The 9th COE Technical Presentation 12/1/2005

Evaluation of Location-based Photo Captioning System

Vision and Media Computing Lab.
IWASAKI Kiyoko

Management of Photos

Spread of digital cameras

⇩ Few effective methods or systems to manage photos easily

Unorganized photos are:
- not used.
- difficult to avail.

Management of photos based on:
- date/time
- event
- people
- location

Need manual entry of information

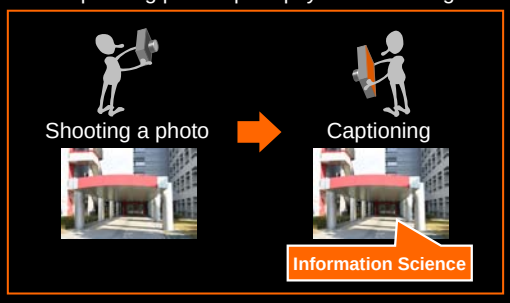
Related Works - Location-based Photo Management

- Acquiring shooting positions of photos
  - via manual entry.
  - from location-aware device.
  - from digital calendar.
  - from surrounding text. [Toyama et al., 2003]
- Organizing photos with geographic coordinates [Naaman et al., 2004]
  → There is a possibility that data appropriate for a photo are not included in a dataset prepared in advance.
- Suggesting location-based metadata of photos with mobile devices [Sarvas et al., 2004]
  → Photos are not processed in practical time.

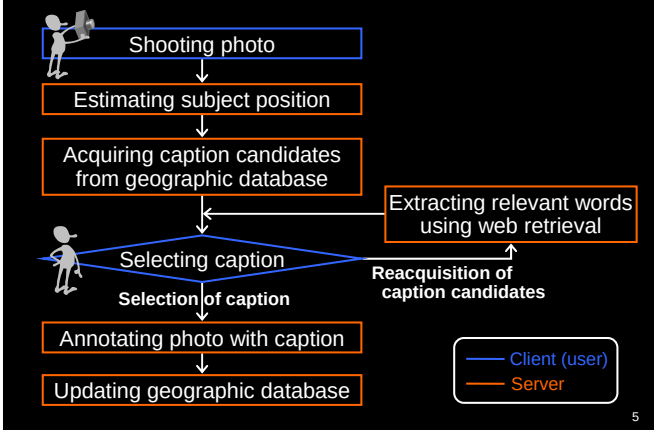
Location-based Photo Captioning

Semi-automatic location-based photo captioning system

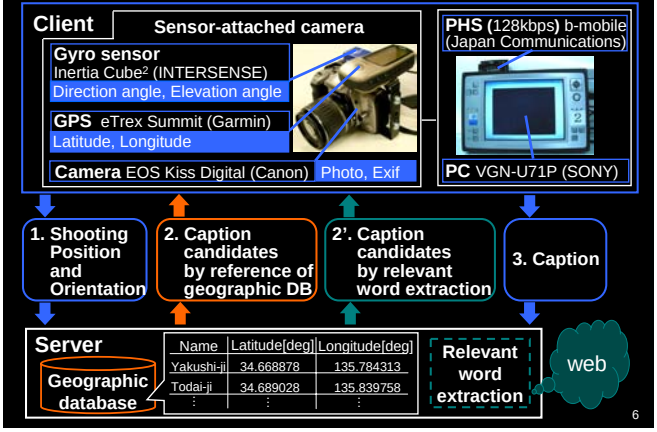
Captioning photos promptly after shooting



