

案内看板文字の傾きの自動補正*

4 N-03

費 立徳 † 長井 隆行 ‡ ‡ 横松 明 ‡ ‡ ‡

電気通信大学大学院 電気通信学研究科 ¶

1はじめに

現代の情報化社会では文字データの処理が急速に増加しており、看板にかかれている文字情報を認識することができれば、翻訳や、通信を介しての関連情報の取り出しなどが可能となる[1]表示看板の文字認識は、大きな技術課題である。本研究の目的は持ち運びできる情報通信端末の画像カメラからの映像を処理して、映像文字パターンの認識結果を音声あるいは文字で表示して高齢者や外国人の案内を支援することである。文字の認識において、各構成文字の補正是認識の精度向上を図る上で必要技術である。実際に撮像した情景画像中の文字は位置によって文字の形が違う。正面に向けて撮った写真は変形がないが、上、下、左、右から撮った写真は看板の形が変形する。よって看板内にある文字も変形する。また、わざわざ文字を傾ける場合もある。もしそのまま認識すれば認識率は低い。そこで看板の中にある文字の傾きを補正する必要がある。

2看板文字の特徴

実際に撮った看板文字に対しては、多様性がある。文字並び方向の歪み、文字形状の傾き、文字の透視の歪みが同時に存在している。また 文字の大きさ、種類、看板にある位置、活字体文字、手書き文字など色々な場合がある。以上の問題に対して、文字の横の線、縦の線また文字行列の輪郭の形によって以下に示す文字補正を行う。

3補正

実際的に撮った写真は位置によって取った写真は变形があるのが普通である。看板の中にある文字を補正する必要がある。まず文字領域を抽出した二値化画像に対して文字のエッジを切り出して次に 0 ~ 179 度まで 1 度づづハフ変換して、回転補正と傾き補正を行う。最後に透視補正をする。

3.1 前処理

普通の文字の二値化画像では背景が黒く文字部分が白い。しかし 黒背景にある白文字と白背景にある黒文字が一緒に入っている場合もある。“白”

背景に黒い文字がある場合に補正しにくい。また、もし混在する場合には正しく補正しても両方同時に文字を認識することが難しい。このような問題を解決するためにまず、二値化文字画像をラベリングして面積の大きさがある値以上になるものを反転候補とする、次にある値以上の穴が付いている候補を反転する。ある値以上の穴が付いている候補だけを処理する理由はある文字が大きくても一つの文字には穴が少なく黒文字の付いている白看板の場合にはたくさんの穴がついていることによる。これにより両者の区別ができる、前処理が行う。

3.2 Hough変換

ハフ変換を用いて直線性の抽出を行う（式 1）。xy 空間での 1 点はハフ空間では正弦曲線で表現される。直線が不連続にある程度ばらついていても比較的安定に分布が顕著な複数の直線を同時に求めることができる。

$$\rho = x \cos \theta + y \sin \theta \quad (1)$$

本研究では θ は 0 度から 179 度までの有限の範囲を細かく分割すればよく、また ρ もある程度の大きさまで取ればよいので、ハフ変換による直線抽出は (θ, ρ) 平面上で実行されるのはエッジを抽出した看板文字のハフ変換の例であるが普通である図 1。

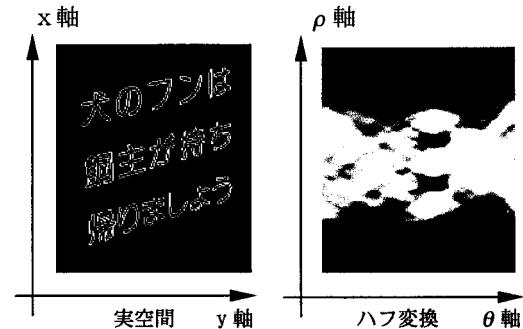


図 1 文字のエッジのハフ変換

3.3 回転補正

回転は θ 方向を考えればよい。従って、0 度から 179 度までの Hough 変換を行う。Hough 平面で各値の自乗を計算し、それぞれの θ に対する ρ の自乗の値の総和でヒストグラムを作り、ピーク位置から回転量を算出することにする。

普通は縦の線または横の線が同じ方向に向かっている。 θ が文字の線の角度と同じになれば、ヒストグラムのピークのところに対応する、従って文

*Automatic Adjustment of Inclination for Characters on Signboard
† Fei Lide

‡ Takayuki Nagai
‡ ‡ ‡ Akira Kurematsu

¶ University of Electo-Communications、Graduate School of Electro-Communications

字列方向が分かる。回転量により文字行列の並び方向の歪みが補正できる。この中で一番大きな総和に対応する角度を α とする。もし α が 45 度から 135 度までの間に入ったら、行補正を行う、それ以外の場合は列補正を行う(表 1)。

