

II-2-7

最適係数を用いた線形色変換による

ノイズの低減された色主成分の抽出法

白井 啓一郎・杉本 拓矢・藤森 和博・宮尾 秀俊・丸山 稔

概 要

本稿では、カラー画像中の主たる色成分、かつ、模様や雑音の少ない平坦な成分の抽出法を提案する。木簡からの文字抽出へ応用する場合は、木目の模様を低減しつつ、文字と木の主な色を抽出する方法となる。本手法は単に RGB カラーの R、G、B 三色を線形結合するいわゆる一般的な色変換法であるが、色変換に用いる変換係数として、上述した特徴を持つ画像を出力する最適変換係数を求めて用いる。本手法は主成分分析 (PCA) と全変動最小化 (Total Variation Minimization) に関連した方法であり、主成分分析で用いられる白色化されたデータ行列を用いて球面多様体上において全変動コストを最小化する問題として表される。求める R、G、B 値が球面上に存在するという制約は非凸であるが、交互方向乗数法 (ADMM) を用いて求解できる。その応用例として、木簡画像からの文字抽出の実験を示す。木目の模様を低減しつつ、墨で書かれたかすれた文字を抽出できる。この文字抽出の目的においては、ユーザーによる補助情報を用いた判別分析法と同等の精度で文字抽出が可能である。

1、はじめに

まずはじめに、本手法は元々は木簡からの文字抽出 [1, 2] を目的として開発を行ったが (図 89 参照)、シミ、サビ、塗装剥がれなどの検出にも利用できると考えている (図 95 参照)。

図 89(a) と (d) に示したような木簡画像の場合、木簡画像は、薄くかすれた文字、及び、文字よりも目につく木目模様で構成されている。図中 (b) と (e) は、画像認識に用いる色成分の一例として明るさ (輝度値) の明暗を表したものである。R、G、B の色を用いて $\frac{1}{3}(R + G + B)$ と計算して求めた (なお、簡略的な計算である)。(b) では、かすれた文字の輝度変化が木目の輝度変化よりも小さく、文字と背景を区別しにくい。一方、(e) では、文字のほ

うが木目よりも輝度変化が大きく、文字を知覚しやすい。このことから、輝度成分を用いた場合は、画像によって認識精度に差が生じると予想される。

同図 (b) のような画像を見やすくするには、通常は平滑化法を用いて雑音(木目によって生じた凹凸)を平坦化するのだが、文字の輝度変換のほうが木目よりも小さいため、文字が平坦化されてより薄れる問題が生じる。よって、平滑化法は有効ではない。別の方法として R、G、B の色同士の差分、すなわち、色差(再度)成分を用いることで、模様を除

去できる場合がある。同図 (c) は色差成分を示しており、(b) に比べて木目が低減され、文字が知覚しやすいことが分かる。ただし、この方法も画像に依存する。(f) は同様の計算を別の画像に施した結果であるが、この場合には悪化している。

色差を求めることで文字が浮き出る理由は次のようになる。まず、画像の色、特に模様には、R、G、B の間に相関があり、同様に変化することが知られている。次に、文字は黒色の墨で書かれており、赤外線吸収する傾向があるため、文字部分からは赤色がカメラへと反射せず、画像の文字領域での赤色成分の観測値は小さくなる。一方、文字以外の領域ではそのような現象は生じない。これらより、R、G、B を比べた際、文字領域のみに変化が生じるため、 $R-G$ 、 $G-B$ 、 $R-B$ などの引き算をし、求まった値を比べれば、値に変化のある領域を文字領域として検出できる。また、可視光線のスペクトルは、赤から緑、緑から青へと並んでおり、色の変化も連続的に変化するため(隣同士の色の値の変化のしかたは似ている)、赤色と青色の差を比べることで文字と背景に差が生じやすい。

上述のように、輝度と色差の成分を抽出すれば、画像に依存はするものの、木目が低減しつつ文字が浮き上がった画像を得ることは可能である。更に、このような成分を得るより一般的な線形変換を考えると、上述の輝度と色差の計算は、それぞれ、 $\frac{1}{3}R + \frac{1}{3}G + \frac{1}{3}B$ または $R - B = (+1)R + (0)G + (-1)B$ などと計算されるが、これらは $\alpha_1 R + \alpha_2 G + \alpha_3 B$ の計算

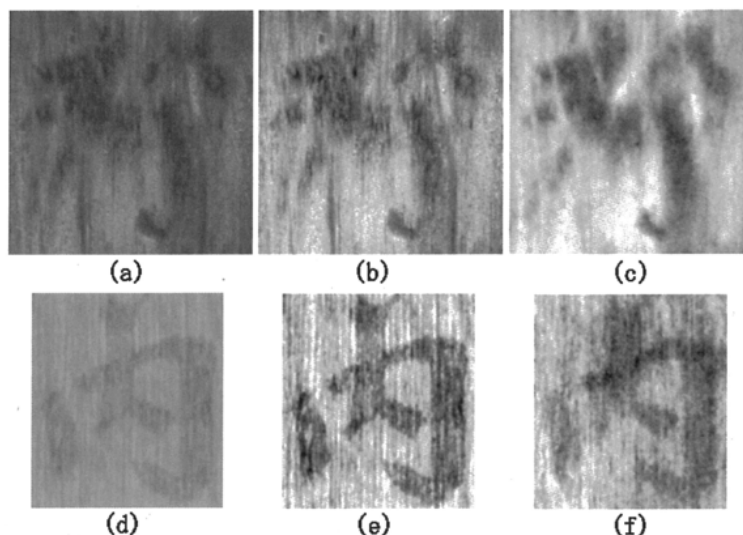


図89 異なる合成係数の組を用いた色の線形結合(線形色変換)の結果。左側から、(a)(d)元の画像、(b)(e)輝度成分($(R+G+B)/3$ として計算)、(c)(f)色差(色彩)成分(一例として $R-B$ として計算)を表す。変換処理後のグレースケール画像の輝度値は $[0,1]$ の範囲となるように線形濃度変換を施し正規化した。係数の組によって、画像中の雑音と木目は低減または強調される。望ましい処理結果を与えているのは(c)と(e)である。

の一部とみなせ、三つの係数 $\{\alpha_1, \alpha_2, \alpha_3\}$ の選び方によって文字の見やすさが変わるのだと見なせる。この計算結果の数値の変化幅が大きい方が、画像のコントラストが高くなるため、そのような係数を得られることが望ましい。この変化幅を最大化する係数 $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ を求める計算は、画像中で用いられている色の分布の広がり幅が最大となる軸方向を解析する計算と関連しており、色データを主成分分析 (Principal Component Analysis: PCA) [3] することで求まる。