Flow Diagram of Photo Captioning

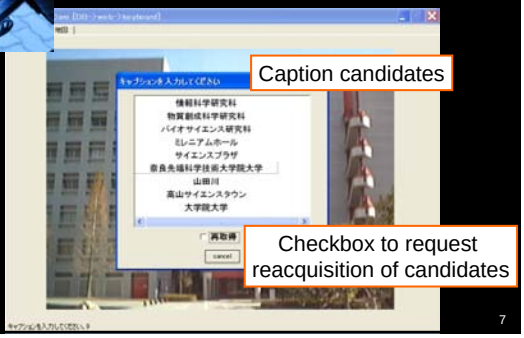


Prototype System



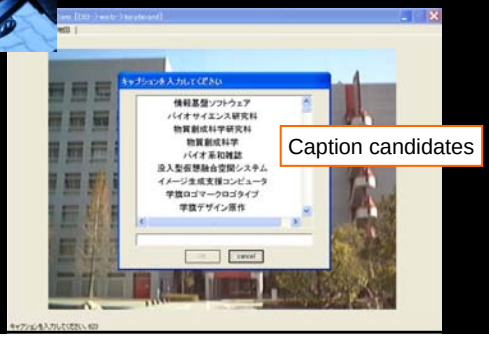
Captioning Interface - DB

Selecting a caption from candidates acquired by reference of geographic DB



Captioning Interface - web

Selecting a caption from candidates acquired by relevant word extraction using web retrieval



Captioning Interface - key

Inputting a caption using a keyboard

Overview of Experiments

- Evaluating 3 types of captioning

[DB]	Selecting a caption from candidates acquired by reference of geographic DB
[web]	Selecting a caption from candidates acquired by relevant word extraction using web retrieval
[key]	Inputting a caption using a keyboard

- Evaluation points
  - Used captioning type (number of photos captioned by each captioning type)
  - Inputting time to caption a photo
  - Presented rank of selected caption in candidates
  - Position of user registered data in geographic DB

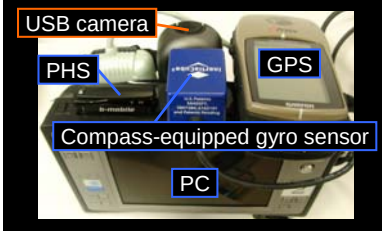
Configuration of Evaluated System Variations

6 system variations

System	Camera	Captioning type	Database**
1-1	Digital Still Camera (DSC)	DB → web →key	Used
1-2		DB → key	Used
1-3		key*	Not used
2-1	USB camera	DB → web → key	Used
2-2		DB → key	Used
2-3		key*	Not used

*key: Adding a function presents candidates using history and makes a short list of the candidates along with key inputting
**Database: Registering data included in a map software (Alps "Proatlas W2" facility data)

Prototype System Using USB Camera



Using shooting positions instead of estimated subject positions when referencing the geographic DB.

Configuration of Evaluated System Variations

6 system variations

System	Camera	Captioning type	Database**
1-1	Digital Still Camera (DSC)	DB → web → key	Used
1-2		DB → key	Used
1-3		key*	Not used
2-1	USB camera	DB → web → key	Used
2-2		DB → key	Used
2-3		key*	Not used

*key: Adding a function presents candidates using history and makes a short list of the candidates along with key inputting
**Database: Registering data included in a map software (Alps "Proatlas W2" facility data)

13

User Tasks

12 men and women age 22 to 34

1. Shooting 4 photos using each system in campus
   - A) Information Science : 2 photos
   - B) Biological Sciences : 1 photo
   - C) Materials Science
   - D) Millennium Hall
   - E) Guesthouse Sentan
   - F) Science Plaza

   } : 1 photo selected by user freely

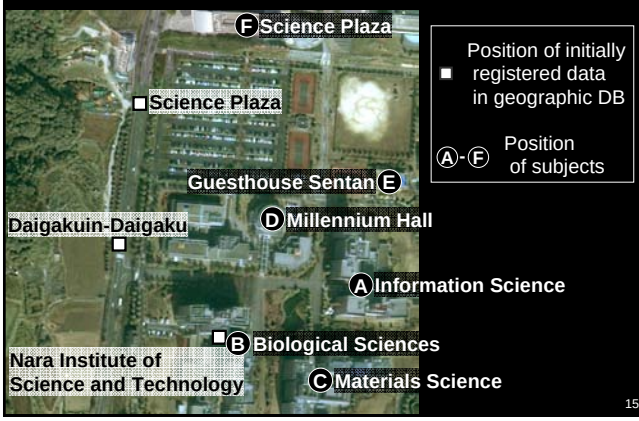
2. Captioning photos

   Captions of the buildings are specified in advance.

