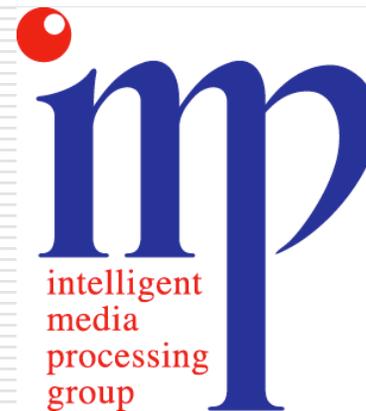


Large-scale pattern recognition and its application to recognition / retrieval of characters, documents and objects

Koichi Kise

Osaka Prefecture University, Japan



Self-introduction

- Koichi Kise
 - Osaka Pref. Univ. ; Professor
 - Visiting Professor at DFKI, Germany, 2000-2001
 - Research Area
 - Document Image Analysis
 - Object Recognition
 - Information Retrieval
-

Outline

- Introduction of the univ.
 - Introduction of the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

Outline

- Introduction of the univ.
 - Introduction of the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

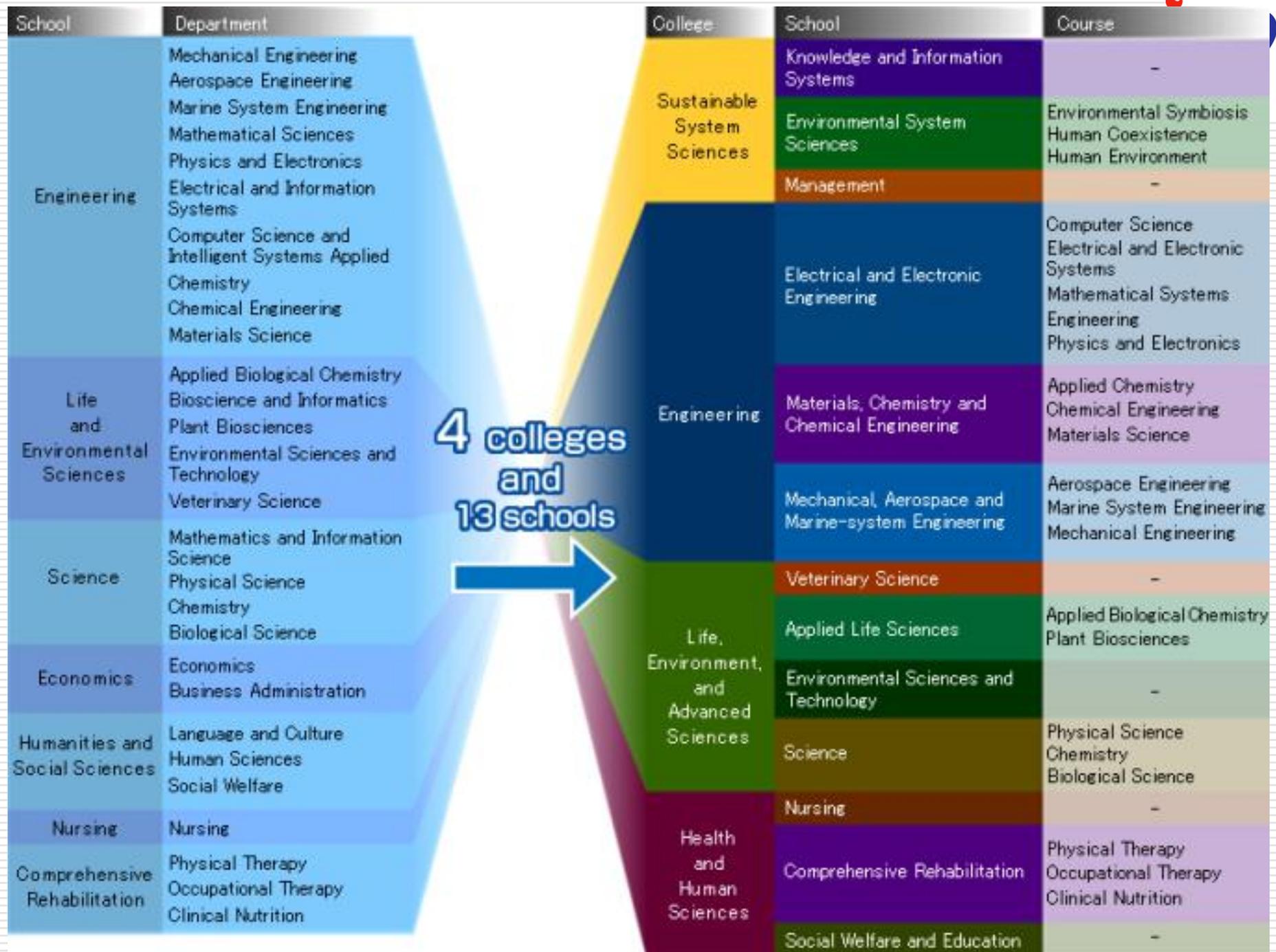


close to
Mozu Ancient
Tombs
(46 tombs)

Osaka Prefecture University

- 7 Schools & Graduate Schools
 - Engineering
 - Life & Environmental Science
 - Science
 - Economics
 - Humanities & Social Science
 - Nursing

They will be reorganized into 4 colleges and 13 schools in April 2012



Outline

- Introduction of the univ.
 - Introduction of the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

Introduction of my lab.

- Intelligent Media Processing Group
 - Since 2005
 - Logo

 - Members
 - Projects
 - Groups
 - Computational environment
-



Members

- Prof. Koichi Kise
 - Assoc. Prof. Masakazu Iwamura
 - Assist. Prof. Yuzuko Utsumi
 - Students (31)
 - Ph.D.: 3 (Japanese, German, Chinese)
 - Master: 11
 - Undergrad : 9 (B4) + 8 (B3)
-

More about students

□ Ph.D.

- Japanese student is JSPS research fellow
- Two foreign students supported by Japanese Gov. Scholarship
 - One will be JSPS research fellow from 2012

□ Masters

- Two are now staying in Germany

□ In addition

- One master course student finished his Master study in Germany
 - now hired by German Research Center for AI
-

Projects

- CREST
 - Strategic Funds for Promotion of Science and Technology
