

テロップ認識のための映像からの文字部抽出法

堀 修

(株) 東芝 研究開発センター 情報・通信システム研究所
〒 210 川崎市幸区小向東芝町 1
osamu.hori@toshiba.co.jp

本稿では、映像中に現れるテロップの文字部をOCRで判読できるように、精度良く切り出す方法を提案する。従来の研究では、映像中のテロップを背景から切り出す方法としてテロップの輝度が背景に比べて高いことを利用し、閾値を決めて二値化する方法がとられていた。閾値の決め方として、背景と図の2つの輝度の分布を仮定した大津の方法や大津の方法を局所ブロックに用いた塩の方法を用いていた。しかし、映像のように背景に様々な輝度を含むものでは、必ずしも仮定が成り立たず、良い閾値が得られないという問題があった。また、実際には、雑音や処理変換の影響で文字の周辺で必ずしも輝度が高くなく、OCRで判読可能な十分な精度で文字を切り出すことができなかった。そこで、文字部の輝度の分布をロバストに推定し、文字部として信頼度の高い領域を最初に抽出し、推定された分布に基づいて領域拡張を行い、文字部のみを切り出す方法を提案する。実験の結果、従来手法より精度良く文字部を検出できることを確認した。

A video text extraction method for character recognition

Osamu Hori

R & D Center, TOSHIBA Corporation
1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki, Kanagawa 210, Japan

This paper presents a method to precisely extract video text portions from a video text rectangle region for making a readable image for OCR. In the conventional methods, the gray image binarization processing with a given threshold is employed to extract high intensity video text portions. Two methods are usually used to decide a threshold for the gray image binarization. One is the Ohtsu method and the other is the Shio method using the Ohtsu method locally in a video text region divided into small blocks. A video has a complex background with various intensities so that proper thresholds are not always obtained by the two methods. Furthermore, the character outskirts of a video text are not bright because such regions are blurred by image conversion and noise. Reliable high intensity portions in a video text region are extracted at first and expanded for making the whole video character regions. The experiments show the proposed method is superior to the conventional methods.

1 はじめに

デジタル映像時代に入り、テレビ放送において数百チャンネルを享受できるようになり、多数の番組から所望の番組を瞬時に選り出す必要が出てきた。それに伴いビデオオンデマンドシステムやビデオデータベースなどのデジタル映像に対して、映像の中に含まれる種々の情報を用いて検索のためのキーとする研究がなされてきた [1]。

その中で、映像の内容を表す密度の濃い情報としてテロップが目目されている [2, 3, 4, 5, 6, 7, 8, 9]。米国の放送の多くは、障害者保護の立場からクローズドキャプションと呼ばれる文字情報の付帯を義務づける法律があるため、会話されるほとんどの情報が文字コードとして送付され、それを簡単に取り出すことができる [1]。しかし、日本にはそのような法律がないために、映像中に書き加えられるテロップ情報は唯一の文字情報である場合が多い。そのため、米国の放送よりも日本の放送は、テロップを多用する傾向にある。テロップは、映像のシーンを端的に短い言葉で形容する重要な情報で、これを抽出し認識すれば検索に極めて有用な情報となる。

本稿では、映像中に現れるテロップの文字部をOCRで判読できるように、精度良く切り出す方法を提案する。以下では、従来の研究についてサーベイし、文字領域を精度良く取り出すアルゴリズムについて述べる。

2 従来の研究

映像中の文字列を切り出す技術は、自然画像の中の看板等に現れる文字の切り出し・認識 [10, 11, 12]、ナンバープレートの切り出し・認識 [13, 14] 及びレコードのジャケットや本の表紙に書かれた見出し文字の切り出し・認識 [15] の研究からはじまっており、近年のテロップの切り出し・認識の研究は、それらの要素技術を利用したものが多い。自然画像からの文字列の切り出しには、文献 [10, 13, 11, 12, 15, 14] などがある。また、映像からのテロップ切り出しを扱った研究は、文献 [1, 2, 3, 4, 5, 7, 8, 9, 16] などで紹介されている。

