

顕著性マップ再構築のための文字に対する誘目性 Visual Attractiveness of Character for Improvement of Saliency Map.

長嶋 太良[†]
Taira Nagashima

高野 博史[†]
Hironobu Takano

中村 清実[†]
Kiyomi Nakamura

1. はじめに

近年、従来の広告形態における一方向的な情報提供に対して、ユーザーが能動的に行動するインタラクティブな広告形態は、より深い印象をユーザーに与えることができるとして注目が集まっている。そのため、人間の視覚的注意が入力インタフェースとして注目されており、視覚的注意がどの領域に対して働くかが推定できれば、有用な情報となると考えられる。この人間の視覚的注意を推定する方法として、視覚的顕著性マップという計算モデルがある[1]。しかし、顕著性マップは主にボトムアップ型の視覚的注意を推定しているため、トップダウン型の視覚的注意が働きやすい広告などの画像においては、その推定精度が低下してしまう問題が挙げられる。現代の広告は多種多様だが、多くの広告に共通して挙げられるのは、文字という特徴が含まれていることである。従って、本研究では様々な分野で応用可能な顕著性マップを再構築するために、文字が持つ誘目性について検証を行った。

2. 視覚的注意と文字

人間の注意については様々な定義があり、一意では決まらない。そこで、本研究では視覚認知における選択的注意のことを視覚的注意と定義する。この視覚的注意にはボトムアップ型の注意とトップダウン型の注意がある。ボトムアップ型の視覚的注意は、色やエッジ、輝度変化などの物理的な特徴に基づいて受動的に働く。それに対してトップダウン型の視覚的注意は、個人の興味や関心、経験などに基づいて能動的に働く。このような注意が働く静止画像や動画像においては、顕著性マップは精度よく注意推定をおこなえない。

一般的に文字という特徴は人の視覚的な注意を引くとされている[2]。また、認知心理学などの分野では、意味がわかる文字とそうでないときの文字で、左右の視野差が異なるといった報告もある[3]。しかし、選択的な注意が働いている状況下で、画像中に存在する文字に対してどのように注意が行われるかについての報告は少ない。Wang らは、アルファベットや漢字を使用して、文字についての注意を検証している[4]。しかし、普段使用している言語体系が他の言語と大きく異なる日本人被験者においては、Wang らの結果とは異なる結果が得られる可能性がある。そこで、本研究では、ひらがなや記号などを画像中に挿入した画像を被験者に対して呈示し、そのときの視線計測結果からどのように文字へ注意が向けられるかを分析した。

3. 実験内容

本研究では、作成した視覚刺激を被験者に呈示し、そのときの視線を計測した。被験者は、本研究の概要を知らない健康な大学生 10 名である。実験は 3 種類あり、実験 1 では、意味がないひらがな 2 文字の文字列と記号が同時に画像中に挿入されている 50 枚のターゲット画像と、何も挿入されていないダミー画像 50 枚を被験者に対して呈示した。実験 2 では、ひらがなの文字列をタイ文字の文字列に変えた 50 枚のターゲット画像と、実験 1 と同じダミー画像を被験者に対して呈示した。実験 3 では、意味のないひらがな 2 文字の文字列とタイ語の文字列を実験 1、実験 2 と同様に挿入したターゲット画像 50 枚とダミー画像 50 枚を被験者に呈示した。被験者には特別なタスクは与えず、自由に注視してもらうよう教示した。1 セッションは 25 枚のターゲット画像と 25 枚のダミー画像で構成され、被験者 1 人あたり 2 セッション実験を行った。視覚刺激はランダムに呈示し、呈示時間は 2 秒とした。なお、10 名の被験者は、タイ語に対しての知識はなかった。

