

図形輪郭除去によるテロップ文字領域抽出

Telop Character Extraction Using Figure Contour Removal

平松 義崇[†] 関本 信博[†] 新庄 広[†] 丸川 勝美[†]
Yoshitaka Hiramatsu, Nobuhiro Sekimoto, Hiroshi Shinjo, Katsumi Marukawa

1. 緒言

近年、EPGによる番組単位の検索だけでなく、映像コンテンツ中の特定のシーンだけを検索したいというニーズが高まっている。シーン検索を実現するにはシーンの内容を表すデータが必要であるが、それらが付加されていることは少ない。そこで、そのデータを自動で抽出する技術が求められている。

この自動抽出を実現するために用いる情報として、映像コンテンツ中に画像として埋め込まれたテロップがある。テロップの文字にはシーンの概要を表す言葉が含まれる場合が多く、その文字を認識することで有効なシーン情報が得られる。テロップの文字の認識は、映像からテロップの文字領域を抽出し、文字認識することで実現できる。

本報告では、近年のテレビ番組で頻繁に見られるようになってきた図形上に文字が描かれたテロップに対して、文字領域をロバストに抽出する方法を提案する。

2. 従来方法の問題点と解決策

テロップから文字領域を抽出する方法として、文献[1][2]では、テロップ矩形領域を抽出し、得られた領域から背景部分を除くことで領域を絞り込み、その領域から文字・背景分離により文字領域を抽出している。文献[1]では、絞込み領域として、テロップ矩形領域内のエッジを膨張させた領域を用い、また、文献[2]では、テロップ矩形内のエッジ画像から横方向に投影をとって文字行矩形を抽出し、その文字行矩形内で縦方向に投影をとって抽出した文字矩形を用いている。しかし、図形上に文字が描かれたテロップでは図形と背景の境界がエッジとして抽出されるため、どちらの方法も領域を絞り込めない。特に図形が複数の行に連なっているテロップでは、複数の行が統合された領域が抽出されてしまい、ロバストな文字・背景分離が困難になる。

文献[3]では、輝度、色相、彩度を用いてフレーム画像を領域分割し、分割された各領域が次の条件を満たすとき文字とする方法が提案されている。(1)次のフレームでサイズと位置の変化が小さい、(2)サイズが妥当、(3)フレームに接触しない。図形上に文字が描かれたテロップでは、上記3つの条件を満たす図形部分があるため、図形部分が誤抽出される。

また、図形部分に着目した研究としては、エッジ画像を抽出し、その外接矩形から文字画像を求め、さらに文字認識させて得た確信度を用いて図形輪郭を除去する方法が提案されている[4]。この方法は、文字認識する必要があるため、処理時間が大きいという問題がある。

図形部分が誤抽出される問題点を解決するために、提案方法では、テロップ矩形領域内のエッジ領域を文字輪郭と図形輪郭に分類し、図形輪郭を除去する処理を追加した。以下で、処理全体の流れと図形輪郭を除去する方法について述べる。

3. 図形輪郭除去によるテロップ文字抽出

処理全体の流れは図1に示すように、テロップ矩形抽出、文字列矩形抽出、文字領域抽出の3ステップにて行われる。テロップ矩形抽出では、フレーム画像からテロップ矩形領域を抽出する。ある時刻 t のフレームの輝度画像からエッジ検出を行い、エッジ強度がしきい値より大きい画素を求める。次に時刻 $t+1$ のフレーム画像の輝度画像の前述の全ての画素に対してエッジ検出を行い、エッジ強度がしきい値より大きい画素を求める。この処理を N 回繰り返し、最終的に残った画素の連結成分を求めテロップ矩形領域とする。

次に、文字列矩形抽出処理により、テロップ矩形領域から文字が含まれる領域を抽出し、文字列矩形領域を求める。まず、テロップ矩形領域からエッジ検出により、テロップ領域内のエッジ領域を求める。得られたエッジ領域を図形輪郭と文字輪郭に分類し、図形輪郭を除去する。図形輪郭除去後のエッジ領域からノイズを除去し、残ったエッジ領域から連結成分を求め、その連結成分を統合して文字列矩形領域とする。