前者の自然画像からの文字列の切り出しは、背景は単色であり文字も単色を仮定している。よって、影や雑音の影響があるものの文字部と背景は、二値化の問題として扱われている。後者の映像から文字列の抽出に関しては、背景から文字部を切り出す以前に文字列の位置を検出する方法に焦点が置かれている研究が多い。この場合、自然画像から看板文字を抽出する手法と同様の方法を用いている研究 [2] もあるが、テロップ文字が自然画像中の看板文字と異なり、後から人工的にスパークインポーズされている特徴を積極的に利用

し、テロップ文字の単色性や文字と背景の境界にできるエッジに着目し、文字列位置を検出する手法が多い [1, 3, 4, 7, 8, 9]。

しかし、複雑な背景から文字部をOCRのために精度良く切り出す方法についての研究は、ほとんどない。自然画像から看板文字を切り出す方法としては、文献 [11] が動的な二値化法によって文字部を背景から精度良く分離している。しかし、これは背景と文字部が局所的に二値化の問題として扱えることを利用しており、その閾値の決め方として大津 [17] の方法を用いている。大津の方法は輝度ヒストグラムが背景と文字で二つの山ができ、二つの山を分離する最適な閾値を自動的に計算する方法である。文献 [2] は、背景からの文字部の切り出しを塩 [11] の方法を用いて行っている。塩の方法は画像をメッシュに区切り、メッシュ毎に大津の方法で閾値を求め、文字と背景の境界を含むメッシュの閾値をそうでないメッシュに伝播することで、すべてのメッシュでの二値化の閾値を求めている。この方法を映像にスパークインポーズされたテロップ文字に適用すると、かえって文字と関係ない背景の雑音が抽出されてしまう傾向がある。それは、背景が看板文字のように単色でなく様々な輝度を含むからである。

佐藤 [6] は、映像の中のテロップ文字を切り出す方法として文字幅に合わせたブリッジフィルタを用いることで、文字を強調させ切り出している。アルファベットのような単純な文字で、太さが既知のものは良いが、漢字のように太さが様々で込み入った文字に関しては、文字を均一に強調することは困難である。

よって、漢字を含む複雑な形状をもった文字を複雑な映像背景から切り出すためには、背景が多数の輝度を含むことを前提に処理を考えなくてはならない。以下では、簡単に文字部の領域方法について触れ、提案する文字部を精度良く切り出す方法について詳細に述べる。

3 処理の流れ

まず、処理の概要を説明する。処理の対象となるデータはMPEG2のデータを用いる。それは、MPEG2がデジタル放送やDVDで用いられる世界標準フォーマットで、今後デジタル映像のほとんどがMPEG2データ形式を採用するからである。また、漢字を認識する観点からMPEG2程度の解像度(704X480)が必要である。

図1に処理の流れの概要を示す。まず、MPEG2データから処理するための画像データを取り出す。次に、文字候補位置を検出しその文字候補位置に基づいて、文字列候補位置を検出する。文字列候補位置の画像から文字領域と背景部を分離し、分離された文字領域をOCRで読みコード化する。

