

大阪府立大学  
OSAKA PREFECTURE UNIVERSITY



# レイアウト非依存な 実時間カメラベース文字認識

岩村雅一 辻智彦 堀松晃 黄瀬浩一

# 実時間カメラベース文字認識システム

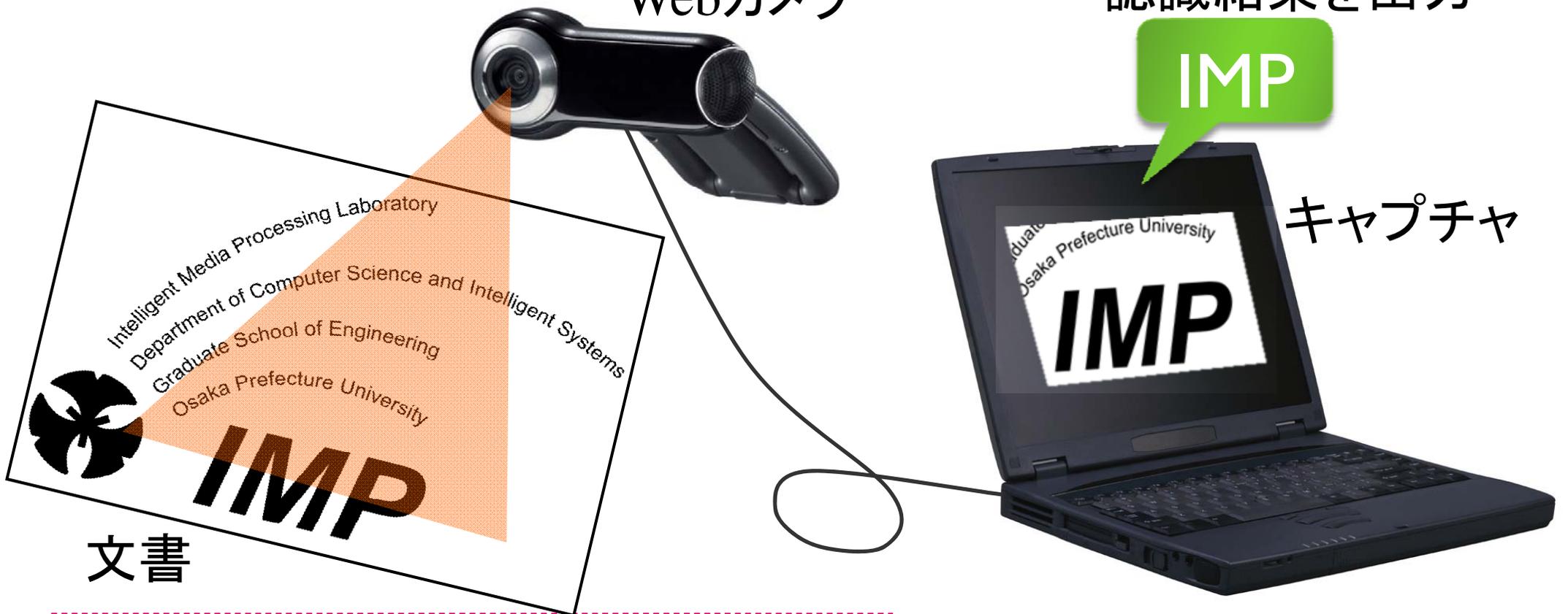
1秒間に200～250文字程度認識可能

リアルタイムに  
認識結果を出力

Webカメラ

IMP

キャプチャ



文書

# 応用例

環境中の全ての文字を認識して、  
必要な情報のみを提供することができる

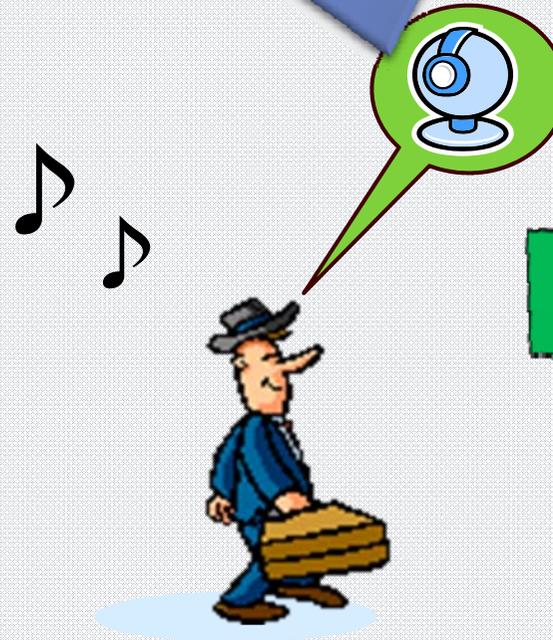
## 視覚障害者への音声案内

『押ボタン信号が  
あります』



## 翻訳システム

Car-free mall



歩行者天国



# 下記3要件を同時に実現した 初めての手法

## 提案手法の特長

### 1: 高速

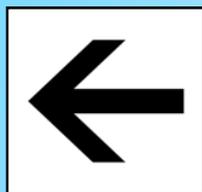
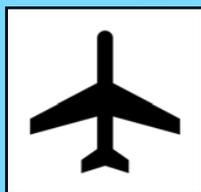
- ・200文字を1秒以内に認識

### 2: 射影歪みに頑健

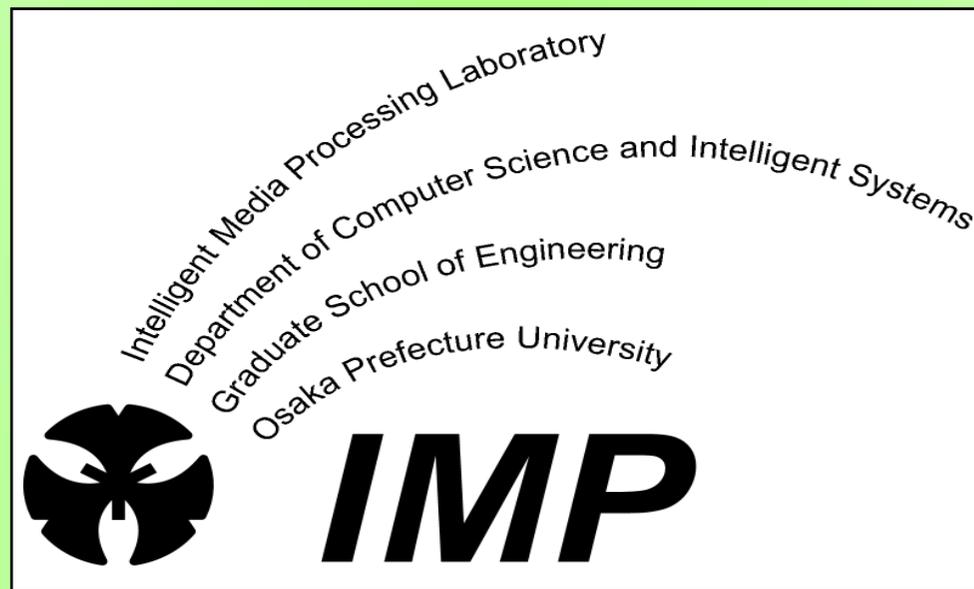
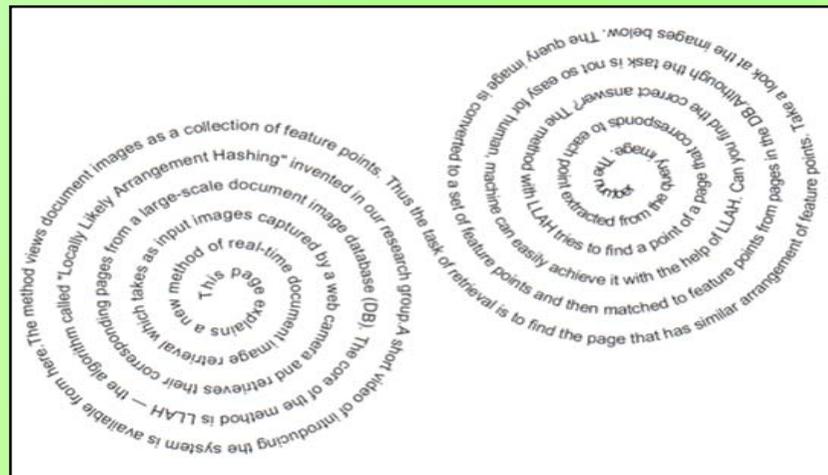
- ・斜め45度から撮影しても  
8割以上の認識率