輝度と色差 ($R-B$) もしくは前述の PCA によって得られたい成分のうち、いずれの成分を用いるかは、見やすさを考慮して選択する。見やすさの尺度のうち、コントラストについては上述したが、その他として白黒の変化の緩急がある。鮮明な画像では、文字と背景の境界部分において急峻に輝度値が黒から白へと変化し、それ以外の場所ではなるべく同じ色を保つのが望ましい (白黒のメリハリがあり、知覚しやすい) と考えられる。この特徴を表す数値尺度に全変動値 (Total Variation: TV) [4] があり、この値が小さいほうが、同じ色が多用されているように見える。例えば、自然風景の画像よりはアニメーションの画像のほうが TV 値は小さい。

以上に述べた PCA と TV 最小化に基づき、本論文ではこれらを組み合わせ、文字部分と背景部分の色のコントラストが大きく、かつ、木目模様が低減された画像を生成できる色変換方法について述べる。カラー画像の R 、 G 、 B 色成分に対して乗算する最適な色変換係数 (合成係数) $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ を求めることを目的とする。提案法で扱う数値最適化の問題は、球面多様体上の Least Absolute Deviations (LAD) 最小化問題 [5] に帰着する。非凸な問題であるため、凸な問題に比べて最適解を得られにくいだが、本稿では相互方向乗数法 (ADMM: Alternating Direction Methods of Multiplier) [6-9] を用いて大域的な最適解が求まることを示す。

提案法はユーザー補助情報を必要としない方法であり、文字部分と背景部分の一部に予めラベルを付けるような手間は生じない。他方、画像中から特定色の領域を抽出する方法、例えば、線形判別分析 (LDA: Linear Discriminant Analysis) や前景・背景の分離法 (ImageMatting) は、良い抽出精度は期待できるが、ユーザー補助情報を必要とする。提案法は LDA と同等の精度を有することを示す。

2. 諸定義及び関連手法

本章では画像データの表現のしかたについて述べた後、PCA を用いた基本的な方法について述べる。

2.1 R 、 G 、 B 成分の線形結合による色成分の抽出法及び画像のコントラストの正規化

入力画像はRGBカラー画像（またはより色数の多い値をベクトルとして各画素にもつ画像）であり、各画素 i での画素値を (r_i, g_i, b_i) と表す。簡素化のため、画像中の N 個の全ての画素値を並べて列ベクトルとして表す。すなわち、赤色成分を $\mathbf{r} = [r_1, \dots, r_N]^T \in R^N$ と表し、 $\mathbf{g}$ と $\mathbf{b}$ も同様に表す。これらのベクトルと結合係数（色変換係数） $\{\alpha_1, \alpha_2, \alpha_3\} \in R^1$ を用いて、輝度や色差を得るための線形色変換を次のように表す。

$$\mathbf{x} := \alpha_1 \mathbf{r} + \alpha_2 \mathbf{g} + \alpha_3 \mathbf{b} = \mathbf{Y} \boldsymbol{\alpha} \quad (1)$$

ここで行列形式で式を表すため、データ行列 $\mathbf{Y} := [\mathbf{r}, \mathbf{g}, \mathbf{b}] \in R^{N \times 3}$ 、及び、係数ベクトル $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]^T$ を導入する。上式において、係数ベクトルが $\boldsymbol{\alpha} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]^T$ と与えられるとき、結果として得られる画像 $\mathbf{x}$ （実際には画素値を並べたベクトル）は輝度成分を表す。一方、係数ベクトルが $\boldsymbol{\alpha} = [1, 0, -1]^T$ と与えられるときは結果画像は色差を表し、特にこの係数の場合は、YCoCg変換[10]のオレンジ色の色差を表すCo成分が得られる。

成分画像 $\mathbf{x}$ の輝度値の範囲、及び、コントラストは係数の値によって変化する。例えば、輝度値は全て加算で表されるが、色差は減算で表され、一般的に色差のほうが輝度よりも値の範囲が狭くなる、すなわち、コントラストが低くなるゆえに、画像認識などで用いるならば、値の範囲を一定の範囲へと正規化する必要が生じる。本稿では正規化を、まず平均値 $\mu(\mathbf{x}) := \frac{1}{N} \sum_i x_i$ を $\mathbf{x}$ から差し引き、次に標準偏差 $\sigma_i(\mathbf{x}) := \frac{1}{N} \sum_i (x_i - \mu(\mathbf{x}))^2$ で除算することで行う。正規化後の画像 $\hat{\mathbf{x}}$ は次のようにも求まる。

$$\begin{aligned} \hat{\mathbf{x}} &:= \frac{1}{\sigma(\mathbf{x} - \mu(\mathbf{x}))} (\mathbf{x} - \mu(\mathbf{x})) \\ &= \frac{1}{\sqrt{\frac{1}{N} \|\mathbf{M}\mathbf{x}\|_2^2}} \mathbf{M}\mathbf{x} \end{aligned} \quad (2)$$

ここで行列形式で表すため、中心化行列 $\mathbf{M} := \mathbf{Id} - \frac{1}{N} \mathbf{1}\mathbf{1}^T \in R^{N \times N}$ 、単位行列 $\mathbf{Id}$ 、すべての要素が1である列ベクトル $\mathbf{1} \in \{1\}^N$ を導入する。 $\|\mathbf{x}\|_p := (\sum_i |x_i|^p)^{1/p}$ は ℓ_p ノルムを表す。 $p=2$ のときにユークリッド距離を表し、 $p=1$ のときに絶対値距離を表す。

2.2 主成分分析を用いた色分布の主軸の解析

前述の合成係数のベクトルは、一般的に色分布の主たる広がりの方角を表すベクトルと関係する。すなわち、式(1)は、色の座標 (r_i, g_i, b_i) （R、G、B三次元空間中の点群）と軸を表すベクトル $[\alpha_1, \alpha_2, \alpha_3]^T$ の内積とみなすことができ、この計算の結果として得られる値は軸方向への変化量を表す。

データ点の分布の主たる軸を解析するには、主成分分析 (PCA) [3] が用いられる。まず、データ行列 $\mathbf{Y}$ の共分散行列 $\mathbf{C} \in R^{3 \times 3}$ を求めた後、その固有値分解を行う。

$$\mathbf{C} := \frac{1}{N} \bar{\mathbf{Y}}^T \bar{\mathbf{Y}} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \quad (3)$$

ここで $\bar{\mathbf{Y}} := \mathbf{M}\mathbf{Y}$ は平均値を差し引いたデータ行列を表し、各列において、各要素から各列の平均値を引いた行列となる。固有値分解を行うことで、固有ベクトルによって構成される直交行列 $\mathbf{U} := [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] \in R^{3 \times 3}$ 、及び、固有値を対角に並べた対角行列 $\mathbf{\Lambda} := \text{diag}(\lambda_1, \lambda_2, \lambda_3) \in R^{3 \times 3}$ へと共分散行列は分解される。ここでは固有値を降順に定義する $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ 。 i 番目に大きな固有値 λ_i に対応する固有ベクトル $\mathbf{u}_i$ (単位ベクトルである) は、色の分布の広がり i 番目に大きな軸方向を表す。また、固有値 λ_i はその軸方向に沿った分布の広がり幅、すなわち、分散を表す。RGB の色と固有ベクトルの内積は、 $\mathbf{x}_i := \bar{\mathbf{Y}}\mathbf{u}_i$ と計算でき、この主成分軸にそった変化量を表す。