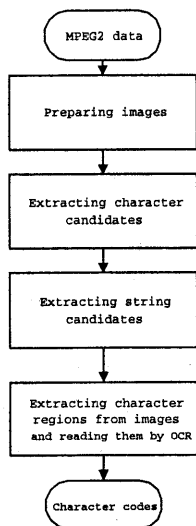


図 1: 処理全体の流れの概要を示す図

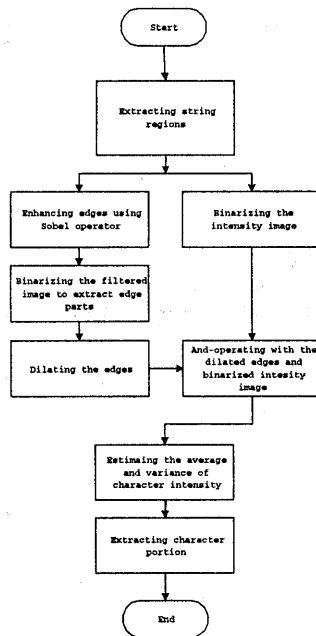


図 2: 提案手法の処理の概要を示す図

4 テロップ文字の位置・領域の検出

テロップ文字の位置及び領域の検出は、MPEG2の映像フレームを0.5秒に一回サンプリングし4枚の画像の平均したものを用いた。2秒以上静止しているテロップを対象にし、テロップ位置の抽出方法は、従来手法 [1] と同じように、テロップ文字がエッジを持っていて輝度が高いことを利用し、エッジ抽出を行って密集している部分を抽出し、その場所が輝度の高い領域かどうかで文字部かどうかを判定し、テロップ領域を抽出した。

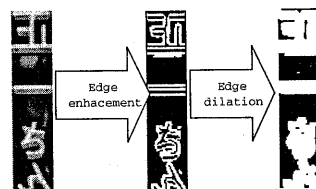


図 3: エッジ膨張によるテロップ文字部領域の推定

5 提案手法によるテロップ文字部の切り出しと認識

5.1 処理の概要

処理の概要を図2に示す。テロップ文字部のみを映像背景から、OCRが判読可能な精度で切り出すには、文字部の輝度分布を正しく推定する必要がある。そのためには、背景の影響をできるだけ排除し、文字部の輝度を推定する。テロップ文字部が存在する可能性の高い部分のみを抽出するため、テロップ文字が輪郭の部分で鋭いエッジを持つことを利用し、エッジ抽出を行い、エッジを膨張させた領域をテロップ文字を多く含む領域とする(図3)。その領域は、主にテロップ文字部、輪郭部、背景の3つの輝度から成り立っていると仮定する。そして、大津方法を多値に拡張した方法

で、3峰性の輝度ヒストグラムを分離する2つの閾値を決定する。ここで、文字部は輝度の高い方であると仮定し、輝度の高い領域を取り出し分布(平均と分散)を推定する。しかし、文字部以外の輝度値が外乱として含まれているので、精度を上げるために、ロバスト推定を用いて文字部の平均と分布を推定する。テロップ文字部の背景からの切り出しは、テロップ文字が文字部の輝度よりも低い輝度を持つ輪郭に囲まれている性質を利用し、輝度が高くかつエッジを膨張させた領域に含まれる部分をまず最初に抽出し、推定されたテロップ文字部の輝度値に基づいて、その領域から膨張させてテロップ文字全体を抽出する。

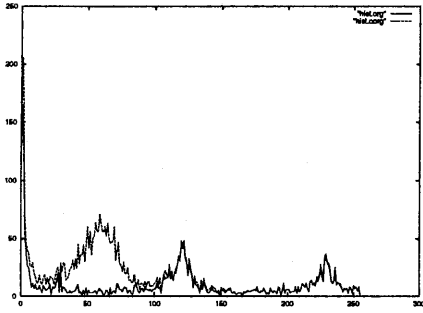


図 4: 破線が領域全体の輝度ヒストグラム。実線がエッジ膨張領域の輝度ヒストグラム

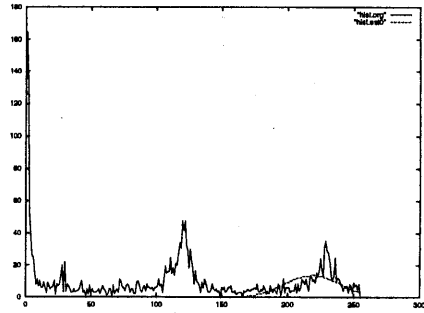


図 5: エッジ膨張領域の輝度ヒストグラムと文字領域を推定して正規分布と仮定して求めた曲線

5.2 文字部輝度分布の大まかな領域推定

図 4 に示すように、破線は領域全体の輝度ヒストグラムを示している。これに対して、エッジを膨張させた領域のみを輝度ヒストグラムを作成すると背景の輝度の多くが対象外になり、実線で示すように、背景、輪郭、文字部の三峰性のヒストグラムになる。

