

重要度を考慮した情景画像中における文字情報抽出

大平 康太[†] 宮崎 智[†] 菅谷 至寛[†] 大町 真一郎[†]

[†] 東北大学工学研究科通信工学専攻

1. はじめに

近年、画像処理技術が我々の生活の様々な場面で利用されており、その中でも文字検出や文字認識といった文字情報抽出に関しては様々なサービスに応用されている。近年、文字情報抽出はかなりの精度でできるようになっているが、ユーザに有益な情報を提供するためには、環境中に存在する文字のうちどれが重要なかを判断することが大切であり、そのような研究はほとんど行われていない。本稿では文字の重要度を計算することで文字の重要性を考慮した文字情報抽出を実現する手法について検討する。

2. 重要な文字の特徴

まず初めに重要な文字の特徴を調査するために主観評価実験を行った。被験者10人に45枚の情景画像中の重要だと思った文字を回答してもらうことで調査を行った。調査の結果、サイズが大きかったり、はっきりとした色で書かれた文字が重要な文字となる傾向があるということが明らかとなった。これは、元々文字とは人間の目に付くように作られたシンボルであるために、重要な文字ほど人間の注目がいきやすいようにデザインされているためだと考えられる。したがって、文字の重要度計算において、人間が画像中のどの部分を注目しやすいかを表す顕著性マップ[1]の利用が有効であると考えられる。

また、重要だと回答された文字は「見出し」や「店名」といったものが多かった。したがって4.の実験において使用する画像は「店名」や「見出し」といった文字が存在するものを用いる。

3. 情景画像中における重要文字情報抽出

重要文字情報抽出は図1のような流れで行う。図1における文字認識の部分は既存の手法[2]を用いるので、本稿では重要文字を検出する処理について述べる。まず初めに MSER 特徴量[3]を用いて文字候補を抽出した後に、文字候補がどの文字列に属するかを判別する。そして検出された文字領域の重要度を計算し、重要文字領域を決定する。重要度計算は機械学習のブースティングを用いる。学習では、情景画像中の文字領域から計算する特徴量を入力とする。前節の調査結果及び考察より、重要な文字ほどサイズが大きく、視覚的顕著性が高い傾向があると考えられるため、検出された文字領域における文字の大きさ、フォントの太さ(ストローク

幅)、及び[1]の顕著性マップの値を特徴量とした。ここで顕著性マップは輝度、色相、方向の3つの成分で構成されているが、今回はそれぞれの値を特徴量に設定した。

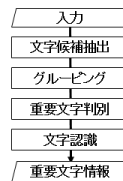


図1. 処理手順 図2. 検出成功例 図3. 検出失敗例

4. 重要文字情報検出実験

初めに重要な文字を検出できるかどうかの実験を行った。実験で使用する画像は54枚で、2.で述べたとおり店名や見出しといったような文字を重要文字とした。また、画像54枚中、44枚を訓練データ、10枚をテストデータとした。実験結果、10枚中4枚の画像で正しく重要文字が検出された。図2, 3において赤枠で囲まれた文字が重要文字、青枠で囲われた文字が今回の手法で検出された文字である。図2は成功しているが図3は重要でない文字が重要であると判断されてしまった。このことから、文字の重要度を計算する際により有効な特徴量を考える必要があると考えられる。

5. まとめ

情景画像中の重要文字情報抽出に関する手法を提案した。今後の課題としては、重要文字判別手法の改善が挙げられる。

謝辞 本研究の一部は JSPS 科研費, 16H02841, 26540077 の助成を受けた。また図 2, 3 の画像は flickr から取得し、各画像の題名と投稿アカウントは順に, "Caution", upvernoz, "商標看板", Miha TAMURA である。

参考文献

- [1] L.Itti, C.Koch, and E.Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 20, no. 11, pp 1254-1259, 1998
- [2] M. Jaderberg, A. Vedaldi, A. Zisserman, "Deep Features for Text Spotting," European Conference on Computer Vision, 2014
- [3] D.Nister, and H.Stewenius, "Linear Time Maimally Stable Extremal Regions," European Conference on Computer Vision, 2008