デザイン文字や  
ピクトグラムも認識可能

A B C D 1 2 3 4



### 3: レイアウトフリー



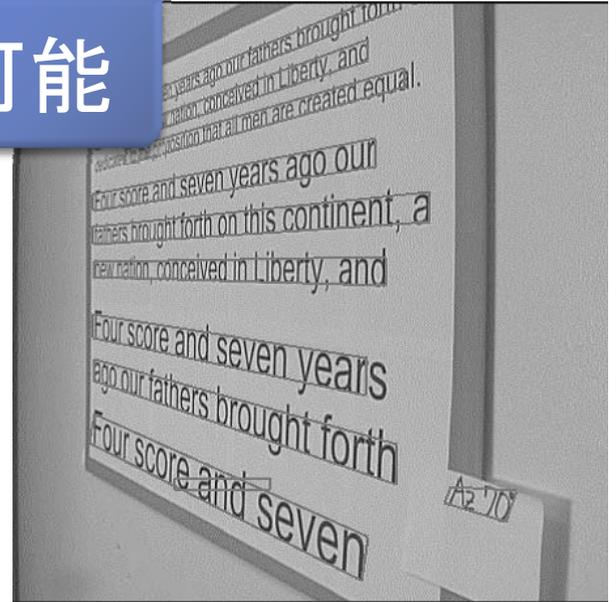
# 従来手法と問題点

1. 実時間認識可能だが、行を成す文字しか認識できない



認識不可能

認識可能



2. 複雑なレイアウトも認識可能だが、実時間で認識できない



# 従来手法 vs 提案手法

	1: 高速	2: 射影歪み	3: レイアウトフリー
Myers 2004	○	○	×
Kusachi 2004	×	○	○
Li 2008	×	○	○
Proposed method	○	○	○



文字単位の認識

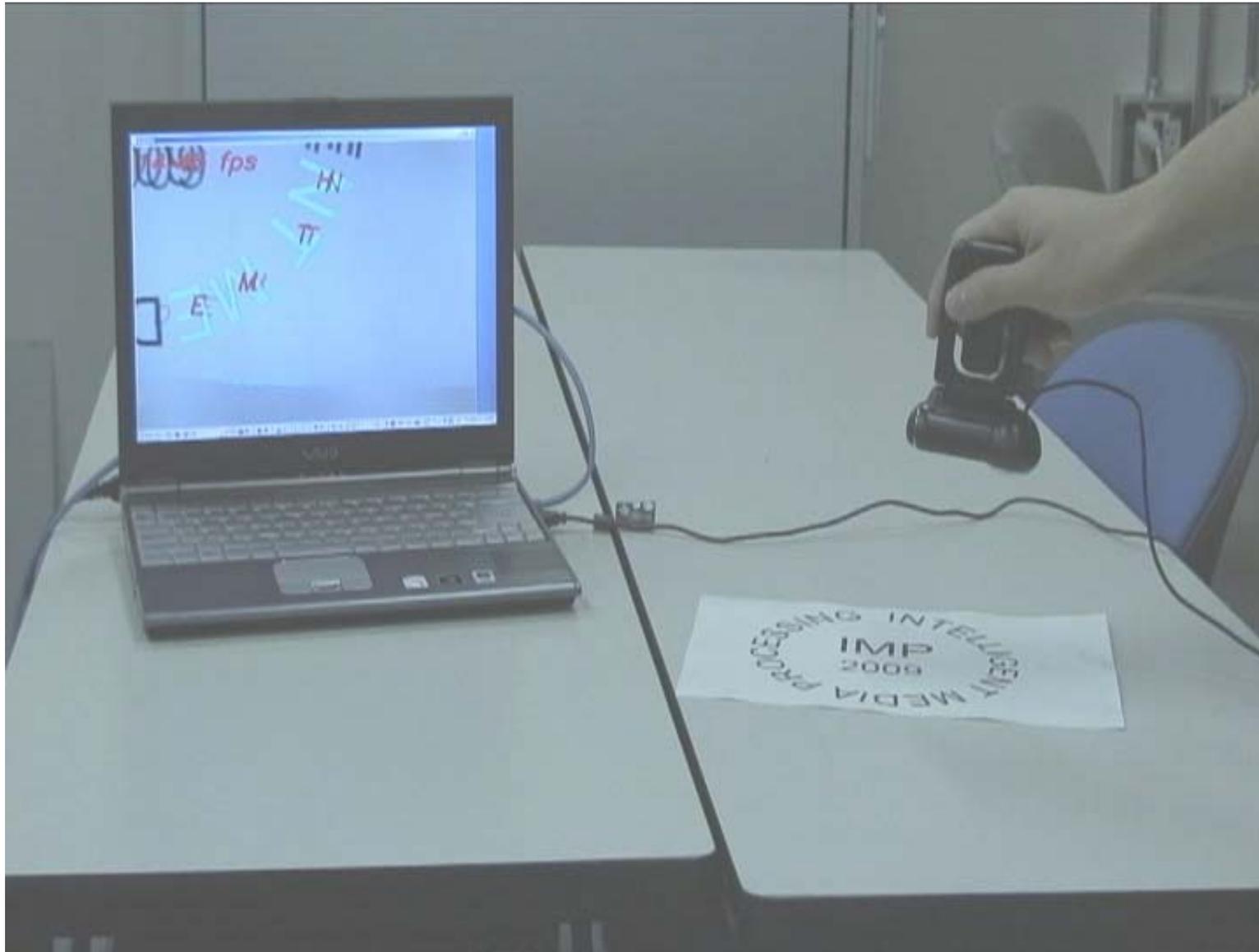


実時間処理



# DEMO

---



# 目次

---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. 姿勢推定
  5. 実験
  6. まとめ
- 



