

## 視覚認知に基づくWebインタフェースデザインに関する考察

吉田 香, Mario Koeppen  
九州工業大学 情報工学部

概要：Webサイトやアプリケーションにおいて、ユーザの目標を正しく達成するためには、見た目の華やかさだけでなく、ユーザの視覚的な特性を考慮した適切なインタフェースデザインがなされなければならない。本稿では、ゲシュタルト則に基づく画像処理研究について調査した内容を報告するとともに、視覚認知に基づくWebインタフェースデザインについて考察する。

### A Discussion of Web Interface Design based on Gestalt Theory

Kaori Yoshida and Mario Koeppen  
Kyushu Institute of Technology

Abstract: When we search information on a web site or by using a specific application, the interface design should follow our perceptual characteristics to satisfy our needs. In this paper, we provide a critical review of the contributions of Gestalt laws in visual perception and discuss about a method to analyze web page design by considering the displayed web page as an image.

#### 1. はじめに

近年Webページのデザインが多彩になり、ポップアップ画像やフラッシュ動画等を用いた見た目が華やかなページが増えてきている。これまでのWebページ解析では、HTML文書をテキスト処理することにより構造を解析[3][4]したり、DOMに基づいてレイアウトを解析[5]することが多かった。しかし、これらの方法では、実際にユーザが閲覧しているブラウザ上における見た目通りのレイアウトを解析するのは困難である。そこで、Webページを画像として取り扱うことにより、ブラウザ上の見た目通りのレイアウトを解析する手法[6][7][8]が提案されている。これらの先行研究では、Webページの自動分類や類似検索を目的とし、文字や画像の画像特徴や配置を利用して、実験および評価を行っている。

一方、視覚認知心理における研究のひとつとしてゲシュタルト則に基づく画像処理研究がある。これらの研究で根底となる考え方は、部分の総和としてとらえられない合体構造に備わる特有の全体的構造に着目するという点である。例えば、知覚現象におけるプレグナンツの原理[1]では、全体として形態的にもっとも優れ、もっとも秩序だっており、もっとも簡潔なまとまりをなそうとする傾向があることが指

摘されている。この考え方をWebインタフェースデザインに適用し、Webページを画像としてとらえ、ゲシュタルト則に基づく画像解析的なアプローチにより、Webページの解析を行う。

本稿では、ゲシュタルト則に基づく画像処理研究について調査した内容を報告するとともに、視覚認知に基づくWebインタフェースデザインについて考察する。

## 2. ゲシュタルト則に基づく画像処理研究

### 2.1 ゲシュタルト則とは

ドイツ語の「ゲシュタルト (die Gestalt)」には、形状、骨組み、構造などの意味がある。ゲシュタルト則とは、近いものをグループとして知覚する（近接の要因）、同じ性質のものをグループとして知覚する（類同の要因）、閉じた図形・領域を見出してグループとして知覚する（閉合の要因）、いくつかの曲線になり得るなめらかな曲線として知覚する（よい連続の要因）などから、図形・領域の形態に存在する単純な規則性や意味の法則を見いだそうとするゲシュタルト心理学的な機能の総称を指す。また、上記の要因をまとめて群化の要因と呼ぶこともある。

