

アナログテレビ放送からのテロップの抽出

尾関 孝史* 倉狩 沙織* 福永 裕也** 小林 富士男*

Extraction of Superimposition from Analogue Television Broadcast

Takashi OZEKI*, Saori Kurakari*, Yuya Fukunaga*, Fujio KOBAYASHI*

ABSTRACT

Superimposition in videos is one of the most important information to make indexes of image contents automatically. In the past, many obstacles have put this into practice. In the case of signboards, it is easy to distinguish characters and the background by using the brightness. Because signboards are usually made by assigning different colors to characters and the background. On the other hand, in the case of superimposition, by using only brightness, it is difficult to separate characters from the background. In this paper, we propose a new method to extract superimposition from videos based on the information of the edge and the labeling of images. Applying the method to real videos and extracting superimposition from 19 images, we obtained exactly almost every characters.

キーワード: 画像処理, テレビ放送, テロップ, アナログ放送

Keywords: Image Processing, Television Broadcast, Superimposition, Analogue Television

1. まえがき

近年, 衛星放送やブロードバンド, あるいはビデオやDVD等の普及によって, 多くの映像が一般家庭に氾濫している. 今後, このような膨大な映像情報から必要な映像を抽出する技術が必要となることが予想され, 数多くの研究がなされている. 例えば, 米国 Carnegie Mellon 大学における Informedia プロジェクト [1] を始めとして, これまでに蓄積された映像資源の効率的な再利用や検索を行ない, 映像データベースを構築するための自動索引付けに関する研究が内外で盛んに行なわれている. これらの多くの研究では, 索引付けの際に映像の内容を示す重要な情報として, 映像中の音声やテロップが使われている. 特に, テロップはタイトル等に使われており, 映像の内容を示す重要な情報源となる. そこで, 映像中からテロップを自動抽出するための研究が, 重要になっている.

情景画像から, 文字領域を抽出する研究は昔から多く行なわれている [2-7]. しかしながら, これらの研究で

は, 抽出する文字領域はナンバープレートや看板のように文字部と背景の色が明確に分かれるものを対象にしており, 輝度ヒストグラムにおいて, 適当なしきい値でその2つの部分を分離している. その結果, テロップのような背景が存在しない画像ではうまく文字部を抽出できない. 一般にテロップには, 背景は存在しないが, 文字の回りに輪郭線のように別の色で囲まれている. そこで, テロップを含む局所領域で濃度ヒストグラムを作成し, それに対して, しきい値を決めて文字部を取り出す試みがなされている [8,9]. ところが, これらの論文で提案された手法では, 最終的に, しきい値のみでテロップを抽出するため, 映像の背景がテロップと同色の場合, 背景の一部が除去されずに必ず残ってしまう. そこで, 背景とテロップの色が似ている映像中に含まれるテロップを抽出する方法を提案する. 本論文で提案する手法では, エッジの情報とラベリングの情報を組み合わせることで, テロップと同色の背景を削除し, 文字部みの検出を試みている.

* 情報処理工学科

** 大学院情報処理工学専攻

2. テロップの抽出法

本節では、背景とテロップを分離し、テロップのみを抽出する方法を述べる。なお、対象とする映像のテロップは、白色で、黒色の輪郭線があるものとし、テロップは数秒間、変化しないものと仮定する。また、提案手法は256階調の濃淡画像を対象にしているが、カラー画像に対しては、例えば変換式