PCA に基づく方法は、レイリー商の最大化として知られる最適化問題 [11] として書き表せ、第一主成分ベクトル (最大固有値に対応した固有ベクトル) の抽出は次の式で表される。

$$\max_{\alpha \in \{\alpha: \|\alpha\|_2=1\}} \left\{ \frac{1}{N} \|\bar{\mathbf{Y}}\alpha\|_2^2 = \alpha^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \alpha \right\} \quad (4)$$

ここで解を単位ベクトルとするために制約 $\|\alpha\|_2 = 1$ を加えている。 $\alpha^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \alpha$ を最大化する解は $\alpha = \mathbf{u}_1$ のときであり、最大値は λ_1 である。 $\alpha^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \alpha$ ただし $\|\alpha\|_2 = 1$ の計算は、原点を中心とする楕円において、原点から楕円表面までの距離を求める計算と関連する。すなわち、直交する軸を $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ としてもち、原点からの径の長さが $\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \frac{1}{\sqrt{\lambda_3}}$ である楕円において、 α の向への原点と楕円表面の長さ r の $\frac{1}{r^2}$ を求める計算であるため、長軸方向 $\mathbf{u}_1$ に α が向くときに長さは $r = \frac{1}{\sqrt{\lambda_1}}$ となり、代入すると最大値 $\frac{1}{r^2} = \lambda_1$ が求まる。更に、max の代わりに min を用いた場合は、解は $\alpha = \mathbf{u}_3$ となる。このように PCA は式 (4) と関連する。

3. 提案方法

式 (1) と (2) に加えて、処理画像 $\hat{\mathbf{x}}$ における全変動(TV)最小化するような α を求めたい。そのため、提案法を次のように定式化する。

$$\min_{\alpha \in \{\alpha: \|\alpha\|_2=1\}} \left\{ \|\hat{\mathbf{x}}\|_{\text{TV}} := \left\| \frac{1}{\sqrt{\frac{1}{N} \|\bar{\mathbf{Y}}\alpha\|_2^2}} \bar{\mathbf{Y}}\alpha \right\|_{\text{TV}} \right\} \quad (5)$$

ここで $\|\cdot\|_{TV}$ は TV ノルムを表し、後述する。次に、 α は、 $\frac{\bar{Y}s\alpha}{\sqrt{\|\bar{Y}s\alpha\|_2^2}} = \frac{\bar{Y}\alpha}{\sqrt{\|\bar{Y}\alpha\|_2^2}}$ の計算におい

てはスケール不変であり、スケーリング要素 s は通分されキャンセルされる。よって、

制約として α が $\frac{1}{\sqrt{N}\|\bar{Y}\alpha\|_2} = \text{const.}$ で一定になるような多様体に含まれる、すなわち、三次

元球面の表面に存在する制約 $\mathcal{A} \in \alpha$ を追加する。これを用いることで、上記の方程式 (5) は次のように書き表せる。

$$\min_{\alpha \in \mathcal{A}} \|\bar{Y}\alpha\|_{TV}, \mathcal{A} := \{\alpha: \|\bar{Y}\alpha\|_2 = 1, \mathbf{1}^T \alpha \geq 0\} \quad (6)$$

ここでは $\mathbf{1}^T \alpha \geq 0$ を符号の反転した解を除外するために追加する。この定式化を単独で用いる場合は、データ行列の白色化を行うことで更に簡略化ができるため、次節 3.1 において述べ、式 (11) に示す。

TV ノルムについては、記載の簡略化のため、本稿では等方性の TV を使い、微分フィルタ行列と ℓ_1 ノルムを用いて以下のように定義する。TV ノルムは、微分フィルタリング後の画素値の絶対値を計算し、全ての画素値を足し合わせる計算を意味する。

$$\|\hat{\mathbf{x}}\|_{TV} := \|\mathbf{D}\hat{\mathbf{x}}\|_1 = \|\mathbf{D}_h\hat{\mathbf{x}}\|_1 + \|\mathbf{D}_v\hat{\mathbf{x}}\|_1 \quad (7)$$

ここで $\mathbf{D} := [\mathbf{D}_h^T, \mathbf{D}_v^T]^T \in \mathbb{R}^{2N \times 3}$ はフィルタ行列の組であり、水平方向と垂直方向への微分フィルタリングを行う。微分フィルタは、長さ 2 の前方差分フィルタ（もしくは後方差分フィルタ）であり、フィルタ係数としてよく知られた $[-1, 1]$ をもつ。

参考用に、後方の頁に示す図 90 は、後述の式 (11) から求まる解と網羅的な探索によって得られたコストマップを表す。各コストマップには一つのコストが最小となる点があり、その点に向けて滑らかにコストが変化しているのが分かる。なお、木簡画像、及び、それ以外のカラフルな画像であってもこのような傾向をもつ。

3.1 データ行列の白色化による問題の簡略化

上述の問題は、データ行列を白色化することで、より簡単な式として書き直せる。式 (5) において、分母の共分散行列は、式 (3) と (4) の記述を用いて、次のように書ける。

$$\frac{1}{N} \|\bar{Y}\alpha\|_2^2 = (\alpha^T \mathbf{U} \Lambda^{\frac{1}{2}}) (\Lambda^{\frac{1}{2}} \mathbf{U}^T \alpha) = \hat{\alpha}^T \hat{\alpha} \quad (8)$$

ここで新たに $\hat{\alpha} \in \mathbb{R}^3$ を導入し、次の関係 $\hat{\alpha} = \Lambda^{\frac{1}{2}} \mathbf{U}^T \alpha$ と $\alpha = \mathbf{U} \Lambda^{-\frac{1}{2}} \hat{\alpha}$ をもつ^{*1}。これを式 (5) へ代入し、 $\hat{\mathbf{x}}$ を以下のように書き換える。

^{*1} $\bar{Y}$ は稀にランク落ちしており、その場合、 Λ は対角値に 0 を含む。よって、 Λ^{-1} の計算をする際に生じるゼロ割を防ぐため、対角値の最小値を定め、これを下回る値を最小値に修正する。

$$\hat{\mathbf{x}} = \frac{1}{\|\hat{\alpha}\|_2} \bar{\mathbf{Y}} \mathbf{U} \Lambda^{-\frac{1}{2}} \hat{\alpha} = \frac{1}{\|\hat{\alpha}\|_2} \hat{\mathbf{Y}} \hat{\alpha} \quad (9)$$

ここで $\hat{\mathbf{Y}} := \bar{\mathbf{Y}} \mathbf{U} \Lambda^{-\frac{1}{2}} \in R^{N \times 3}$ は白色化されたデータ行列を表し、このデータ行列から得られる共分散行列は単位行列となる。つまり、このデータ行列に含まれる要素の分布は球状である。