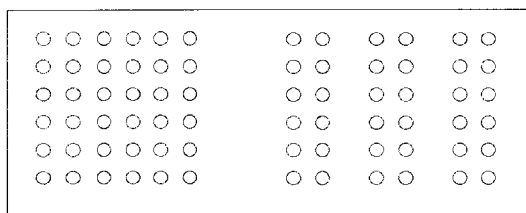


図1 近接の要因

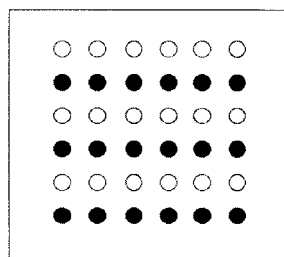


図2 類同の要因

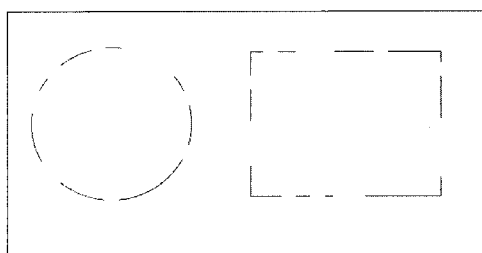


図3 閉合の要因

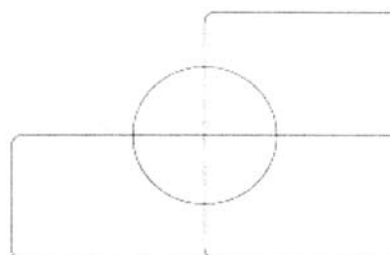


図4 よい連続の要因

ここで、ゲシュタルト則とは、知覚における現象を説明しているだけであり、なぜそのような現象が起こるのかについては言及していないことに注意したい。つまり、画像解析的にゲシュタルト則を見出すことは非常に困難であるといえる。

## 2. 2 先行研究

本節では、ゲシュタルト則に関連した画像処理研究論文について調査した結果を述べる。調査は、SCOPUS[2]により提供されている1980年から現在に至るまでの研究論文約200件を対象に行った。これらの研究論文は、画像の構成要素解析、初期レベルの知覚メカニズム、部分画像と全体画像の統計的特徴に大きく分類することができた。ここでは、本稿の主題である、Webインタフェースデザインに適用可能であると考えられる20件の研究論文を対象に、特徴毎に要約する。

まず、手法としてゲシュタルト則を用いている研究論文について述べる。画像処理研究であることから明らかなように、すべての論文は画像処理過程におけるゲシュタルト則に関連したアルゴリズム[12][13][15][18][19][25][27]、もしくは処理の枠組み[9][14][24][26][29]を示している。また、文献[10][11][14][16][20]はゲシュタルト則に従う知覚の生理モデルに基づく画像処理モデルを提案しており、文献[18][22]は実世界アプリケーションへの適用を試みている。さらに、文献[9][26]では、画像検索の特徴としての適用に言及している。一般的に、これらの論文は、単一のゲシュタルト則に着目している。ほとんどの論文は、プレグナンツの原理に従うまとまりを検出することを目的としていたが、対称性の検出[10]や閉合の検出[13]を目的とした論文もみられた。

次に、ゲシュタルト則を画像の構造解析に応用した研究論文について述べる。最も多かったのは、エッジ検出に適用した研究論文[12][13][18][20][22][25][27]であり、領域分割に適用した研究論文[20][21][26]はあまり多くなかった。領域分割に適用した研究論文は、相互に研究論文を参照しており、共通の参考文献を引用していたことも特徴であった。そして、あまり数は多くないが、視覚プロセスの概念としてゲシュタルト則を用いた研究論文も見られた[23][27]。また、初期知覚と実際の内容の間のセマンティックギャップに着目した研究論文の見られた[29]。さらに、1件だけであるが、ゲシュタルト則とガボール関数を関連づけ、境界領域の検出を目的とした研究論文[28]も見られた。

最後に、論文中に確率に関する知見を述べている文献[12][16][15]、確率に関わる参考文献を引用している文献[21][28]など、多くの研究論文が確率の概念に関連していたことを述べておく。例えば、ゲシュタルト則に従う知覚とある視覚現象が起こる確率とを関連づけた文献[15]が見られる。しかし、文献[24]で主張されているように、このようなアプローチは適切ではないと考える。実際、プレグナンツの原理に従うまとまりとしての構成要素の数は、問題とならない。

これらの調査結果より、単一のゲシュタルト則であれば画像解析的アプローチで画像検索等に適用可能な特徴を抽出できる可能性があること、ゲシュタルト則をエッジ検出や領域分割に適用し構造を解析できる可能性があること、プレグナンツ

の原理に従うまとめりとしての構成要素は画像中に複数あってもよいこと、などがわかった。このことを、Webインタフェースデザインへ適用することを検討する。

### 3. Webインタフェースデザインへの適用

#### 3.1 Webページ構造解析

まず最初に、Webページの構造解析について述べる。一般に広く用いられているのは、Webページのソースファイルからタグ等の情報より意味を持つ領域を得て構造を解析する手法[3][4][5]である。つまり、HTML、CSS、DOM等の情報より、それぞれの構造の意味（例えば、タイトル、メニュー、ロゴ等）を解析することができる。例として、図5に示すWebページをコメカミエディット[3]を用いて構造解析した結果を図6に示す。また、得られた構造を図5中に追加した結果を図7に示す。しかし、我々が取得したい構造は、図8に示すような見た目の構造である。

