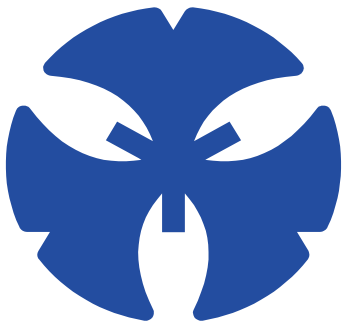


リアルタイム単語認識技術を利用した カメラベース情報取得システム

辻 智彦 岩村 雅一 黄瀬 浩一

大阪府立大学大学院工学研究科



大阪府立大学
OSAKA PREFECTURE UNIVERSITY

発表の流れ

- 背景
 - 文字認識手法の概要
 - 単語認識の提案手法
 - 実験
 - まとめ
-

提案システム

リアルタイムで動作
認識結果
関連情報

カメラ



"University"

ノートPC



翻訳

・大学

画像



音声



Intelligent Media Processing Laboratory
Department of Computer Science and Intelligent Systems
Graduate School of Engineering
Osaka Prefecture University



IMP

文書

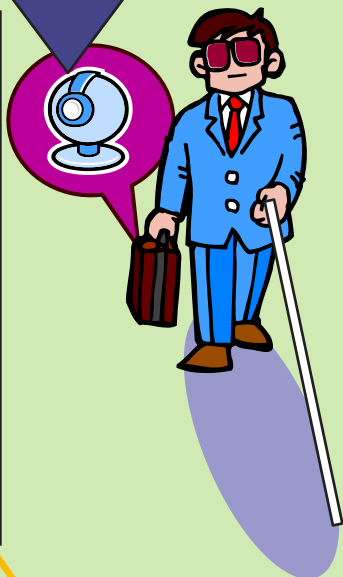
環境中文字列のリンクアンカー化

インタラクティブデモ 16:15～ (本発表のすぐあと)

有望な応用例

視覚障害者への音声案内

『押ボタン信号が
あります』



関連情報取得システム

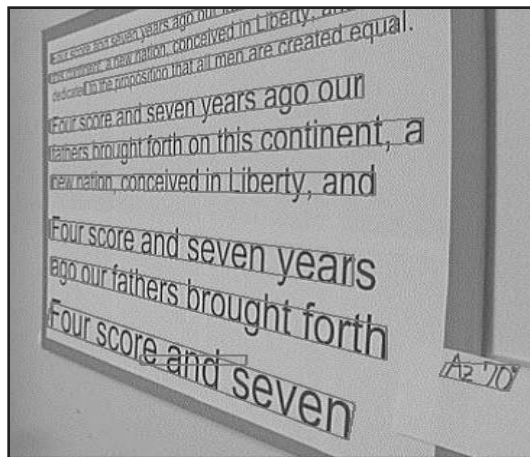
タンブラー
新デザイン発表!!



単語の認識が必要

多フォント対応・レイアウトフリー

単語の認識



認識対象

平行な文字行

行を成さない

既存手法 (Myers 2004)

行ごとに認識



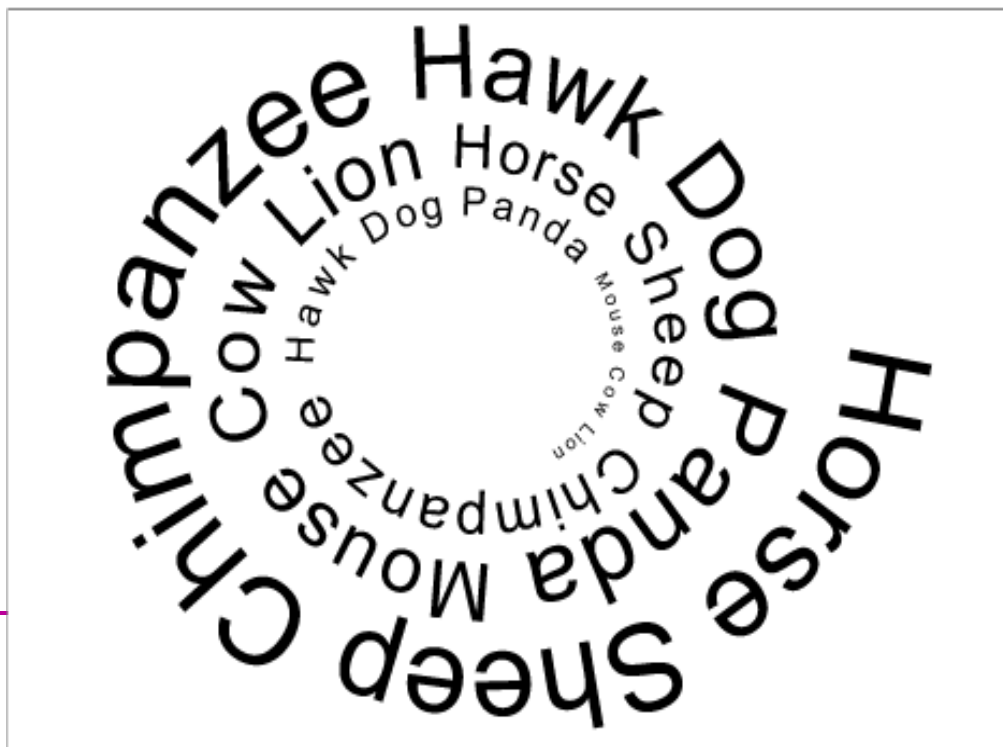
提案手法

1文字ごとに認識



問題設定

- 同一平面上の紙面
- 白紙に黒色の文字
- 連結成分を容易に切り出すことが出来る
- 単語間がスペースで区切られる
- 左から右に記される
- 文字行を成さない



認識の流れ

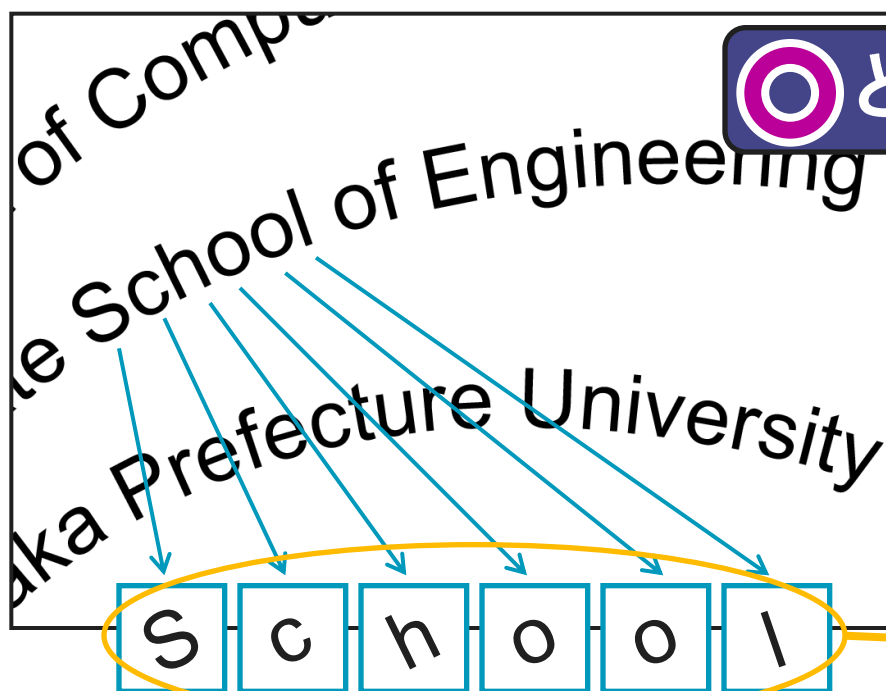
岩村 雅一, 辻 智彦, 黄瀬 浩一:
“カメラ撮影文字の事例に基づく実時間認識”
テーマセッション4 19日 11:45予定

