

文字らしさを総体的・局所的な視点で捉える ポスター画像中のイベント情報抽出システム

柳井 美保^{1,a)} 河野 恭之^{1,b)}

概要: 本研究では告知に用いられるポスター画像中からイベント情報を抽出する手法を提案する。イベント情報はイベントの特定に重要であるイベント名と日時とする。携帯端末の普及に伴い、告知に用いられるポスターのイベント情報を画像情報として保管する人が増加している。保管した大量の画像情報から、目当てのイベント情報を探すことは手間がかかる。この手間を省くために保管したポスター画像を OCR エンジンを用いてテキストデータへ変換し、電子カレンダーに記載する手法がある。しかし既存の OCR エンジンではポスター画像を対象としていないため、正確なテキストデータを得られないことが多い。そこで文字と文字周辺の背景の HSL の差に着目し、ポスター画像中から文字領域を抽出することにより、OCR エンジンの文字認識精度を向上させ、ポスター画像のイベント情報をテキストデータへ変換可能にする。

キーワード: 画像処理, OCR, ラベリング, HSL

Event information extraction system in a poster image capturing character likeness from a overall and locally viewpoint

YANAI MIHO^{1,a)} KONO YASUYUKI^{1,b)}

Abstract: In this paper we propose a method for extracting event information from the poster image used for announcement. The event information is defined as the event name and the date and time which are crucial for identifying the event. The number of people who store event information of posters as image information is increasing with the spread of portable terminals. It is troublesome to search for targeted event information from a large amount of stored image information. There is a method to convert the poster image to text data using the OCR engine and describe it in the electronic calendar in order to save this trouble. However, since the existing OCR engine does not cover poster images, accurate text data can not be obtained in many cases. Therefore, attention is paid to the HSL difference between of the character and the background around the character, and a character region is extracted from the poster image. This improves the character recognition accuracy of the OCR engine and makes it possible to convert the event information of the poster image into text data.

Keywords: Image processing, OCR, labelling, HSL

1. はじめに

携帯端末の普及に伴い、告知に用いられるポスターのイベント情報を画像情報として管理する人が増加している。

例えば公共の場に掲示されているポスターを携帯端末に搭載されているカメラで撮影することにより、短時間で手軽にイベント情報を保管・管理する人が挙げられる。保管した大量の画像情報から、目当てのイベント情報を探すことは手間がかかる。この手間を省くために保管したポスター画像を OCR エンジンを用いてテキストデータへ変換し、電子カレンダーに記載する手法がある。しかし既存の

¹ 関西学院大学
Kwansei Gakuin University
a) dov97462@kwansei.ac.jp
b) kono@kwansei.ac.jp

OCR エンジンではポスター画像を対象としていないため、正確なテキストデータを得られないことが多い。これはポスター画像の複雑なレイアウト構造により、文字領域がポスター画像中から検出できないことが原因である。

岡本ら [1] は情景画像を対象に文字を検出している。文字列は同じ色とフォントであると仮定し、文字を識別するためにまずエッジに基づき、画像をブロックに分割を行う。各ブロックに対してクラスタリングと色の類似性を捉えて更に領域分割する。各領域に存在する外接最小矩形の縦横比とピッチなどを捉えてそれらを統合し、文字列パターンの候補を得る。これらの候補に対して幾何学的特徴を求めて SVM により文字列パターンを識別する。この手法は複雑な背景と文字で構成されるポスター画像には対応できない。そこで本研究ではポスター画像でも対応可能にする。

平野ら [2] は広告画像を対象に、テキスト領域の特徴をノード間の重みとして与え、グラフカットを用いることで文字領域とそれ以外に 2 値化した。この手法はフォントサイズが一定であることが多い広告画像には有用である。しかしポスター画像のような 1 枚の画像中に様々なフォントサイズの文字が存在する画像には対応できない。そこで本研究では文字と文字周辺の背景は HSL の差が大きい特徴に着目し、ポスター画像中から文字領域を抽出することにより、様々なフォントサイズに対応可能にする。

本研究では文字領域と背景領域を分類する際、ポスター画像中の文字らしさを 2 つの視点で捉える。まず総体的な視点では文字列は 1 文字の集合体であり、1 文字が正方形に類似している特徴を捉える。次に局所的な視点では文字列はコーナーと線分を多く含んでいる特徴を捉える。