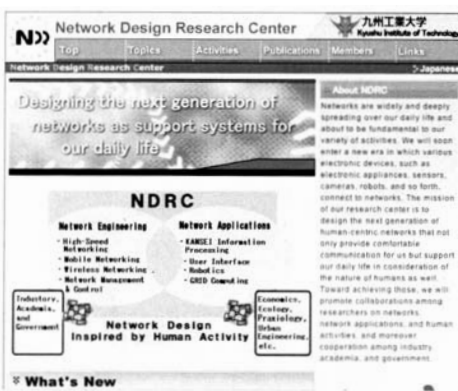


図5 サンプルWebページ

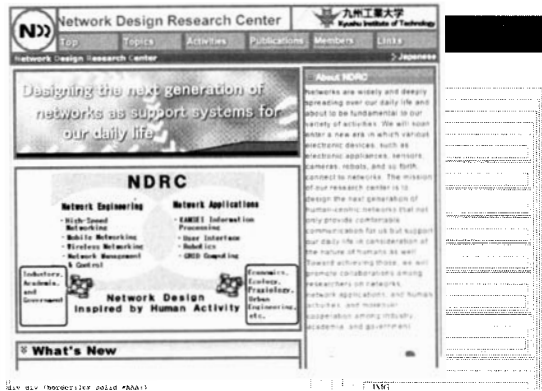


図6 ソースファイルからの構造解析例

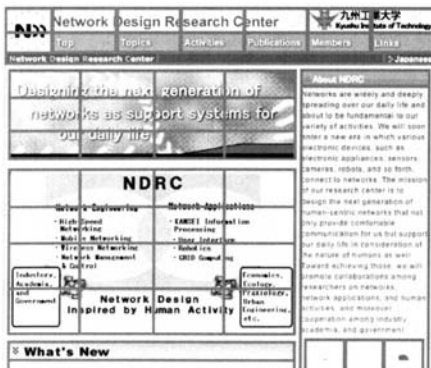


図7 ソースファイルから得られた構造

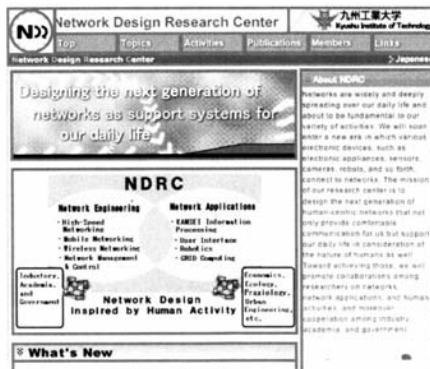


図8 取得したい構造

Webページを画像としてとらえ、画像解析的に構造を解析したい場合、(1) 意味を持つ領域を得ることが困難、(2) 背景と意味を持つ領域を区別することが困難、(3) 文字列が含まれているかどうかを判断することが困難、という問題が考えられる。(1) および(2)の問題については、画像解析的手法と従来の構造解析的手法を組み合わせることにより解決できる可能性がある。(3)の問題については、OCRなどに代表される文字列認識エンジンを用いることにより解決できる可能性がある。しかし(3)の問題については、見た目の構造のみを取得できればよい場合は、文字列が含まれているかどうかを判断する必要がないため、本研究では扱わないこととする。

### 3.2 ゲシュタルト則の発見

Webページを画像としてとらえ、画像解析的に構造を解析するために、前処理としてぼかし処理およびエッジ検出処理を行う。ぼかし処理は、文字列を文字列としてではなくひとつの領域として認識するために、文字の大きさにあわせてぼかし範囲を設定するのが望ましい。エッジ検出は、文字列領域および図形領域の分割を行うためであり、一般的な領域分割法を用いてもよい。例として、図5の画像に対し、最も多くあらわれるフォントサイズである12ピクセルをぼかし範囲としてぼかし処理を行い、一般的なガウシアンフィルタによるエッジ検出を行った結果を図9に示す。図9では、特にプレグナンツの原理より、近接の要因(図1)、類同の要因(図2)、閉合の要因(図3)を確認することができる。



