

RGB 色成分の線形和による木簡画像からの墨文字の抽出法

杉本 拓矢¹ 藤森 和博¹ 白井 啓一郎¹ 宮尾 秀俊¹ 丸山 稔¹

¹ 信州大学工学部 (〒 380-8553 長野県長野市若里 4-17-1)

Character Extraction from Wooden Tablets using Linear RGB Color Composition

Takuya SUGIMOTO¹, Kazuhiro FUJIMORI¹, Keiichiro SHIRAI¹,
Hidetoshi MIYAO¹, and Minoru MARUYAMA¹

¹Shinshu Univ., 4-17-1 Wakasato, Nagano, Nagano 380-8553

Abstract: We present a character extraction method of degraded faint characters written on wooden tablets. Characters are written with Japanese black ink (SUMI) and absorb near-infrared light, and therefore differences among R, G, and B color components indicates the principal color component of characters. Our method emphasizes the color of characters by finding optimal linear combination of R, G, and B colors. Its blending coefficients are calculated so as to not only extract the principal color component but also reduce the total variation of gradients of the component. In short, our method is performed by principal component analysis (PCA) with total variation regularization (TV). In the experiment, we show how the blending coefficients are automatically calculated, and characters is emphasized by linear combination with the blending coefficients.

Key Words: color transform, PCA, total variation, color extraction

1 はじめに

本稿では、古代の木簡に書かれた文字の解読支援システム¹⁾における文字抽出方法の一検討について述べる。木簡上の文字は経年劣化によりかすれ、背景の木目模様の影響もあり、文字部分のみを抽出することが難しいが、文字は黒色の墨で書かれており近赤外光を吸収しやすいため、近赤外光で撮影すれば、文字部分が黒く写り、文字を強調できる。ただし、これまでの全ての撮影画像が近赤外光で撮影されたわけではなく、可視光や輝度値のみで撮影された画像が大半を占める。

上述の理由から、可視光で撮影された一般的な RGB カラー画像からの文字抽出法を検討する。墨で書かれた文字は赤色光を吸収しやすく、カラー画像の赤色成分と他の緑色と青色成分を比べると、文字部分において光の吸収量の差による色差が生じる。そのため、簡単かつ有効な抽出方法としては赤色と青色の色差 $R - B$ が利用でき、文字色と背景色のコントラストを高められ、文字領域を抽出しやすくなる (図 1 参照)。ただし、他の色差 (クロマ成分) や輝度成分に変化が現れる場合があり、現在のシステム¹⁾では、各種色変換 (YCbCr や YCoCg 色変換) を施して複数のクロマ成

分を得た後、対話的に文字抽出に用いる成分を決めており、この自動決定を行いたい。

主要な色成分の自動抽出には主成分分析 (PCA) を利用できると考えている。RGB 色空間での全画素色の分布とその広がり (分布軸と軸に沿った分散) のうち、分散の大きな軸およびそれに直交する軸 (主成分軸) に沿った色の変化は、通常、輝度や色差などの主成分を表す。そのため、PCA により代表的な主成分軸を解析して主成分を求めれば、前述の汎用的な色変換を用いるよりも成分の候補数を絞り込める。これに加え、より文字として知覚しやすい成分を得る条件を PCA に施せば、一意に成分を決定できる可能性がある。

知覚的にノイズが少なく見やすい画像のもつ特徴としては、被写体ごとに多くの領域が単一色で平坦であり、領域の境界で急峻に色に変化する特徴がある。数値的には輝度の変化量の総和 (Total Variation : TV) が小さくなるため、この TV 値を最小化するという条件は、ノイズ除去やぼけ除去をはじめとする画像の復元処理において一般的に用いられる。

提案法では、上述の PCA と TV を合わせ、画像の主たる色成分かつ文字形状が鮮明に現れる成分の自動取得を目的とする。色成分を R,G,B の基本的な色成分の線形結合によって表し、その線形結合に用いる最適

な合成係数（三つのスカラー値）を求めることを問題とする。この問題は、求める合成係数の二乗和が一定となる制約付きの最小絶対値誤差（LAD）の最小化問題²⁾に帰着するが、ADMM³⁾ (alternating direction methods of multiplier) と呼ばれる凸最適化のアルゴリズムを用いて解けることを示す。