 - Grant-in-Aid for Scientific Research
 - A-Step
 - Joint project with Spain and Germany
-

CREST

□ Title

- Development of Fundamental Technologies for Innovative Use of **Character/Document Media** and Their Application to Creating **Human Harmonized Information Environment**

□ Partners

- Tohoku Univ. and Kyushu Univ.

□ Period: 5.5 years (now 2nd year)

Strategic Funds for Promotion of Science and Technology

□ Field:

- R&D program for implementation of anti-crime and anti-terrorism technologies for a safe and secure society

□ Partners:

- Osaka Univ., Waseda Univ., Wakayama Univ., Toshiba and National Institute of Police Science

□ Period: 5 years (2nd year)

Grant-in-Aid for Scientific Research

□ Scientific Research B

- Title: Large-scale specific object recognition and its application to real-world oriented web.
- Period: 3 years (2nd year)

□ Specially Promoted Research

- Title: Intelligent Camera-Pen
 - Period: 2 years (1st year)
-

A-Step

- Title: Intelligent camera-pen
 - Period: 1 year (has just started)
-

Joint Projects with Spain and Germany

- start from January 2012
 - Partners
 - Computer Vision Center, UAB, Barcelona
 - German Research Center for AI
 - Topics
 - Scene text recognition
-

Research Groups

□ Kise

- Image recognition group
- Document analysis group

□ Iwamura

- Character recognition group
- Nearest neighbor search group
- Computational photography

□ Utsumi

- Face recognition group
 - Behavior understanding group
-

Computational environment

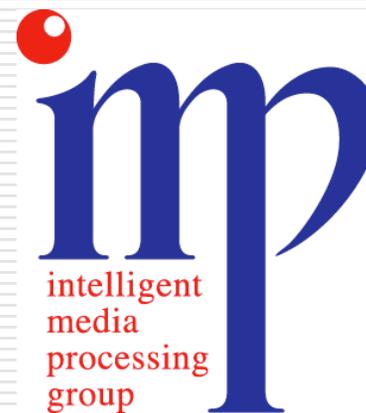
- For each student
 - Desktop (Core i5 or 7) + 2 display
 - optionally a laptop
- Servers
 - 760 cores
 - 4TB memory
 - 650TB disk
 - 10Gbps network



Outline

- Introduction of the univ.
 - Introduction of the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