これらの関係を用い、主たる方程式 (6) を次のように書き換える。

$$\min_{\hat{\alpha} \in \mathcal{A}} \|\mathbf{D} \hat{\mathbf{Y}} \hat{\alpha}\|_1, \mathcal{A} := \{\hat{\alpha}: \|\hat{\alpha}\|_2 = 1, \mathbf{1}^\top \mathbf{U} \Lambda^{-\frac{1}{2}} \hat{\alpha} \geq 0\} \quad (10)$$

更に、予め微分フィルタリングを施したデータ $\hat{\mathbf{Y}}' := [(\mathbf{D}_h \hat{\mathbf{Y}})^\top, (\mathbf{D}_v \hat{\mathbf{Y}})^\top]^\top \in R^{2N \times 3}$ を用い、表記 $\mathbf{1}' := \Lambda^{-\frac{1}{2}} \mathbf{U}^\top \mathbf{1} \in R^3$ と書き換えることで、最終的に次式を得る。

$$\min_{\hat{\alpha} \in \mathcal{A}} \|\hat{\mathbf{Y}}' \hat{\alpha}\|_1, \mathcal{A} := \{\hat{\alpha}: \|\hat{\alpha}\|_2 = 1, \mathbf{1}'^\top \hat{\alpha} \geq 0\} \quad (11)$$

3.2 ADMM を用いた式 (11) の解法

式 (11) に示した問題は Least Absolute Derivations (LDA) と呼ばれ、これに球面制約が加わる [5, 12]。もしくは、 ℓ_1 ノルムに基づく主成分分析 (ℓ_1 -PCA) と呼ばれる [13, 14]。本稿では、この問題を前者の最小化問題として扱い、その求解を相互方向乗数法 (ADMM: Alternating Direction Method of Multipliers) [6-9] を用いて行う。同様の解法は [12, 15] で述べられている。

方程式 (11) に含まれる制約をより扱いやすくするため、指示関数を用いて正則化項として表す。すなわち、 $\min_{\hat{\alpha} \in \mathcal{A}} \|\hat{\mathbf{Y}}' \hat{\alpha}\|_1 + \iota_{\mathcal{A}}(\hat{\alpha})$ 。ここで制約 $\mathbf{1}'^\top \hat{\alpha} \geq 0$ については一旦取り除き、最終的に加える。指示関数 $\iota_{\mathcal{C}}(\cdot)$ は次のように定義される。

$$\iota_{\mathcal{C}}(\hat{\alpha}) = \begin{cases} 0 & \hat{\alpha} \in \mathcal{C} \\ \infty & \text{otherwise} \end{cases} \quad (12)$$

次に、変数を書き換え、 $\hat{\mathbf{Y}}' \hat{\alpha} \rightarrow \mathbf{z}_1 \in R^{2N}$ と $\hat{\alpha} \rightarrow \mathbf{z}_2 \in R^3$ と書き換える。これより、

$\min_{\hat{\alpha} \in \mathcal{A}} \|\mathbf{z}_1\|_1 + \iota_{\mathcal{A}}(\mathbf{z}_2)$ s.t. $\hat{\mathbf{Y}}' \hat{\alpha} - \mathbf{z}_1 = \mathbf{0}, \hat{\alpha} - \mathbf{z}_2 = \mathbf{0}$ が得られる。また、これらを用いた拡張ラグランジアン関数を、双対変数 $\mathbf{u}_1 \in R^{2N}$ と $\mathbf{u}_2 \in R^3$ を用いて、以下のように定義する。

$$L(\hat{\alpha}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{u}_1, \mathbf{u}_2) := \|\mathbf{z}_1\|_1 + \frac{\rho_1}{2} \|\hat{\mathbf{Y}}' \hat{\alpha} - \mathbf{z}_1 + \mathbf{u}_1\|_2^2 + \iota_{\mathcal{A}}(\mathbf{z}_2) + \frac{\rho_2}{2} \|\hat{\alpha} - \mathbf{z}_2 + \mathbf{u}_2\|_2^2 \quad (13)$$

ここでパラメータ $\rho_1 > 0$ と $\rho_2 > 0$ は反復求解時の解の収束を制御する。 ρ_2 には大きな値を指定する必要がある、 ρ_1 の関わる ℓ_2 ノルムとのバランスを取るには、 ρ_2 へは画素数に関連した値を指定する必要がある。この関数を用いることで、ADMM での i 回目の反復で行う計算は以下のように導出される。

$$\begin{aligned}
\mathbf{z}_1^{t+1} &:= \arg \min_{\mathbf{z}_1} L(\hat{\alpha}^t, \mathbf{z}_1, \mathbf{u}_1^t) = S_{1/\rho_1}(\hat{\mathbf{Y}}' \hat{\alpha}^t + \mathbf{u}_1^t), \\
\mathbf{z}_2^{t+1} &:= \arg \min_{\mathbf{z}_2} L(\hat{\alpha}^t, \mathbf{z}_2, \mathbf{u}_2^t) = \Pi_{\hat{\mathcal{A}}}(\hat{\alpha}^t + \mathbf{u}_2^t), \\
\mathbf{u}_1^{t+1} &:= \arg \min_{\mathbf{z}_2} L(\hat{\alpha}^t, \mathbf{z}_1^t, \mathbf{u}_1) = \mathbf{u}_1^t + \hat{\mathbf{Y}}' \hat{\alpha}^t - \mathbf{z}_1^{t+1}, \\
\mathbf{u}_2^{t+1} &:= \arg \min_{\mathbf{z}_2} L(\hat{\alpha}^t, \mathbf{z}_2^t, \mathbf{u}_2) = \mathbf{u}_2^t + \hat{\alpha}^t - \mathbf{z}_2^{t+1}, \\
\hat{\alpha}^{t+1} &:= \arg \min_{\mathbf{z}_2} L(\hat{\alpha}, \mathbf{z}_1^t, \mathbf{z}_2^t, \mathbf{u}_1^{t+1}, \mathbf{u}_2^{t+1}) \\
&= \left(\rho_1 \hat{\mathbf{Y}}'^T \hat{\mathbf{Y}}' + \rho_2 \mathbf{Id} \right)^{-1} (\rho_1 \hat{\mathbf{Y}}'^T (\mathbf{z}_1^{t+1} - \mathbf{u}_1^{t+1}) + \rho_2 (\mathbf{z}_2^{t+1} - \mathbf{u}_2^{t+1})),
\end{aligned} \tag{14}$$