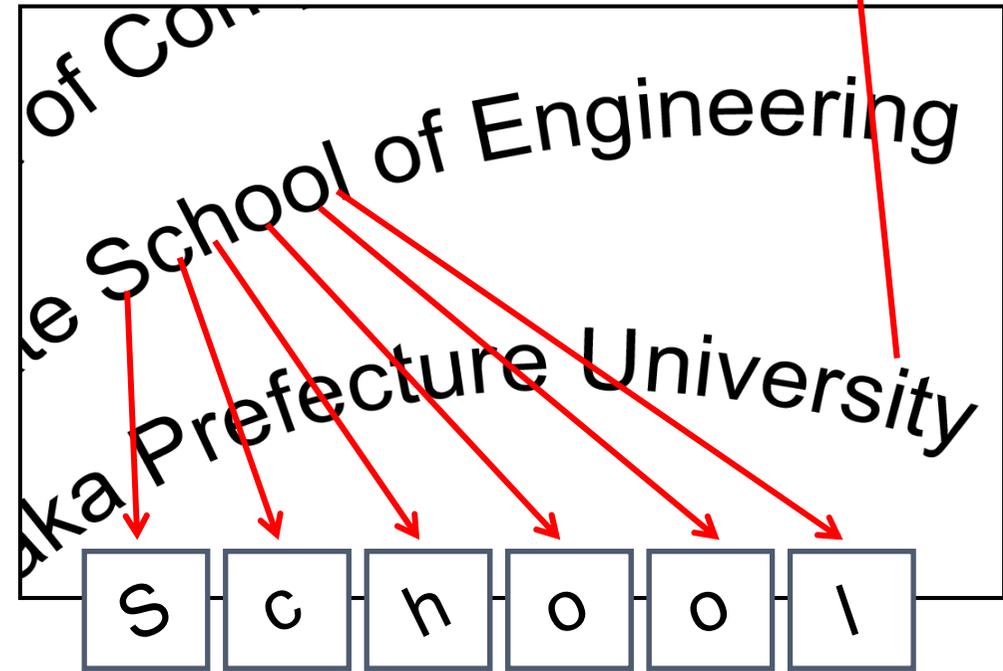
# 提案手法のアプローチ 1

- ▶ 連結成分単位の認識

3: レイアウトフリー の実現

- ▶ 問題設定
  - ▶ 文字は同一平面上に存在
  - ▶ 文字は二値化で簡単に抽出可能

後処理へ i

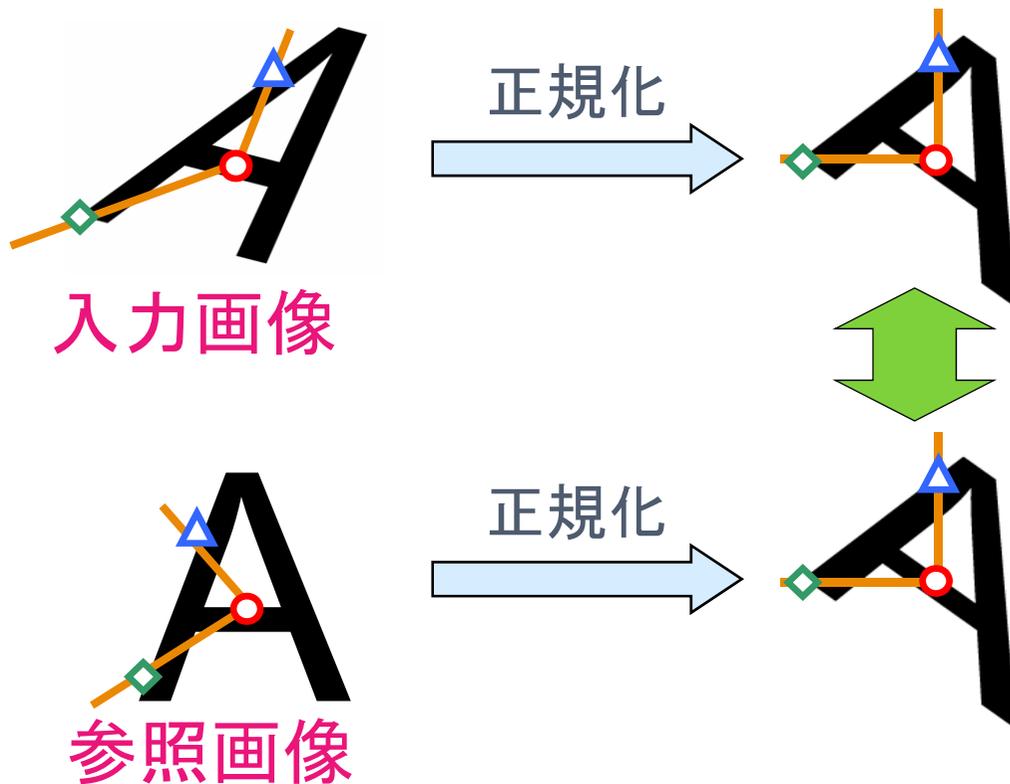


切り出した後の文字を高速処理に特化

# 提案手法のアプローチ2

## ▶ アフィン不変な認識

- ▶ 同一の3点が選択できれば、照合可能



2: 射影歪み

に頑健な認識の実現

# 提案手法のアプローチ2 輪郭版GHのアイデア

従来手法:

Geometric Hashing (GH)