呈示した視覚刺激は、IAPS (International Affective Picture System) の画像データベースより選出した。このデータベースには、Valence (誘発性) と Arousal (覚醒度)、そして Dominance (支配的主観) という次元でスコアがつけられている。本研究では、Valence のスコアが 3~7、Arousal のスコアが 1~5 という基準を設け、これをもとに興味が変わりにくいとされる画像を 100 枚選んだ。文字列等を挿入した画像には、文字や記号を挿入することで生じる可能性がある「違和感」を軽減させるため、挿入領域に対してぼかしをいれた。その後、既存の顕著性マップを用いて、ターゲット画像を評価し、顕著性が高いとされる上位 25% の領域を算出した。この領域内に、挿入した文字列等が含まれないものを実験の視覚刺激として利用した。文字列等を画像中に挿入する位置はランダムである。

4. 解析方法と解析結果

本研究では、30Hz でサンプリングされる Tobii 社の Eye-tracker X2-30 を用いて視線を計測した。この計測から、視線が停留したところを注視と定義した。停留の定義は、サンプリングされた視線が 50 ピクセル内に 100ms 以上留まったときとした。次に、挿入領域の最小矩形を算出し、各被験者のキャリブレーション誤差をその矩形の縦方向、横方向に考慮した。この矩形を最終的な文字列等の挿入領域とし、この領域内に視線が停留したときに、文字列や記号を注視したと判断した。

図 1~3 に、実験 1、実験 2、実験 3 の結果をそれぞれ示す。図 1~3 の横軸は、文字列や記号に初めて視線が停留したときの順序を示している。縦軸は、視線の停留割合を示している。図 1 より、ひらがなと記号の視線の総停留割合に顕著な差が見られる。この差について、対応ありの t

[†]富山県立大学院 工学研究科 知能デザイン工学専攻
Graduate School of Engineering, Toyama Prefectural
University