図9 サンプルWebページのエッジ検出結果

### 3. 3 ゲシュタルト則の要因と定量化

ゲシュタルト則の要因を定量化するためには、抽出した領域が類似した形状であるかどうかを判断する必要がある。形状特徴は、距離、面積、重心等の幾何学的特徴、および、画像の濃度分布の平均濃度、分散値、カラー等の統計的特徴から判断できる可能性がある。しかし、例えば、近接の要因を定量化しようとする場合は、どの程度近ければまとまって見えるのか、その閾値が必要となる。同様に、類同の要因を定量化しようとする場合は、その程度類似していれば同じまとまりに見えるのか、類似度の閾値が必要となる。このような閾値は、主観評価実験により求めるのが主流であるが、汎用的な画像に対して一律に用いることができる閾値は得られていないのが現状である。

我々は、ゲシュタルト則に従う構造を持つロゴを用いて、構造を崩すことによってどの程度構造が崩れたことを知覚できるかという主観評価実験を行った[30]。主観評価実験では、構造を崩すパラメタだけでなく、構造を崩さないパラメタも準備し、それぞれにパラメタの値を変化させ、ロゴが変更されたと知覚するかどうかを調査した。主観評価実験の結果、すべてのパラメタにおいて、値の変化が大きいほどロゴが変更されたと知覚する割合が高いこと、構造を崩す場合は構造を崩さない場合に比べて値の変化が小さくてもロゴが変更されたと知覚すること、構造を崩す場合だけでなく、構造を崩さない場合でもロゴが変更されたと知覚することがあることを確認した。しかし、その閾値に関しては検証することができなかった。

## 4. おわりに

ゲシュタルト則に基づく画像処理研究について調査した内容を報告するとともに、Webインタフェースデザインへの適用可能性について考察した。ゲシュタルト則に基づく画像処理研究の調査結果より、画像解析的にプレグナンツの原理に従うまとまりを検出し、構造を解析できる可能性があることを述べた。また、Webページを画像としてとらえ、ゲシュタルト則に基づく構造のまとまりを見た目の構造として抽出可能かどうかを検討した。画像解析的アプローチによりゲシュタルト則の要因を定量化することができれば、ユーザの視覚的特性を考慮したインタフェースデザインに役立つと考えられる。今後も主観評価実験を行い、汎用的に利用可能な法則を抽出したいと考えている。

## 謝辞

本研究の一部は日本学術振興会科学研究費補助金（特別研究員奨励費18.06094）の助成による研究成果である。

## 参考文献

- [1] Wertheimer: Gestalt theory, Social Research, 11, 1924.
- [2] SCOPUS: <http://www.scopus.com/>
- [3] コメカミエディット : <http://la.ma.la/misc/tmp/ledit/>
- [4] 橋本, 五十嵐 : レイアウトによるWEBページ検索, インタラクシオン2004, 情報処理学会ヒューマンインタフェース研究会, 2004.
- [5] 辰巳, 森口, 旭 : DOMの画面レイアウトに基づくWebページのセグメンテーション手法, 電子情報通信学会第8回Webインテリジェンスとインタラクシオン研究会, 2007.
- [6] 渡井, 中平, 山崎, 相澤 : 視覚的特徴に基づくWebページ閲覧履歴検索, 情報科学技術フォーラムFIT2006, I-066, 2006.
- [7] 石川, 細川, 高橋 : 色とその配置位置に基づいた視覚的印象によるWebページ検索手法の実現方式, 電子情報通信学会データ工学ワークショップDEWS2005, 4B-i6, 2005.
- [8] 三橋, 山口, 高間 : 視覚的類似性に基づくwebページ検索手法の提案, 人工知能学会全国大会第17回論文集, 1C4-01, 2003.
- [9] S. Alwis and J. Austin. : An integrated framework for trademark image retrieval using gestalt features and cmm neural network. In Image Processing And Its Applications, 1999. Seventh International Conference on, volume 1, pages 290–295, 1999.
- [10] A. Assadi, S. Palmer, and H. Eghbalnia. : Learning gestalt of surfaces in natural scenes. In Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop, pages 380–389, 1999.
- [11] Ohad Ben-Shahar and Steven W. Zucker. : Continuation of general 2d visual features: Dual harmonic models and computational inference. In Computer Vision, Tenth IEEE International Conference on, volume 2, pages 1643–1650, 2005.
- [12] F. Cao. : Good continuations in digital image level lines. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, volume 1, pages 440–448, 2003.
- [13] Yung-Sheng Chen and Tsay-Der Lin. : Simulation of closure process for line patterns. In Pattern Recognition, 1996., Proceedings of the 13th International Conference on, volume 2, pages 215–219, 1996.
- [14] R.W. Connors and C.T. Ng. : Developing a quantitative model of human preattentive vision. IEEE Transactions on Systems, Man and Cybernetics, 19(6): 1384–1407, 1989.
- [15] A. Desolneux, L. Moisan, and J.-M. More. : A grouping principle and four applications. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 25(4):508–513, 2003.
- [16] Cheng en Guo, Song-Chun Zhu, and Ying Nian Wu. : Towards a mathematical theory of primal sketch and sketchability. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, pages 1228–1235, 2003.