① 1文字ごとに文字認識 概説

② 文字を連結して単語を推定 本発表の主題

○どこにどんな文字があるか

○どんな単語があるか



"School"

発表の流れ

- 背景
 - 文字認識手法の概要
 - 単語認識の提案手法
 - 実験
 - まとめ
-

文字認識の特徴

実時間処理が可能

射影歪みに頑健

レイアウトフリー

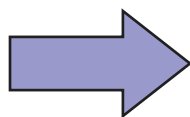
多種フォント登録可能



0123ABCDabcd
0123ABCDabcd
0123ABCDabcd
0123ABCDabcd
0123ABCDabcd

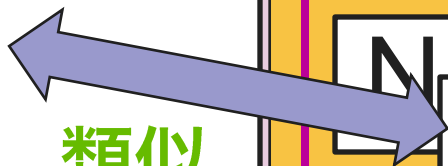
クラスレベルの識別

- 互いにアフィン変換の関係にある文字群は識別できない

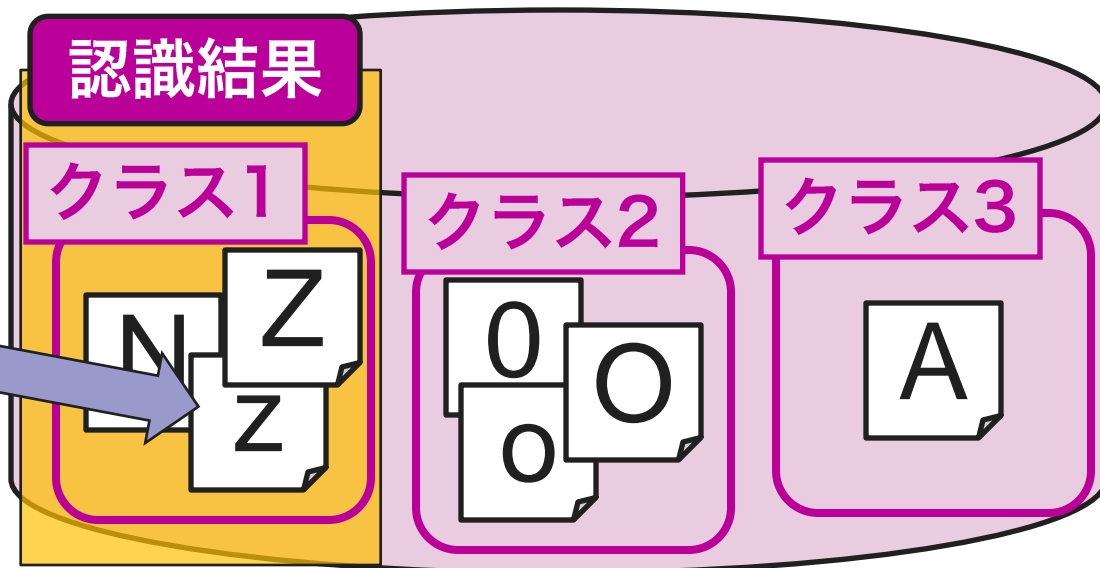


1つのクラスに統合

キャプチャ画像



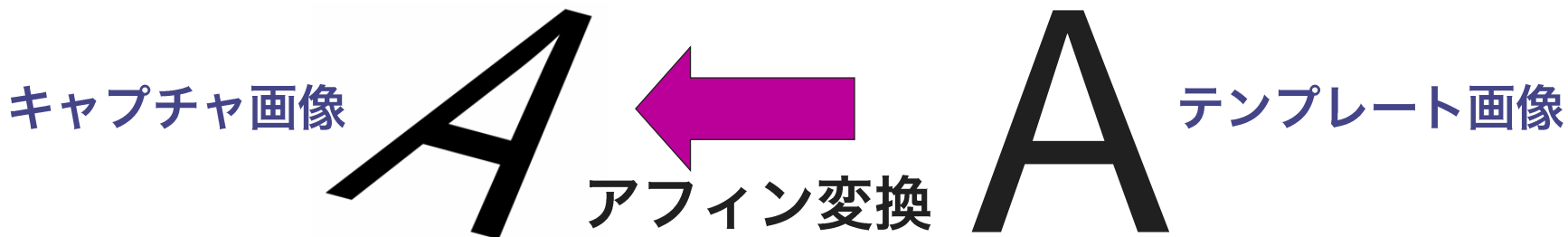
類似



データベース

姿勢推定

- 認識時にアフィン変換パラメータが計算可能



歪みを取り除く

紙面の姿勢

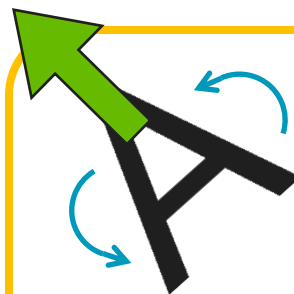


独立変倍

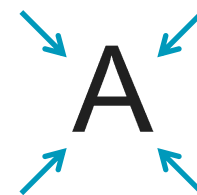


シアー

文字の姿勢



回転



拡大・縮小

単語推定で利用

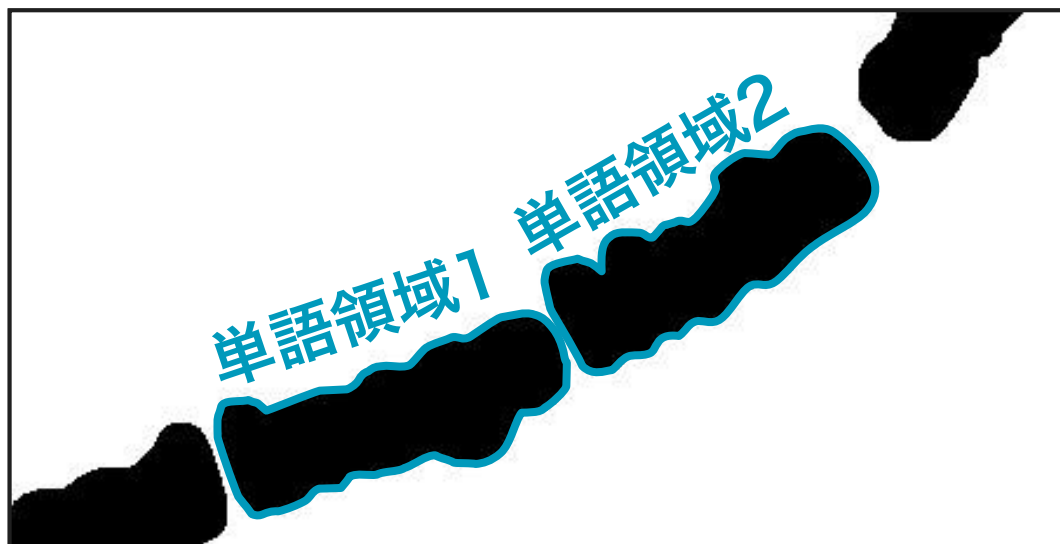
発表の流れ

- 背景
- 文字認識手法の概要
- **単語認識の提案手法**
- 実験
- まとめ

単語領域の推定

- 画像をぼかして隣接文字を結合させる
- ぼかし画像の連結成分を単語領域とする

文字間距離・連結成分面積によって
逐次的にぼかし強度を決定



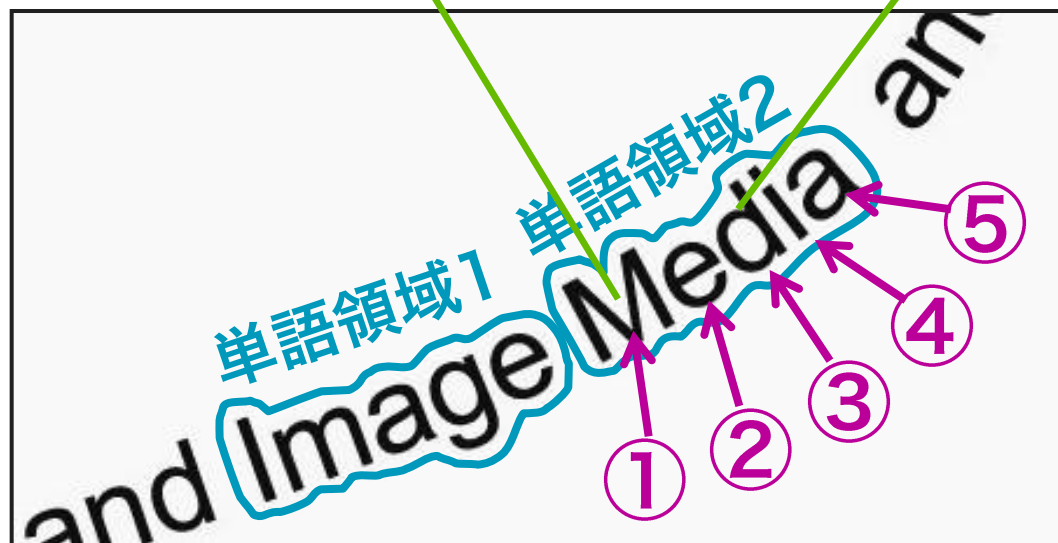
構成する文字の決定

- 単語領域内に含まれる文字を調べる

クラスレベルの
識別結果

M or W ?

d or p ?