最後に、文字列矩形領域内から文字領域を抽出する。その手順は、文字列矩形領域からエッジ検出によりエッジ領域を抽出し、矩形領域を求め、その矩形内で二値化し、文字候補領域を求める。次に文字候補領域に対して、微小領域と矩形の端に接触する領域を除去する。

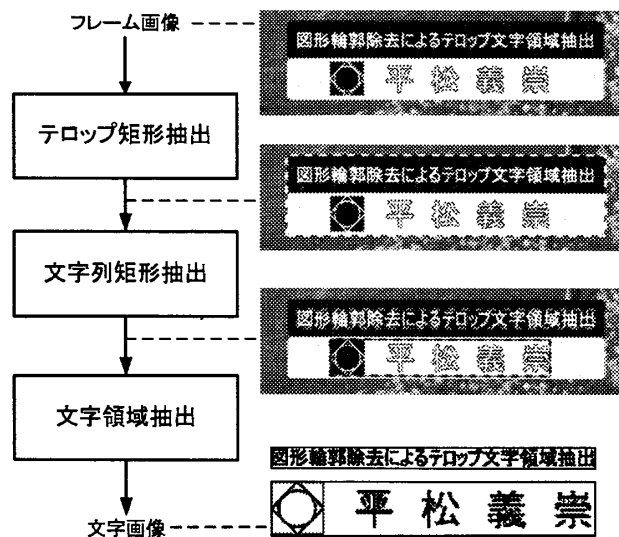


図1: 処理全体の流れ

4. 図形輪郭除去

次に、図形輪郭と文字輪郭を判別し、図形輪郭を除去する手法を述べる。

4.1. 図形輪郭と文字輪郭の分析

まず、テロップ矩形領域内のエッジ領域の形状を調査した。その結果、図形輪郭の場合、図2に示すような(1a)

[†] 株式会社 日立製作所 中央研究所

直線状, (1b)かぎ括弧状, (1c)大括弧状, (1d)四角形, (1e)文字輪郭と一体化, が存在した. 一方, 文字輪郭の場合は, (1a)直線状, (1b)かぎ括弧状の輪郭は見られなかった. これらの形状に基づいて図形輪郭と文字輪郭を判別する.

また, (1e)文字輪郭と一体化した輪郭に対応するため, 横一行毎, または, 縦一列毎に図形輪郭か否かを判定・除去する方法をとることとした. 以降, 横一行毎の場合を例にして説明を行う. 横一行毎の輪郭の特徴は, 図形輪郭の場合, 少数の長いランで構成されるのに対して, 文字輪郭の場合, 多数の短いランで構成されることが判明した. ただし, '国'の上部のように, 文字幅と同じ長さの水平な部分では図形輪郭と同様のラン構成になるため, 一定長以上のランの有無だけでは対応できない.

そこで, 一定長以上のラン付近の他の輪郭成分の存在を用いることとした. 図形輪郭の場合, 図2に示すように, (2a)上下に一文字以上の文字輪郭が並ぶ, (2b)一文字以上の文字輪郭を囲む, が存在した. 一方, 文字輪郭の場合, 文字の付近に複数の輪郭が存在する場合は見られなかった. 付近に存在する輪郭成分に基づいて図形輪郭と文字輪郭を判別する.

また, 一定長以上のラン付近に存在する輪郭が1つだけの場合, その輪郭は文字の縦棒や横棒の輪郭の可能性もある. これを区別するため, 横ランの対応関係を利用することとした. 図形輪郭の場合, 図3(a)に示すように, 対応する横ランは文字輪郭をとり越した位置に存在する. 一方, 文字輪郭の場合, 図3(b)に示すように, 付近に対応する横ランが存在する. この対応関係に基づいて図形輪郭と文字輪郭を判別する.

以上の分析に基づいて構成した図形輪郭除去方法を次節で概説する.

