

採点ミス発見支援システム内の 抽出処理の改良

村本 幸次郎[†] 松尾 賢一^{††}

[†] 奈良工業高等専門学校 専攻科 システム創成工学専攻 情報システムコース

1. はじめに

筆者らは、採点者の負担軽減のための採点済答案画像内の採点記号と部分点に対する文字認識結果に基づき、採点記号の配点と部分点間に不整合があると判定した箇所を採点ミスとして通知するシステム[1]を改良中である。

現在、本システムの抽出処理部では、特に採点記号パターンが枠線や解答文字等の黒色領域と重畳箇所においてパターンの欠損(断線、欠落等)が生じるため、膨張処理によって欠損を補完していた。しかし、膨張処理は、閉じたループを持つ数字パターンに対して、ループを穴埋めすることがあり、数字パターンの品質を低下させることがあった。

本研究では、重畳箇所の欠損を生じさせない新たな改良手法の提案ならびに提案手法の効果を確認する。

2. 採点済答案の色分布の分析

RGB 画像の青(blue)の輝度値を横軸、赤(red)のそれを縦軸とした色空間における赤ペンで採点済の模擬答案の色分布を図1に示す。

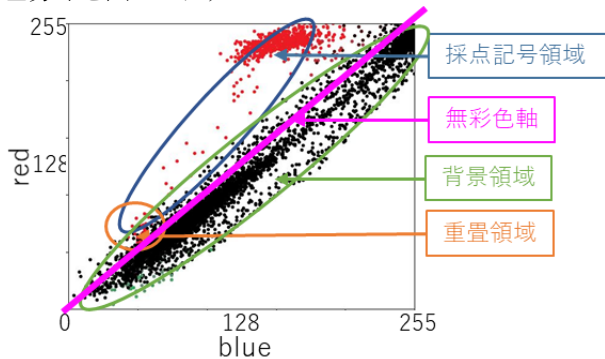


図1 赤ペン採点での色分布

背景領域である答案色の白色、枠線や解答文字の黒色は、図中の対角上の無彩色軸にクラスタが分布しているが、無彩色軸を中心に大きな分散が見られた。特に、輝度が低い黒色領域付近で採点記号の色クラスタ同士の混在が確認された。このことから、本システムの抽出処理では、採点記号領域と背景領域を線形分離していたために、重畳箇所での欠損が生じてしまうことが明らかとなった。

3. 提案手法

前述の問題を解決するために、図内の背景領域のクラスタの分散を減少させ、採点記号領域と背景領域を明確にクラスタリングし、さらに重畳領域のクラスタを背景領域と分離するように、背景領域の分散を減少させる。これを実現させるために、採点領域抽出以前にノンローカルミン(NLM)フィルタを用いたフィルタリング処理を導入することを提案する。NLM は、画像中のパターンの冗長性を利用し、ノイズ信号の類似した実測値を平均化することで高品質な画像ノイズ除去を実現するものである。[2] NLM は類

似色を用いてフィルタリングをするため、穴の欠損を発生させることなく、採点・背景・重畳の3つのクラスごとに分けられるのではないかと考えられる。

$$\hat{x}(i) = \frac{\sum_{j=\Omega} w(i,j)y(i)}{\sum_{j=\Omega} w(i,j)} \quad w(i,j) = e^{-\frac{\|x_i - x_j\|}{h}} \quad (1)$$

ここで \hat{x} : 求めた注目画素の値, x_i : 注目画素での元の値, x_j : 探索領域内の画素, h : ノイズ除去パラメータ, $w(i,j)$: 類似度である。

4. 処理結果

図1で示した模擬答案にNLMフィルタをかけた時の色分布を図2に示す。

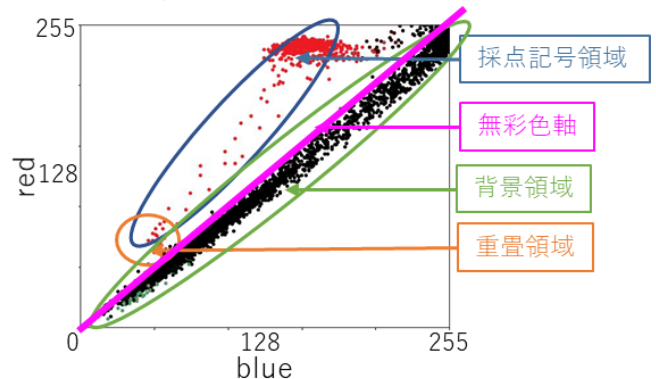


図2 赤ペン採点での色分布(NLM適用後)

答案画像にNLMを適用することで、特に無彩色領域の色クラスタの分散を小さくさせることを確認できた。

5. おわりに

提案手法であるNLMの答案画像への適用によって、特に無彩色軸を中心とした色クラスタの分散を減少させることができた。これによって、採点記号領域と黒色領域の色クラスタを明確に分離することができた。今後の課題として、NLM適用後の画像に対する線形分離での欠損発生の有無の確認、様々なペン色や線幅が異なった採点記号パターンを含む答案画像に対する提案手法の有効性の検証を進めていく。

参考文献

- [1] 西川雅清, 松尾賢一, "採点ミス発見支援システムの開発～部分点と採点記号の認識を用いた採点ミス発見手法～", 情報科学技術フォーラム(FIT:Forum on Information Technology), 2015
- [2] New methods for MRI denoising based on sparseness and self-similarity. José Manjón, Pierrick Coupé, Antonio Buades, D. Louis Collins, Montserrat Robles

謝辞: 本研究は JSPS 科研費(課題番号:20K03143)の助成を受けたものである。