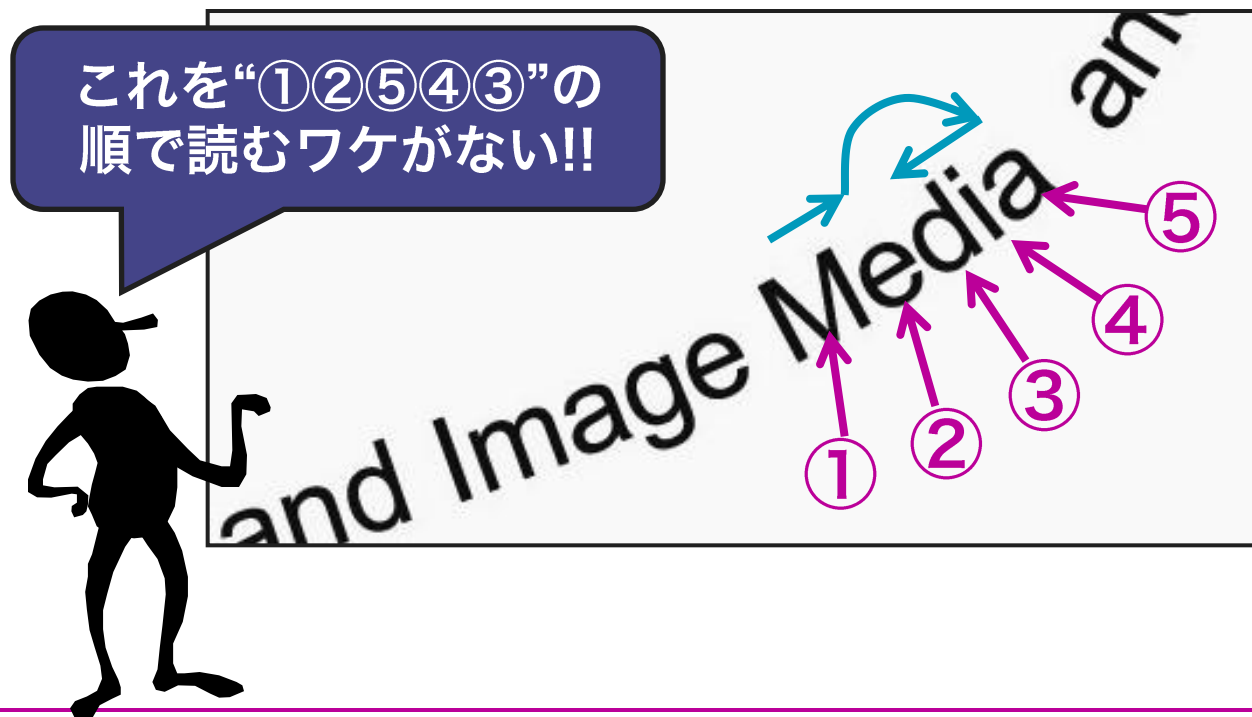
各文字の字種は一意に決定していない

文字の並び順の推定

仮定

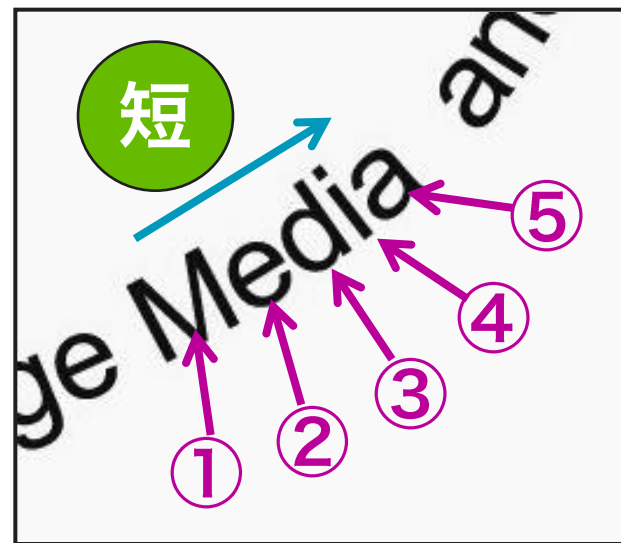
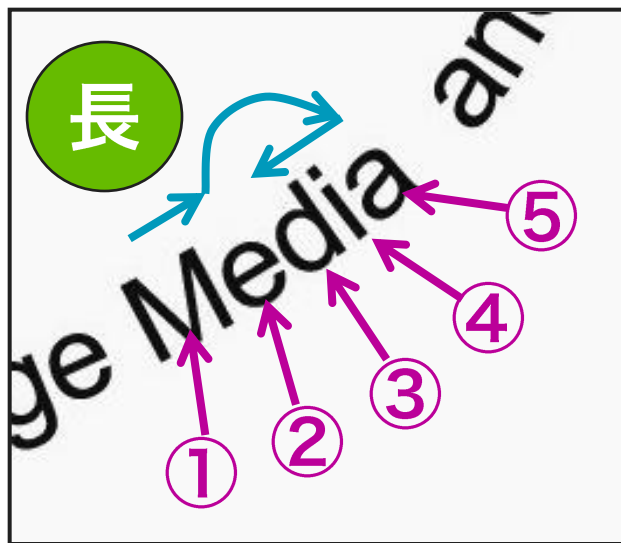
読む方向が急激に変化したり

文字を飛ばして読んだりするデザインは存在しない



文字の並び順の推定

- 全ての文字を最短で通る経路を求める

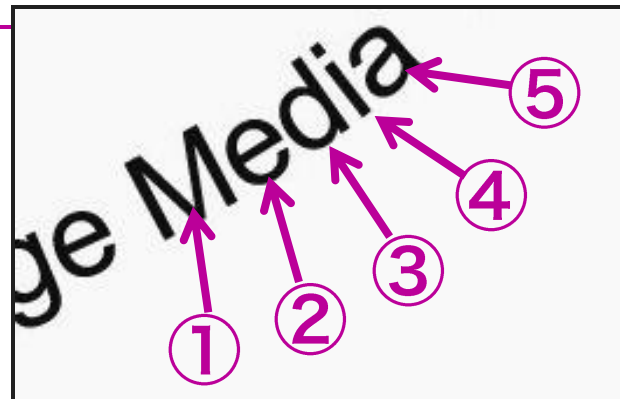


最短は“①②③④⑤”
および“⑤④③②①”

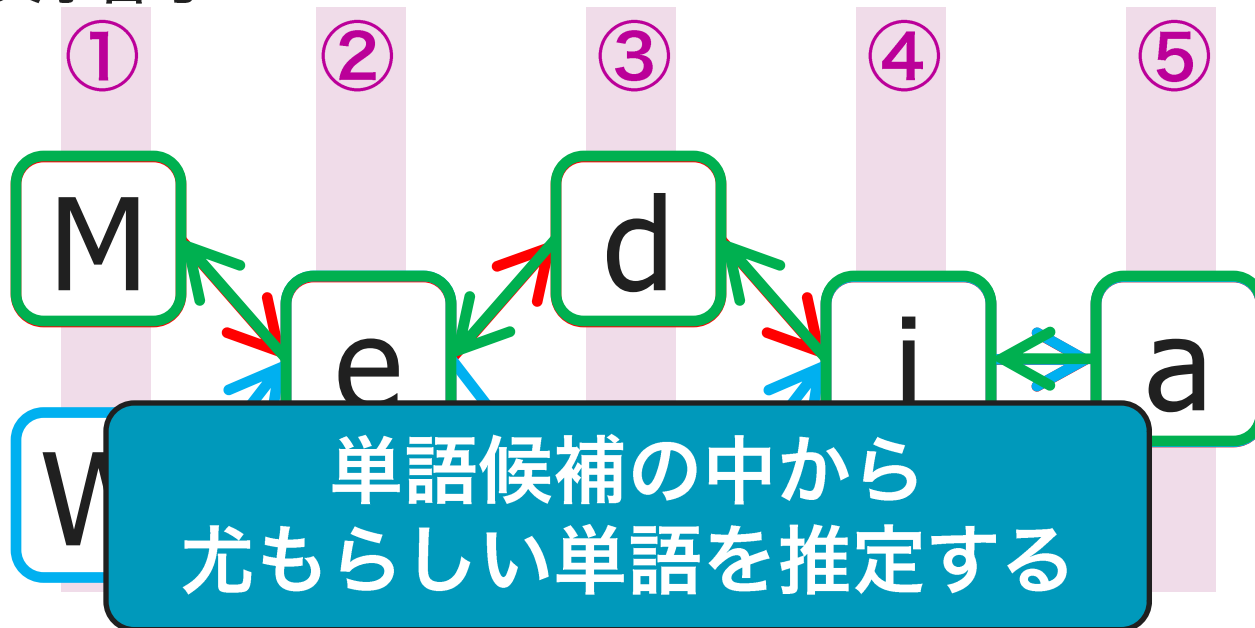