$$0.299 * R + 0.587 * G + 0.114 * B$$

を用いて、濃淡画像に変換して適用する。ここで、赤成分をR、緑成分をR、青成分をBで表している。

2.1 処理全体の流れ

始めに、テロップ抽出法の大まかな流れを図1に表す。

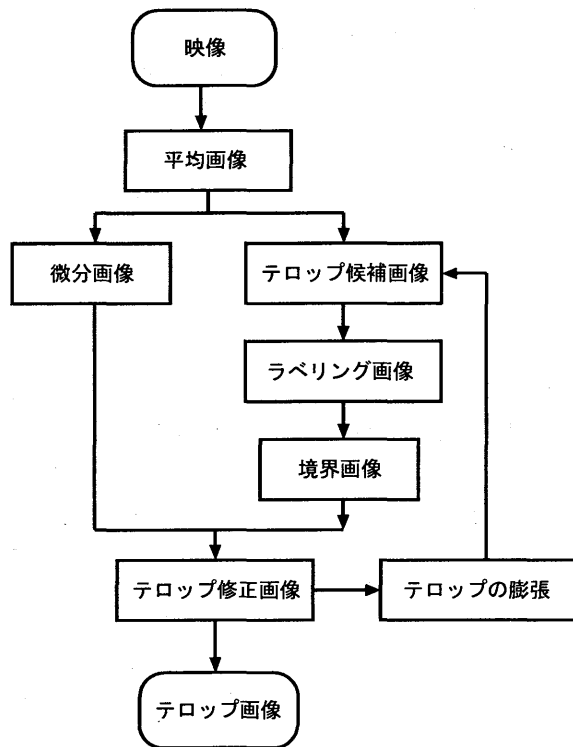


図1 処理全体の流れ

2.2 平均画像の作成

日本のアナログ放送は、1秒間の映像に約30フレームが使用されている。また、テロップは通常、数秒間映像中で変化しない。そこで、論文[9]に使われたのと同じように、同じテロップが続くフレームの平均画像を用いる。このようにすることで、雑音を除去し、安定した

テロップを得ることが可能となる。また、テロップ以外の動きのある背景部分は、モーションブラーの影響のため、ぼけやすくなり、テロップと背景を分離しやすくなる。なお、本論文では、同じテロップを含むフレームは目視によって得ている。図2は、約2.5秒間75フレームの平均カラー画像である。



図2 平均画像

2.3 テロップ候補画像の作成

次に、この平均画像に対して、輝度があるしきい値 T_h 以上の画素をテロップの候補とするテロップ候補画像を求める。ここで、しきい値はテロップの輝度分布のうち、高めの値を取る。

図3は、図2の平均画像の濃淡画像に変換した画像に対して、その輝度が閾値 $T_h = 230$ 以上となる画素を白(255)、その他を黒(0)として得られたテロップ候補画像である。



図3 テロップ候補画像

2.4 微分画像の作成

平均画像の微分画像では、テロップ以外の背景が移動するため、テロップ部分だけのエッジが鮮明に現れる。

そこで、平均画像に対して、ソーベルフィルタを適用し、微分画像を求める。その後、得られた微分画像を適当なしきい値 T_d で2値化し、2値化画像を求める。

図4は、図2の平均画像にソーベルフィルタを適用し、しきい値 $T_d = 400$ で2値化した画像である。

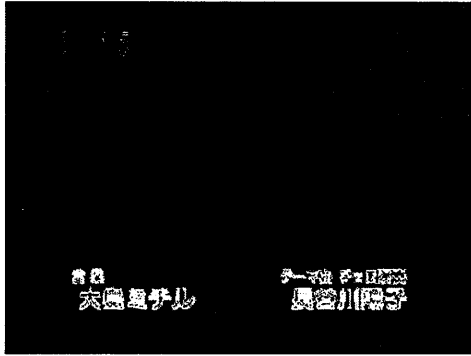


図4 微分画像の2値化画像

2.5 テロップ候補画像の不要点の削除

テロップ候補画像には、テロップに関係ない背景画素が多く含まれているので、これらを除去する。

始めに、テロップ候補画像のラベリングを行ない、更にその境界画素を抽出する。あるラベリング番号の各境界画素に対し、その8方向近傍に微分画像のエッジを含むかどうかの割合を調べる。もし、その割合がある値 R より小さい場合は、そのラベリングに対応する白画素をテロップ候補画像から削除する。これを全てのラベリング番号に行なった結果、残された画素からなる画像をテロップ修正画像とする。



