The role of context in image interpretation

Prof. Dag Elgesem and Assoc. Prof. Joan C. Nordbotten
{Dag.Elgesem, Joan.Nordbotten}@infomedia.uib.no

Department of information science and media studies, Univ. of Bergen

Abstract.
With the increasing availability and use of mobile phones with camera functionality and Internet access, it becomes feasible to use these devices to query image databases for information. However, the photo/image query from the phone-camera commonly depicts several objects and can thus be open to a number of interpretations. This leads to ambiguities when determining the intention of an image used as the basis for an information retrieval query.

We suggest that this problem is structurally similar to the problem of how to interpret an ambiguous sentence, and that the task can be modeled in a similar way. Though the role of context is a key factor in the solution of the problem of disambiguation of text, we argue that existing accounts of context do not explain the role that context can play in image interpretation.

In this paper, we propose a definition of image *context* based on the role of *common knowledge* for successful communication. We then show how disambiguation of images as queries can be modeled as a game with partial information. On the basis of this, a more precise account of the role of context in image interpretation is proposed, which we believe can function as the basis for development of a Multi-modal Information Retrieval System, MIRS.

The role of context in image interpretation

Dag Elgesem
Joan Nordbotten

Dept. of Information and Media Science
University of Bergen

1 Mobile Phone Visual Queries

Image + GPS

2 But, what is the question?

Foreground/background:

The sailboat? Statsraad Lemkul

The wharf? Bryggen i Bergen

Whole/component:

The cathedral? Nidaros Domkirken

A statue? Apostle Peter

3 Resolving Ambiguity

based on:

- Object identification
- Image & object CONTEXT
- Interpretation probabilities

4 CONTEXT

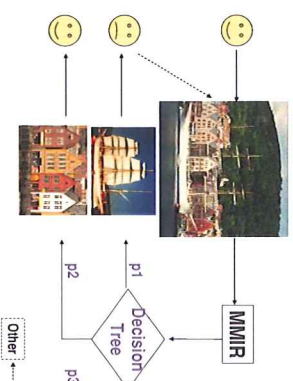Def: the information that must be

common knowledge

for communication to succeed.

Examples:

- Location & time
- Object representation
- Interpretation options & probabilities

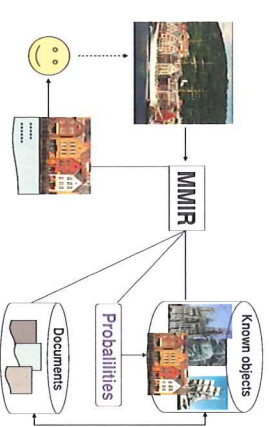**5 Visual query interpretation**

6 Determining user intention

Based on:

- Object relationships
- Camera settings (focus, perspective, ...)
- User location & feedback
- MMIR settings & 'learning'

7 Mobile Multi-modal Information Retrieval

MMIR:

8 Summary

Given

- A set of non-ambigous images
- Location(s) for source photos
- Object probabilities

Then

- Map query object(s) to stored images
- Determine probability of query intention

http://caim.uib.no