次に、このグラフを利用して輝度の高い文字部の部分の大まかな領域を取り出す。一般に大津の方法は、二峰性のヒストグラムを分ける閾値を決定する方法として知られているが、多峰性のヒストグラムを分離する方法にも容易に拡張が可能である [17]。今回は、背景、輪郭、文字部の三峰性の輝度ヒストグラムを 3 つに分離する 2 つの閾値を選択する問題として解いた。図 5 に対して 2 つの閾値は、それぞれ 64 と 167 であり、背景、輪郭、文字部を分離した閾値となっている。しかし、一番高い方の輝度領域の輝度分布を単純に切り出して、平均と分散を推定すると曲線のようになる。これは、裾の部分で他の分布からの要素が混じっていることと、外乱のために正しく分布が推定されていない。正しくは、EM 法などで 3 つの正規分布として推定すべきであるが、図からわかるように、外乱の影響によって、精度良く推定できるとは思えない。そこで、ひとつの分布のみを推定するのであれば、裾の分布自身も広い意味で外乱と考え、外乱に強いロバスト推定を用いて高い輝度の分布のみを見積もることとする。

5.3 ロバスト推定で文字部輝度を推定する手法

5.3.1 M 推定法

ロバスト推定の一つである M 推定法を用いて、外乱にロバストな輝度分布の推定を行う。以下に最小二乗誤差の拡張である M 推定法について説明する。誤差分布が正規分布に従うとき、最小二乗誤差を最小にするようにパラメータを推定することが、最尤推定であることが知られている。しかし、現実の世界では観測値に、観測したい事象以外のサンプルが含まれ、それが外乱 (アウトライヤー) となることが多い。最小二乗誤差はそのような外乱に敏感で、たとえ外乱が少数でも大きく誤った推定を行ってしまうため、外乱に強い推定方法としてロバスト推定が考案された。M 推定は、二乗誤差最小を基準とした推定法の自然な拡張になっている。まずは、最小二乗誤差を基準とした推定方法について述べる。

5.3.2 最小二乗誤差基準

$$S(\mathbf{x}) = \sum_{i=1}^n (y_i - f_i(\mathbf{x}))^2 / \sigma_i^2 = \min \quad (1)$$

ここで、 n 個の測定値 $y_i (i = 1..n)$ は、真値 y_i^0 で分散 σ_i^2 とする。誤差の確率分布は $P(y_i; y_i^0, \sigma_i) = P((y_i - y_i^0) / \sigma_i)$ と書くことができ、実際は、これが正規分布となるため $(2\pi\sigma_i^2)^{-1/2} \exp[-(y_i - y_i^0)^2 / 2\sigma_i^2]$ となる。また、ここで、測定値から平均値を引いたものを以下の式で表し残差と呼ぶことにする。

$$v_i = y_i - \hat{y}_i = y_i - f_i(\hat{\mathbf{x}}) \quad (i = 1..n) \quad (2)$$

$$S(\mathbf{x}) = \sum_{i=1}^n (v_i(\mathbf{x})/\sigma_i)^2 = \min \quad (3)$$

また、重み $w_i = 1/\sigma_i^2$ ($i = 1..n$) を導入し、書き換えると、

$$S(\mathbf{x}) = \sum_{i=1}^n v_i^2(\mathbf{x})w_i = \min \quad (4)$$

と表される。すなわち、誤差 σ_i で規格化された残差の二乗誤差和を最小にする条件、または重み付きの残差二乗誤差を最小にする条件である。このような最小二乗誤差を満たすパラメータ $\hat{\mathbf{x}}$ を求めるのが最小二乗誤差法である。

5.3.3 線形モデルでの最適パラメータの決定法

ここで、 $f_i(\mathbf{x}) = \sum_{j=1}^m A_{ij}x_j$ の線形モデルで表すと最小二乗誤差の式は以下になる。

$$S(\mathbf{x}) = \sum_{i=1}^n (y_i - \sum_{j=1}^m A_{ij}x_j)^2/\sigma_i^2 = \min \quad (5)$$

