boldsymbol{E} \in \mathbb{R}^{1024\times40}$.

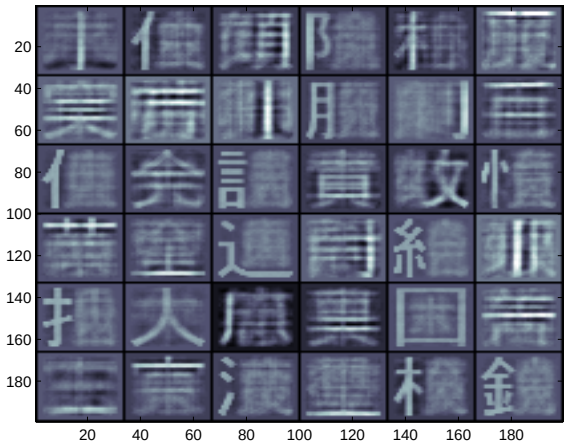

図 9: 1 文字を 32 × 32 の画素値の画像にして得られた辞書 $\boldsymbol{E} \in \mathbb{R}^{1024\times40}$.

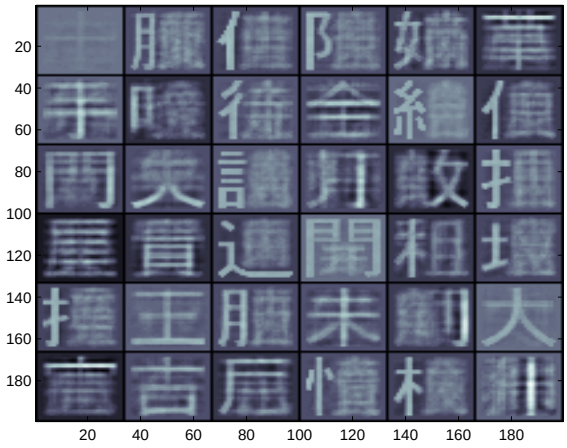

図 10: 1 文字を 32 × 32 の画素値の画像にして得られた辞書 $\boldsymbol{E} \in \mathbb{R}^{1024\times40}$.

参考文献

[1] 坂井邦夫, 入江文平, 水谷博之, “文字・文書の認識と理解, ”, IPSJ Magazine Vol.44 No.11, Nov. 2003.

[2] 黄瀬浩一, 大町真一郎, 内田誠一, 岩村雅一, “ディジタルカメラによる文字・文書の認識・理解, ” 電子情報通信学会誌 Vol.89, No.9, 2006.

[3] 草地良規, 鈴木章, 伊藤直己, 荒川賢一, 安野貴之 “景観画像中の文字候補群による画像インデクシング及び検索技術, ” 電子情報通信学会論文誌 D Vol.J90-D No.9 pp.2562-2572, 2007.

[4] 内田誠一, “文字・文書の認識・理解に関するグランドチャレンジ私案, ”, PRMU, パターン認識・メディア理解 Vol.108, No.432, pp.49-54, 2009.

[5] 長井隆行, 影広達彦, 金子正秀, 榑松明, “情景画像中の文字及び看板領域の抽出, ”CAS2000-125, DSP2000-183, CS2000-145, 2001.

[6] 國重康弘, 馮尭楷, 内田誠一, “環境コンテキスト利用による情景画像中文字検出, ”PRMU2009, パターン認識・メディア理解 Vol.109, No.418, pp.81-86, 2010.

[7] M.Aharon, M.Elad, and A.Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, ” IEEE Trans. Signal Processing, Vol 54, No.11, pp.4311-4322, 2006.

[8] J.Mairal, F.Bach, J.Ponce, and G.Sapiro “Online dictionary learning for sparse coding, ” Proc ICML 2009, pp.689-696, 2009.

[9] 全書体共通第一水準漢字テキスト検索, http://zinsta.jp/font/a-hyou-txt.html

[10] Chitrakala Gopalan, D.Manjula, “Contourlet based approach for text identification and extraction from heterogeneous textual images, ” 2008.

[11] Angshul Majumdar, “Bangla Basic Character Recognition Using Digital Curvelet Transform, ” 2007 JPRR, 2007.

[12] 金谷健一, “これなら分かる応用数学教室ー最小に情報からウェーブレットまで, ” 共立出版, 2003.

[13] 石川孝明, 渡辺裕, “方向性フィルタバンクによる動画像符号化方式に関する基礎検討, ”IE, 画像工学 Vol 106, No.424, pp.45-49, 2006.