なお、頁数の都合上、PCA と TV を組み合わせるに至った背景は省略し、実装に必要な情報のみ記載する。

2 諸定義

本章では、画像のデータ表現や提案法の基となる主成分分析（PCA: principal component analysis）および全変動（TV: total variation）の最小化について簡単に説明する。

2.1 線形結合による画像成分の抽出

画像は RGB カラー画像とし、各画素 i での基本的な色成分を (r_i, g_i, b_i) と表す。更に、表現の簡素化のため、画像の二次元配列を列優先で一次元に並べ直したベクトルとして扱い $\mathbf{r} = [r_1, \dots, r_N]^\top \in \mathbb{R}^N$ 、 \mathbf{g} と \mathbf{b} についても同様に定める。 N は総画素数を表す。輝度や色差などの成分を得る線形結合を、合成係数 $\{\alpha_1, \alpha_2, \alpha_3\} \in \mathbb{R}^1$ を用いて次式で表す。

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

また、PCA で用いられる表記法にあわせ、データ行列 $\mathbf{X} := [\mathbf{r}, \mathbf{g}, \mathbf{b}] \in \mathbb{R}^{N \times 3}$ と係数ベクトル $\boldsymbol{\alpha} := [\alpha_1, \alpha_2, \alpha_3]^\top$ を用いて表す。上式では、 $(\alpha_1, \alpha_2, \alpha_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ ならば \mathbf{y} は輝度成分となり、 $(1, 0, -1)$ ならば色差（橙色）成分となる。

結合結果 \mathbf{y} であるが、値の変化幅（分散や標準偏差）が適切でないと抽出した文字が見にくくなるため、平均値 $\mu(\mathbf{y}) := \frac{1}{N} \sum_i y_i$ を差し引いた後、標準偏差 $\sigma(\mathbf{y}) := \frac{1}{N} \sum_i (y_i - \mu(\mathbf{y}))^2$ で除算して正規化して用いる。

$$\hat{\mathbf{y}} := \frac{1}{\sigma(\mathbf{y} - \mu(\mathbf{y}))} (\mathbf{y} - \mu(\mathbf{y})), \quad \|\boldsymbol{\alpha}\|_2 = 1 \quad (2)$$

なお、この場合は $\boldsymbol{\alpha}$ をスカラー倍した値も解となり、解が定まらないため、合成係数の二乗和が定数 $\|\boldsymbol{\alpha}\|_2 = 1$ となるように制約を施す。

2.2 主成分分析による色成分の抽出

画像内で用いられる色データ (r_i, g_i, b_i) の分布は、陰影の影響を受けて黒色方向へと伸びる傾向があり、

この分布軸に沿った色の変化量は輝度成分を表す。ここで分布軸を単位ベクトル $\hat{\boldsymbol{\alpha}} := \boldsymbol{\alpha} / \|\boldsymbol{\alpha}\|_2$ として定義すると、軸方向への変化量は $\hat{\boldsymbol{\alpha}}$ と (r_i, g_i, b_i) の内積として与えられる。ここで $\|\mathbf{x}\|_p := (\sum_i |x_i|^p)^{1/p}$ はベクトルの l_p ノルムを表す。全画素に対する一括の内積計算は $\mathbf{X} \hat{\boldsymbol{\alpha}}$ と表わせ、式 (1) と関連する。

R, G, B 三次元空間での色分布の広がり（分布軸と軸に沿った分散）を求めるには PCA が利用できる。色の分布をガウス分布（楕円形の分布）とみなすと、その共分散行列の固有値ベクトルが分布の軸に相当し、固有ベクトルに対応する固有値が分散を表す。共分散行列を求めるには、データ行列の各列の色から平均色を差し引いた行列 $\bar{\mathbf{X}}$ を以下のように乗算する。

$$\text{cov}(\mathbf{X}) := \frac{1}{N} \bar{\mathbf{X}}^\top \bar{\mathbf{X}} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top \quad (3)$$