*** The order of evaluating the systems is randomly changed.**

14

Initial Database and Subjects



15

Examples of Photos

288 photos were acquired in whole experiments.

Shot with different distances and angles

16

Evaluation of Captioning Photos

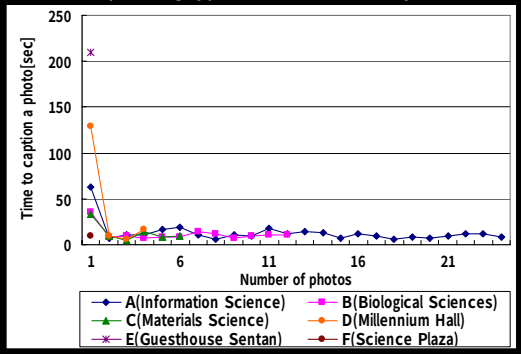
System No.	1-1			1-2		1-3	2-1			2-2		2-3
Captioning type	DB	web	key	DB	key	key	DB	web	key	DB	key	key
Number of photos	43	3	2	44	4	48	44	3	1	45	3	48
Average time of captioning* [sec]	10	44	113	10	52	26	8	46	108	9	38	22
Average rank of caption	1.8	66.3	-	2.9	-	-	2.4	66.3	-	2.8	-	-

*Captioning time: elapsed time from shooting a photo to finishing the captioning

17

Captioning Time – System 1-1

Captioning type: DB → web → key



F: initially registered data **A,B,C,D,E: user registered data**

18

Registering Data in DB by Captioning
- Shooting and Estimated Subject Position

Subject: Building A (Information Science)
System: 1-1

■ Shooting Position ■ Estimated subject position
× Position of user registered data
Ⓐ Actual position of Building A (Information Science)

The registered position is close to the shooting positions rather than the actual position.

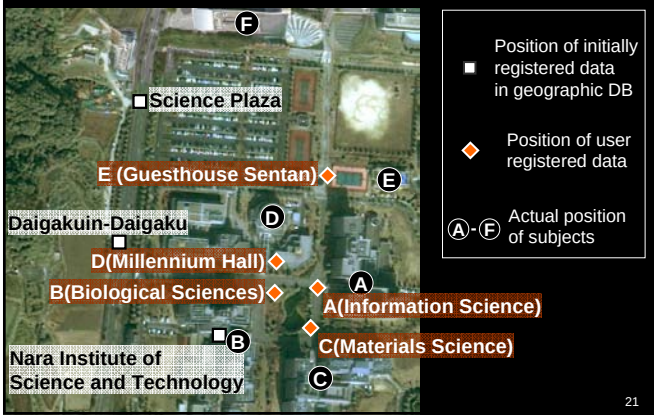
Registering Data in DB by Captioning
- Shooting Position of System 1-1 and 2-1

Subject: Building A (Information Science)
System: 1-1 **System: 2-1**

■ Shooting Position
× Position of user registered data
Ⓐ Actual position of Building A (Information Science)

The registered position of system 1-1 is a little closer to the actual position than the position of system 2-1.

Registered Data in DB by Captioning
- Using System 1-1



Summary

Evaluation of semi-automatic location-based photo captioning system

- Captioning type DB: 10 [sec]
  -> Captioning photos promptly after shooting
- Captioning type web: 45 [sec]
- Positions of user registered data are closer to their shooting positions than the actual positions of subjects.

Future Work

- Improvement of registering position of new data
- Experiments in densely built-up area