“Media” か “e!aM” か分からない

単語候補の生成

- 各文字のクラスが属する字種を組み合わせる



文字番号



単語候補

Media

Wepia

aideM

... etc.

文字の向きを利用した単語推定

“Wepia” の場合



“aideM” の場合



一部だけ180度回転しているのはオカシイ!!

左向きには読まないと言ってたのに!!

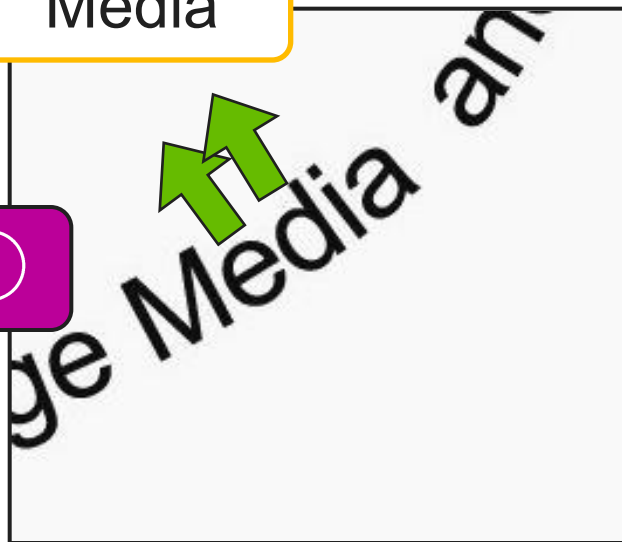
文字の向きに関するペナルティを設ける

ペナルティの計算①

文字の向きに関するペナルティ

- 「急激に変化しない」という仮定を利用

“Media”