図5 図3の境界画像

図5は、図3のテロップ候補画像に対して、その境界画素を抽出した画像である。図6はその境界画素がエ

ジに含まれる割合が、 $R = 0.6$ より小さいラベリング画素を図3から削除して得たテロップ修正画像である。

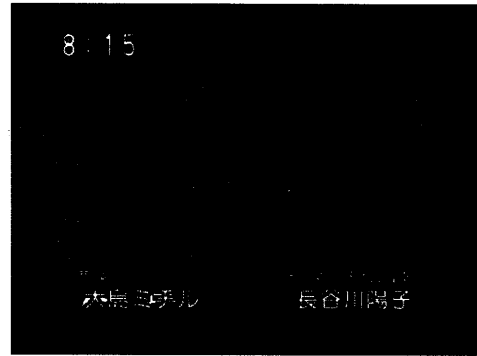


図6 テロップ修正画像

2.6 テロップ修正画像の膨張

図6のテロップ修正画像には、テロップに関係ない背景画素が若干含まれているので、これらを除去する。

今度は、図6のテロップ修正画像の画素の8方向近傍画素で、テロップ対象の画素より輝度が高々 S だけ小さい値、またはあるしきい値 T_l 以上の画素を1画素分、テロップ候補画像に追加する。これを数回繰り返す、図6のテロップ修正画像を成長させる。この時、しきい値 T_l は、 $T_h \geq T_l + S$ を満たすように決める。

図7は、図6のテロップ修正画像の各画素に対して、輝度がそのテロップ候補の画素より高々 $S = 10$ 小さい画素または、輝度が $T_l = 185$ 以上の画素を付け加える操作を5回行なった結果である。



図7 膨張したテロップ候補画像

2.7 不要点の削除および膨張の繰り返し

先ほどと同様に膨張したテロップ候補画像に2.5節のテロップ候補画像の不要点の削除の方法と2.5節のテロ

プ修正画像の膨張の方法を交互に繰り返し、テロップ修正画像が膨張できなくなった時、最終的なテロップ画像とし、反復を終了する。

図8は、以上の反復を21回繰り返した結果、得られた最終的なテロップ画像である。

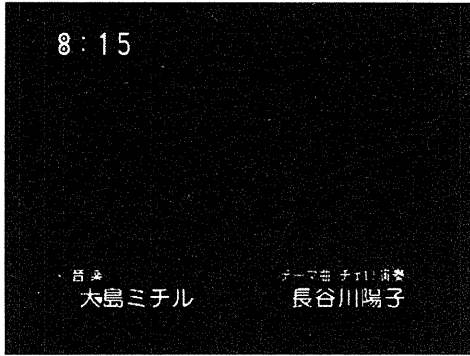


図8 最終的に得られたテロップ画像

3. 実験結果

提案手法を19種類のテロップに使用してみた。映像の大きさは、720×540のカラー画像である。各種パラメータは、前節と同様に $T_d = 400$, $T_h = 230$, $T_l = 185$, $S = 10$, $R = 0.6$ としている。図9は、テロップの抽出がうまくいった例である。これらの例からわかるように、テロップと背景の色が大きく異なる場合は、良好な結果を得ることができている。

これに対して、図10は、テロップの抽出が完全にはできなかった例である。背景にテロップと同様に白色が比較的多く現れる画像である。

図11は、図10のテロップ627-701の拡大図(一部)である。この例に見られるように、数字の「8」の中や「大島」の2つの文字の間にテロップと無関係な不要点の画素が残ってしまっている。また、図12は、図10のテロップ1647-1766の拡大図(一部)である。この例では、「吉崎」の「崎」の一部や、「写真」の一部が完全に欠落してしまっている。また、「(写真)」にあるはずの括弧は完全に消えてしまっている。その他、テロップ702-776では、「語」のごんべんの一部が欠けたり、テロップ1647-1766では「レコード」の濁点、テロップ1767-1886では「ピアノ」のような半濁点が欠けてしまった。

