

レンダリング情報と役割に基づく Web 画像分類

Classification of Web Images Based on Theirs Function and Rendering Information

中平 浩二十 渡井 康行† 山崎 俊彦† 相澤 清晴†
 Koji Nakahira Yasuyuki Watai Toshihiko Yamasaki Kiyoharu Aizawa

1. まえがき

今日、ブロードバンドの普及に伴い、Web 上の画像は飛躍的に増加し続けており、視覚的な Web サイトも非常に多くなっている。従って、Web 画像は Web における情報として、非常に重要な要素となっている。しかしながら、一般の検索システムは Web 情報の解析において、画像を用いていない。画像はデータ量が大きいので、ウェブページの解析において、画像を無視し、テキスト情報のみを用いているのである。従って、Web 画像を解析し、その結果を情報として利用した新しい Web 解析やアプリケーションの提案を本研究の目的とする。Web 画像は自然画像と異なり、Web サイト内で画像が持つ意味や果たす役割といった独特の情報を持つ。我々はこの情報が Web 画像にとって特に重要な情報であると考え、この情報に基づく分類や、これを用いた種々のアプリケーションを提案してきた[1]。特徴量として画像サイズや色ヒストグラム、DCT 係数などを導入してきたが、その分類精度はまだ不十分であった。Web 画像は単独でユーザに露出するのではなく、Web ページ上に配置される形で表示される。したがって、Web ページにおける Web 画像の配置情報、レイアウト情報といったものが Web 画像やその他の視覚構成要素における重要な情報になるのではと考え、我々の研究グループではレンダリング情報を用いた、視覚的要素の抽出手法を提案してきた[2]。また Web 画像の役割に基づく分類においても有用な情報になると考え、レンダリング情報を用いた分類を行った。

2. レンダリング情報に基づく視覚的要素の抽出

我々の研究グループではレンダリング情報を用いた、DOM(Document Object Model) ツリーベースの視覚的要素の抽出手法を提案してきた[2]。HTML より DOM 木構造を抽出し、レンダリング情報を元に、視覚構成要素、及び内容構成要素の抽出を行う。抽出された結果を図1に示す。

3. レンダリング情報を用いた Web 画像の分類

3.1 意味や役割に基づく Web 画像の分類

自然画像と異なり、Web 上の画像は役割を持っている。すなわち、その画像がその Web ページでどのように使われているか、という事である。我々はこの役割に基づく分類について提案してきた[1]。"Photo", "Illustration", "Main-title", "Section-title", "Menu", "Icon", "Logo", "Advertisement" の9つの役割に基づくグループを定義し、学習分類には、学習アルゴリズムとして Support Vector Machine を使い、また特徴量として、画像に基づく特徴量と、テキストに基づく特徴量の2種類を用い、