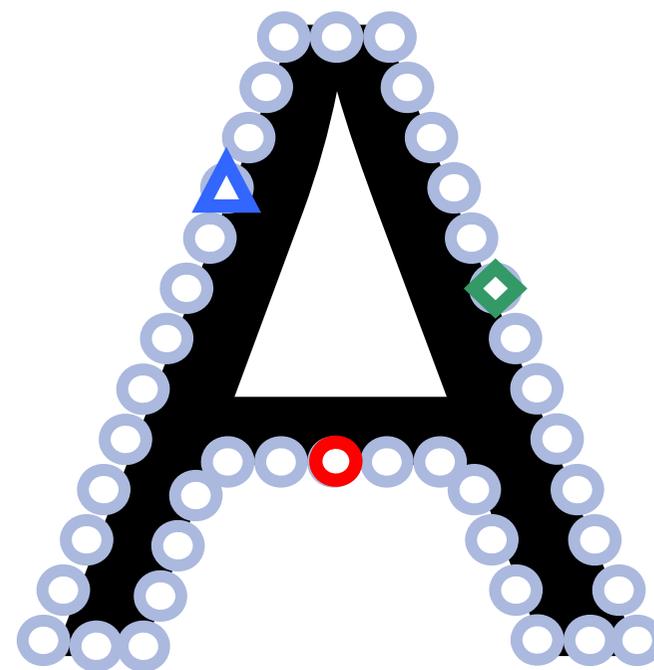
輪郭版GH

提案手法の出発点

連結成分に  
GHを適用

特徴点数:  $P$

特徴点の配置の照合

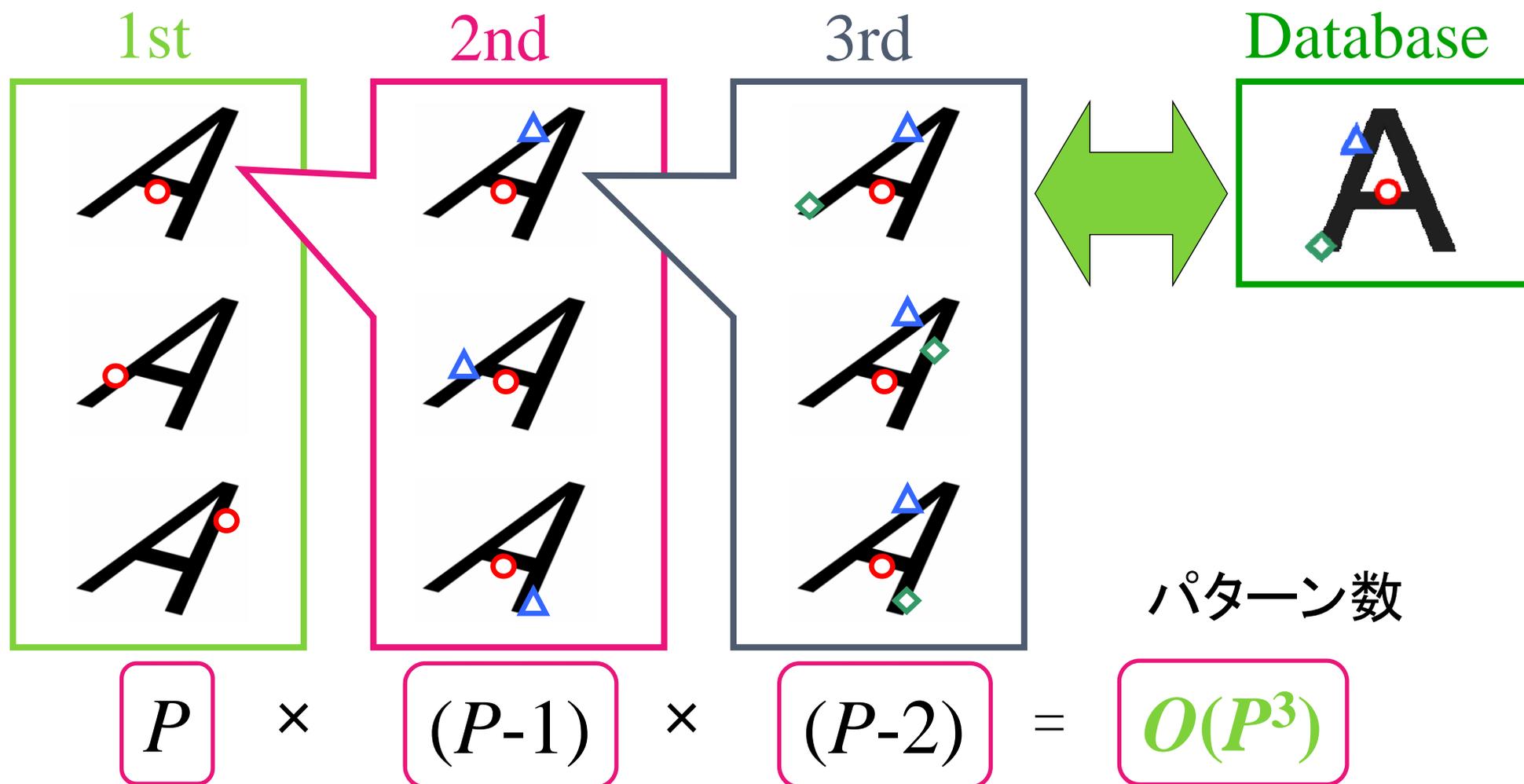


図形の照合

# 提案手法のアプローチ 3

## 輪郭版GHが作る 3点の配置

- ▶  $P$  点から3点を選択する全ての組み合わせを試す



# 提案手法のアプローチ 3

## 提案手法が作る 3 点の配置

1: 高速 の実現

- ▶ 存在しない組み合わせを計算しない

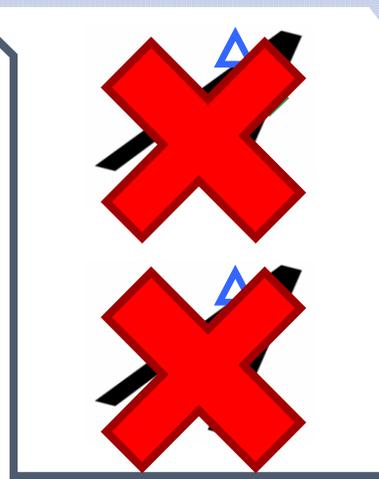
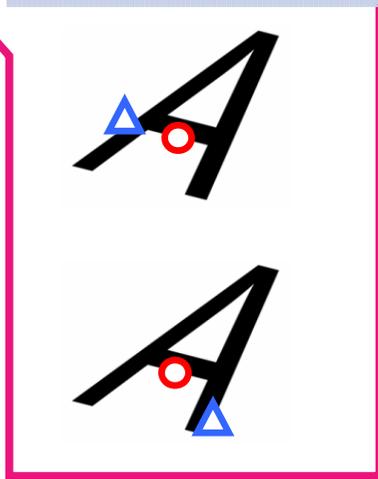
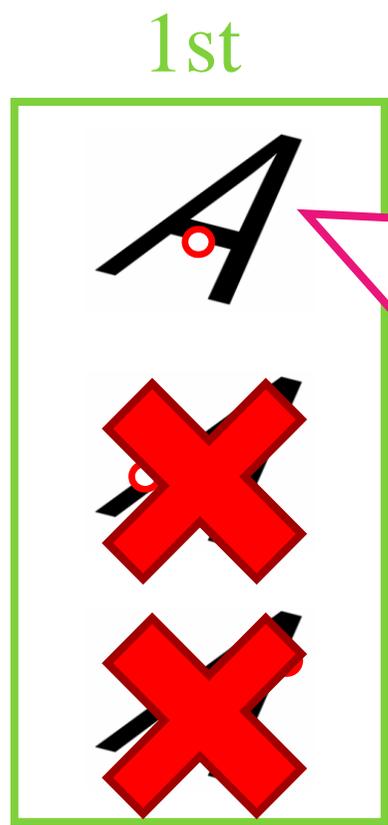
$P=100$  の場合

輪郭版GH  
提案手法

970,200

100

実時間認識を実現



1

×

P

×

1

=

$O(P)$

$O(P^3)$

パターン数



# 目次

---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. 姿勢推定
  5. 実験
  6. まとめ
- 

