

歴史資料における解説表示設定のための領域抽出法の検討

西郷 智気[†] 徳永 幸生[†] 杉山 精[†] 安達 文夫[‡]

[†]芝浦工業大学 [‡]国立歴史民俗博物館

1. はじめに

博物館で所蔵する資料の中には、屏風や絵図のような大型の資料がある。これらの資料には、対象物が非常に細かく描かれたり、文字が小さく記載されたものも多い。国立歴史民俗博物館では、このような資料の画像を細部まで読み取れるよう高精細にデジタル化し、資料中の所望の箇所を自由に移動し、適切な倍率で読み取ることのできる画像閲覧システムを開発して、展示や研究支援に使用してきた[1]。

この画像閲覧システムには、表示される資料画像の箇所に応じて、解説を表示する機能を具備している。これは実物資料の展示では、展示パネルによる資料全体の説明しか与えられないことに対し、描かれた個々の場面や対象に応じた説明を与えることができる点で有効である。

この解説を表示する機能は、専門家が歴史的意味を考慮して解説を付与する。その後、解説付与した対象の位置や大きさならびに閲覧者が解説の付いていることを想定する領域を考慮してから解説を表示する領域を設定する。

しかし、実際に資料は多量にあり、すべての資料に対してアンケートなどで閲覧者が解説の付いていることを想定する領域を抽出することは非常に困難である。また、閲覧者が解説の付いていることを期待する領域は、閲覧する倍率で変化するなどといった問題もある。

そこで本研究では、解説表示設定のために閲覧者が解説の付いていることを期待する領域を自動抽出することを目的とする。

2. 領域抽出の手順

領域抽出は以下の手順で行った。まずは、被験者による領域抽出実験を行って被験者がどのような領域に解説を期待しているかを明らかにする。そして、被験者の選択した解説の付いて

いることを期待する領域の自動抽出を行う。

2-1. 被験者による領域抽出実験

画像閲覧システムを用いてあらかじめこちらで用意した3つの資料の中からランダムに20枚の画像を用意して閲覧者が解説の付いていることを期待する領域の抽出する実験を行った。具体的には被験者10名にそれぞれの画像を見せ、「この倍率でこの画像を見たときに説明が付いている、もしくは欲しいと思う領域を1つ以上指定してください。また、そこを選んだ理由を述べてください。」という設問に対する回答を収集した。

結果として当然ではあるが、歴史資料の特徴である金雲などをはじめとした背景は選択されないことがわかった。また、アンケートの記述欄の選択理由からは、被験者は色や大きさなどで領域を決定していることがわかった。色では鮮やかなもの、大きさでは一定以上の面積の領域を選んでいる。

また、画像閲覧システムを用いた際の特徴として、被験者が指定された領域の大きさは用意した画像により面積が大きく異なったが、一つの画像で複数の領域を選んだ場合では領域の大きさはあまり変化することはなかった。これは同じ画像内でみた場合、一番に目についた領域のものと比較して一定以内のサイズの領域のみを選択したからだと考えられる。

そこで本研究では領域の大きさに着目し、閲覧者が解説の付いていることを期待する領域を求めることにした。

2-2. 領域の自動抽出

被験者による領域抽出実験の結果から抽出する閲覧者が解説の付いていることを期待する領域を決定する。今回は半数以上の被験者が選択した領域を閲覧者が解説の付いていることを期待する領域として抽出することにした。

まず、背景は閲覧者が解説の付いていることを期待する領域に入らない。歴史資料では背景には何も描かない、もしくは金雲などを描くこ

Examination of area extraction method for setting displaying explanation of historical documents

Tomoki Saigo† Yukio Tokunaga†

Kiyoshi Sugiyama† Fumio Adachi‡

†Shibaura Institute of Technology

‡National Museum of Japanese History

とが多いという特徴があげられる(図1)。このことから背景では色の大きな変化はなく、エッジは存在しないと考えた。そこでエッジを用いて領域の分割を行い、背景部分の除去を行った(図2)。また、残った領域のうち、一番に目についた領域と比較して大きさが一定以内の領域に被験者は解説が付いていることを期待することから収縮膨張処理を用いて領域の抽出を行った(図3)。この際に一番大きな領域のものと小さな領域のものを、収縮処理を行って消える回数で比較してみると相関係数 0.85 で比例関係にあることがわかった。このことから回帰直線を用いて収縮処理のパラメータを決定した(図4)。