なお、アルゴリズムの順序であるが、ベクトルの初期値として $\mathbf{u}_1^0, \mathbf{u}_2^0, \mathbf{z}_1^0, \mathbf{z}_2^0$ を用いた場合に $\hat{\alpha}^{t=1} = \mathbf{0}$ となることを避けるために、一部変えている。その他の初期値については、 $\hat{\alpha}^0 \neq \mathbf{0}, \hat{\alpha}^0 \leftarrow \frac{\hat{\alpha}^0}{\|\hat{\alpha}^0\|_2}$ を用いる。関数については、 $S_{1/\rho_1}(\cdot)$ は縮退処理（ソフト閾値処理）を表す。

$$[S_{1/\rho_1}(\mathbf{x})]_i := x_i(1 - 1/\max(1, \rho_1 |x_i|)) \tag{15}$$

ここで $[\cdot]_i$ はベクトルの要素ごとの計算を表す。関数 $\Pi_{\hat{\mathcal{A}}}(\cdot)$ は球面多様体 $\hat{\mathcal{A}}$ への射影を表し、ベクトルを単位ベクトルに修正する処理を表す。

$$[\Pi_{\hat{\mathcal{A}}}(\mathbf{x})]_i := x_i / \|\mathbf{x}\|_2 \tag{16}$$

逆行列 $(\cdot)^{-1} \in \mathbb{R}^{3 \times 3}$ は、反復処理内では変化しないため、事前に計算可能である。最後に、一旦取り除いた制約 $\mathbf{1}'^T \hat{\alpha} \geq 0$ を加え、一つの全域最適化を得る。

$$\hat{\alpha}^* := \text{sign}(\mathbf{1}'^T \hat{\alpha}) \cdot \hat{\alpha}, \quad \hat{\mathbf{x}}^* := \hat{\mathbf{Y}}' \hat{\alpha}^* \tag{17}$$

球面上のコストマップは球面の原点を基準として点対称であり（球面の正反対に同一コスト値をもつ）、球面上に全域最適解は二つ存在し、互いの符号を入れ替えたものとなる。そのため、どちらかの全域最適解に収束させた後に、その解を表すベクトルの向きがユーザーが与えた基準ベクトルの方向を向くように、解のベクトルの符号を修正すれば良い。

4. 実験結果

木簡画像データベース [2] に含まれる画像に対して提案法を用いた結果を示す。画像サイズはおおよそ 200h×200w～500h×500w である。また、シミ検出のテスト画像を用いた実験結果もあわせて示す。画像の輝度値は [0,1] の範囲となるように正規化して用いる。

4.1 収束の評価

図90に図89の画像を用いた場合の解の収束の様子を示す。ADMMのパラメータは $\rho_1 = 10$ と $\rho_2 = 10^6$ と設定し、反復回数は200とした。コストマップ(b)と(d)は網羅的に計算を行って生成した。黒色の領域が小さなコスト値をもち、求めたい解が存在する領域である。コストマップ上の(赤色の)点列は、収束する解を表し、表示のため、初期解 $\hat{\alpha}$ は大域最適解から離れた場所に設定した。この結果からは、赤い点列がコスト最小の点へと正しく収束している様子が分かる。また、得られた解である変換係数を用いて色変換を行った結果を図89(c)と(e)に示す。この図からも、図89と比べて望ましい色変換が行われたことが分かる。

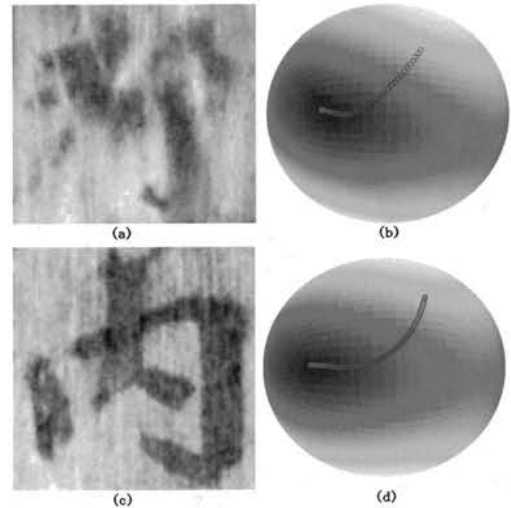


図90 図89に示した画像の最適解。(a)(c) 最適な変換係数の組を用いて色変換を行った結果。(b)(d) 三次元の球面上の座標 $(\hat{\alpha}_r, \hat{\alpha}_g, \hat{\alpha}_b), \|\hat{\alpha}\|_2 = 1$ 上でのコストマップ。赤い点列は解の収束の様子を表す。

ADMMの収束を制御するパラメータ(ステップ幅) ρ_1 と ρ_2 に関しては、凸な問題に用いるときよりも注意をして設定する必要がある。図91に幾つかのパラメータの組を用いた場合の結果を示す。ステップ幅を大きく設定した場合、例えば、同図中の右側の場合、大域最小値には収束せず、球面状を動き続けている。木簡画像を用いた場合には、最適なパラメータの組は $\rho_1 = 10$ と $\rho_2 = 10^6$ であった。先の図90もこの組を用いている。

4.2 代表的な色変換方法及び色成分分解方法との比較

図92に代表的な色変換方法と色成分の分解方法との比較を示す。見やすさを考慮し、各画像の輝度値を線形濃度変換によって最小値と最大値 $[\min, \max]$ を $[0, 1]$ となるように正規化している。正負の符号をもつ場合は、文字領域が黒色となるように符号反転を行った。PCAの結果については白色化を用いておらず、輝度値の正規化を行った場合、白色化の影響は結果画像には影響しないためである。