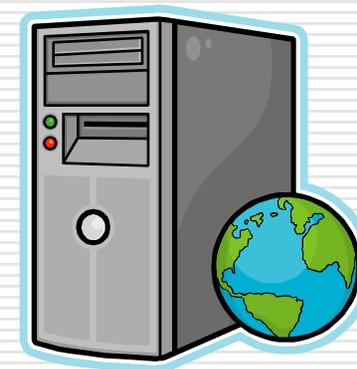
Introduction to Pattern Recognition



The task

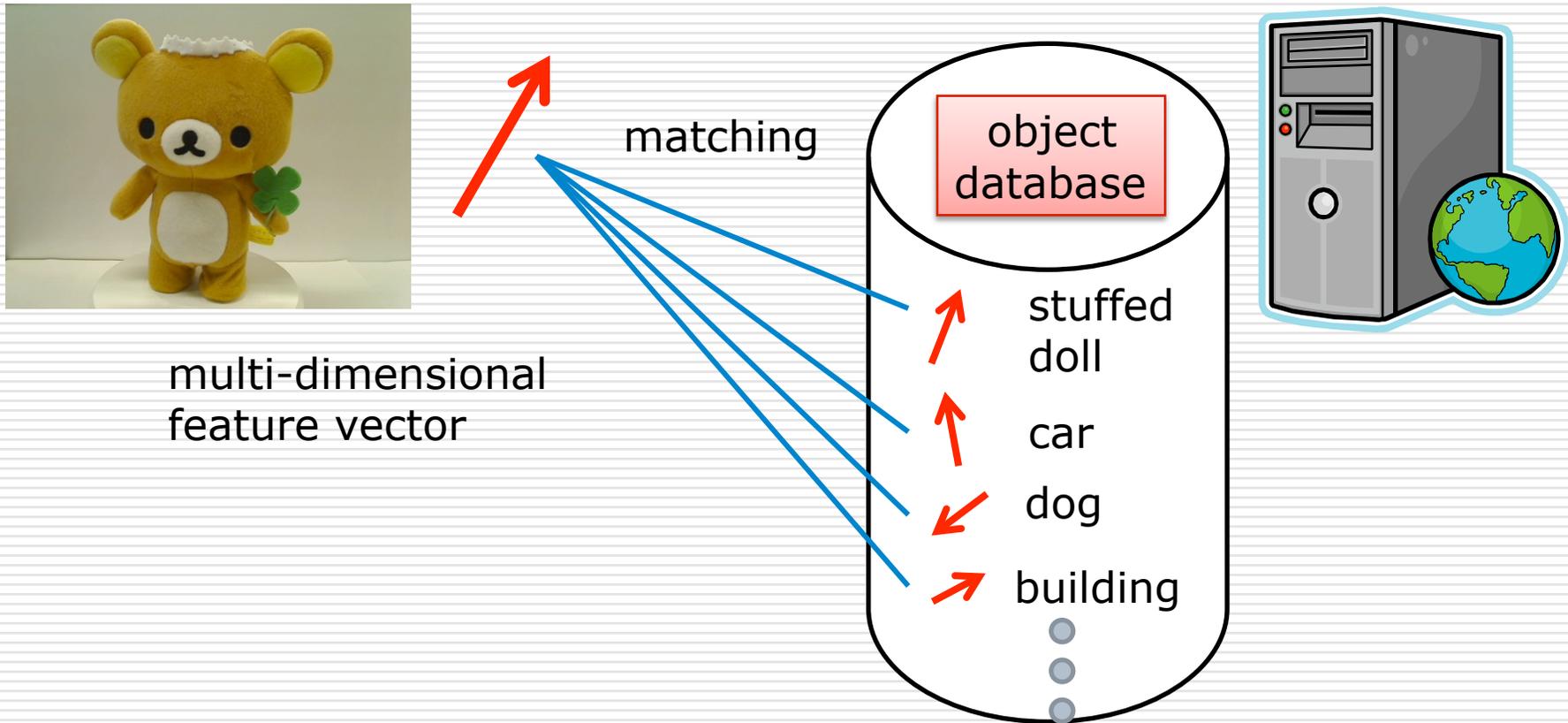


What's this?



This is a stuffed doll.
This is a Rirakkuma.

How can we realize it?



But many images...