3. 評価実験と考察

一つの資料から画像を新たにランダムに 20 枚用意し、被験者 20 人に対して閲覧者が解説の付いていることを期待する領域の抽出実験を行った。実験では 2-1 のときと同じ質問を被験者に行い、解説の付いていることを期待する領域の抽出を行った。そして、被験者が抽出した領域のうち半数以上の人を選択した領域をシステムで抽出することを目標とした。

この領域とシステムで抽出した領域を 1 画素ごとに比較して評価を行った結果、抽出率 58.4%、未抽出率 41.6%、誤抽出率 33.3%となった。未抽出の原因は、小さな領域でも鮮やかな色がある場合や小さな領域が複数集まってできている領域などは大きさを用いた処理では抽出できなかったことが挙げられる。また川などを背景とした場合、エッジを元にした背景除去を行っているため背景の除去ができずに領域の抽出に失敗してしまった。しかし、閲覧者が解説の付いていることを期待する領域の一部は抽出できており、解説表示の設定の改善につながると考えられる。

4. おわりに

画像閲覧システムの解説表示設定のために、閲覧者が解説の付いていることを期待する領域の自動抽出手法の提案をした。今回は領域の大きさという要素に着目し、領域の抽出を行った。しかし、閲覧者が解説の付いていることを期待する領域は大きさだけで決まるものではない。今後はこれらの踏まえたうえで色差などを用いた抽出方法の検討を行う。

本研究は科学研究費補助金(課題番号 18300086)の助成を受けている。



図1. 金雲が描かれた資料(東海道五十三画卷)

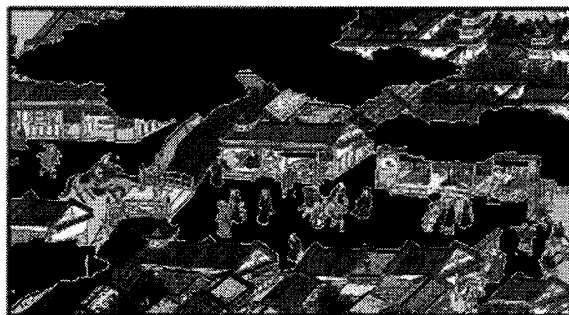


図2. エッジを基に背景を除いた画像

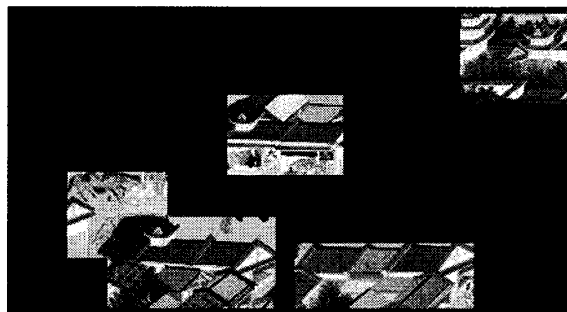


図3. システムで抽出した領域

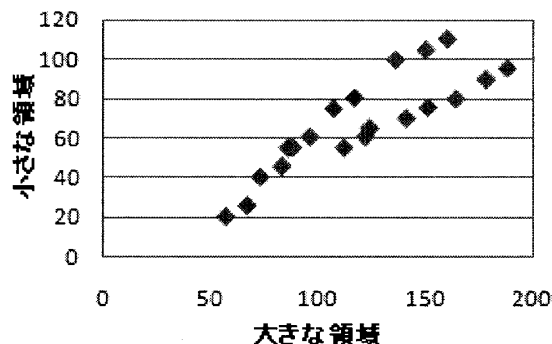


図4. 領域の大きさ比較

参考文献

- [1] 鈴木卓治, 安達文夫, "歴史研究・展示用画像表示システムの機能に関する検討," 情報処理学会シンポジウム論文集, vol.2001, No.18. pp.229-234 (Dec. 2001).
- [2] 早野浩章, 安達文夫, 鈴木卓治, 徳永幸生, 杉山精, "歴史画像閲覧システムにおける解説表示法の検討" 第7回情報科学技術フォーラム講演論文集, J-012, pp.123-128(Sep. 2008)