“Wepia”



小

大

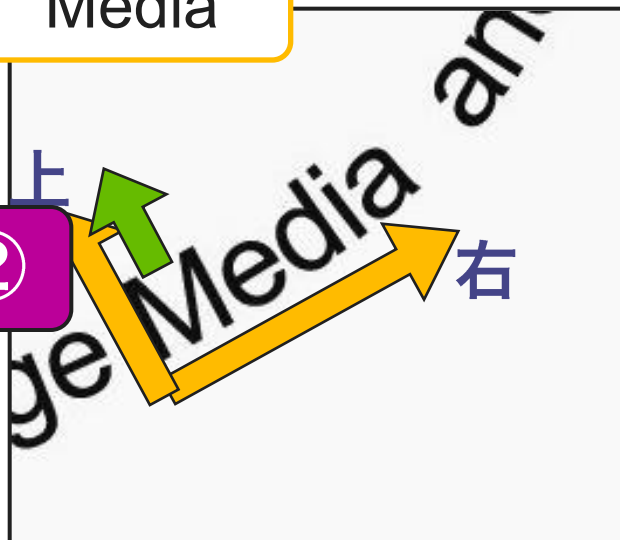
隣接文字間の
角度差

ペナルティの計算②

単語の読む方向に関するペナルティ

- 「右向きに読む」という仮定を利用

“Media”

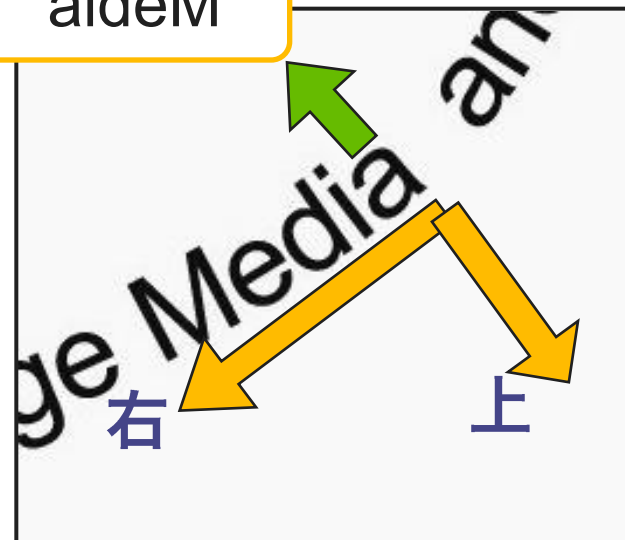


ペナルティ②

仮定した上方向と
文字の上方向との
角度差

小

“aideM”



大

単語の推定

- 候補をペナルティの昇順にソート

“Media”	17
“aideM”	185
“aideW”	359
“Wedia”	361
“Mepia”	365
“aipeM”	533
“aipeW”	707
“Wepia”	709



発表の流れ

- 背景
 - 文字認識手法の概要
 - 単語認識の提案手法
 - 実験
 - まとめ
-

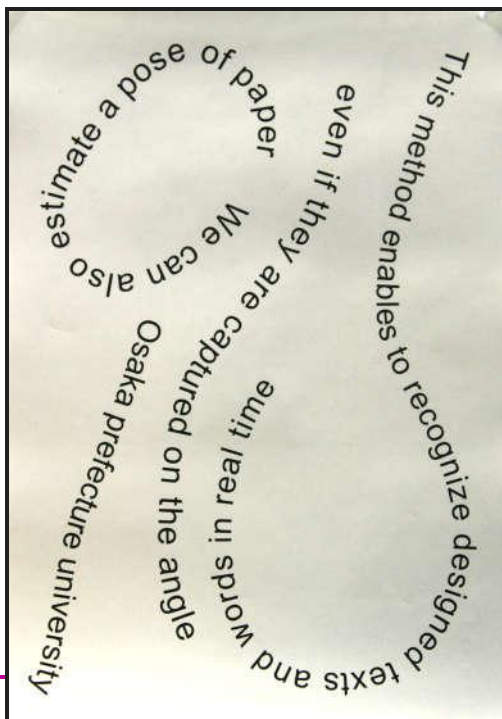
実験対象

- 英字が書かれた文書を3方向から撮影

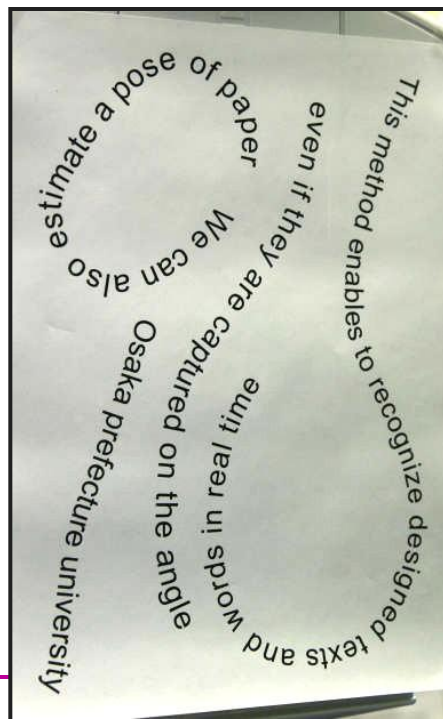
144文字

30単語

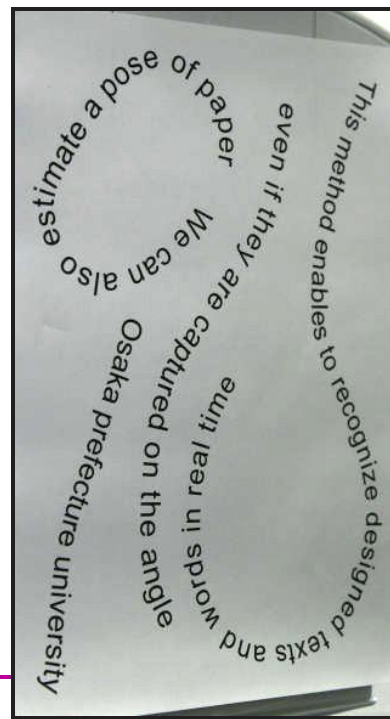
0度



30度



45度

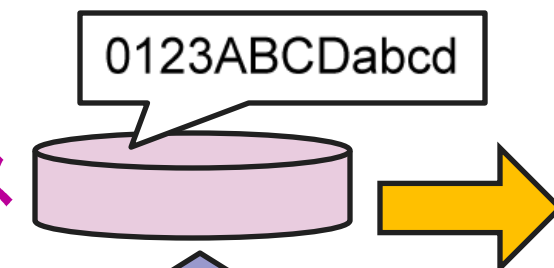


実験条件

- 登録フォント数を増やし、
認識性能を評価

最大10フォント

データベース



使用した計算機

CPU

Opteron 2.8GHz

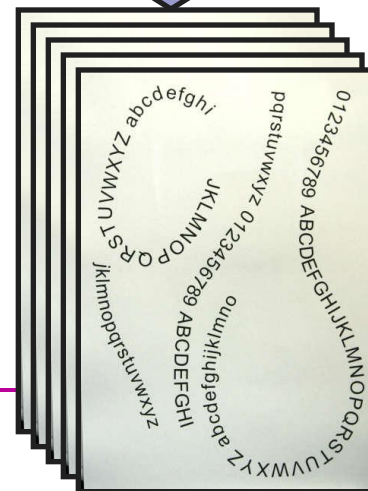
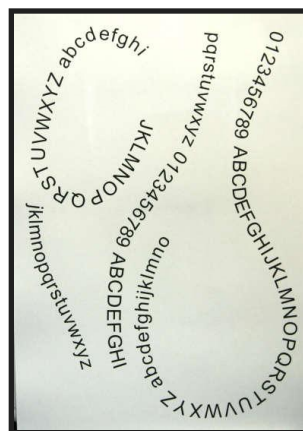
メモリ