α	補正方向	補正角度
$0^\circ \leq \alpha < 45^\circ$	列補正	右回り α
$45^\circ \leq \alpha < 90^\circ$	行補正	左回り $90^\circ - \alpha$
$90^\circ \leq \alpha < 135^\circ$	行補正	右回り $\alpha - 90^\circ$
$135^\circ \leq \alpha \leq 179^\circ$	列補正	左回り $180^\circ - \alpha$

表 1: 文字行列の補正(但し、 α は文字行列に応じる角度である)

3.4 傾き補正

文字傾きの回転量を探すことは文字行列の回転量の探索より難しい。画像によりそれぞれ異なるヒストグラムとなる。あるヒストグラムにはいくつかのピークが存在する。一番高いピークに対応する角度を文字行列の補正角度としてよいのでだが、次にそれ以外にどのピークに対応する角度を文字傾きの補正角度とするのが問題である。もし二番目に高いピークに対応する角度を文字傾きの補正角度とすれば失敗する恐れがある。この問題を解決するために 文字傾きの回転量を β として範囲を設定して β の最高値を探す。実際には 变形のない文字では横線と縦線が垂直であるが、変形のある文字では横線と縦線が変化している。 α と垂直する角度を γ として γ の左右の範囲内(15 度以内)に β を探せばよい。また、 β が確かにその範囲の中に存在するにもかかわらずヒストグラムのピークが α のピークより小さい場合は、正しく β を検出ににくい。このような場合には近似的に γ を β とする、表 2 に傾き補正における文字の行あるいは列の傾き角度に対する文字画像の上下左右の移動方向を示す。

行補正		列補正	
$45^\circ \leq \alpha < 90^\circ$	$90^\circ \leq \alpha < 135^\circ$	$0^\circ \leq \alpha < 45^\circ$	$135^\circ \leq \alpha \leq 179^\circ$
$\beta - \alpha - 90^\circ$	$\beta - \alpha + 90^\circ$	$\alpha - \beta + 90^\circ$	$\alpha - \beta - 90^\circ$
If $\alpha > 0$: Then 上→ 下←	If $\alpha < 0$: Then 上← 下→	If $\alpha > 0$: Then 上→ 下←	If $\alpha < 0$: Then 左↓ 右↑
If $\alpha < 0$: Then 上← 下→	If $\alpha > 0$: Then 上→ 下←	If $\alpha < 0$: Then 左↑ 右↓	If $\alpha > 0$: Then 左↓ 右↑

表 2: 文字傾きの補正(α : 文字行列に応じる角度である、 β : 文字の傾きに応じる角度である)

3.5 透視補正

まず二値画像に対し、膨張を行う、文字領域の二値画像は値が 1 である画素に対して、横方向の前後

それぞれある一定画素の値を全て 1 に設定して、文字行列全体の輪郭を求める。しかし、輪郭の部分はまだぎざぎざしているのでこれに対して膨張と縮退を行う。ここでは、半径が 6 画素という円のようなフィルタを用いて滑らかな輪郭を得る。最後に二値化画像のエッジを切り出す。二値エッジ画像に対し、-70 度から、+70 度まで、5 度づつ 変換して 29 枚の画像を得る、透視変換された画像は幅が狭くなるから、それぞれの角度によって、元の幅に戻す。 $L = R/\cos(E)$ 。ただし、 L は元の幅、 R は透視変換された幅、 E は透視変換の角度である。次に平行性を判断する。透視変換したすべての画像の中から線と線の平衡性が最も高い画像を選択し、それに対応する透視変換角度を透視変換角度とする。

4 実験

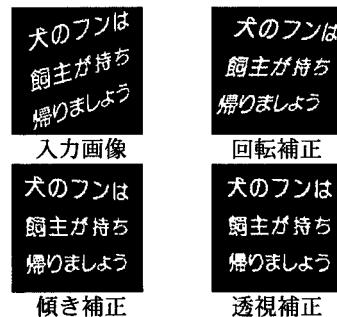


図 2 文字補正の例

2000 枚の看板を含む情景画像を収集し、その内 100 枚をテスト用に用いた、実験例を図 2 に示す。完全に補正できたのは 9.2%、よく補正されたのは 6%、ほぼ変わっていないのは 2% となり、効果がかなりあることがわかった。

5 おわりに

活字のような看板文字、手書きのような看板文字など、様々な例に対して本研究の方法は、有効性がかなりある。複雑な場合では文字以外のものも画像の中に一緒に入ってくる。補正した後でもまだ完全に補正されていない例も存在するので、再度同じような処理をする必要があるが、このような例は少ない。今後はさらにロバスト性を向上させるように改良する必要がある。

謝辞 本研究に関し、有意義なご意見、ご討論をいただきいた研究室各位ならびに関係諸氏に厚く御礼申し上げる。

参考文献:

- [1] H. Fujisawa, H. Sako, Y. Okada, and S-W. Lee, "Information capturing camera and development issues," Proc. ICDAR99, 1999, 9.