各パラメータ x_j に関して最小となるためには、その微分が 0 となればよい。

$$0 = \frac{\partial S(\mathbf{x})}{\partial x_j} = -2 \sum_{i=1}^n v_i A_{ij}/\sigma_i^2 \quad (j = 1..m) \quad (6)$$

となり、この式を解くことになる。

5.3.4 M 推定への拡張

誤差分布が正規分布でない一般の場合を考え、最適なパラメータ推定を最尤度の観点から考える。まず、パラメータ推定値 $\hat{\mathbf{x}}$ に対する尤度は、

$$L(\hat{\mathbf{x}}|y, \sigma^2) = \prod_{i=1}^n P(y_i; f_i(\hat{\mathbf{x}}), \sigma^2) \quad (7)$$

$$= \prod_{i=1}^n P(|y_i - f_i(\hat{\mathbf{x}})|/\sigma) \quad (8)$$

である。尤度最大の条件は、対数尤度の形で表して、

$$\log L(\hat{\mathbf{x}}|y, \sigma^2) = \sum_{i=1}^n \log P(|y_i - f_i(\hat{\mathbf{x}})|/\sigma) = \max \quad (9)$$

となる。これをパラメータ x_j ($j = 1..m$) で微分すると、方程式

$$\sum_{i=1}^n (A_{ij}/\sigma_i) \Psi(v_i/\sigma_i) = 0 \quad (10)$$

が得られる。ただし、 A_{ij} はヤコビアン係数 $\partial f_i/\partial x_j$ であり、 v_i は残差、関数 $\Psi(z)$ は

$$\Psi(z) = -d \log P(z)/dz \quad (11)$$

である。もし、誤差分布が正規分布ならば、 $\Psi(z) = z$ となり、式は、

$$\sum_{i=1}^n A_{ij}v_i/\sigma_i^2 = 0 \quad (12)$$

となり、最小二乗誤差法と同じになる。一方、もし正規分布以外の誤差分布を考えれば、式はその誤差分布に対する最尤推定法の基本式となり、最小二乗誤差法とは異なる。さらに考えを一般化し、誤差分布 $P(z)$ から離れて任意の関数 $\Psi(z)$ をとり上げ、以下のような式に変形する。

$$\sum_{i=1}^n A_{ij}v_i [\Psi(v_i/\sigma_i)]/\sigma_i^2 = 0 \quad (j = 1..m) \quad (13)$$

これは、最小二乗誤差法の重み $w_i = 1/\sigma_i^2$ の代わりに、有効重み

$$w_i^{eff} = [\Psi(v_i/\sigma_i)]/\sigma_i^2 = [\Psi(z_i)/z_i]w_i$$

where $z_i = \frac{v_i}{\sigma_i}$

を用いたものと同等になる。ただし、この有効重みは残差に依存しているので、パラメータの推定値の変化に応じて有効重みを変化させる必要がある。このように有効重みを調節しつつ、重み付き最小二乗誤差法を反復すれば、最尤法をもととする任意の当てはめ法が実現できる。

5.3.5 有効重み調節因子

実際に使用した有効重み調節因子

$$w_i^{adj} = w_i^{eff}/w_i = \Psi(z_i)/z_i \quad (14)$$

として Biweight という関数を用いた。

$$\Psi(z) = \begin{cases} z[1 - (z/c)^2]^2 & \text{for } |z| < c\pi \\ 0 & \text{for } |z| \geq c\pi \end{cases} \quad (15)$$