16GB

認識

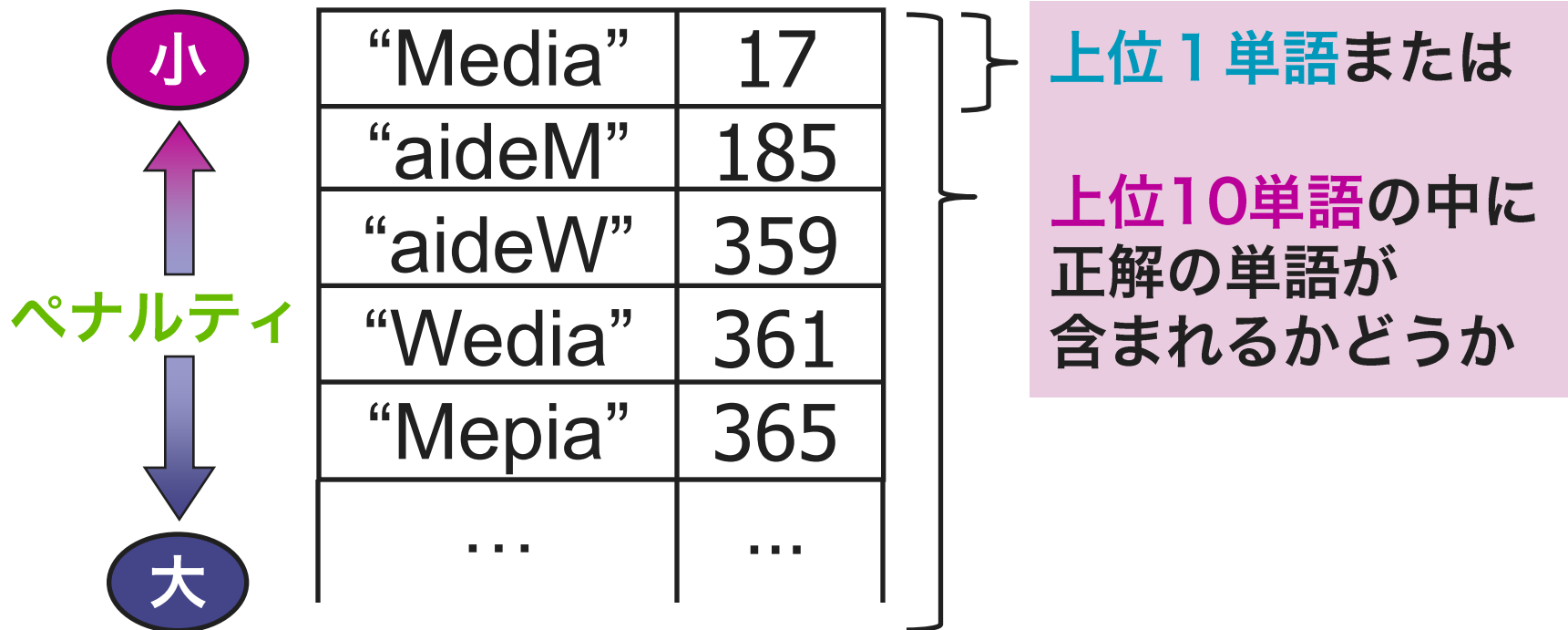
認識

文書

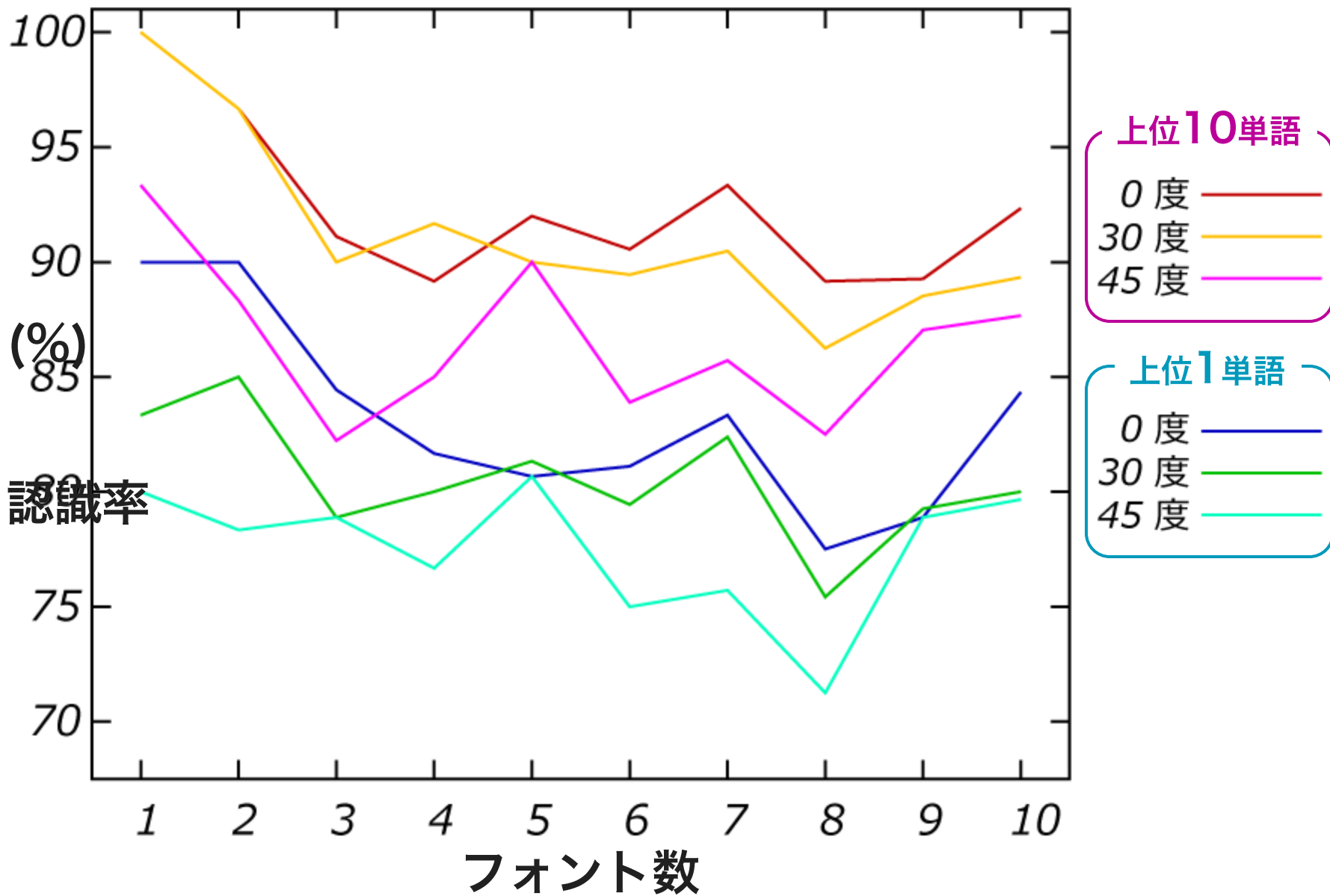


実験条件

■ 単語認識率



単語認識率



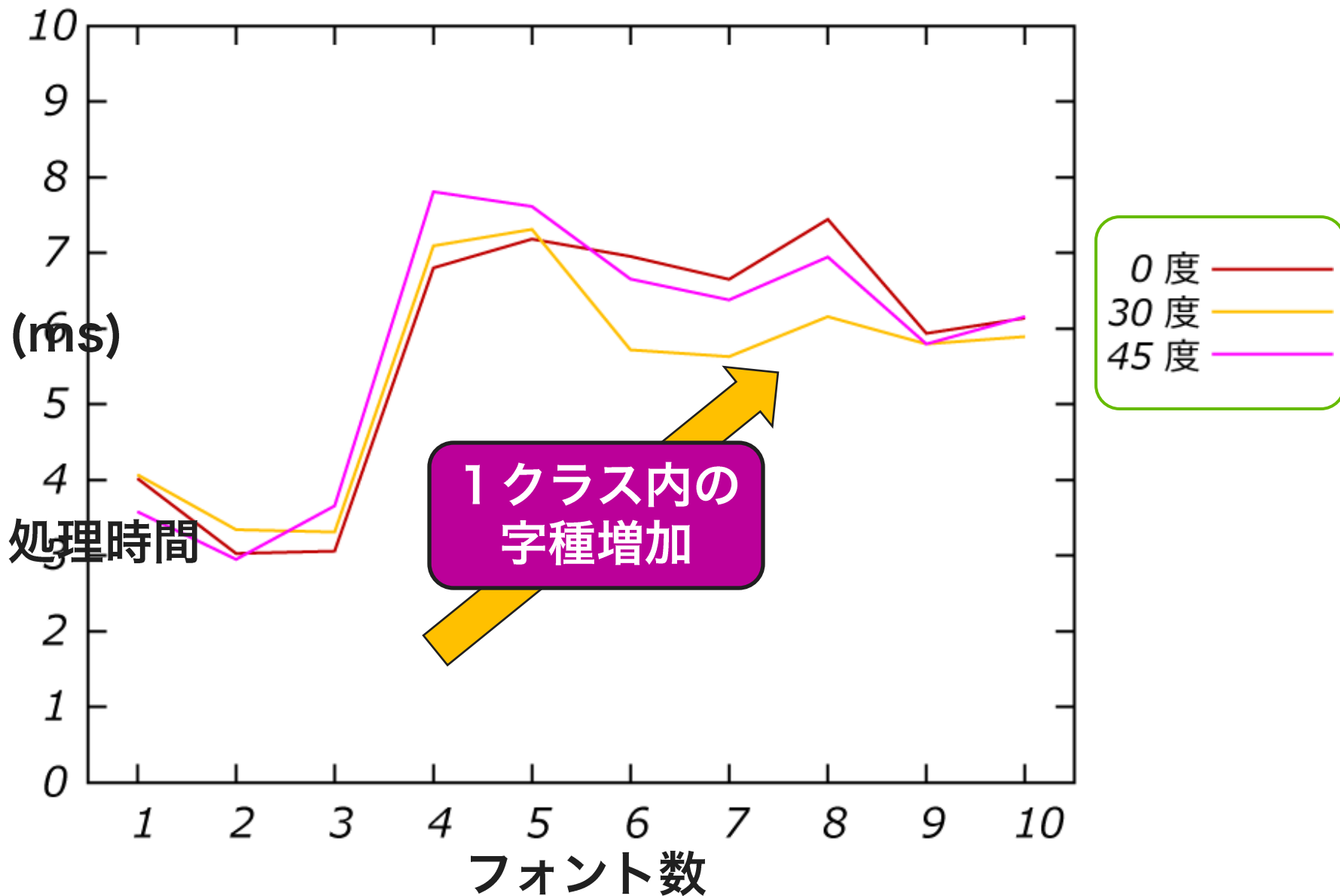
単語認識の失敗例

- クラスレベルでの認識失敗 ⇒ 後で検証
- 単語領域取得の失敗

分離している



1 単語あたりの単語認識時間

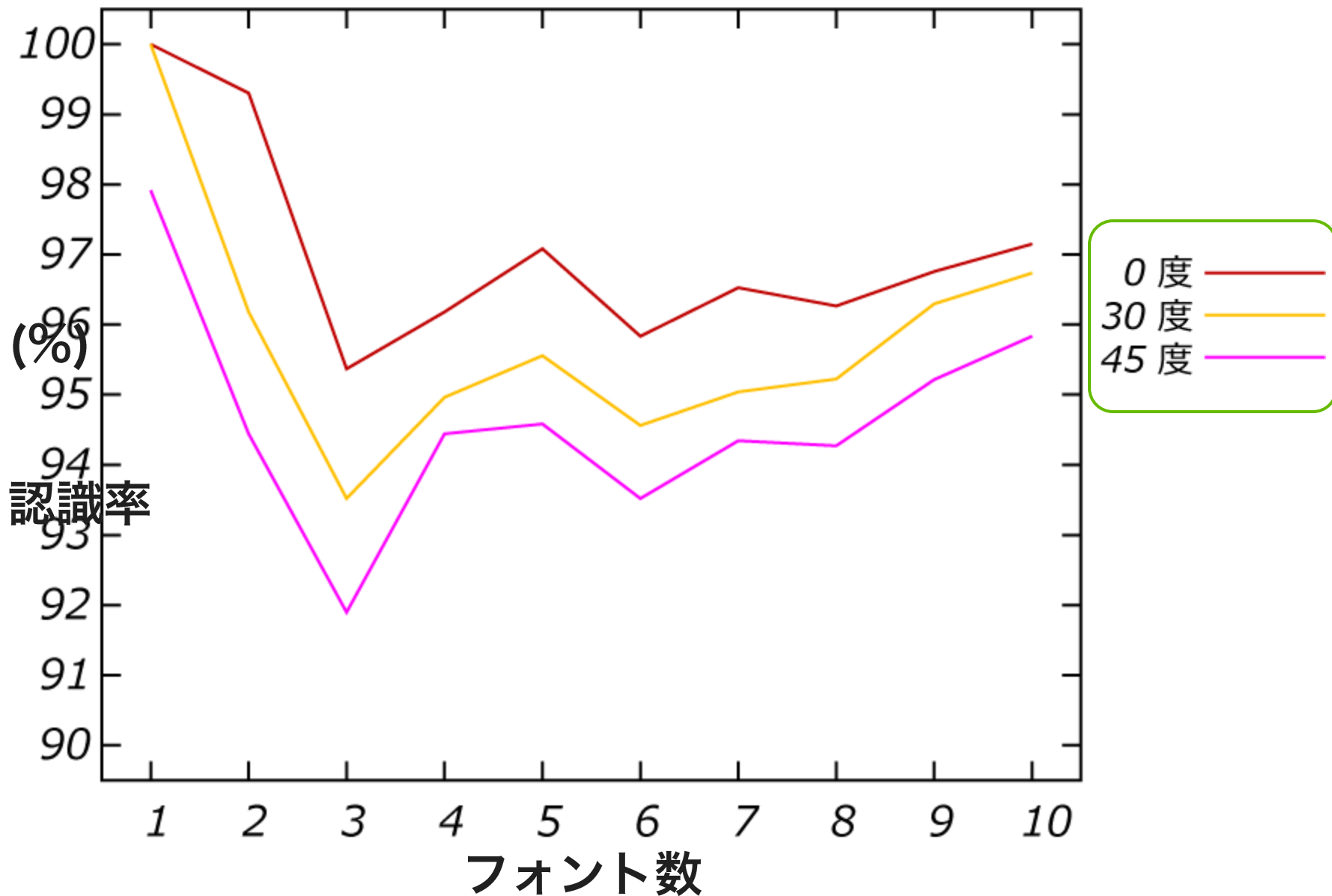


同一クラスに統合された字種

- Arialの場合

0 O o	6 9	7 L	C c