ここで最右項は固有分解を表し、互いに直交する固有ベクトルからなる行列 $\mathbf{U} := [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] \in \mathbb{R}^{3 \times 3}$ と、対応する固有値が対角に並ぶ行列 $\boldsymbol{\Lambda} := \text{diag}(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^{3 \times 3}$ ただし $\lambda_1 \geq \lambda_2 \geq \lambda_3$ が得られ、各固有ベクトルとデータ行列の積がデータの主成分を表す。

$$\mathbf{y}_k := \bar{\mathbf{X}} \mathbf{u}_k, \quad \|\mathbf{u}_k\|_2 = 1 \quad (4)$$

上述のような $\boldsymbol{\alpha} = \mathbf{u}_1$ として第一主成分軸を解として与える最適化の問題は、合成係数の二乗和が 1 となる制約のもと、結合結果 $\mathbf{y} = \bar{\mathbf{X}} \boldsymbol{\alpha}$ の全要素の二乗和を最大化する $\boldsymbol{\alpha}$ を求める問題に相当する。

$$\mathbf{u}_1 = \arg \max_{\boldsymbol{\alpha}} \|\bar{\mathbf{X}} \boldsymbol{\alpha}\|_2^2, \quad \text{subject to } \|\boldsymbol{\alpha}\|_2 = 1 \quad (5)$$

また、式 (2) のように正規化された $\hat{\mathbf{y}}_k$ を得るには、式 (3) の共分散行列が単位行列となるように予め無相関化したデータ行列 $\tilde{\mathbf{X}} := \bar{\mathbf{X}} \mathbf{U} \boldsymbol{\Lambda}^{-1/2}$ を用いれば良い。

$$\hat{\mathbf{y}} := \tilde{\mathbf{X}} \tilde{\boldsymbol{\alpha}}, \quad \|\tilde{\boldsymbol{\alpha}}\|_2 = 1, \quad \tilde{\boldsymbol{\alpha}} := \boldsymbol{\Lambda}^{1/2} \mathbf{U}^\top \hat{\boldsymbol{\alpha}} \quad (6)$$

以降では $\hat{\boldsymbol{\alpha}}$ の代わりに $\tilde{\boldsymbol{\alpha}}$ を用いる。

2.3 全変動最小化による平滑化

画像の TV 最小化は、ノイズ除去やぼけ除去で用いられ、ここではノイズ除去・平滑化での使用例を示す。

隣接する画素の画素値の変化量は、画素値の差分で表される。全画素において隣接する水平方向および垂直方向の画素との差分を計算するには、線形微分フィルタを用い、本稿では各方向に対する微分フィルタ係数を $[-1, 1, 0]$ と $[-1, 1, 0]^\top$ と定義する。また、微分

フィルタリングに相当する行列演算をフィルタ行列 \mathbf{D}_h と $\mathbf{D}_v \in \mathbb{R}^{N \times N}$ を用いて $\mathbf{y}_h := \mathbf{D}_h \mathbf{y}$ と $\mathbf{y}_v := \mathbf{D}_v \mathbf{y}$ と表す. 通常はこれらをまとめ, 行方向に行列を連結して表す: $\mathbf{D}\mathbf{y} = [\mathbf{y}_h^T, \mathbf{y}_v^T]^T$ ここで $\mathbf{D} := [\mathbf{D}_h^T, \mathbf{D}_v^T]^T \in \mathbb{R}^{2N \times N}$.

観測された画素値 \mathbf{y} を保ちつつ, 画素値の変化量を抑えて平坦化した \mathbf{x} を求める最適化問題は次式のように与えられる.

$$\arg \min_{\mathbf{x}} \|\mathbf{D}\mathbf{x}\|_1 + \lambda \|\mathbf{x} - \mathbf{y}\|_2^2 \quad (7)$$

第一項目は平滑化用の TV 正則化項, 第二項目は観測値からの誤差を表すデータ項, λ は調整用パラメータとなる. TV 項には一般的に l_1 ノルムが用いられる. l_2 ノルムの場合は最小二乗法用の解法を用いて計算でき, l_1 ノルムの場合は凸最適化用の解法を用いて計算できる. 汎用的な解法としては, ADMM³⁾ や主双対アルゴリズム⁴⁾ が用いられる. 解法例については次章の提案法の説明の際に示す.