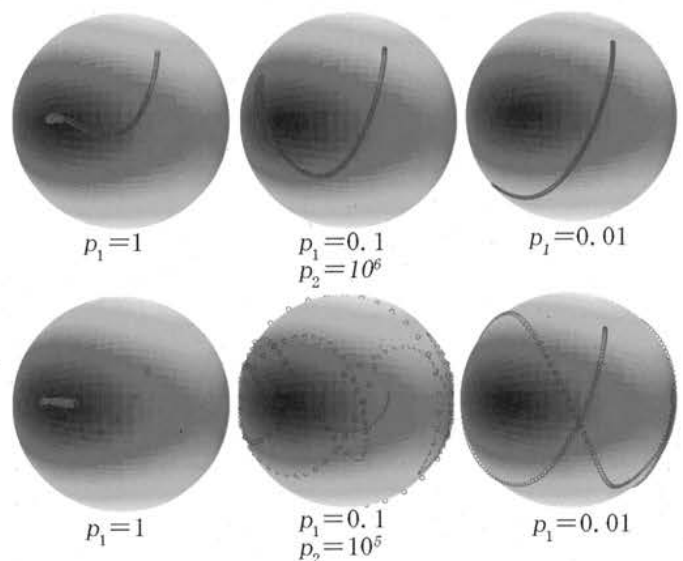


図91 パラメータの設定と解の収束の様子

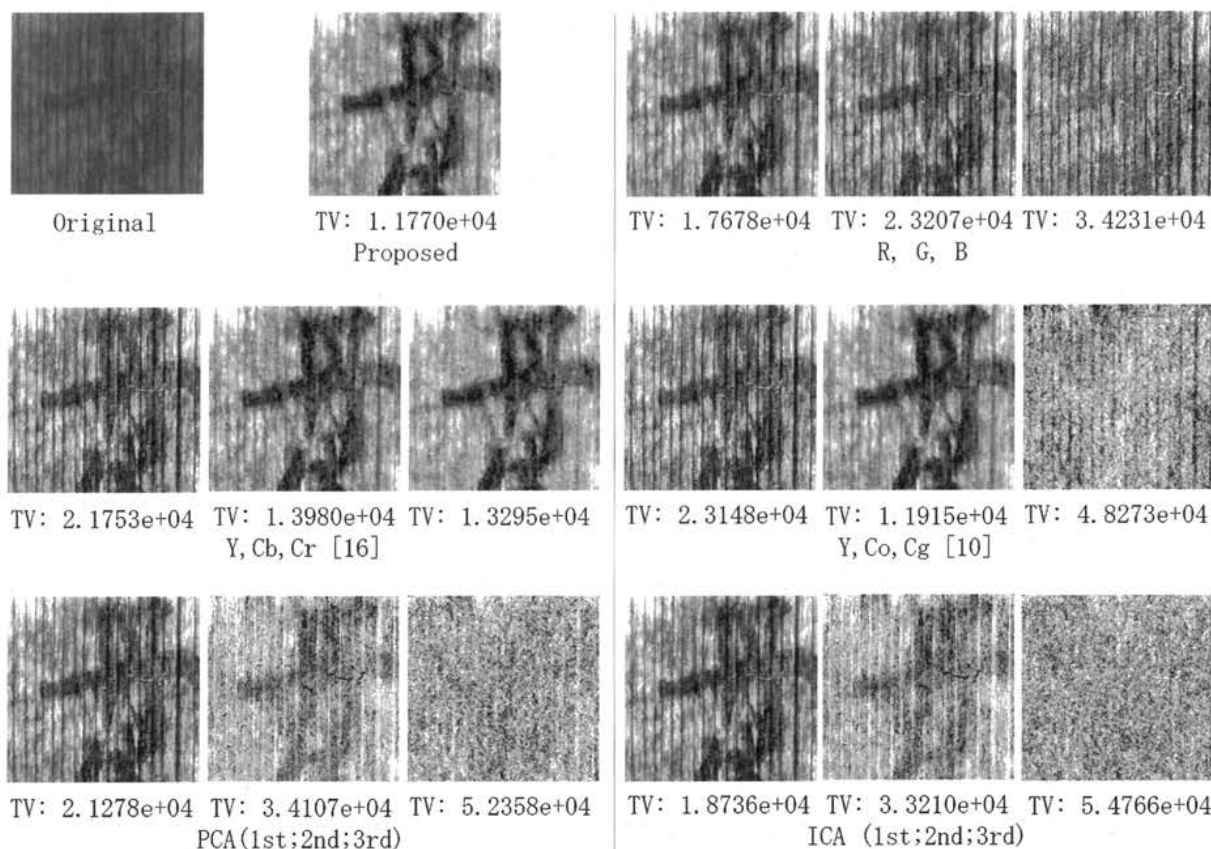


図92 代表的な色変換方法と色成分分解方法との比較。処理結果の輝度値は表示用に正規化している。各画像の下に TV 値を参考用に示した。水平・垂直方向への微分フィルタを施した後、微分値の絶対値を計算し、全ての画素の値を合計した値である。

独立成分分析 (ICA: Independent Component Analysis)^{*2} の結果は、PCA と同様となった。

この変換結果の画像の一覧からは、文字が良く抽出できているといえるのは、提案法、Cr 成分 (Chroma red: 赤色の色差。変換係数はおよそ $\alpha \approx [1, -1, 0]^T$)、Co 成分 (Chroma orange: オレンジ色の色差。変換係数は $\alpha \approx [1, 0, -1]^T$) の三つであることが分かる。

各画像の下に示した TV 値については、見やすいと思われる画像においては、TV 値が小さいことが分かり、上記のように輝度値の範囲を [0,1] に補正した後であれば、TV 値を用いて見やすさを評価できることが示せた。

4.3 判別分析を用いた同色領域分割法との比較

図 93 に、文字色と木目の色の判別を線形判別分析法 (LDA: Linear Discriminant Analysis) [17] を用いて行った結果を示す。この方法はユーザー補助情報を必要とするため、(a-1) に示すような文字部分を黒、背景部分を白、判別したい領域を灰色に設定した

^{*2} ICA については、B.Moore 氏が MATLABcentral 上で配布しているソースコード“PCA and ICA Package”を用いた。