Biweight 関数の図 6 に示す。

Biweight 関数を用いた有効重み調節因子は以下の特徴がある。

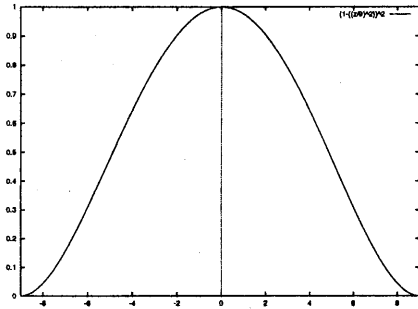


図 6: Biweight 関数

1. 規格化残差 $|z_i|$ が小さき時は、最小二乗法と同等の重み付けを持つ。
2. 規格化残差 $|z_i|$ が極めて大きい時は、重みがゼロとなる。
3. その中間では、重みが連続して変化する。

5.3.6 M推定法を用いた解法手順

M推定法を実際使用するためには、次の手順で行う

1. パラメータの初期値 $\hat{x}^{(0)}$ を選ぶ。初期値は最小二乗法を用いて求めてその値を使用する。
2. 残差 v_i を計算し、測定精度で規格化し z_i とする。
3. 規格化残差 $|z_i|$ の平均的な大きさ s を中央値を使って見積もる。 $s = \text{median}\{|z_i|\} = \text{median}\{|v_i/\sigma_i|\}$
4. 各測定値の残差 z_i に対応して、有効重み調節因子 w_i^{adj} を計算する。このときのスケール s は Biweight を用いる。
5. 有効重み w_i^{eff} を用いて、最小二乗解法によりパラメータの補正項 $\Delta\hat{x}$ を計算する。
6. パラメータの値を $\hat{x}^{(k+1)} = \hat{x}^{(k)} + \Delta\hat{x}$ とし、上記の過程を繰り返し、 \hat{x} を求める。

今回は、輝度ヒストグラムに関してロバスト推定を行うため x は一次元である。

5.3.7 M推定法を用いた文字部の輝度分布の推定例

大津の方法で求めた大体の文字部の部分を初期値として、M推定法を用いて反復処理を行い、より精度の高い分布を求める。図5で示す結果を初期値として、文字部の輝度分布をM推定を用いて処理した結果が図7である。文字部の輝度分布が、裾の外乱の影響を受けずに正しく推定されていることがわかる。

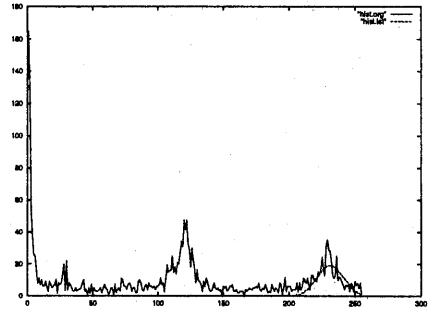


図 7: ロバスト推定によって文字部の輝度分布の推定した結果

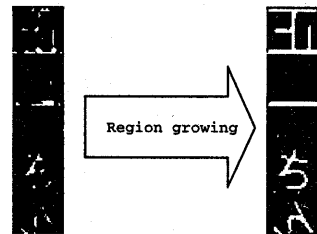


図 8: 本手法による文字領域の抽出例

5.3.8 推定された文字の輝度分布を利用した文字部の切り出し法

次に、求められた文字部の輝度分布を用いて、比較的安定している輝度の高い画素を文字領域の一部と仮定し、その領域を種に文字領域を拡張させて文字を取り出す。よって、推定された輝度の平均 (m) と分散 (σ) から $m + \sigma$ よりも高い輝度を持ちかつエッジを膨張させた領域にある画像領域をまず抽出し、その画素の8近傍を検査し、輝度が $m - 3\sigma$ よりも大きな値をとる場合は、文字の領域として併合していく。この処理は、以下の仮定に基づいている。膨張領域の中では文字領域以外の場所で、 $m + \sigma$ よりも高い輝度を持つ領域がない。文字を構成する画素の輝度は $m - 3\sigma$ 以上の輝度を持つ。文字と背景の境界は $m - 3\sigma$ よりも低い値の輝度で囲まれている。この条件が満たされている文字部は、完全に背景から切り出すことができる。