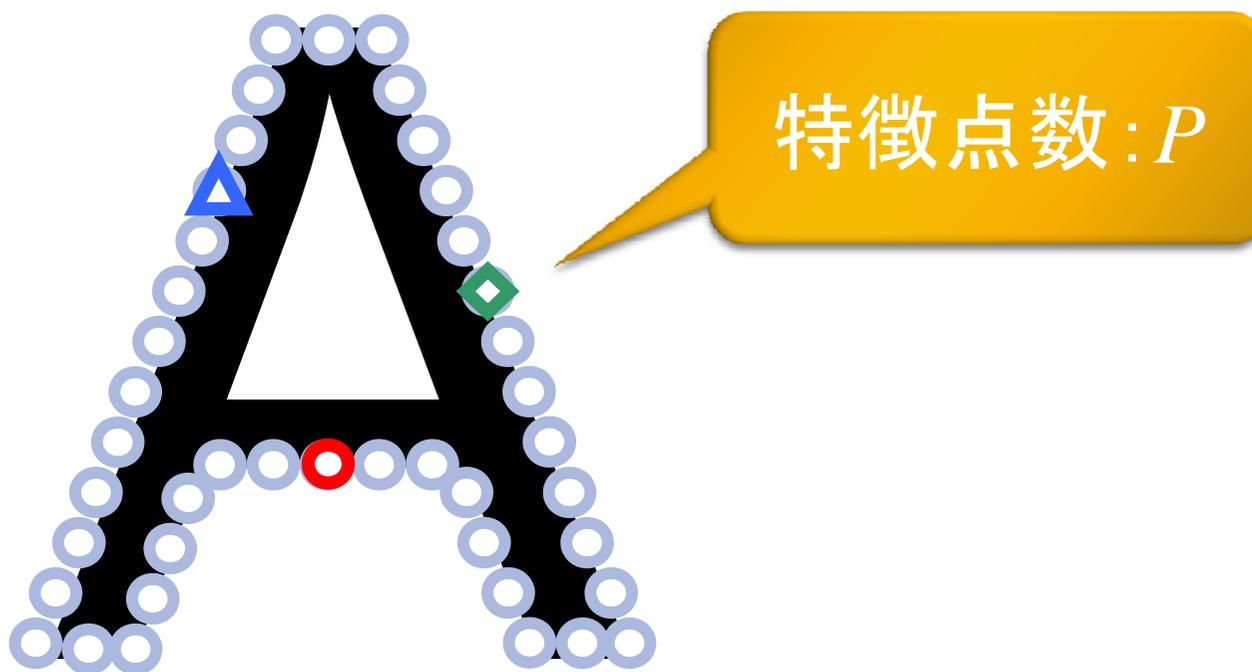


# 輪郭版GH

---

## ▶ GHとの違い

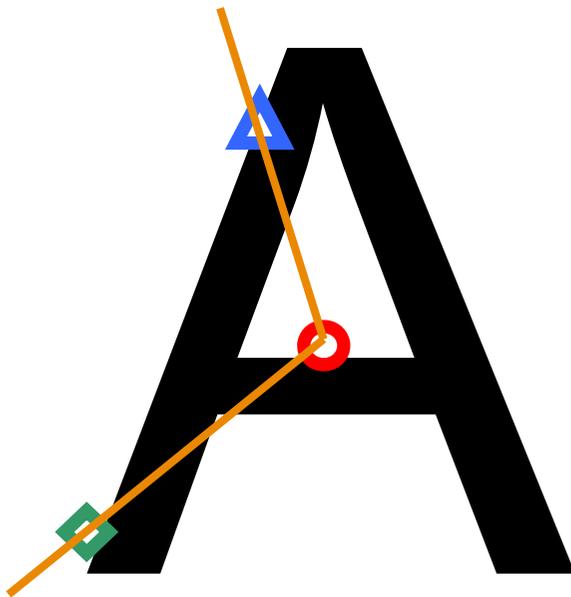
- ▶ 特徴点を外側の輪郭から抽出
- ▶ 照合に図形の特徴を使用



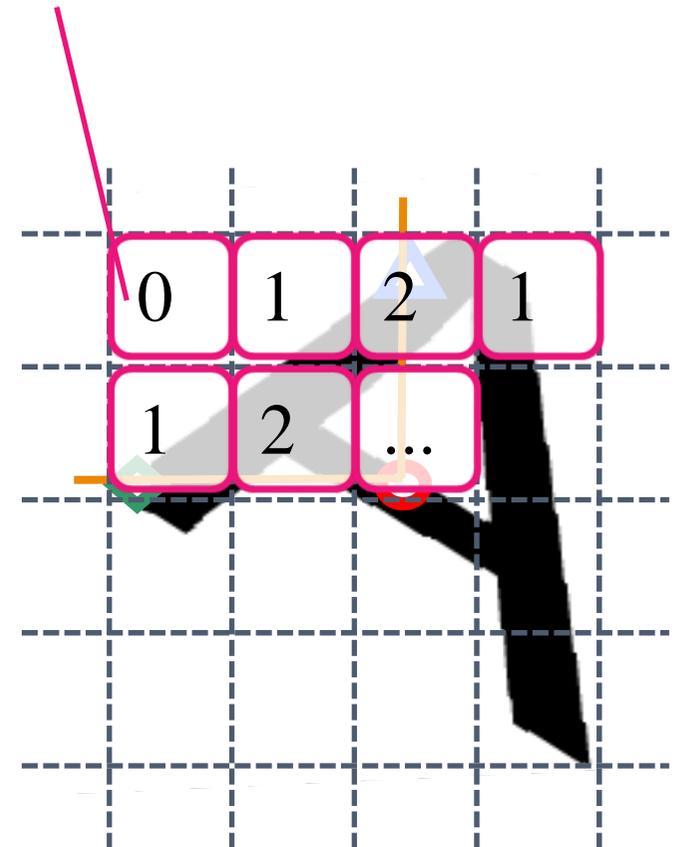
# 輪郭版GH — 図形の照合

## ▶ 特徴ベクトルの計算

1. 正規化
2. 領域分割
3. 黒画素の割合のヒストグラム作成
4. 量子化

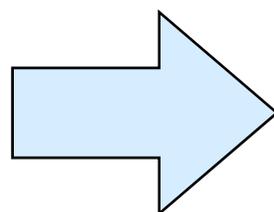
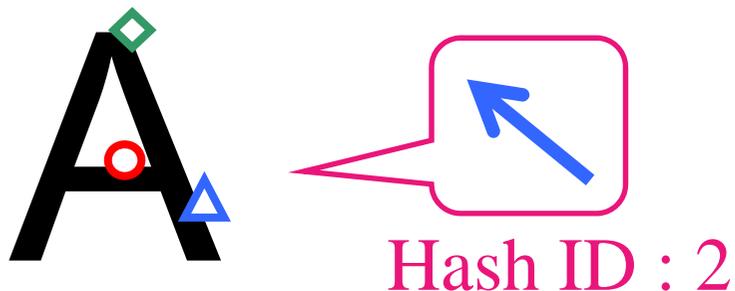


## 特徴ベクトル



# 輪郭版GH — 登録

- ▶ 特徴ベクトルをハッシュテーブルに登録

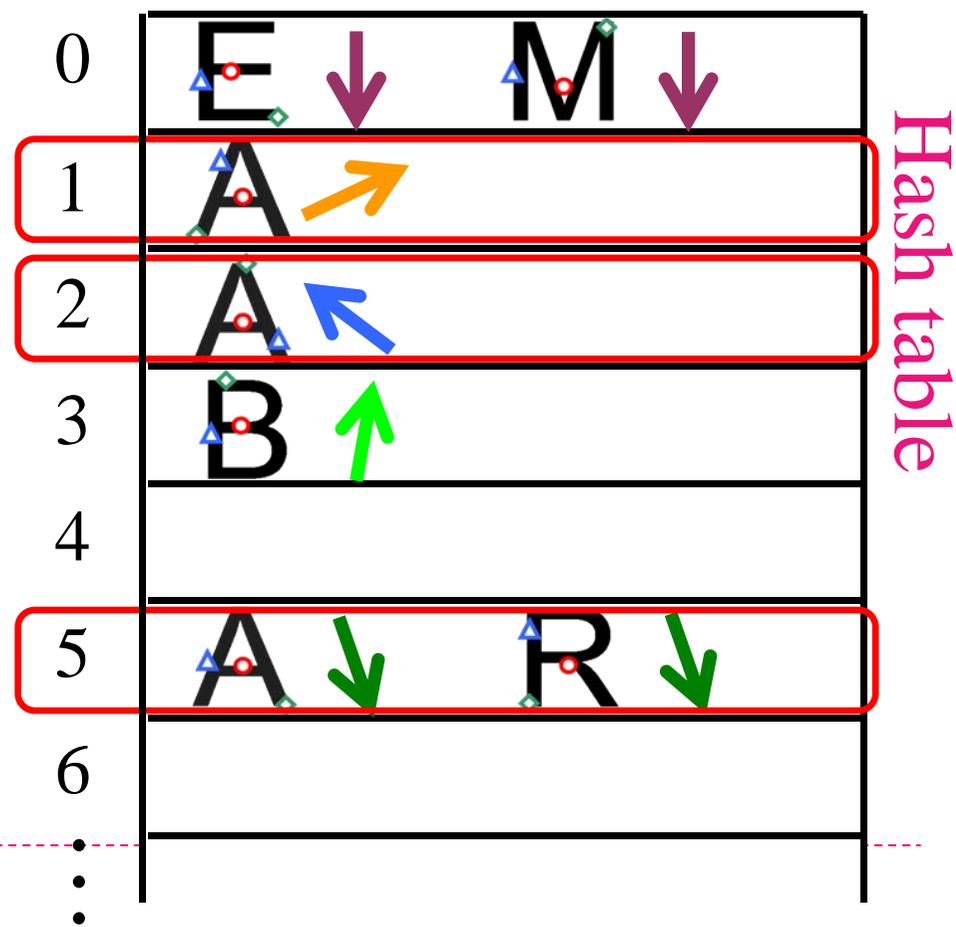
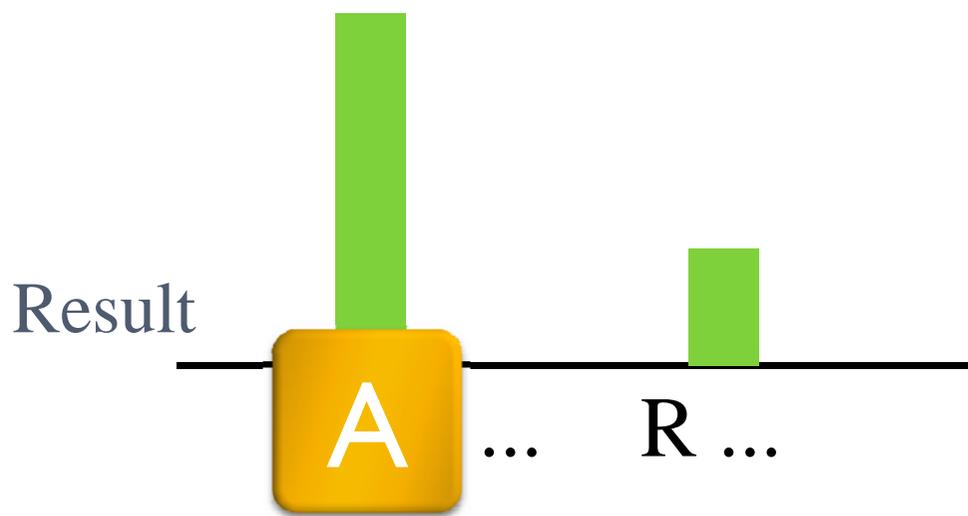
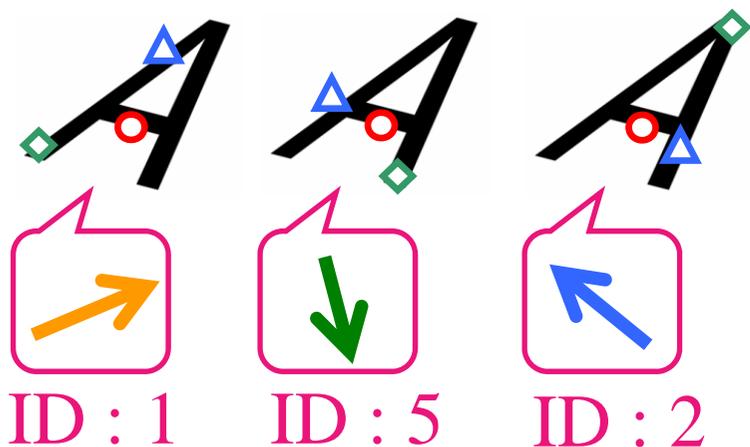