3 提案法

前章 2 で示した PCA と TV をあわせ, 線形結合に用いる最適な合成係数を求める. すなわち, 式 (6) の $\hat{\mathbf{y}}$ について, 更にその TV: $\|\mathbf{D}\hat{\mathbf{y}}\|_1 = \|\mathbf{D}\tilde{\mathbf{X}}\tilde{\boldsymbol{\alpha}}\|_1$ を最小化するような $\tilde{\boldsymbol{\alpha}}$ を求める.

$$\arg \min_{\tilde{\boldsymbol{\alpha}} \in \mathcal{M}} \|\tilde{\mathbf{X}}'\tilde{\boldsymbol{\alpha}}\|_1, \quad \mathcal{M} := \{\tilde{\boldsymbol{\alpha}} \mid \|\tilde{\boldsymbol{\alpha}}\|_2 = 1\} \quad (8)$$

ここで $\tilde{\mathbf{X}}' := \mathbf{D}\tilde{\mathbf{X}} = [(\mathbf{D}_h\tilde{\mathbf{X}})^T, (\mathbf{D}_v\tilde{\mathbf{X}})^T]^T \in \mathbb{R}^{2N \times 3}$ と簡略表記および予め微分計算しておく ($\tilde{\mathbf{X}}$ の各列を表す画像を微分フィルタリングする). また \mathcal{M} は多様体を表し, $\tilde{\boldsymbol{\alpha}}$ は球面上に存在する. 多様体上の解を求める問題となるが, 球面のような単純な多様体であれば ADMM で解けることが報告されており⁵⁾, 以降ではその解法を示す.

式 (8) において, $\|\cdot\|_1$ の内部を別変数で置き換え, また, 多様体の制約を指示関数を用いて書き直す.

$$\begin{aligned} & \arg \min_{\tilde{\boldsymbol{\alpha}}} \|\mathbf{z}_1\|_1 + \iota_{\mathcal{M}}(\mathbf{z}_2) \\ & \text{subject to } \mathbf{z}_1 = \tilde{\mathbf{X}}'\tilde{\boldsymbol{\alpha}}, \quad \mathbf{z}_2 = \tilde{\boldsymbol{\alpha}} \end{aligned} \quad (9)$$

ここで指示関数は次式で表され,

$$\iota_{\mathcal{M}}(\tilde{\boldsymbol{\alpha}}) := \begin{cases} 0 & \tilde{\boldsymbol{\alpha}} \in \mathcal{M} \\ +\infty & \text{otherwise} \end{cases} \quad (10)$$

$\tilde{\boldsymbol{\alpha}}$ が解集合 \mathcal{M} に含まれれば 0 を, そうでなければ $+\infty$ を返す. そのため, 求まる解は必ず $\tilde{\boldsymbol{\alpha}} \in \mathcal{M}$ を満

たさねばならない. 次に, 拡張ラグランジュ関数として制約条件を目的関数へと移項する.

$$\begin{aligned} L(\tilde{\boldsymbol{\alpha}}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{u}_1, \mathbf{u}_2) &:= \|\mathbf{z}_1\|_1 + \frac{\rho_1}{2} \|\mathbf{z}_1 - (\tilde{\mathbf{X}}'\tilde{\boldsymbol{\alpha}} + \mathbf{u}_1)\|_2^2 \\ &+ \iota_{\mathcal{M}}(\mathbf{z}_2) + \frac{\rho_2}{2} \|\mathbf{z}_2 - (\tilde{\boldsymbol{\alpha}} + \mathbf{u}_2)\|_2^2 \end{aligned} \quad (11)$$

ここで $\rho_1 > 0$ と $\rho_2 > 0$ は収束の制御パラメータである. パラメータの値については次章 4 の実験にて示す. 式 (11) を用いた ADMM による解の更新アルゴリズムは次のように与えられる.