図8は、本手法を用いて文字領域を抽出した結果である。原画に対して輝度の高い領域のみを抽出した例が真ん中の図である。その領域を元に $m - 3\sigma$ 以上の輝度を持つ周辺の領域を併合した結果が最右図である。文字部だけ抽出されていることがわかる。検出した文字領域は、OCRへ入力し文字コードとして答えを得る。

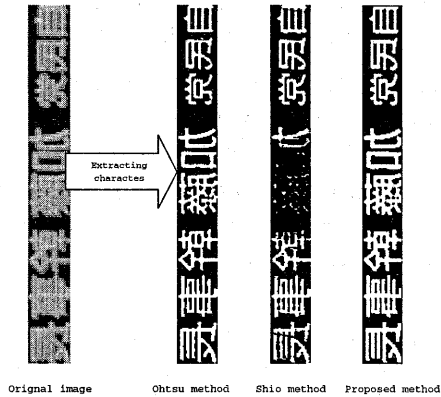


図 9: 背景が単純なテロップに対して 3 種類の文字検出を適用した例

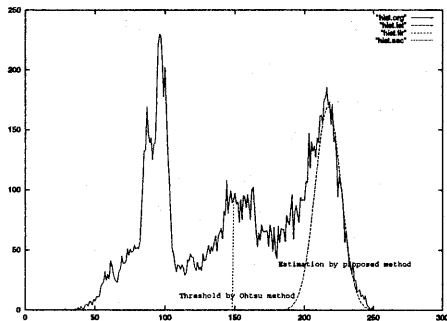


図 10: 原画の輝度ヒストグラムと輝度分布推定

6 実験例

(1) 背景が単純で背景と文字からなるテロップの例
 比較的、背景と文字部がはっきりと分かれた例を対象に大津の方法、塩の方法、提案手法で処理した結果を示す。図 9 では、一番左に原画を示している。順番に大津の方法、塩の方法、提案手法を示した。このように、背景と文字部がはっきり分かっている例では、大津の方法でも良好に文字部を取り出すことができる。一方、塩の方法では局所的に大津の方法を用いて閾値を決めているため、条件が悪いと局所的に閾値が合わずにかえって文字がきれいに二値化できない場合がある。提案手法では、大津の方法と変わらない程度に文字部がきれいに二値化されている。提案手法は、三値を仮定しているが、境界あたりの部分が 3 番目の峰になっているので、抽出ができています。

図 10 は、原画の輝度ヒストグラムである。背景と文字部しかないので比較的大津の方法でも良好な閾値が設定されている。文字領域の輝度分布を推定した結果も合わせて表示されている。この場合も比較的良好的に

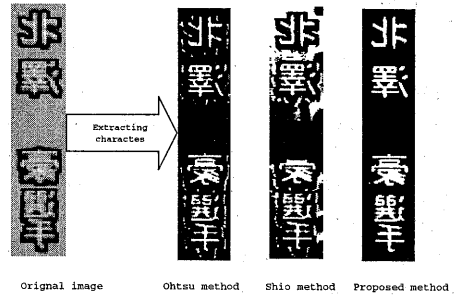


図 11: 3 種類の輝度からなるテロップに対して 3 種類の文字検出を適用した例

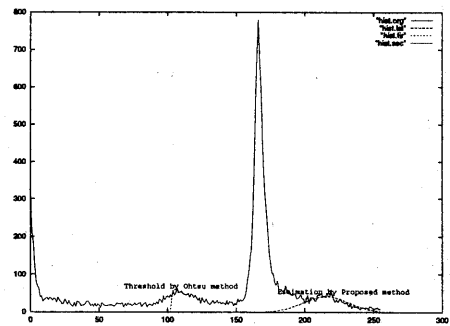


図 12: 原画の輝度ヒストグラムと輝度分布推定

推定されている。

(2) 背景、文字、文字の縁の 3 種類の濃度が混在しているテロップの例

図 11 に示した例は、テロップの文字分布が背景、文字部、文字の縁の 3 種類から構成されている例である。図 12 に示すように輝度分布が 3 つの領域の分割されているため、大津のように二峰性を仮定した方法では誤った二値化を行ってしまう。塩の方法も基本的には大津の方法を局所的に適用しているため、誤った二値化を行う。提案手法では、高い輝度の一峰を検出するので正しく文字部を検出している。