2. 提案手法

図 1 に全体の流れを示す。携帯端末に搭載されたカメラでポスターを撮影し、そのポスター画像を入力画像とする。入力画像に対して文字と文字周辺の背景の HSL の差に着目した領域分割を行い、似通った領域ごとに結合する。分割された領域内で文字の特徴であるコーナーと線分を検出し、どちらもしきい値以上検出されたものを文字領域、それ以外を背景領域に分類する。しかし分類した画像に対して既存の OCR エンジンを用いても、複雑なレイアウト構造により正確なテキストデータを得られない。そこで既存の OCR エンジンが文字を認識しやすい、背景と文字が明瞭に分割された画像に変換する前処理を行う。その後文字認識を行い、認識された文字列からイベント名、日時と判定された文字列をユーザに提示する。提示した内容が正しくない場合はユーザがイベント名と日時の文字列をそれぞれ黒色の実線で囲むことにより補助する。囲まれた領域ごとに上記の全体の流れを基本的にはくり返す。再度イベント名と日時が正しくない場合、提案手法では対応不可能と判断する。提案デモンシシステムでは携帯端末に搭載されたカ

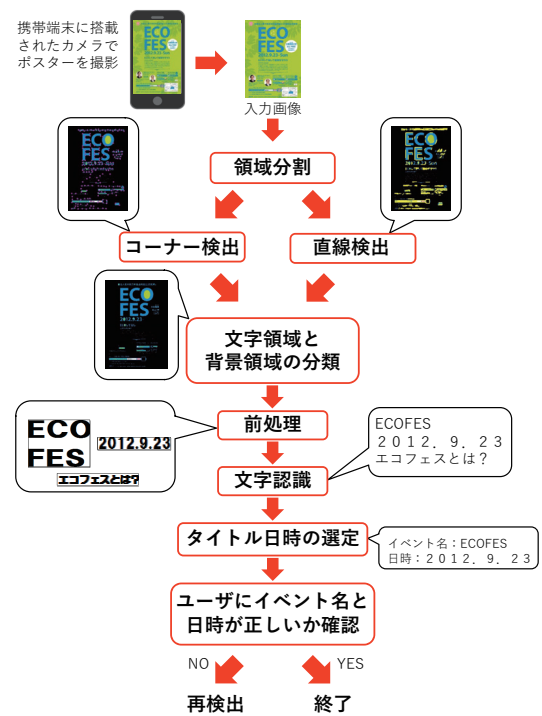


図 1 全体の流れ
Fig. 1 Overall flow

メラでポスターを撮影し、撮影したポスター画像をクラウド上にアップロードする。アップロードされたポスター画像を入力画像とし、PC 上で提案手法を実装する。

2.1 文字領域と背景領域の分類

初めに文字と文字周辺の背景の HSL の差に着目した領域分割を行うことで類似している領域ごとに領域分割し、分割された領域を似通った領域ごとに結合する。結合された領域に対し、文字の特徴であるコーナー検出と直線検出を行い、コーナーと線分どちらもしきい値以上検出されたものを文字領域、それ以外を背景領域に分類する。

本研究ではコーナー検出に垂直、水平方向の変化を重視する特徴を持つ Harris オペレータ [3] を用いており、直線検出はまず途切れたエッジが結合される可能性が高くなる特徴を持つ canny [4] アルゴリズムを用いてエッジを検出し、エッジがノイズにより連続していない場合でも、事前に設定したしきい値以上の特徴点が存在すれば同一直線上に直線を検出できる特徴を持つ、確率的ハフ変換を用いる。

2.2 領域分割

図 2 は領域分割の流れを示す。本研究では色に着目するため、領域分割はカラー画像に対して行う。まず画像全ての画素のラベル番号を -1 で初期化する。領域分割は入力画像の左上からラスタスキャンを行い、注目画素を中心に 8 近傍に対して HSL 空間上のユークリッド距離を求める。この距離がしきい値以下であれば同じラベルを振り、しきい値以上であればラベルを振らない。次にラベルを振られ

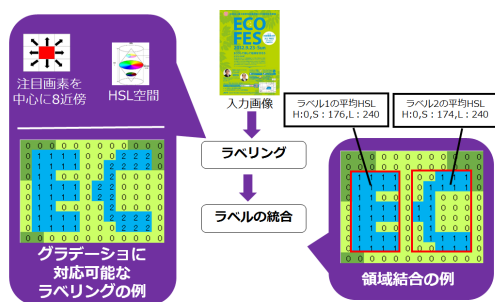


図 2 領域分割の流れ

Fig. 2 Region segmentation flow

なかった画素に対し、左上の画素から同様に繰り返す、-1 以外のラベルが全画素に振られるまで行う。ここでのしきい値はポスター画像中のすべての隣接する画素に対して HSL 空間上のユークリッド距離を求め、最大値に 0.7 を掛けた値とする。これは各ポスター画像に対して自動的に変動するため、それに適したしきい値を与えることが可能である。1 つの文字列に 1 つのラベルを振り分けられることが理想だが、この時点では文字列だけでなく、1 文字にも異なるラベルが振られる。例えばひらがなの「た」の場合、3 つのラベルに分割される。