0	E ↓	M ↓
1	A →	
2	A ←	
3	B ↑	
4		
5	A ↓	R ↓
6		
⋮		

Hash table

# 輪郭版GH — 検索 (認識)

1. 特徴ベクトルを作成
2. 字種に投票



# 目次

---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. 姿勢推定
  5. 実験
  6. まとめ
- 

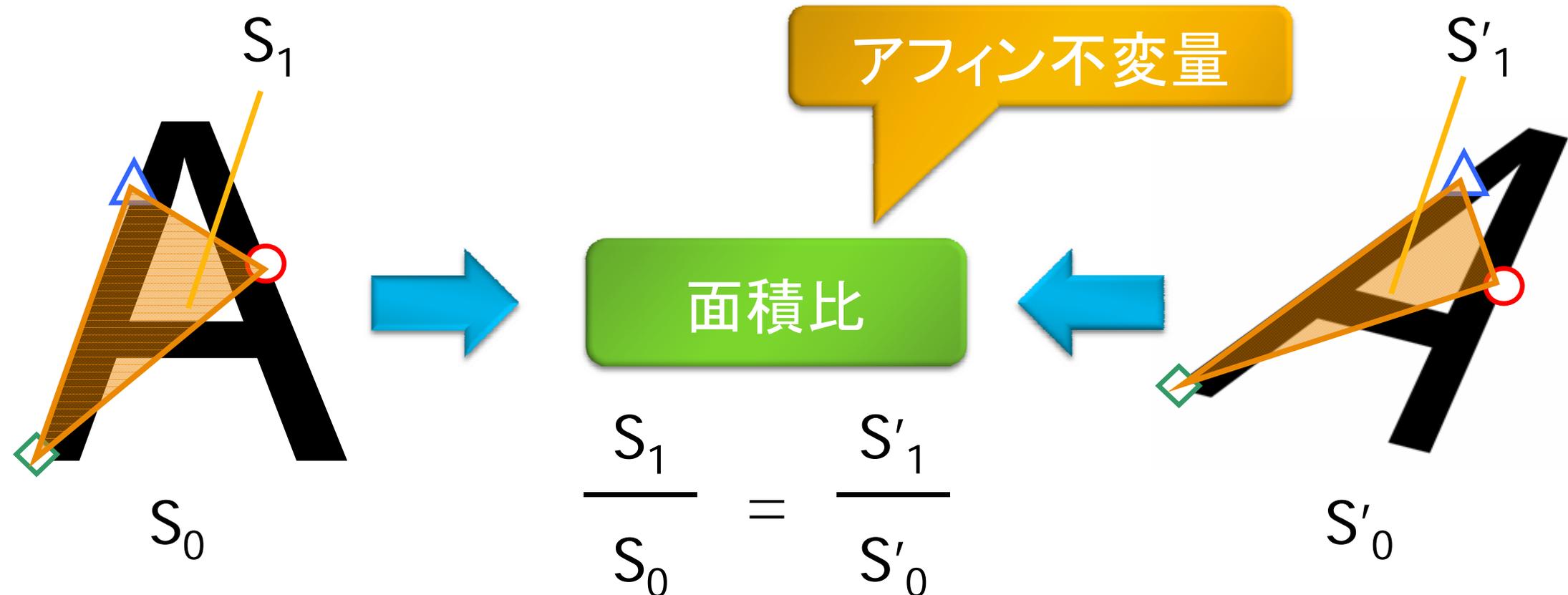


# 提案手法 1：輪郭版GHの高速化 パターンを削減する原理

通常の方法

## ▶ 面積比

▶ 3点の配置 → 面積比

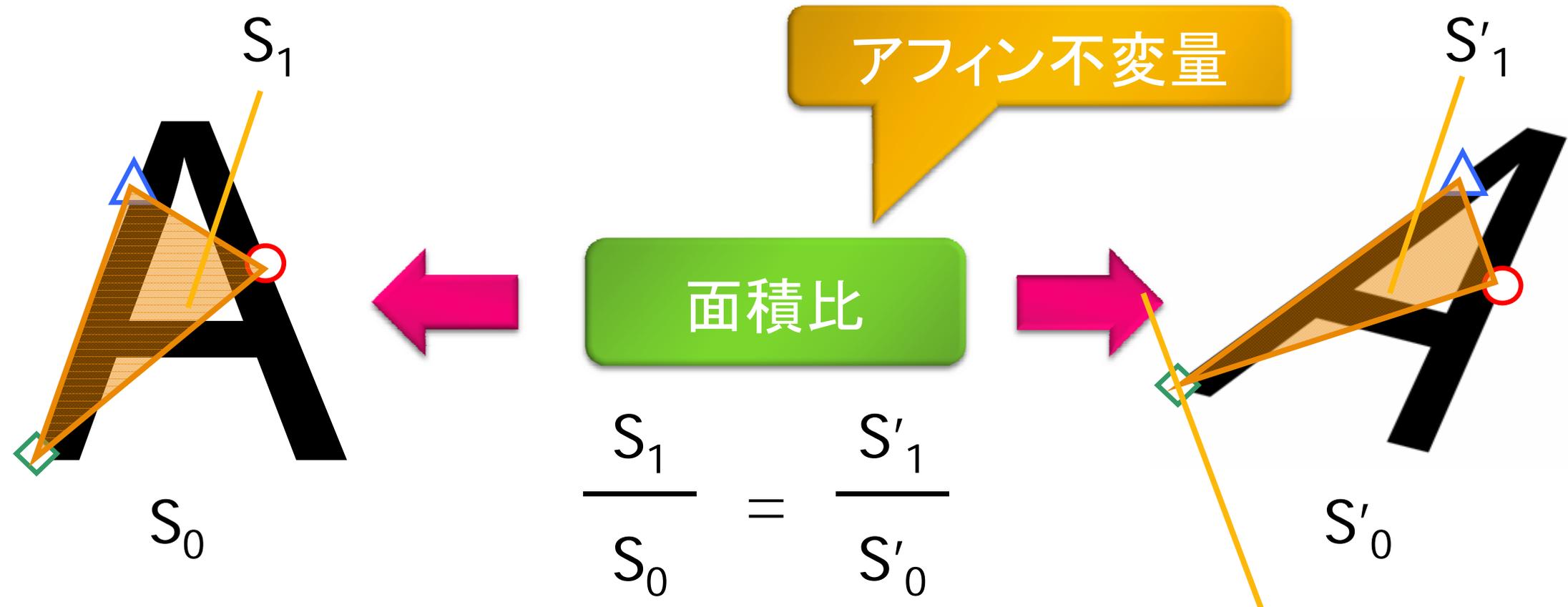


# 提案手法 1：輪郭版GHの高速化 パターンを削減する原理

通常とは逆の方法

## ▶ 面積比

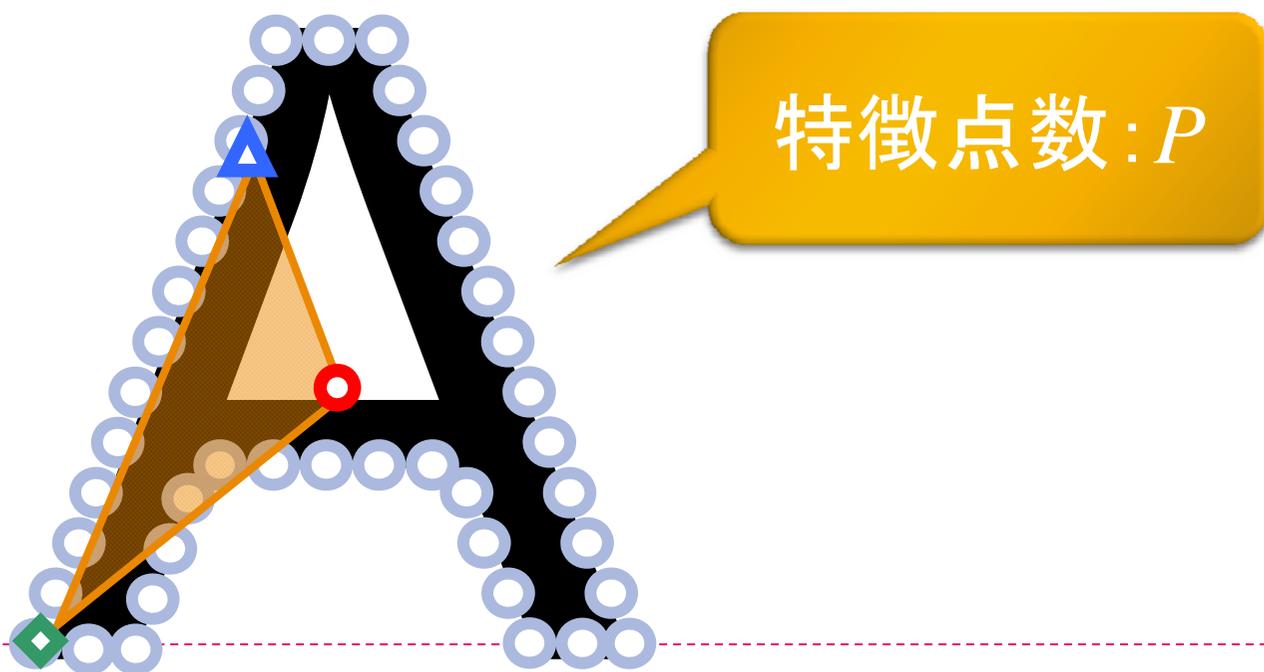
- ▶ 2点の配置 + 面積比 → 3点目の位置



# 提案手法 1：輪郭版GHの高速化

## 提案手法のパターンの生成方法

- ▶ 1点目：図形の重心（アフィン歪みに不変） ← 一意