7 おわりに

本稿では、映像からテロップ領域を抽出後、文字部のみを背景から精度良く分離するアルゴリズムについて述べた。今後の課題は、この方法を拡張し濃度の低い色文字及びグラデーションのかかった文字に対しても適用できるようにすることである。

参考文献

- [1] Michael Simth and Takeo Kanade. Video skimming and characterization through the combination of image and language understanding techniques. In *In proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 775–781, 1997.
- [2] 倉掛正治, 桑野秀豪, 新井啓之, 小高和己. 認識技術を用いた映像中のキーターゲットインデクシングの検討. 信学技報, Vol. PRU95, No. 237, pp. 15–20, April 1994.
- [3] Rainer Lienhart and Frank Stuber. Automatic text recognition in digital videos. In *In proceedings of Soc. Opt. Eng. Image and Video Processing IV*, Vol. 2666, pp. 180–188, 1997.
- [4] Rainer Lienhart. Automatic Text Recognition for Video Indexing. In *In proceedings of ACM Multimedia 96*, pp. 11–20, 1996.
- [5] Y. Ariki and T. Teranishi. Indexing and Classification of TV News Articles Based on Telop Recognition. In *In proceedings of International Conference on Document Analysis and Recognition*, pp. 422–427, 1997.
- [6] Toshio Sato, Takeo Kanede, Ellen K. Hughes, Michael Smith, and Shin'ichi Satoh. Video OCR for digital news archive. In *In proceedings of International Workshop on Content-Based Access of Image and Video Database*, pp. 52–60, 1998.
- [7] Ullas Gargi, Sameer Antani, and Rangachar Kasturi. Indexing Text Events in Digital Video Databases. In *In proceedings of International Conference on Pattern Recognition*, pp. 916–918, 1998.
- [8] Jae-Chang Shim, Chitra Dorai, and Ruud Bolle. Automatic Text Extraction from Video for Content-Based Annotation and Retrieval. In *In proceedings of International Conference on Pattern Recognition*, pp. 618–620, 1998.
- [9] Anil K. Jain and Bin Yu. Automatic Text Location in Images and Video Frames. In *In proceedings of International Conference on Pattern Recognition*, pp. 1497–1499, 1998.
- [10] 滝沢圭, 仙田修司, 美濃導彦, 池田克夫. 動画像からの看板文字パターン列の抽出. 信学技報, Vol. PRU94, No. 133, pp. 25–32, March 1995.
- [11] 塩昭夫. 情景中文字の検出のための動的 2 値化処理法. 信学論D, pp. 863–873, May 1988.
- [12] 松尾 賢一梅田 三千雄. 濃淡及び色情報による情景画像からの文字列抽出. 信学技報, Vol. PRU92, No. 121, pp. 25–32, January 1992.
- [13] 崔享振, 安居院猛, 中島正之, 横前高広. 動画像処理による通行車両とそのナンバープレート領域の抽出. 信学論D, Vol. J71-D-II, No. 3, pp. 392–398, 1990.
- [14] Cui Y. and Huang Q. Character extraction of license plates from video. In *In proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 502–507, 1997.
- [15] 上羽 葵岡田 至弘. 等高線処理によるカラー画像からの文字領域の抽出. 信学技報, Vol. PRU94, No. 28, pp. 9–16, September 1994.
- [16] Jun Ohya, Akio Shio, and Shigeru Akamatsu. Recognizing Character in Scene Images. *PAMI*, Vol. 16, No. 2, pp. 214–220, 1994.
- [17] Jun Otsu. A threshold selection method from gray-scale histograms. *IEEE Trans. Syst Man Cybern.*, Vol. SMC-9, No. 1, pp. 62–66, 1986.