このような複数のラベルを 1 文字と判定するために、1 文字は正方形の矩形にバランスよく記載されることが多い特徴を捉える。まずラベルごとに HSL の平均を求め、図 2 の赤色で示す矩形を作成する。この矩形が重なり、且つ HSL の差がしきい値以下であればラベルを結合する。次に注目した矩形の重心からユークリッド距離が近い重心を持つ矩形を順に結合していく。ここで結合する矩形は HSL の平均の差がしきい値以下の矩形のみとする。結合した矩形が正方形に類似した時点で結合を終了し、結合した矩形のラベルを注目したラベルと同じラベルに結合する。最後に 1 つの文字列に 1 つのラベルを振り分けるため、文字列は同色で且つ等しいフォントサイズで記載される特徴を捉える。この特徴を捉えるために各ラベルの高さと幅の最大値を求め、その高さまたは幅の差がしきい値以下で且つ HSL の差がしきい値以下のラベルは結合する。

2.3 前処理

OCR エンジンが対象としている画像は背景と文字が明瞭に分割されており、複雑なレイアウト構造ではない画像である。そこでポスター画像を OCR エンジンが対象としている画像に変換することで、OCR エンジンが文字を認識しやすくする。複雑なレイアウト構造を 1 つの文字列ごとに分割するため、まずポスター画像中の文字領域と判断されたラベルごとに矩形に分割する。次に分割した各矩形の文字領域と判定された領域で、且つ最も面積を占める色を黒色に、それ以外を白色に 2 値化する。

2.4 イベント名と日時判定

ポスター画像のイベント名は比較的大きく記載されることが多い。この特徴を捉えるために各文字列のフォントサイズを求め、求め方は各矩形に対して文字領域を包含する面積を求め、その面積を文字数で割る。フォントサイズが最も大きい矩形にはイベント名が記載されていると判定する。もしイベント名が記載されていると判定された矩形が複数あった場合はポスター画像中の上、または右に記載されているものを優先する。

またイベント名が記載されていると判定された矩形以外で、2 文字以上数字が含まれている矩形は日時が記載されていると判定する。しかしアドレスや電話番号、値段、定員、住所などのワードが含まれているものは除く。

2.5 再検出

再検出ではユーザーがイベント名と日時をそれぞれ実線で包含することにより補助する。包含された領域に対して基本的には上記の手法を適応するが、ラベルの結合の際の仕様は少し異なっている。2.2 章で述べたラベルの結合では幅と高さ、HSL の差を比較してラベルを結合していたが、包含された領域には必ず 1 つの文字列のみが存在していると仮定できるため、再検出では HSL の差のみを比較することでラベルを結合する。

3. まとめと今後の展望

本研究ではポスター画像から文字領域を抽出する手法を提案した。これにより OCR エンジンのポスター画像に対する文字認識精度を向上させ、ポスター画像のイベント情報をテキストデータへ変換可能にする。今後の展望は Web 上での利用が考えられる。Web サイトにはタイトルや住所などの肝心な情報が画像データになっている場合がある。このとき画像データを自動的にテキストデータに変換することで、検索や入力の手間を省けるのではないかと考える。

参考文献

- [1] 芦田和毅, 永井弘樹, 岡本正行, 宮尾秀俊, 山本博章: “情景画像からの文字抽出”, 電子情報通信学会誌文誌, Vol.J88-D-II, No.9, pp.1817-1824, (2005-9)
- [2] Hiromi Hirano, Makoto Okabe, Rikio Onai: “Detection of Inserted Text in Images”, Mathematical Progress in Expressive Image Synthesis I, Series Volume 4, Part V, 2014, pp. 177-183.
- [3] Chris Harris and Mike Stephens. A COMBINED CORNER AND EDGE DETECTOR. In Proceedings of 4th Alvey Vision Conference, pp. 147151, Manchester, U.K., Aug. 1998.
- [4] John Canny, “A computational approach to edge detection”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(6): 679698, Nov. 1986.