- ▶ 2点目：輪郭上の任意の点
- ▶ 3点目：面積比によって決定 ← 一意



# 目次

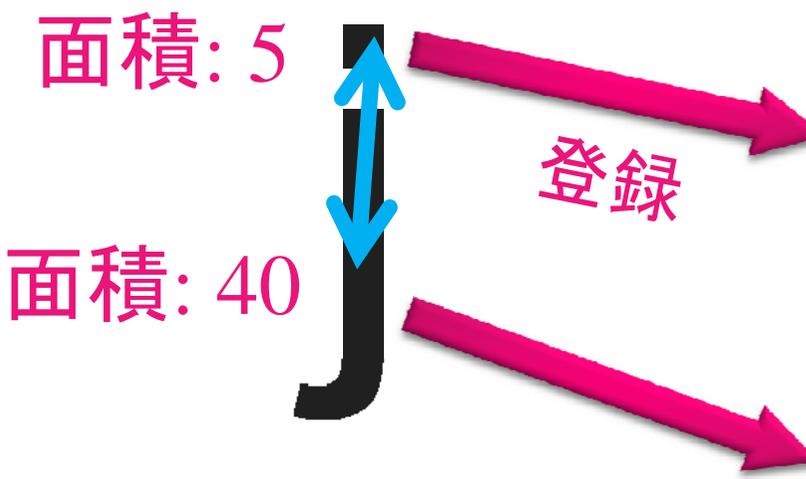
---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. 姿勢推定
  5. 実験
  6. まとめ
- 



## 提案手法2：分離文字の認識

### ▶ 分離文字テーブルを作成

	連結成分	字種	相対位置	面積	相手の面積
	■	i	↓	5	25
	■	j	↓	5	40
	┆	i	↑	25	5
	J	j	↑	40	5

# 目次

---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. **姿勢推定**
  5. 実験
  6. まとめ
- 



# 提案手法 3・姿勢推定

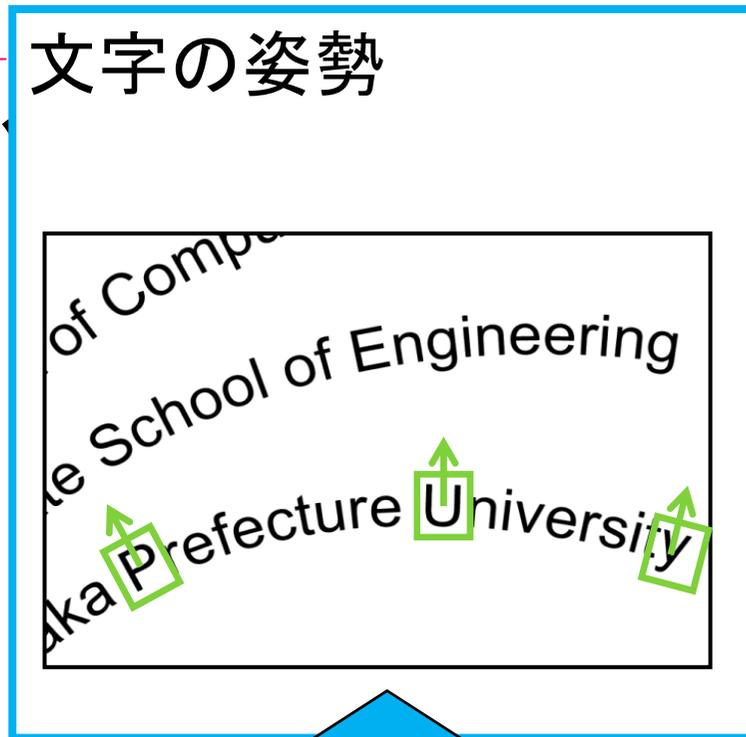
▶ 対

紙面の姿勢



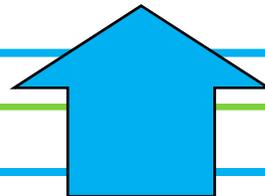
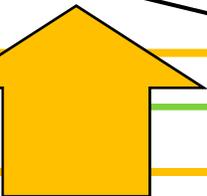
変換

