

ポスター画像中のイベント情報抽出システム

柳井 美保* 河野 恭之*

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。

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 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や飾り文字により、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。平野ら [1] は広告画像に対し、テキスト領域の特徴をノード間の重みとして与え、グラフィック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、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。

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。入力画像から局所的に似た領域の中

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 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 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。提案デモシステムでは携帯端末に搭載されたカ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。

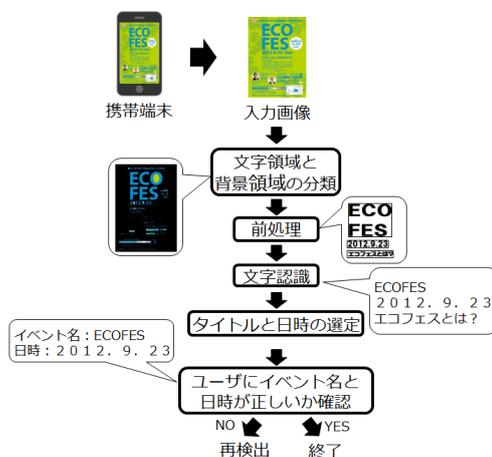


図 1. 全体の流れ

2.1 文字領域と背景領域の分類

初めに局所的に似た領域に基づく領域分割を行うことで文字に近い領域に絞る。絞った領域に対し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。

図2は局所的に似た領域に基づく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。次に注目した矩形からユークリッド距離が近い順に矩形を結合していく。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。結合していく際、結合した矩形が正方形に類似した場合は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。また矩形を結合しても正方形に類似しない場合はラベルの結合を行わない。1つの文字列に1つのラベルを振り分けるために、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の最大値を求め、その高さの最大値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。

2.2 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。そこで文字領域と背景領域に分割されたポスター画像をOCRエンジンが対象に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



図2. 局所的に似た領域に基づく領域分割の流れ

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。

2.3 イベント名と日時の判定

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る。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。

3 まとめ

既存のOCRエンジンではポスター画像中の文字を正確に読み取ることができない。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。

参考文献

- [1] 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.