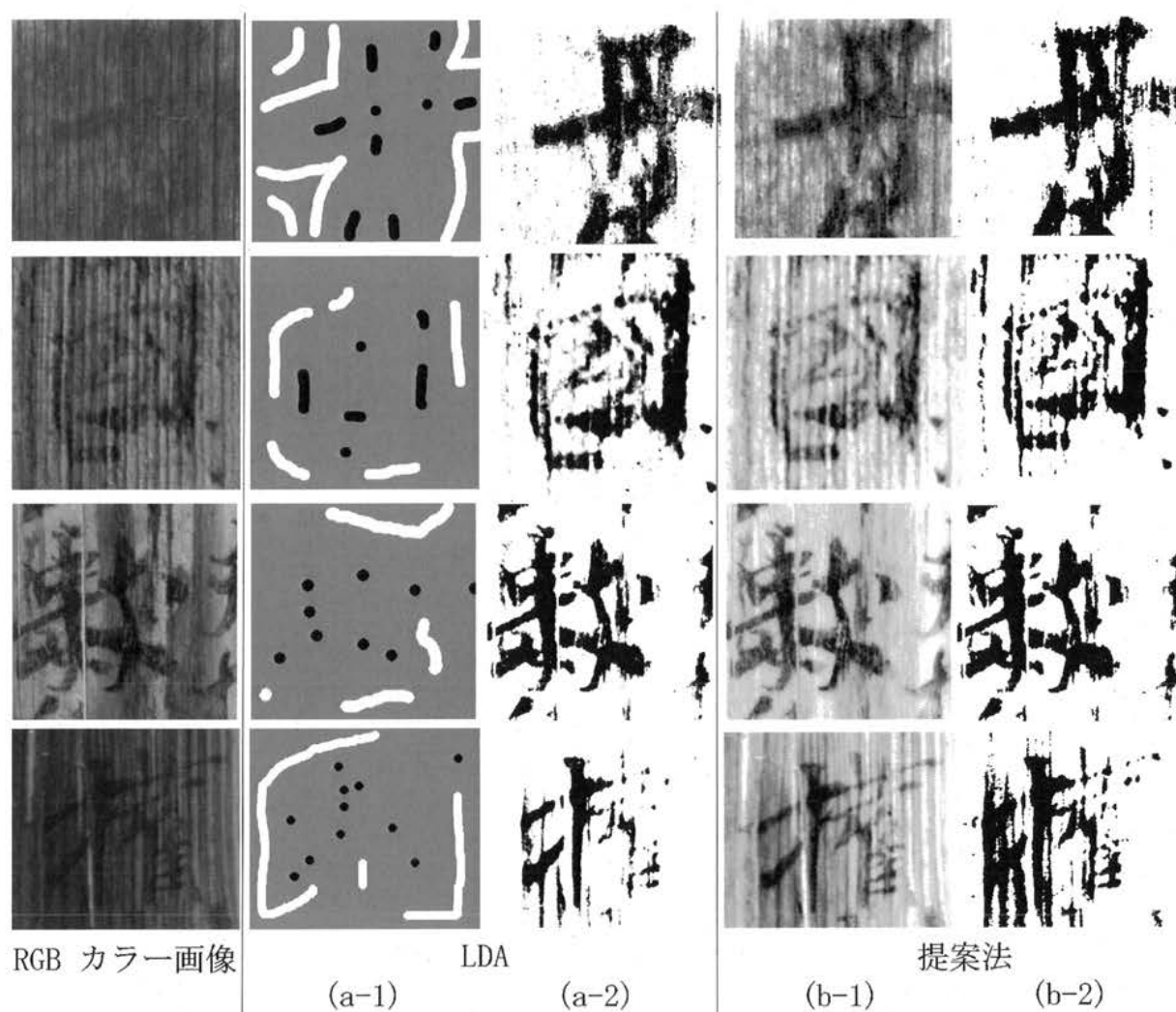


図93 線形判別分析(LDA)[17]を用いた文字色と背景色の分類法との比較。(a-1)はLDAで必要となるユーザー補助情報、(a-2)は結果画像を表す。(b-1)は提案法の色変換結果、(b-2)は色変換後の画像を二値化した結果を表す。二値化にはNiblack[18]の方法を用いた。

三色マップを用意した。このマップで指定した文字部分と背景部分から色を抽出し、その複数色（ここでは文字と背景の二色）の分布を上手く分けるパラメータを学習する。マップ中の灰色領域の画素色を入力した場合、学習結果を元に、その色が文字と背景のどちらに属するかを計算する。この図においては文字に属する確率が高いものを黒で表し、確率が低いものを白で表した。

このLDAを用いた結果では、(a-2)に示すように良好な結果が得られた。しかし、このように望ましい文字形状を得るためには、(a-1)に示したように文字と背景と完全に断定できる画素を注意して選ぶ必要がある。小数の画像であれば、このような対話形式の選択作業を用いても良いが、膨大な画像を処理する場合には手間がかかると予想される。一方、提案法においても同様の結果が得られており(b-2)、(b-1)の色変換結果に対してNiblackの二値化[18]を行うことで、LDAに近い結果が得られる。この二値化では、色変

換後の画素値を x_i 、全画素での平均値と標準偏差をそれぞれ $\text{mean}(\{x_i\})$ と $\text{std}(\{x_i\})$ と求めておき、画素値が $x_i \leq \text{mean}(\{x_i\}) - \gamma \text{std}(\{x_i\})$ を満たすときに文字（黒色）、それ以外では背景（白色）と判断した。調整用のパラメータは $\gamma = 0.7$ とした。提案法では各画素での色の選択を行う手間はかからない。

この図の最下段の結果は、両方の手法において分類が難しかった結果を示す。この原因は主に水平方向へ徐々に輝度が変わるグラデーションが生じていることに由来する。提案法の色変換後の画像 (b-1) では、文字は十分に見えるようになっているため、二値化方法を改良することで文字を抽出できる可能性はある。

4.4 領域抽出の関連手法との比較

図 94 に文字抽出の目的に使用できる他の強力な方法との比較を示す。各結果画像は左から順に、(a) LearningBased Matting 法 [19]（ユーザー補助情報が必要なため、図 93 の三色マップを用いた）、(b) ℓ_0 平滑化法 [20]、(c) MSER 特徴を用いた領域分割法 [21] である。(c) は情景画像からの文字抽出に用いられ、研究分野では標準的に用いられる。なお、(b) と (c) では処理を RGB 成分ではなく YCbCr 成分へと色変換した後、Cr 成分に対して用いた。輝度成分 Y は輝度の変化量が大きく、所望の結果を得ることは困難である。

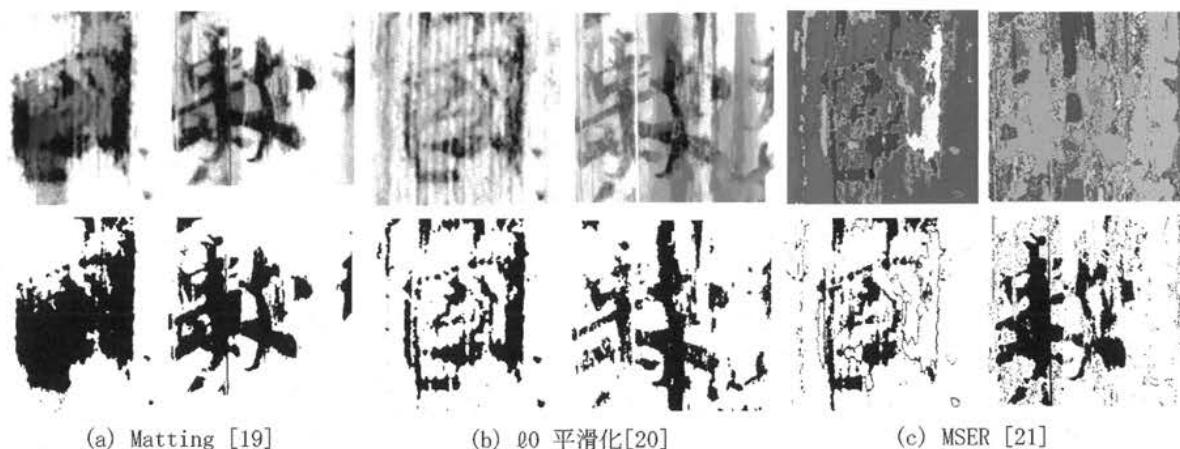


図94 強力な雑音除去能力や領域分割性能をもつ標準手法との比較結果。下段は二値化の結果を表す。

同図の結果からは、(b) の ℓ_0 平滑化法が良い結果を与えている。しかし、いずれの手法も LDA と提案法の性能には及ばなかった。ここに示した方法はいずれの方法も二値化を必要とするが、二値化前の画像において木目を低減できていない。提案法と組み合わせることで性能を改善できる可能性がある。特に (b) は提案法でも用いた TV 値を用いており、 ℓ_1 ノルムの代わりに ℓ_0 ノルムを用いる。ただし、 ℓ_0 ノルムの計算は ℓ_1 よりも計算を収束させることが難しくなると考えられる。今後の課題として検討したい。

4.5 シミの検出への応用

文字抽出以外への応用として、図 95 にシミ検出の実験結果を示す。目的は布地に染み込んだコーヒーのシミを検出することである。参考用に (a) に LDA を用いた検出結果を示す。(a) の画像の右上部の画像はユーザー補助情報として与えた三色マップである。なお、シミの色が何色かが予め分かっていると仮定している。次に、(b-1) と (b-2) は単純な PCA を用いた結果である。(b-1) に示すように PCA は知覚的には良い結果を与えるが、画像全体として輝度が一様ではないため、二値化処理において失敗する (b-2)。(c-1) と (c-2) は提案法の結果であり、色変換後の画像の輝度値は背景領域では一様で平坦である (c-1)。二値化処理では、左側のシミは大きく検出されているが、背景領域での誤検出は低減されている (c-2)。