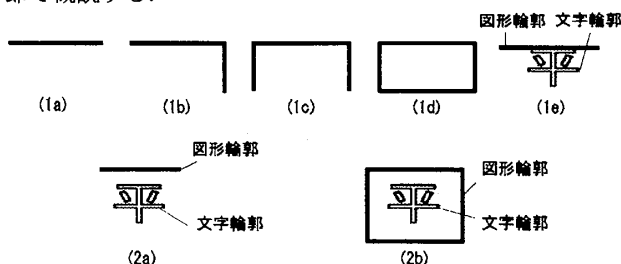


図2: 図形輪郭の形状と位置関係

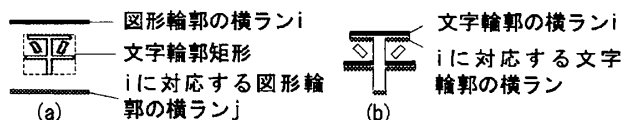


図3: 横ランの対応関係

4.2. 図形輪郭除去方法

処理の流れは, (1)形状による判定, (2)付近の輪郭成分の存在による判定, (3)輪郭成分間の対応関係による判定, の3つのステップにて行われる.

(1) 形状による判定

直線状とかぎ括弧状の輪郭を除去する. これらの形状の判定条件は, 一定長以上の横ランに対して両端に縦方向に伸びる画素がないときを直線状と判定し, 片端に縦方向に伸びる画素があるときは, かぎ括弧状と判定する.

(2) 付近の輪郭成分の存在による判定

形状による判定で除去されなかった一定長以上の横ランに対して, 付近に別の輪郭があるか否かに基づいて図

形輪郭を除去する. 図形輪郭と判定する条件は, 別の輪郭が複数存在した場合である.

(3) 輪郭成分間の対応関係による判定

上記判定(2)で, 別の輪郭が1つまたは存在しなかった横ランに対して, 対応する横ランの存在に基づき図形輪郭を除去する. 2つの横ランの対応付け方法は, 輪郭の勾配方向が 180° 異なるもの同士を対応付ける. 図形輪郭と判定する条件は, 対応する横ランが付近にない場合と, 対応する横ランが別の輪郭の横ランである場合とした.

5. 実験

本提案手法の効果を想定するために文字列抽出実験と文字領域抽出実験を実施した.

5.1. 実験条件

野球1番組とゴルフ1番組から図形上に文字が描かれたテロップ領域を用いて提案手法の効果を確認した.

評価は, 図形輪郭除去がある場合と図形輪郭除去がない場合で次の3つの指標で比較した. (a)複数行から構成されるテロップ272個に対して一行ごとに抽出される率(一行抽出率), (b)1385個のテロップ文字に対して, 文字領域の一部が抽出できなかった文字数(見落とし), (c)テロップ領域内で文字でない場所を誤って文字領域として抽出した場合(誤抽出), である. いずれの指標も目視でカウントした.

5.2. 実験結果

実験結果を表1に示す. 表1より, 図形輪郭除去がある場合は, 複数行で構成されるテロップに対して, より高精度に一行ごとの抽出ができるようになった. また, 見落とし, 誤抽出共に改善された. 以上の結果から, 図形輪郭除去を追加する有効性が示された.

表1: 文字領域抽出精度

#	一行抽出率	見落とし	誤抽出
図形輪郭除去なし	64.7%(176/272)	89	388
図形輪郭除去あり	98.9%(269/272)	51	161

6. 結言

本報告では, 図形上に文字が描かれたテロップに対して文字領域をロバストに抽出する方法を提案した. 本方法は, テロップ矩形領域内のエッジ領域を, 輪郭の形状, 付近の輪郭成分の存在関係, 輪郭成分間の対応関係, に基づいて図形輪郭か文字輪郭かを判定し, 図形輪郭を除去することが特徴である. これにより, 図形部分が誤抽出される従来の問題を解決し, ロバストな文字抽出を実現できた. 野球, ゴルフ番組の映像を用いた評価実験により本方法の有効性が明らかになった.

参考文献

- [1] 堀修他, "テロップ認識のための映像からのロバストな文字部抽出法", 信学論, Vol.J84-D-II, No.8, pp.1800-1808, 2001.
- [2] X. Tang et al., "A Spatial-Temporal Approach for Video Caption Detection and Recognition," IEEE Transactions on Neural Network, Vol.13, No.4, July 2002.
- [3] H. Kuwano et al., "Telop Character Extraction from Video Data," In Proc. of DIA'97, pp. 82-88, 1997.
- [4] 藤本克仁他, "低解像度カラー文書画像から高品質な文字画像を抽出する二値化方式", 信学技法, PRMU99-89, pp.7-14, 1999.