これらの主な失敗の原因は次の二つである。

- テロップの背景が白く、文字の間の背景が残ってしまった。すなわち、テロップの膨張をしても、文字中または文字間の背景のため、膨張することができ

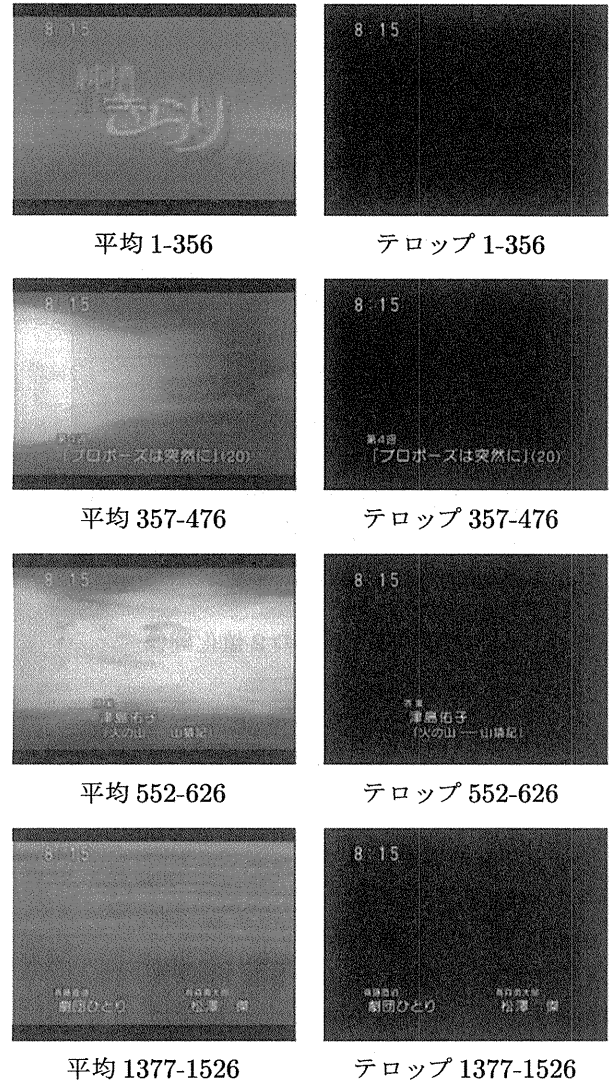


図9 テロップの抽出がうまくできた例

なかつたため、不要点として取り除くことができなかった。

- しきい値 T_h の値以上の画素が存在しなかつたため、最初のテロップ候補画像の段階で、テロップの一部が完全に欠落してしまつたため、その後何度、膨張を行なつても、画素が全く追加されなかつた。

後者の問題を改善するため、しきい値 T_h の値を下げて実験してみた。 $T_h = 195$ とし、平均1647-1766に提案手法を適用した結果が図13で、その拡大図(一部)が図14である。この例では、「吉崎」の「崎」の一部は相変わらず欠けているが、「(写真)」の括弧が現れている。また、テロップ1647-1766では「レコード」の濁点も抽出することに成功している。しかし、図15およびその拡大図(一部)の図16にあるように、逆にテロップ627-701では、「口」の中などに不要点が現れてしまつており、 T_h