$$\begin{cases} \mathbf{v}_1^{(t)} := \tilde{\mathbf{X}}'\tilde{\boldsymbol{\alpha}} + \mathbf{u}_1^{(t)} \\ \mathbf{v}_2^{(t)} := \tilde{\boldsymbol{\alpha}} + \mathbf{u}_2^{(t)} \\ \mathbf{z}_1^{(t)} := \arg \min_{\mathbf{z}_1} L(\tilde{\boldsymbol{\alpha}}^{(t)}, \mathbf{z}_1, \mathbf{u}_1^{(t)}) = \mathcal{S}_{1/\rho_1}(\mathbf{v}_1^{(t)}) \\ \mathbf{z}_2^{(t)} := \arg \min_{\mathbf{z}_2} L(\tilde{\boldsymbol{\alpha}}^{(t)}, \mathbf{z}_2, \mathbf{u}_2^{(t)}) = \Pi_{\mathcal{M}}(\mathbf{v}_2^{(t)}) \\ \mathbf{u}_1^{(t+1)} := \mathbf{v}_1^{(t)} - \mathbf{z}_1^{(t+1)} \\ \mathbf{u}_2^{(t+1)} := \mathbf{v}_2^{(t)} - \mathbf{z}_2^{(t+1)} \\ \tilde{\boldsymbol{\alpha}}^{(t+1)} := \arg \min_{\tilde{\boldsymbol{\alpha}}} L(\tilde{\boldsymbol{\alpha}}, \mathbf{z}_1^{(t)}, \mathbf{z}_2^{(t)}, \mathbf{u}_2^{(t+1)}, \mathbf{u}_2^{(t+1)}) \\ = (\rho_1 \tilde{\mathbf{X}}'^T \tilde{\mathbf{X}}' + \rho_2 \mathbf{Id})^{-1} \\ (\rho_1 \tilde{\mathbf{X}}'^T (\mathbf{z}_1^{(t)} - \mathbf{u}_1^{(t+1)}) + \rho_2 (\mathbf{z}_2^{(t)} - \mathbf{u}_2^{(t+1)})) \end{cases} \quad (12)$$

ここで, t は反復回数, \mathbf{v}_1 と \mathbf{v}_2 は一時的なベクトル, $\mathbf{Id} \in \mathbb{R}^{3 \times 3}$ は単位ベクトルを表す. 関数 $\mathcal{S}_{\gamma}(\mathbf{x})$ はソフト閾値処理関数であり, \mathbf{x} の要素のうち 0 に近い微小な値を除去する.

$$\mathcal{S}_{\gamma}(\mathbf{x}) := \text{sgn}(\mathbf{x}) \cdot \max(0, |\mathbf{x}| - \gamma) \quad (13)$$

ここで $\text{sgn}(\cdot)$ は符号関数で, 要素ごとの符号を返す. 関数 $\Pi_{\mathcal{M}}(\mathbf{x})$ は射影関数であり, \mathbf{x} に最も近い集合 \mathcal{M} に含まれるベクトルを返す. 今回は単位球面上に存在する単位ベクトルとなる.

$$\Pi_{\mathcal{M}}(\mathbf{x}) := \mathbf{x} / \|\mathbf{x}\|_2 \quad (14)$$

初期化が必要な変数の初期値は, $\mathbf{u}_1^{(0)} = \mathbf{0}$ と $\mathbf{u}_2^{(0)} = \mathbf{0}$, $\tilde{\boldsymbol{\alpha}}^{(0)} = [\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}]^T$ と設定した.

4 実験結果

奈良文化財研究所が公開している木簡画像⁶⁾の一部を用いて, 提案法の効果を示す. 画素値を $[0, 1]$ に正規化し, ADMM のパラメータは $\rho_1 = 30$, $\rho_2 = 10^6$ 反復回数を 100 回とした. $\boldsymbol{\alpha}$ の初期値には, 解が正し



図 1: R, G, B の色成分の線形結合の例。(左) 原画像, (中央) 輝度成分: $(R+G+B)/3$, (右) 色差成分: $R-B$. 合成係数によりノイズや木目の模様が低減される. 上段では色差成分, 下段では輝度成分において文字が現れている

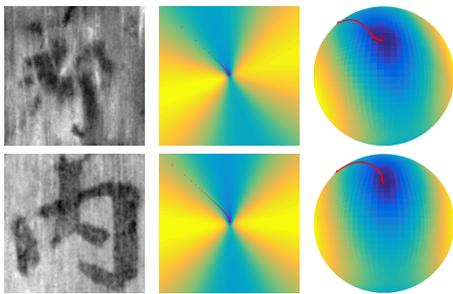


図 2: 最適な合成係数を用いた結合結果 (左) とコストマップ (中央), (右) は球面上に表示したもの. コストマップの計算では $\hat{\alpha}_1$ を固定し ($\hat{\alpha}_2, \hat{\alpha}_3$) を変化させた

く収束するかを調べるため, 最適値から離れた値を意図的に設定した.

