

## 色情報の付加によるテキストの可読性向上の試み

4N-6

内田友幸 田中英彦 東京大学 工学部

## 1 はじめに

近年、ネットワークや大規模文書データベースの普及などによって電子化された自然言語文書が大量に流通するようになってきた。しかし、これらの自然言語文書に対するユーザーインターフェースはハイパーテキストによる構造化、他のメディアとの統合化がなされるなどの工夫は施されているものの、文書そのものは依然として白黒であり、従来の紙メディアと同様の表現が主流である。

カラープリンタ、カラーディスプレイが普及し、色を利用したユーザーインターフェースが主流になってきている現在、このような白黒の文書ではなく色を利用する文書という選択肢も選び得ると考えている。適切に文書を彩色する事でより表現力を豊かにしたり、より読みやすい文書にできる可能性は充分にあると考えられる。

本稿では自然言語文書に色情報を付加する事による可読性の向上を検討し、その可能性を提言する。また予備実験として心理実験を行なったのでその結果も併せて報告する。

## 2 彩色による利点と欠点

## 2.1 期待できる彩色の効果

色を付けることによる心理的效果は以下のものが期待できる。

- 文書中要となる要素に対するアクセスを速く行なえる
- 色彩に特定の意味を持たせることにより、文字、単語を読まなくても色彩の分布を見るだけで概要を把握できる
- 意味構造に色彩構造を関連させて組み込むことで意味が把握し易くなる
- 色彩は感情に訴える心理的效果があるため、これを利用して文書の表現力を増す事ができる

これらの効果の結果、読者は文書の概要を速く掴めたり、文書の内容をより理解できたり、より文書を楽しむ事ができるようになると期待される。具体的には、文書検索の際、キーワード周りの重要と思われる部分を素早く読める。複数の人数の会話文に対して発言者を色によって表現すれば会話の流れを自然に汲み取れる。接続関係などに対して特定の彩色を規定しておけば文の流れの概要を素早く把握できる。小説などで色鮮や

かな情景描写を行なう際により効果的な表現ができる。などの応用が考えられる。

## 2.2 彩色による弊害

しかし、彩色による影響には好ましくない点もある。以下に、彩色による弊害を列挙する。

- 彩色が偏っていたり、乱用されている場合、意図した効果が得られない
- 読み手が彩色に慣れていなかったり、彩色が目的としている表現に対して読み手が興味がない場合、意図した効果が得られなかったり可読性を損なったりする
- 色表現はデバイスによって異なる事があるので、色に依存した表記にすると、意図した効果が得られない危険性がある
- 色表現による心理的效果は文化によって異なるなど個人差が大きいので読む人によっては意図した効果が得られない危険性がある
- 誤った彩色方法によって元の文書の表現内容を損なう危険性がある

このような弊害に対しては、読み手側の文書に対するスタンス、色に対する嗜好によってカスタマイズを施すなどの工夫を含めた対応が必要になってくると考えられる。

## 3 彩色の方法

## 3.1 彩色表現上の制約

有効な彩色を行なうには、彩色表現上の制約と効果の期待できる彩色のポイントの2点を考える必要がある。多彩な自由度の中から文字色、文字の背景色の設定を仮定すると彩色表現上の制約としては次のものがある。

- 混同を避けるため、利用する色間に適切な距離が必要である
- 文字色と背景色の組合せによっては色が異なって見えてしまう事があるのでそのような事の起こらない組合せを利用する
- 色にはそれぞれ心理的影響を与える効果があるので、色に対する定義をその影響に合ったものにする
- 読み手に不快感などの心理的悪影響を及ぼさないような配色にする
- 一定の系統に色が偏ったり、色が多過ぎたりしないようにする

続いて効果的な彩色のポイントは次のようなものがある。

- 読み手が着目しがちな単語（検索のキーワード等）

- 意味内容が周りと異なるポイント (例や会話文)
- 節などの意味単位毎の区切り
- 文の構造 (接続関係やかかり受け)
- 色を連想させる意味表現 (赤信号や青空など)
- 書き手が特別な意味を持たせたいポイント