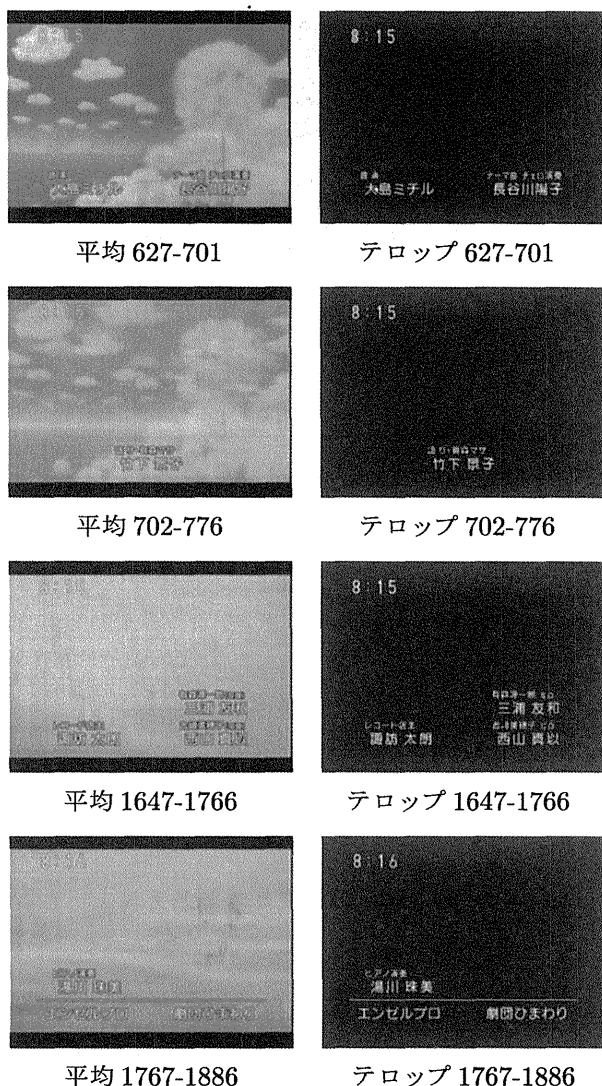


図 10 テロップの抽出に一部失敗した例

の値を下げるのが必ずしも改善になるとは限らないことがわかる。

また、図 8 のテロップの濃度ヒストグラムをグラフで表示したものが図 17 である。図 17 の実線は、抽出されたテロップ部分のヒストグラムを表示したもので、最初に $T_h = 230$ のしきい値で、後に $T_l = 185$ のしきい値で膨張している。なお、 $T_l = 185$ 以下の部分は、テロップ候補の回りの画素で、高々 $S = 10$ 以内だけ輝度が小さい結果、付け加えられた画素の部分である。一方、グラフの点線は、最終的に得られたテロップの画素領域を無条件に 1 画素分膨張した画素領域の濃度ヒストグラムである。1 画素分、追加しただけで、ヒストグラムの分布範囲が大きく広がってしまっている。このことから、提案手法によって、かなりの精度でテロップ部分が抽出できたことがわかる。また、このヒストグラムから、従来手法で行なわれているような輝度だけを利用してテロップ

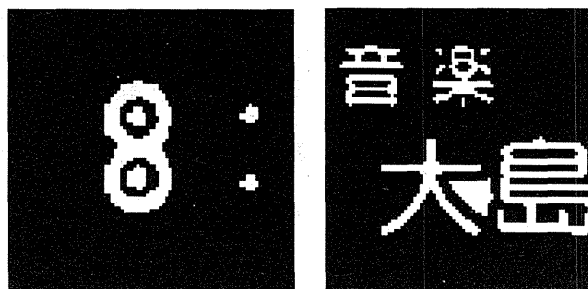


図 11 テロップ 627-701 の拡大図 (一部)

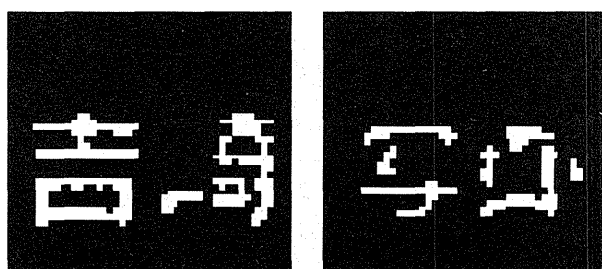


図 12 テロップ 1647-1766 の拡大図 (一部)

を抽出する方法ではうまくいかないことがわかる。

4. むすび

背景とテロップの色が似ている映像からテロップを抽出する方法を提案した。提案手法では、エッジの情報とラベリングの情報を組み合わせることで、テロップと同色の背景を削除し、文字部の検出を行なった。実際のテレビ画像に提案手法を適用し、テロップの抽出を行なったところ、19 種類の画像に対してほぼ全ての文字部を正確に抽出することができた。しかし、提案手法では、背景がテロップと同色の場合に、文字中や文字間にある不

