

ズーム倍率変化が情景画像中の文字認識に及ぼす影響の調査

向井田一平[†]松尾賢一[‡][†] 奈良工業高等専門学校 専攻科 電子情報工学専攻[‡] 奈良工業高等専門学校 情報工学科

1 はじめに

カメラで撮影した情景画像中の文字認識では、カメラと文字間の距離による文字パターンの解像度不足によって、認識精度が低下する問題 [1] がある。この解像度不足を回避する方法として、文字パターンに対する光学的なズームアップが有効と考えられる。

本研究では、光学カメラの倍率が文字認識に及ぼす影響について調査し、光学的ズームアップが認識結果に与える有効性を明らかにする。

2 倍率による認識精度の調査

情景画像中の文字パターンに対する倍率変化が文字認識の精度に与える影響を認識精度調査により明らかにする。まず、情景中に配置した文字列に光学ズームカメラを向け、ズーム倍率を変化させた複数枚の情景画像を取得する。次に取得した情景画像中から、看板領域を長井らの手法 [2] を改良した看板検出エンジンで検出する。その後、検出した看板領域内に存在する各文字パターンを切り出し、予め用意した辞書画像とのテンプレートマッチングを行う。以上の手順によって、文字パターンに対し倍率の異なる認識結果を得る。

情景中に配置した文字及び辞書の文字は、大文字アルファベットの 'A'~'Z'、フォントは 150pt の MS ゴシックとした。またカメラと文字間の距離は 4m、ズーム倍率は 1 倍と 10 倍の 2 種類とする。これらの条件で撮影した画像中の文字パターンは、1 倍時で幅が約 10pixel、高さが約 14pixel となる。またズームアップ後の文字パターンは、10 倍時で幅が約 90pixel、高さが約 150pixel となる。

3 調査結果・考察

調査の結果として、各倍率での認識率を表 1 に示す。

表 1: 各倍率における認識率 (認識率順にソート) [%]

倍率	A	C	D	E	F	J	K	L	O
1	100								
10	100								
倍率	P	R	S	V	Z	B	I	X	Y
1	100					67			
10	100								
倍率	G	N	Q	T	U	H	M	W	
1	33						0		
10	100								

調査の結果、全ての文字種において、倍率 10 倍時の認識率は 1 倍時より平均 29%、最大 100% もの向上が見られた。この認識精度向上の要因を明らかにするため、テンプレートマッチングで算出される類似度を、各倍率で比較する。'A' から 'Z' までの第 1 位類似度、第 2 位類

似度の平均値と、第 1 位と第 2 位との差を表 2 に示す。

表 2: 類似度の比較 (A-Z の平均)

倍率	正解文字の類似度	不正解文字の類似度 (最大)	差
1	0.511	0.429	0.117
10	0.913	0.582	0.331

テンプレートマッチングでは、正解文字と不正解文字との間において類似度の差が大きいほど、誤認識の可能性が低下する。そして調査の結果、倍率 10 倍時では、正解文字と不正解文字との類似度との差が 0.214 大きくなった。つまり、10 倍時では認識における正解文字と不正解文字との分離性が高まり、認識精度が向上したといえる。

また 1 倍時では、'M'、'N'、'W' など斜めのストロークがある文字の認識率が、縦横のストロークのみで構成された文字に比べて低かった。これはズームアップ前では解像度が不足しており、斜めのストロークが他のストロークと結合してしまったことが原因と考えられる。'G' や 'Q' も 1 倍時では文字ストロークの結合により、'C' や 'O' など似た形状の文字に誤認識してしまった。しかし、10 倍では解像度が高まったため、これらの文字の認識率も 100% に達したと考えられる。

4 まとめ

ズーム倍率を変化させたときの認識精度の調査結果から、高倍率化によって文字認識精度の向上が明らかになった。したがって、情景画像中の文字を対象とした文字認識には、光学ズームカメラによる文字パターンの高解像度化が有効だといえる。

また、本研究では情景画像から文字が確実に切り出せることを前提とした調査を実施した。しかし、実際は文字切り出しの精度も情景画像中の文字認識の精度に影響する。この文字切り出し精度についても、光学ズームが有効 [3] であるため、今後の予定は、光学ズームを用いて、新たな文字スポッティングシステムを開発する。

参考文献

- [1] 新村貴彦ら, "撮影距離の変化に対する情景内文字の認識特性", 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解, 101 巻 (303 号), p.21-27, 2001 年
- [2] 長井隆行ら, "情景画像中の文字及び看板領域の抽出", 電子情報通信学会技術研究報告 CS, 通信方式, 100 巻 (721 号), p.103-108, 2001 年
- [3] 向井田一平, 松尾賢一, "ズーム倍率変化に伴う文字着目精度の調査", 電子情報通信学会 関西支部学生会 第 19 回学生会研究発表講演会, 2014 年