文字の姿勢

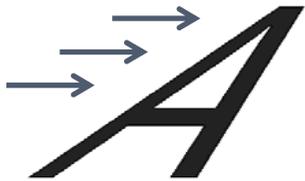


変換

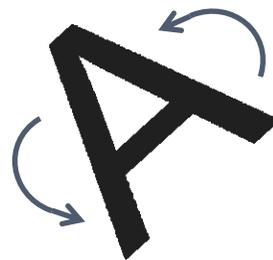
パラメータ



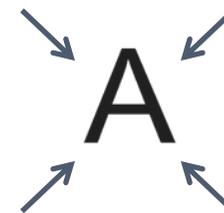
独立変倍



シアー



回転



拡大・縮小

# 目次

---

1. 背景
  2. 提案手法のアプローチ
  3. 輪郭版GH
  4. 提案手法
    1. 輪郭版GHの高速化
    2. 分離文字の認識
    3. 姿勢推定
  5. **実験**
  6. まとめ
- 



# 認識対象

3フォント

236文字



Meeting on Image Recognition and Understanding 2009  
Organized by Forum for Image Informatics in Japan,  
IPSJ SIG Computer Vision and Image Media, and  
IEICE Pattern Recognition and Media Understanding

**MIRU**

**MIRU2009**  
**KUNIBIKI MESSE**  
**MATSUE**



# 認識実験

- ▶ 3方向から撮影した画像を認識
- ▶ 計算サーバー (Opteron 2.6GHz) を使用



撮影角度: 0度



撮影角度: 30度



撮影角度: 45度

# 実験条件

---

- ▶ アフィン変換を受けると類似する文字は同一クラスとした

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



# 実験結果

- ▶ 高い認識率と高速性を実現

1秒間に約200文字

- ▶ S: 精度と速さをコントロールするパラメータ

高精度

高速

試行回数 $S$	200			20		
角度 (度)	0	30	45	0	30	45
処理時間 (ms)	7990	7990	7020	1300	1260	1140
認識率 (%)	94.9	90.7	86.4	86.9	81.8	76.3
リジェクト率 (%)	0.4	3.0	6.4	6.4	9.3	16.5
誤認識率 (%)	4.7	6.4	7.2	6.8	8.9	7.2

# 目次

---

1. 背景
2. 提案手法のアプローチ
3. 輪郭版GH
4. 提案手法
  1. 輪郭版GHの高速化
  2. 分離文字の認識
  3. 姿勢推定
5. 実験
6. **まとめ**



# 実時間カメラベース文字認識システム

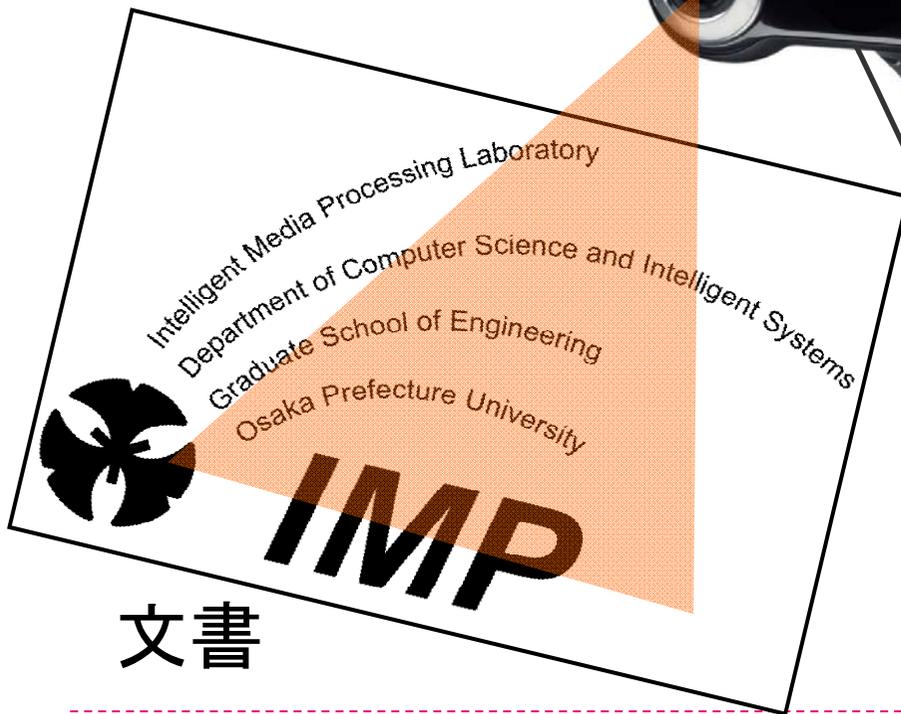
1秒間に200～250文字程度認識可能

リアルタイムに  
認識結果を出力

Webカメラ

IMP

キャプチャ



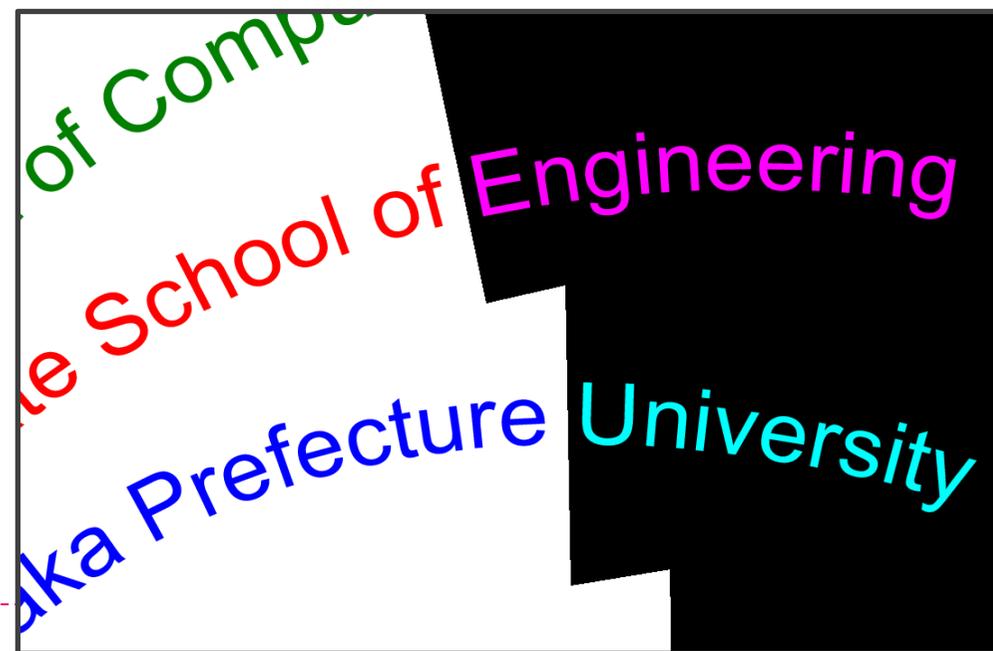
文書



# 今後の課題

---

- ▶ 漢字への対応
- ▶ 切り出し方法の改良
  - ▶ 連結成分の欠損への対応
  - ▶ 着色された文字への対応





大阪府立大学  
OSAKA PREFECTURE UNIVERSITY



# レイアウト非依存な 実時間カメラベース文字認識

岩村雅一 辻智彦 堀松晃 黄瀬浩一