- Same obj. with different viewpoints



- Same category, but different obj.



- We focus here on the first case.
-

Traditional Way

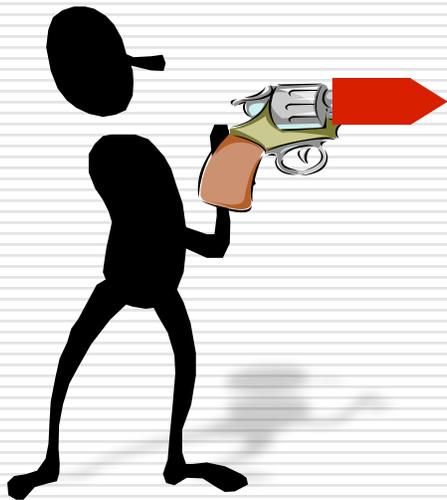
- Small amount of data are available
 - Single query and single feature vector for each object
 - Robust features
 - valid for many variants
 - Sophisticated matching
 - for not missing correct match
 - Problem
 - Limited accuracy achieved
 - Computationally expensive
-

New Way

- Huge data are available
 - Memory-based
 - Store all variants in the database
 - Employ multiple queries for better accuracy
 - Match each query to the stored variants
-

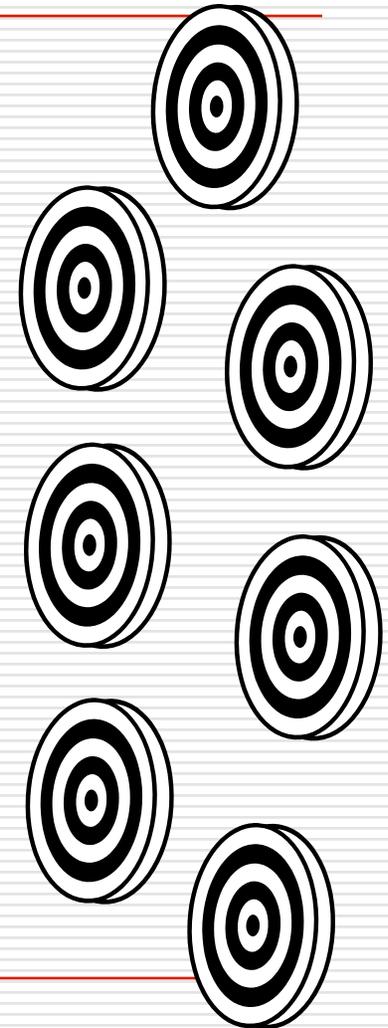
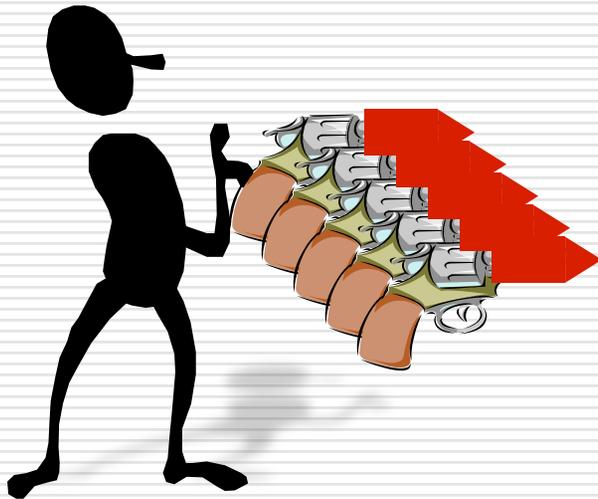
Comparison: Traditional Way

Single



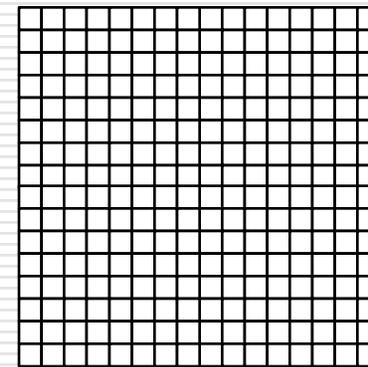
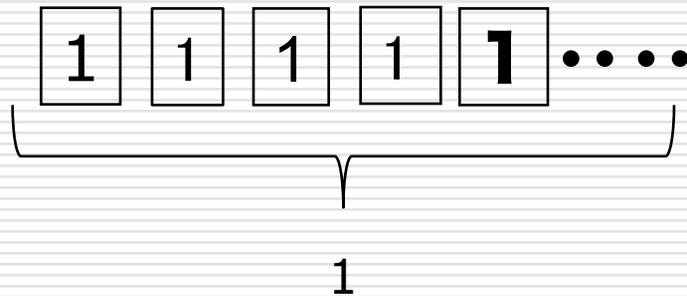
Comparison: New Way

multiple



Ultimate Pattern Recognition

- Store all possible patterns with their labels



for a 16 X 16 binary image,
how many possible patterns?