E m	I l	N Z z	S s
V v	W w	b q	d p
n u			

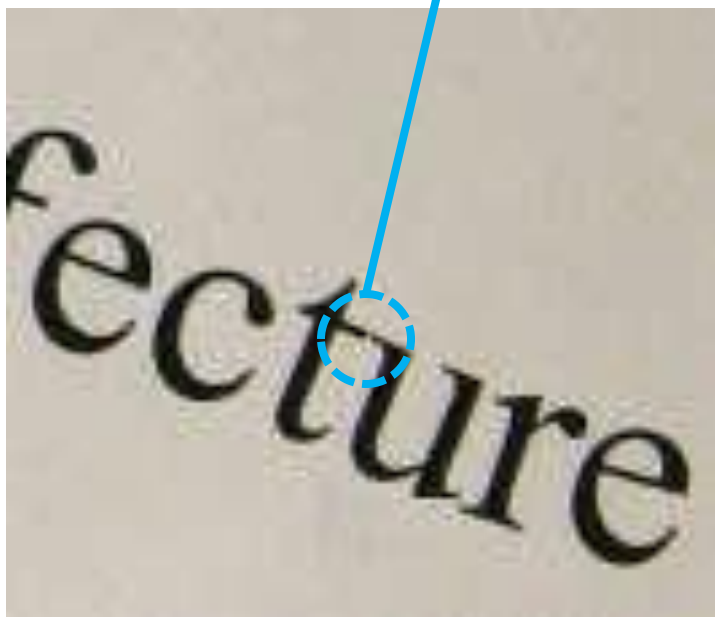
検証 - 1文字あたりのクラス識別率



文字認識の失敗例

- 誤認識
- 連結成分取得の失敗

結合している



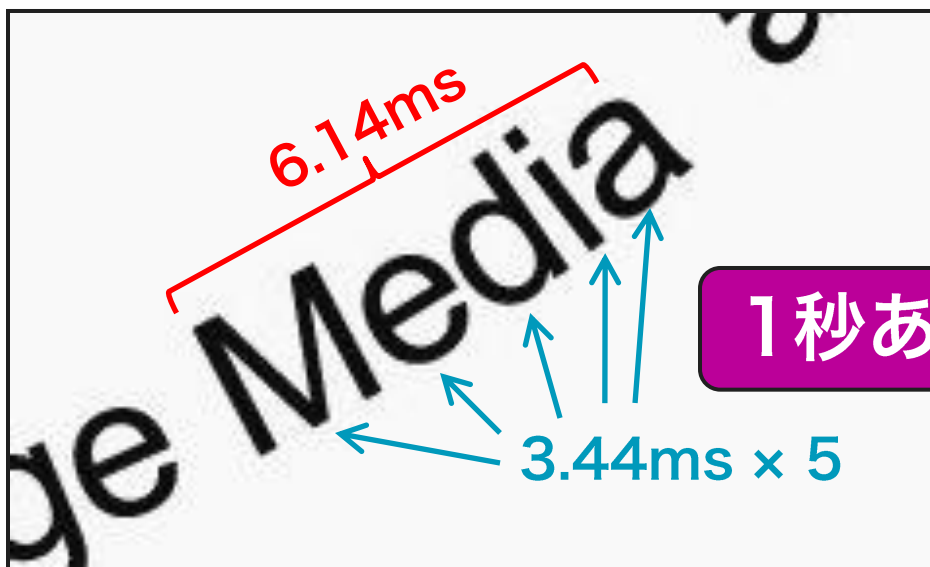
実用性に関する性能評価

■ 10フォント登録, 正面から撮影の場合

単語認識率(上位10個を見た場合)	92.3%
1文字あたりの文字認識時間	3.44ms
1単語あたりの単語認識時間	6.14ms
メモリ使用量	397MB

100フォントの場合
約4GB (実測値)

メモリ削減が課題



1秒あたり42単語の認識が可能

発表の流れ

- 背景
 - 文字認識手法の概要
 - 単語認識の提案手法
 - 実験
 - まとめ
-

まとめ

提案手法

- リンクアンカー化システムに有効な単語認識
 - 文字の向きを用いた単語認識

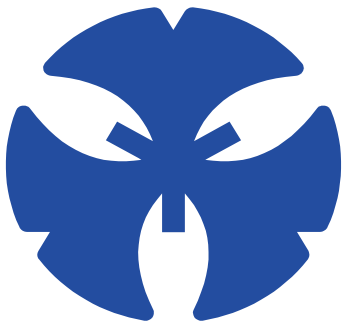
今後の課題

- 文字領域・単語領域の抽出法の改良
 - メモリ削減
-

リアルタイム単語認識技術を利用した カメラベース情報取得システム

辻 智彦 岩村 雅一 黄瀬 浩一

大阪府立大学大学院工学研究科



大阪府立大学
OSAKA PREFECTURE UNIVERSITY