- [17] Gerald M. Edelman and Giulio Tonini. : A Universe of Consciousness: How Matter Becomes Imagination. Basic Books, 2001.
- [18] P. Gamba, M. Lilla, and A. Mecocci. : Extraction of discontinuous chains of symbols by means of perceptual grouping. In Image Processing, 1997. Proceedings., International Conference on, volume 2, pages 422–425, 1997.
- [19] J. Huart and P. Bertolino. : Similarity-based and perception-based image segmentation. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 3, pages 1148–1151, 2005.
- [20] A. Rosenfeld. : Computer vision: a source of models for biological visual processes? Biomedical Engineering, IEEE Transactions
- [21] T.R. Reed and H. Wechsler. : Segmentation of textured images and gestalt organization using spatial/spatial-frequency representations. Pattern Analysis and Machine Intel ligenge, IEEE Transactions on, 12(1):1–12, 1990.
- [22] F. Saitoh. : Boundary extraction of brightness unevenness on lcd display using genetic algorithm based on perceptive grouping factors. In Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, volume 2, pages 308–312, 1999.
- [23] S. Sarkar. : Learning to form large groups of salient image features. In Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on, pages 780–786, 1998.
- [24] Sudeep Sarkar. : An introduction to perceptual organization. In Integration of Knowledge Intensive Multi-Agent Systems, 2003. International Conference on, pages 330–335, 2003.
- [25] M. Teranishi, N. Ohnishi, and N. Sugie. : Subjective contours are useful for extracting contours with very weak contrasts. In Neural Networks, 1993. IJCNN '93 Nagoya. Proceedings of 1993 International Joint Conference on, volume 1, pages 139–142, 1993.
- [26] A. Wardhani and R. Gonzalez. : Perceptual grouping of natural images for cbir. In Signal Processing and Its Applications, 1999. ISSPA99. Proceedings of the Fifth International Symposium on, volume 2, pages 923–926, 1999.
- [27] Song Wang, Toshiro Kubota, Jeffrey Mark Siskind, and Jun Wang. : Salient closed boundary extraction with ratio contour. Pattern Analysis and Machine Intel ligenge, IEEE Transactions on, 27(4):546–561, 2005.
- [28] Song-Chun Zhu. : Embedding gestalt laws in markov random fields. Pattern Analysis and Machine Intel ligenge, IEEE Transactions on, 21(11):1170–1187, 1999.
- [29] N. Zlatoff, B. Tellez, and A. Baskurt. : Image understanding and scene models: a generic framework integrating domain knowledge and gestalt theory. In Image Processing, 2004. ICIP '04. 2004 International Conference on, volume 4, pages 2355–2358, 2004.
- [30] Y. Kinoshita, M. Koeppen, and K. Yoshida. : Perception of image similarity considering Gestalt theory, International Conference on Soft Computing and Human Sciences, 2007.