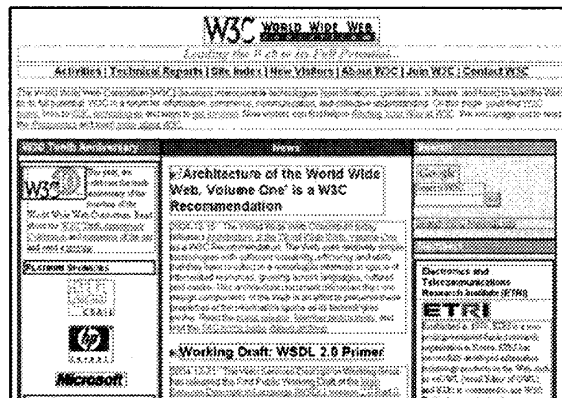


図1. Web ページからの視覚構成要素の抽出

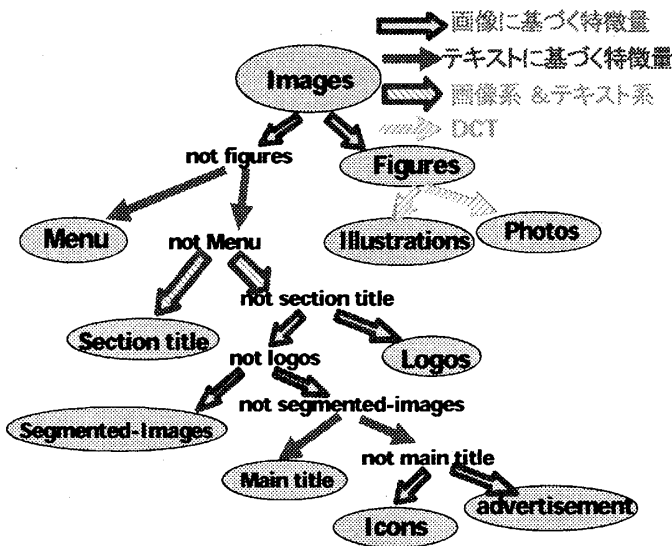


図2. Web 画像の分類

図2の分類木に沿って分類を行った。

3.2 レンダリング情報を用いた分類実験

2章で述べた手法を用いて、各画像のレンダリング時の Web ページにおける位置情報を取得し、これを特徴量として用いる。各分類段階毎で、画像の位置情報を用いた場合と用いなかった場合で精度の比較を行った。実験対象には、複数の企業サイトからランダムに選択した 3,334 ファイルを使い、そのうち 2,423 ファイルを教師データに、911 ファイルをテストデータとした。結果を表1に示す。

†東京大学大学院新領域創成科学研究科

表1 レンダリング情報を考慮した時の分類精度の比較

	今までの特徴量のみ	レンダリング情報を加えた場合
Images→ Figure, nonFigure	0.932	0.909
Figure→ Photo, Illustration	0.629	0.610
nonFigure→ Menu, nonMenu	0.787	0.756
nonMenu→ Sec_title, nonSec_title	0.685	0.715
nonS_title→ Logo, nonLogo	0.760	0.640
nonLogo→ Seg, nonSeg	0.721	0.740
nonSegment→ M_title, nonM_title	0.811	0.779
nonM_title→ Icon, Advertise	0.955	0.931

表の中列がレンダリング情報を用いなかったときの分類精度であり、右列がレンダリング情報を加えたときの分類精度である。“not Menu”から“Section title”とそれ以外を分類するとき、及び“not Logos”から“Segmented Image”とそれ以外を分類するときは、レンダリング情報に基づく位置情報を用いることで精度の向上が確認できた。しかしながら、それ以外の分類では、精度はやや悪化するという結果になった。

3.3 考察

“Section title”は、コンテンツが幾つかのサブコンテンツに分けられるときのタイトルの役割を果たす画像であり、たいていの場合においてx軸座標が揃えられている。従ってレンダリング情報に基づく位置情報がその分類に寄与したと言える。また、“Segmented Image”は複数の画像が集まって一つの画像を構成するものであり、互いの位置が非常に近い事が多い。従って位置情報を特徴量に追加することにより精度が向上したと考えられる。しかしながら、その他の分類においては精度がやや下がっている。これはレンダリング情報に基づく位置情報の分類への寄与が小さいことと、位置情報を追加したことにより、特徴量空間の次元数が増加し、その結果学習分類の精度に影響を与えたためと考えられる。

4. まとめ

レンダリング情報は、人間の視覚に即する情報であり、Web ページ上の画像やその他の視覚要素に対し、非常に重要な情報となる。本論文で示した実験により、Web 画像の分類において、レンダリング情報に基づく位置情報を特徴量に加えることで、“Section title”と“Segmented image”の分類精度に寄与する事が確認された。

今後は他の画像との相対的な位置関係や、テキスト領域との関係といった情報を解析することにより、他の分類精度の向上に繋がらないか検討を行う予定である。

5. 参考文献

- [1] K.Nakahira, T.Yamasaki, K.Aizawa “Accuracy enhancement of function-oriented Web image classification,” International World Wide Web Conference (WWW2005) pp.950-951, May.10-14, 2005, Chiba, Japan.
- [2] 渡井康行, 上野智史, 中平浩二, 山崎俊彦, 相澤清晴 “レンダリング情報に基づく Web ページの視覚的構造抽出” 電子情報通信学会 総合大会, D-12-69 Mar. 21-24 2005, Osaka.