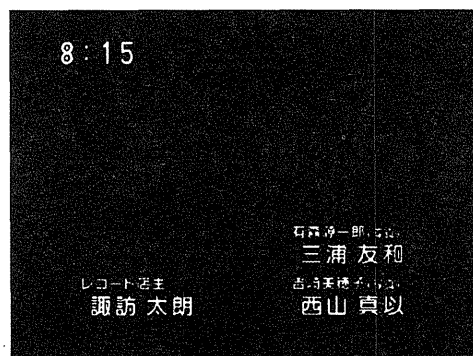


図 13 $T_h = 195$ で得られたテロップ画像 1647-1766

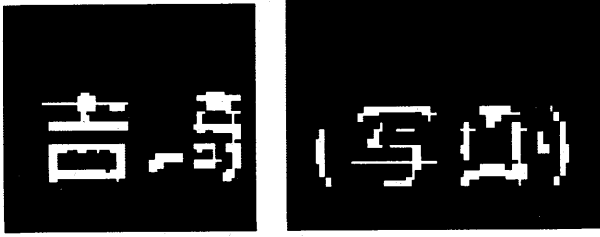


図 14 $T_h = 195$ のテロップ画像 1647-1766 の拡大図

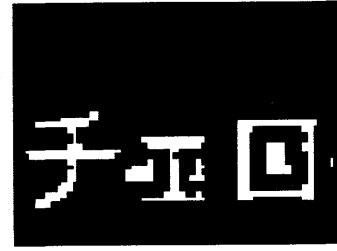


図 16 $T_h = 195$ のテロップ画像 627-701 の拡大図

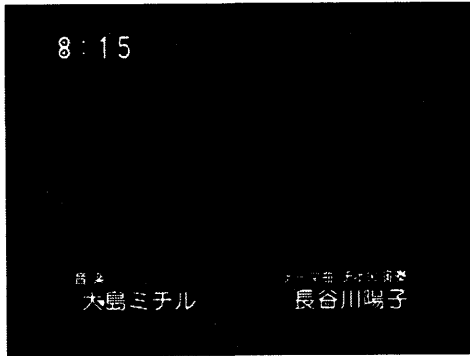


図 15 $T_h = 195$ で得られたテロップ画像 627-701

要画素を抽出してしまったり、濁音や半濁音といった文字の一部の抽出を失敗したりした。また、数種類のパラメータを今回は経験的に決定したが、これらを自動的に決定することはできていない。これらの問題を解決することが今後の課題である。

参考文献

- [1] H. D. Wactler, A. G. Hauptmann, M. G. Christel, R. A. Houghton, and A.M. Olligschlaeger, "Complementary video and audio analysis for broadcast news archives," *Commun. ACM*, vol. 43, no. 2, pp. 42-47, Feb. 2000.
- [2] 塩 昭夫: "情景中文字の検出のための自動2値化処理法", *信学論 (D)*, vol. J71-D, no. 5, pp.863-873, May 1988.
- [3] 大谷 淳, 塩 昭夫: "情景画像からの文字パターン抽出と認識", *信学論 (D)*, vol. J71-D, no. 6, pp.1037-1047, Jun. 1988.
- [4] 松尾 賢一, 上田 勝彦, 梅田 三千雄: "適応しきい値法を用いた情景画像からの看板文字列領域の抽出", *信学論 (D-II)*, vol. J80-D-II, no. 6, pp.1617-1626, Jun. 1997.

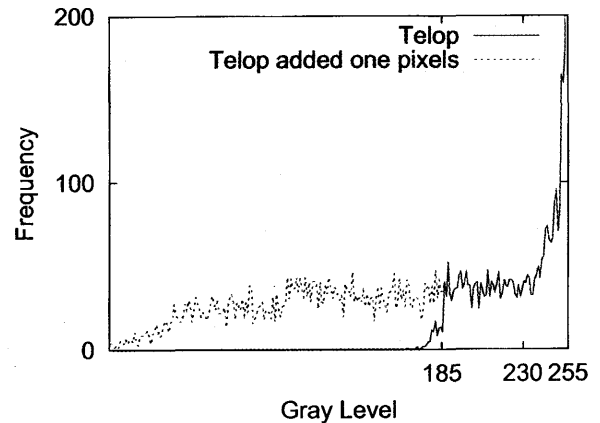


図 17 図 8 の濃度ヒストグラム

- [5] 劉 詠海, 山村 毅, 大西 昇, 杉江 昇: "シーン内の文字列領域の抽出について", *信学論 (D-II)*, vol. J81-D-II, no. 4, pp.641-650, Apl. 1998.
- [6] 高橋 裕樹, 金 大祐, 中島 正之: "ハングルの構造特徴を用いた情景画像からのハングル文字列領域抽出", *信学論 (D-II)*, vol. J88-D-II, no. 9, pp.1808-1816, Sep. 2005
- [7] 芦田 和毅, 永井 弘樹, 岡本 正行, 宮尾 秀俊: "情景画像からの文字抽出", *信学論 (D-II)*, vol. J88-D-II, no. 9, pp.1817-1824, Sep. 2005
- [8] 森 稔, 倉掛 正治, 杉村 利明, 塩 昭夫, 鈴木 章: "背景・文字の形状特徴と動的修正識別関数を用いた映像中テロップ文字認識", *信学論 (D-II)*, vol. J83-D-II, no. 7, pp.1658-1666, Jul. 2000.
- [9] 堀 修, 三田 雄志: "テロップ認識のための映像からのロバストな文字部分抽出法", *信学論 (D-II)*, vol. J84-D-II, no. 8, pp.1800-1808, Aug. 2001.