IMPOSSIBLE

$$2^{256} = 10^{79}$$

How many patterns we see?

- Suppose human vision is 30 fps
- Life would be 80 years
- How many frames?

$$80 \times 365 \times 24 \times 60 \times 60 \times 30 = 10^{11}$$

- less than 1 T ($=10^{12}$) frames

Space of realistic patterns are much smaller than that of possible patterns

Challenges

- How to obtain realistic patterns
 - How to store them
 - How to match a query to those stored patterns
-

Trade-off

- Better accuracy with more time
- Better accuracy with more memory

$$\frac{\text{Accuracy}}{\text{Time} \times \text{Memory}} \leq \text{Limit}$$

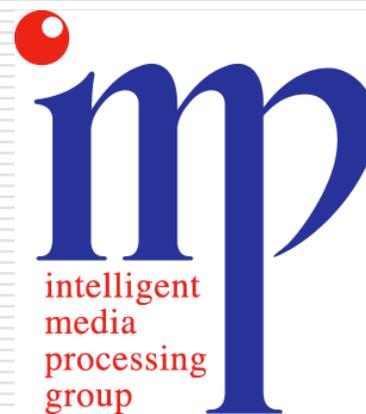
Find the theoretical limit

Develop a method for exchanging A, T and M

Outline

- Introduction of the univ.
 - Introduction of the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

Demos



Demos

- Large-Scale Document Image Retrieval
 - Large-Scale Object Recognition
 - Camera-Based Alphabet Recognition
 - Camera-Based Kanji Recognition
-