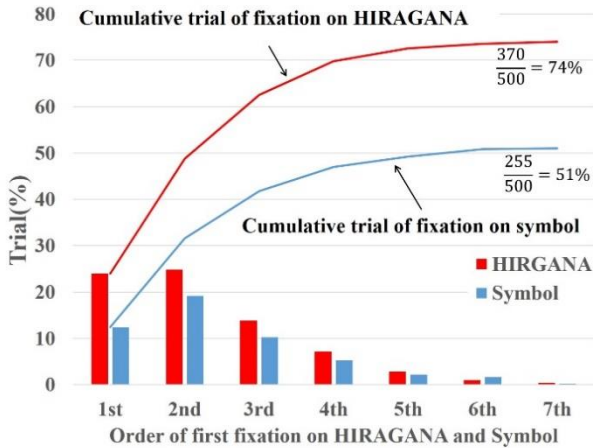


図1：ひらがなと記号に対する視線停留割合

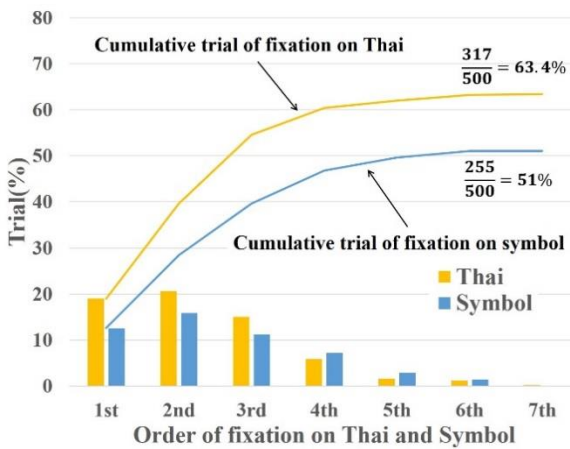


図2：タイ文字と記号に対する視線停留割合

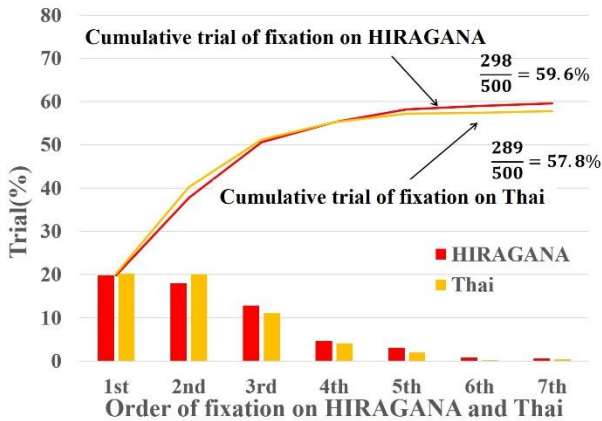


図3：ひらがなとタイ文字に対する視線停留割合

検定を行った結果、 $p < 0.001$ で有意差があった。次に、図2よりひらがなと記号ほど顕著ではないが、タイ語の文字列と記号との間の視線停留割合に差が見られた。これに対しても対応ありの t 検定を行った結果、 $p < 0.001$ で有意差があった。また、実験1と同様に第一停留において最も顕著な停留差が生じるといった注視傾向が見られた。しかし、ひらがなとタイ文字の結果については、これまでの実験結果とは異なる結果になった。図3より、ひらがなとタイ文字の視線停留割合にほとんど差が現れなかった。各停留順序間に対しても、日本語とタイ語では同様な傾向が見られ、視線停留割合間の対応ありの t 検定では $p > 0.657$ で有意

差はなかった。

5. 考察

本研究では、画像中の文字に対する人間の視覚的注意について分析するために、3種類の実験を行った。その結果、実験1と実験2では、挿入した両領域に物理的な顕著性の差は存在しないにもかかわらず、文字特徴と記号への視線停留割合の間に有意な差があった。このことから、少なくとも記号と比較したときに、文字の方が注意を引きやすい特徴を有していることが示された。次に、実験1で使用した日本語には、単語としての意味が存在せず、また、実験2で使用したタイ語の文字列は、被験者にとって、何の予備知識もない文字であったにもかかわらず、記号との停留割合間に有意な差があった。このことから、その注意は文字を読もうとするようなものではなく、文字自体に特殊な顕著性が存在する可能性が考えられる。このため、実験3においては、視線停留割合に差が生まれず、日本語とタイ語両方に注意が向けられたと考えられる。以上のことから、文字には、例えば意味がなく初めてみるような文字であっても、その文字という特徴自体に注意を引かせる性質がある可能性が高いと考えられる。

一方で、Wangらは親近性の高い文字に対して注意が向きやすいと報告しており、本研究の実験3の結果とは異なる結果を報告している[4]。だが、今回使用したひらがなには単語としての意味はなく、また日本人の普段の生活環境に着目すると、ひらがなだけを用いて情報を伝達することは希で、ほとんどの場合、漢字と仮名を併用して用いている。従って、親近性の低いタイ文字と比較しても有意な差が現れなかったのではないかと考えられる。

6. まとめと今後の課題

本研究では、様々な分野で応用可能な顕著性マップを再構築するために、文字の誘目性について検証を行った。その結果、文字には人の注意を引き付ける特殊な誘目性がある可能性が高いことがわかった。その誘目性は、例えば被験者が初めて見るような文字であっても有効であり、文字の特徴それ自体に注意が引き付けられている可能性が示唆された。

今後は、日本人における文字の親近性についての検証と、今回得られた実験結果を基に、文字の視覚的顕著性を従来の顕著性マップに考慮したモデルを開発していく。

参考文献

- [1] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.20, No.11, pp.1254-1259, 1998.
- [2] M. Cerf, E. P. Frady, and C. Koch, "Faces and text attract gaze independent of the task: Experimental data and computer model," *Journal of Vision*, Vol.9, No.12, pp.1-15, 2009.
- [3] K. Yoshizaki and T. Hatta, "Shift of visual field advantage by learning experience of foreign words," *Neuropsychologia*, Vol.25, No.3, pp.589-592, 1987.
- [4] H. C. Wang and M. Pomplun, "The attraction of visual attention to texts in real-world scenes," *Journal of Vision*, Vol. 12, No. 6, pp.1-17, 2012.