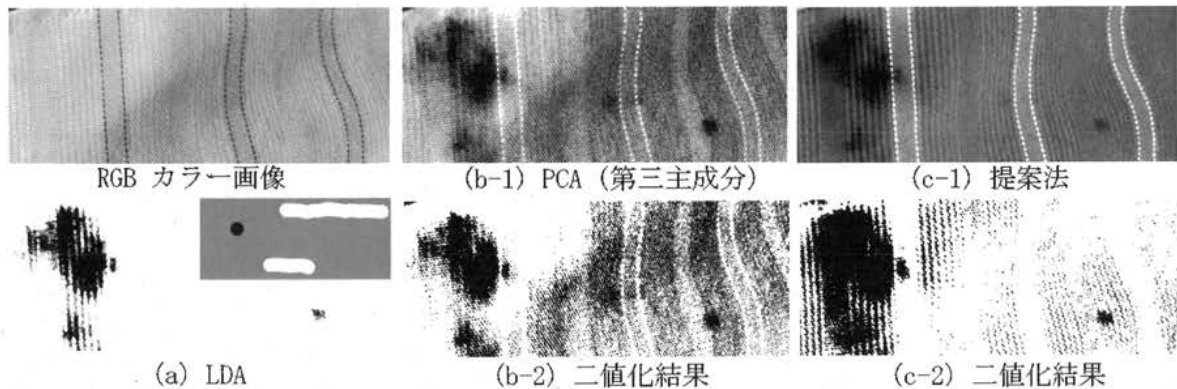


図95 シミ検出：布地の上のコーヒーのシミの検出。(b-2)と(c-2)の二値化は図 93 に示した結果と同様の計算を用いて行った。

この画像において、提案法では青い線が主たる色領域として誤って検出される傾向があった。この問題を解決するには、提案法の PCA 関連部分で求めている最大固有値に対応する固有ベクトル（変換係数）の他に、第二、第三に大きい固有値に対応する固有ベクトルを求めるのが良いと思われる。PCA の結果 (b-1) においても、シミが浮き上がっている色成分を与える変換係数は最小固有値（三番目に大きな固有値）に対応した固有ベクトルであった。提案法の制約部分にグラムシュミット直交化を加えることで求められるかもしれない。

4.6 実行時間

実験に用いた PC 環境は、Intel Core i7@2.7GHz、実装は C++ と MATLAB を用いて行った。提案法での変換係数を求めるのに必要な実行時間は画素数に比例する。ADMM のアルゴリズムの一回あたりの必要時間は 30 (msec/Mpixs) であった。例えば、図 2 (c) の画像は、画像サイズが 200h×180w であり、100 回の反復を必要とし（大域最適解から離れた初期解を用いた）、このときの処理時間は 70 (msec) であった。この実行時間は主に式

(14)での行列 $\hat{\mathbf{Y}}'$ と $\hat{\mathbf{Y}}'^T$ の計算にかかる。この計算は並列処理できるため、高速化が可能である。

5. まとめ

本稿では、画像の主たる、かつ、見やすい色成分を抽出する方法について述べた。最適な線形色変換係数を求めるため、PCAとTV最小化を組み合わせで定式化を行い、その求解法を示した。すなわち、方程式が球面多様体上でのLAD最小化に帰着することを示し、ADMMを用いて求解を行った。この形式の定式化は、非凸な問題の最適化や二値の二次計画問題に利用でき、近年研究がすすめられている。今回は色変換後に二値化を行って文字形状を取得したが、0-1二値（文字と背景のラベル）を求める問題として定式化し、直接二値画像を得ることも今後の課題として検討したい。

謝辞

本研究を遂行するにあたり御意見を賜った木簡からの文字抽出システム[1]の著者ら、及び、北九州市立大学の奥田正浩教授、東京工業大学の小野俊介らに感謝申し上げます。

参考文献

- [1] A. Kitadai, K. Shirai, E. Yuki, M. Nakagawa, H. Baba, A. Watanabe, S. Watanabe, and N. Kurushima, "Adaptation of normalization process to historical mokkan images and update of support system to support reading historical mukkans (in japanese)," in *Proc. JINMONKON*, 2013, pp. 65-70.
- [2] H. Baba and A. Watanabe, "Ancient Japanese wooden tablets (mokkan) image database," <http://jiten.nabunken.go.jp/index.html>.
- [3] I.T. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics. Springer, 2002.
- [4] L.I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259-268, 1992.
- [5] Z. Wen and W. Yin, "A feasible method for optimization with orthogonality constraints," *Math. Program.*, vol. 142, no. 1, pp. 397-434, 2013.
- [6] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite elements approximations," *Computers & Math. with Appl.*, vol. 2, no. 1, pp. 17-40, 1976.
- [7] J. Eckstein and D.P. Bertsekas, "On the Douglas-Rachford splitting method and proximal point algorithm for maximal monotone operators," *Math. Pro-gram.*, vol. 55, pp. 293-318, 1992.

- [8] T. Goldstein, and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, pp. 323-343, 2009.
- [9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, 2011.
- [10] R. Malvar, G.J. Sullivan, and S. Srinivasan, "Lifting-based reversible color transformations for image compression," in *SPIE Apps. Digital Image Process.*, 2008, pp. 1-10.
- [11] B.N. Parlett, *The symmetric eigenvalue problem*, vol. 20 of *Classics in Applied Mathematics*, SIAM, 1998.
- [12] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," Tech. Rep., Xi'an Jiaotong Univ., 2015.
- [13] N. Kwak, "Principal component analysis based on l1-norm maximization," *IEEE Trans. Pattern Anal. Machine Intelli. (TPAMI)*, vol. 30, no. 9, pp. 1672-1680, 2008.
- [14] P.P. Markopoulos, G.N. Karystinos, and D.A. Pados, "Optimal algorithms for L1-subspace signal processing," *IEEE Trans. Signal Process. (TSP)*, vol. 62, no. 19, pp. 5046-5058, 2014.
- [15] B. Wu and B. Ghanem, " ℓ_p -box ADMM: A versatile framework for integer programming," arXiv.
- [16] C.A. Poynton, *Technical introduction to digital video*, John Wiley & Sons, 1996.
- [17] R.A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals. of Human Genetics*, vol. 7, no. 2, pp. 179-188, 1936.
- [18] W. Niblack, *An introduction to digital image processing*, Prentice Hall, 1986.
- [19] Y. Zheng and C. Kambhamettu, "Learning based digital matting," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2009, pp. 889-896.
- [20] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via ℓ_0 gradient minimization," *ACM Trans. Graph. (TOG)*, vol. 30, no. 5, pp. 174:1-174:12, 2011.
- [21] H. Chen, S.S. Tsai, G. Schroth, D.M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2011, pp. 2609-2612.