Large-Scale Document Image Retrieval

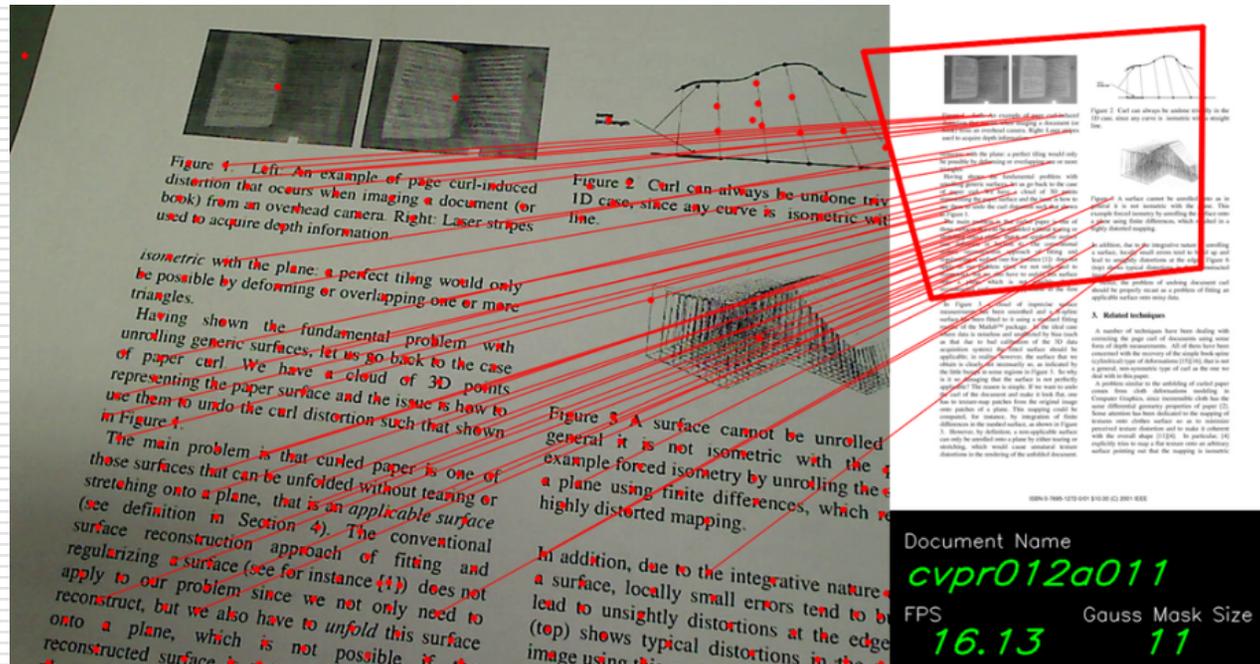


Figure 1. Left: An example of page curl-induced distortion that occurs when imaging a document (or book) from an overhead camera. Right: Laser-stripes used to acquire depth information.

Figure 2. Curl can always be undone in a 2D case, since any curve is isometric with a straight line.

Figure 3. A surface cannot be unrolled in general; it is not isometric with the plane. For example, forced isometry by unrolling the surface of a sphere onto a plane using finite differences, which results in highly distorted mapping.

Figure 4. The main problem is that curled paper is one of those surfaces that can be unfolded without tearing or stretching onto a plane, that is an applicable surface (see definition in Section 4). The conventional surface reconstruction approach of fitting and regularizing a surface (see for instance [1]) does not apply to our problem since we not only need to reconstruct, but we also have to unfold this surface onto a plane, which is not possible if the surface is not applicable.

Figure 5. A surface cannot be unfolded in general; it is not isometric with the plane. This example forced isometry by unfolding the surface of a sphere onto a plane using finite differences, which results in a highly distorted mapping.

Figure 6. In addition, due to the integrative nature of a surface, locally small errors tend to be amplified and lead to unsightly distortions at the edge of the reconstructed surface (top) shows typical distortions in the image using this approach.

A. Related techniques

A number of techniques have been dealing with unrolling the page out of documents using some form of depth measurements. All of them have been concerned with the accuracy of the depth measurements and the type of distortions of the 2D image. In a general, non-rigid type of curl as the one we are dealing with in this paper.

In general, non-rigid type of curl as the one we are dealing with in this paper.

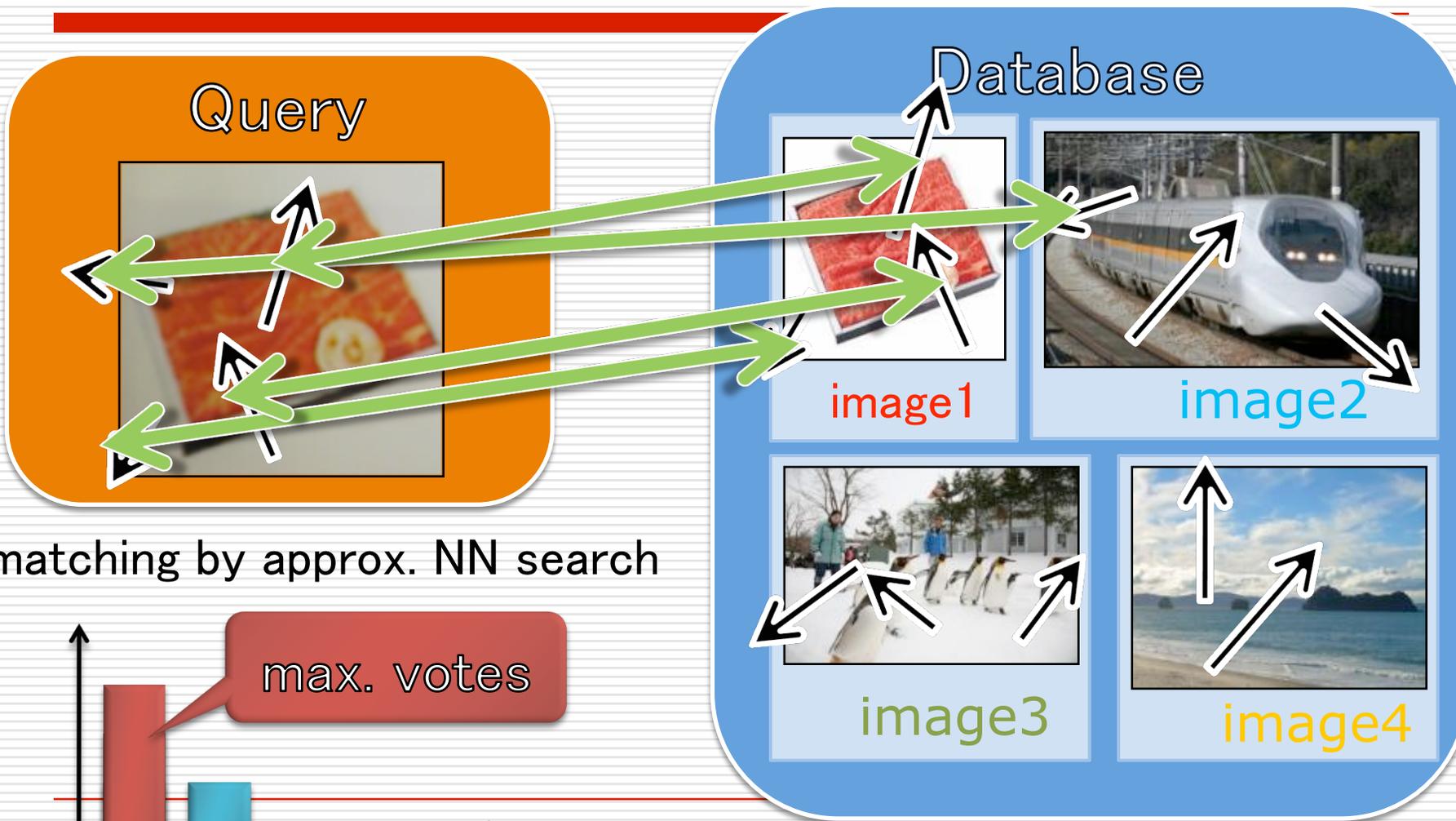
In general, non-rigid type of curl as the one we are dealing with in this paper.