これらのポイントに対して適切な彩色を施す事で前述した有効性を引き出す。

## 4 予備実験

### 4.1 実験の概要

文書の彩色によって実際にどの程度効果があるのかを調べるために予備実験を行なった。この実験では、より速く文書の概要を掴めるという効果の可能性に着目し、それを心理実験により評価した。心理実験では被験者に白黒の新聞記事と彩色した新聞記事それぞれについて一定の基準で分類を行なってもらい、それにかかった時間を測定した。この時間が短ければ彩色の効果があったものとする。

### 4.2 実験方法

この実験では、被験者が今後の景気の動向に興味があり、「景気」という単語で新聞記事データベースからいくつかの新聞記事を得た、という状況を設定している。そして「景気の今後の予想に具体的に役に立つと思われる」ものとそうでないものの2つになるべく速く分類してもらった。

分類する新聞記事は「景気」という単語を含むものを、内政、経済、文化、外国情勢という分野の中から20記事を無作為抽出した。これを2つに分けて記事群を2群作り、それぞれについて色付きと白黒のものを用意した。

彩色は以下の7種類のポイントについて行なった。

表1. 彩色のポイントと指定色

ポイント	背景色	文字色
記事のタイトル	薄い青	-
会話文、発言内容	薄い黄色	-
キーワード	-	赤
述語中の主要単語	-	紺
人名以外の固有名詞	-	水色
数詞	-	紫
人名	-	緑

ここで「述語中の主要単語」とは、文末や節末に存在する、動詞、動詞「する」に繋がるサ変動詞性名詞、形容詞、形容動詞の終止形、体言止めの名詞、助動詞に繋がる名詞としている。

また、これらの色づけは構文解析ツール等を組合せて半自動で生成し、細部を手動で調整したものを用いた。

心理実験に当たってはまず、色付きの新聞記事のサンプルを本番同様に分類してもらい、続いて実験を行った。実験は各群よりそれぞれ色付き、色なしの2つを選び、それぞれについて分類にかかった時間を計測した。

## 4.3 実験結果

記事群1の白黒と記事群2の彩色したものを計った人が7名、その逆の組合せが7名の計14名の結果を図1に示す。

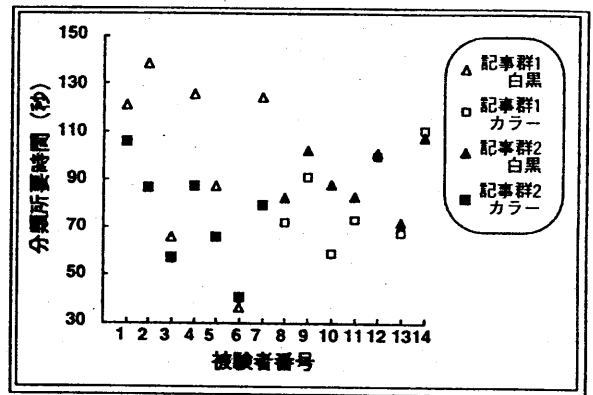


図1: 被験者からの回答の分布

各条件の計測時間の平均を表2に示す。

表2. 平均分類所要時間 (秒)

	記事群1	記事群2
白黒	99.5	81.6
カラー	90.6	74.5

これを見ると記事群1、2のそれぞれにおいて彩色したものが平均分類所要時間が短くなっている。このことから彩色する事により新聞記事の分類を速くする効果を得られる可能性があると言える。

また、被験者の感想を実験後に聞いたので、その主要なものを以下に示す。

- 色付きの場合飛ばし読みをする場合にポイントが分かって良い
- 色があると読み方が変わる (内容にまで目が行ってしまったりする)
- 色を使いこなすには慣れが必要だ
- 白黒の文書に慣れているので色があるとちらついて読みにくい

色を全く気にしなかった人から色のおかげで読むのが楽だったという人まで人それぞれの回答があり、彩色文書への対し方はかなり個人差がある様である。

## 5 考察

予備実験の結果から彩色によって文書をより有効に利用できる可能性がある事がわかった。

情報検索が発達したとしても興味のある文書が大量に残るといった状況は残ると考えられる。そのような際にこのような手法は有効であると考えられる。

このように彩色の可能性は非常に将来性のあるものと考えており、今後より有効な彩色の規則の導出について検討していきたい。

## 参考文献

- [1] 松岡 武: "色彩とパーソナリティ" 金子書房
- [2] (財) 新世代コンピュータ技術開発機構: "TRIE 構造日本語辞書とその形態素分類体系の概要" Distributed Free Software