まず, 提案した式 (8) を最小化するような最適な合成係数すなわち解が一意的に存在することを示し, ADMM の反復により解が収束する様子を示す. 図 2(中央) と (右) に, 図 1 の画像を対象として, 予め用意した合成係数 ($\alpha_1, \alpha_2, \alpha_3$) を用いて式 (8) を網羅的に計算した結果を示す. この図からは, 球面上に一箇所, コストの低い (青黒い) 箇所が存在することが分かる. コストマップ上の赤い点が ADMM による解の収束経路であり, コスト最小の点に収束しているのが分かる.

次に, 得られた合成係数を用いての結合結果を示す. 前述の図 2(左) が結合結果となる. 図 1 に示した輝度成分や色差成分に比べて文字が鮮明に現れているといえる. 一方, 図 3 に提案法が不得意とする画像を示す. 画像内に木簡と墨以外の茶色系以外の撮影台の色が写っており, この色が外れ値となり, 合成係数の推定に失敗する. そのため, 過度に白色となる箇所が現れている. 対処法としては, 茶色以外の色のクラスターに属する画素を予め除去しておく方法が考えられ, 今後の検討課題とする.



図 3: 撮影台の色を外れ値として含む画像の例

5 まとめ

本稿では, 古代の木簡に記載された文字の抽出方法として, RGB カラー画像の R,G,B の色成分を線形結合して文字が鮮明に現れる成分を求める方法について述べ, その線形結合の際に用いる最適な合成係数を求めるため, 主成分分析と全変動最小化を組み合わせた最適化問題を提案した.

今後の課題としては, 色の分布から大きく値が外れる色 (外れ値) を含む画像に対しての耐性を高める必要がある. 例えば, 木簡と文字は茶色系の色分布をもつが, 撮影時に用いた撮影台などの茶色から外れた色が写り込むと, その外れ値の影響を受ける問題がある. 対話的に背景領域を除外すれば良いが, 自動的に排除できればユーザーの手間を低減できる.

謝辞

本研究は, 奈良文化財研究所 (NNRICP) と桜美林大学の末代らと共に開発している “古代木簡解読支援システム” の成果の一部であり, 科学研究費 JP25220401 の支援を受けている. 研究をすすめる上で協力いただいた NNRICP の馬場氏・渡邊氏, 桜美林大学の末代氏をはじめとする諸氏らに感謝の意を表する.

参考文献

- 1) 末代 誠仁, 白井 啓一郎, 遠藤 友樹, 中川 正樹, 馬場 基, 渡辺 晃宏, 井上 聡, 久留島 典子: 古代木簡に対する平滑化処理の適用および古代木簡解読支援システムのアップデート, *じんもんこん* 2013 論文集, 65-70 (2013).
- 2) Z. Wen, W. Yin: A feasible method for optimization with orthogonality constraints, *Mathematical Programming*, 142(1), 397-434 (2013).
- 3) S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein: Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends in Machine Learning*, 3(1), 1-22 (2011).
- 4) A. Chambolle, T. Pock: A first-order primal-dual algorithm for convex problems with applications to imaging, *J. of Math. Imaging Vision*, 40(1), 120-145 (2011).
- 5) Y. Wang, W. Yin, J. Zeng: Global convergence of ADMM in nonconvex nonsmooth optimization, in *Tech. Rep. on ArXiv*, 1-29 (2015).
- 6) 馬場 基, 渡辺 晃宏, “木簡画像データベース,” 奈良文化財研究所, <http://jiten.nabunken.go.jp/index.html>.