Document Name
cvpr012a011

FPS 16.13 **Gouss Mask Size** 11

1 Million page DB

Large-Scale Object Recognition



1

2

3

4

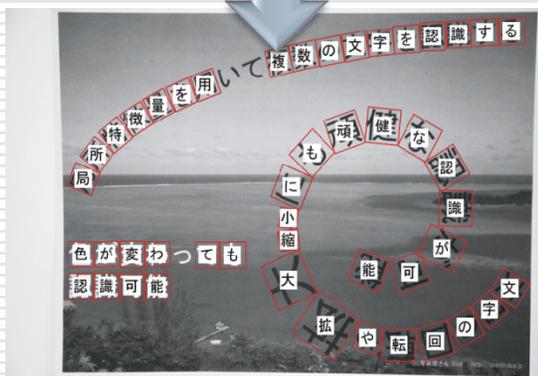
Camera-Based Alphabet Recognition



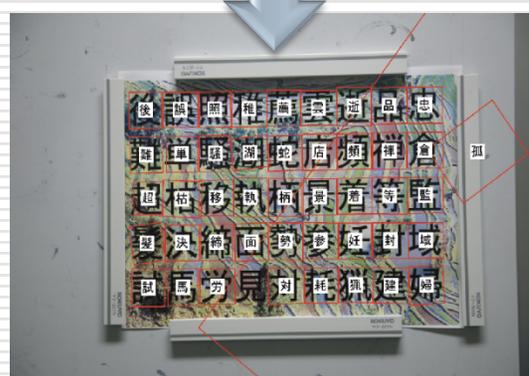
Camera-Based Kanji Recognition



Various layout



Complex. Bg



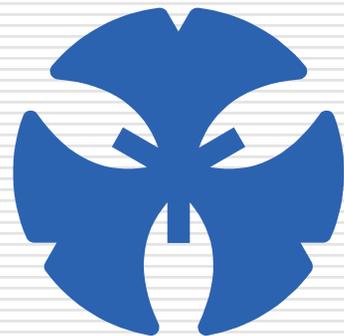
Outline

- Introduction to the lab.
 - Introduction to Pattern Recognition
 - Demos
 - Selected Topics
-

Please select:

- Large-Scale Document Image Retrieval
 - Large-Scale Object Recognition
 - Camera-Based Alphabet Recognition
 - Camera-Based Kanji Recognition
-

Selected Topics



大阪府立大学
OSAKA PREFECTURE UNIVERSITY

