Eliciting Concepts of Place for Text-based Image Retrieval

Alistair J. Edwardes
Department of Geography
University of Zurich
Winterthurerstrasse. 190
CH-8057 Zurich
+41-44-635 52 56
aje@geo.unizh.ch

Ross S. Purves
Department of Geography,
University of Zurich
Winterthurerstrasse. 190
CH-8057 Zurich
+41-44-635 65 31
rsp@geo.unizh.ch

ABSTRACT
This paper describes research being employed in the Tripod project to improve the retrieval of photographs through a comprehensive knowledge of where they were taken. The methods described here attempt to elicit semantics, and thus terms for use in indexing, related to the idea of place from theoretical geography. Three approaches are outlined: literature-based, empirical experiments, and data mining.

Categories and Subject Descriptors
H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing – Abstracting methods, dictionaries, Indexing methods, Linguistic processing, thesauruses

General Terms
Experimentation, Human Factors, Standardization, Languages, Theory.

Keywords
place, semantics, image retrieval, basic levels, concept ontology, natural language description.

Geographic Information Retrieval (GIR) tends to focus on queries which include a toponym, a spatial relationship and some contextual information – for example, “castles near Edinburgh”. Toponyms are identified within documents and indexed through geoparsing and geocoding such that spatial relationships may be exploited. All other terms are generally stored in a simple textual index and can be retrieved through standard IR techniques [9]. However, where a resource is stored with only an associated location (e.g. a set of coordinates) it is possible to imagine exploiting spatial information in a different way – by building descriptions of the location from geographic data (e.g. a sandy beach or a steep mountain). Such descriptions could then be added to the indexing terms for a given resource, based not on the resource’s textual content, but rather the resource’s geographic footprint.

The importance of geographic location to image description has been frequently highlighted by researchers working in image retrieval. For example, Shatford [12] identifies specific, generic and symbolic locations, associated with her ‘where’ facet, as amongst the most important descriptors for indexing images. This work was applied in practice by Armitage and Enser [1] and Choi and Rasmussen [3], who used Shatford’s classification to analyse requests for images sent to various picture libraries.

The Tripod project has taken up such observations to examine how the synergy of geographic information science (GIScience) with image retrieval might improve access to digital image resources. A key hypothesis in this project is that much of the semantics describing a picture can be derived through knowing where a picture was taken. In seeking to exploit this proposition, the project draws on a broad interpretation of where relating not only to a specific location [7], but also to its characteristics as a setting, such as its constituent features (e.g. buildings and landforms), and cultural, conative and affective meaning. Within geographic theory such aspects are commonly evoked through the concept of place (c.f. [5]).

The research described here reports on three different methods that are being employed in the Tripod project to elicit place semantics in order to generate a concept-ontology for use in image retrieval. These are:

1. A review of literature to identify methodologies and dimensions for describing place and environmental conceptualisation.
2. A set of empirical experiments eliciting a database of terms in response to photographs of different types of European landscapes.
3. Analysis of a database of image captions describing photographs that document the geography of the United Kingdom.

1. PREVIOUS RESEARCH

1.1 Dimensions of place
Within geography, the subject of place was introduced in its modern form by the seminal work of the phenomenologists Tuan [14] and Relph [10]. These researchers focused on describing how place comes about when people ascribe meaning to otherwise ‘undifferentiated space’. Tuan, in particular, considered how places are created by direct experience and sensation. Relph focused on describing the essence of place as well as more existential factors such as feelings of belonging. He highlighted three elements of places, their physical setting (their location,
physical appearance and changing character over time), the activities performed in them, and the meanings they have for people.

### 1.2 Descriptions of place

A number of researchers have also been interested in how places are conceptualised within natural language. Early work included that of Craik [4] who developed a list of adjectives describing landscapes elicited through experiments stimulated by photographs and Kasmar [8], who produced a similar lexicon for the built environment using descriptive terms identified in architectural and interior design magazines. Tversky and Hemenway [15] looked at how people categorise scenes in language and memory using so-called basic levels [11]. They identified a number of different basic level scenes (e.g. ‘mountain’, ‘park’, ‘beach’ and, ‘city’) based on the range of terms that differently described scenes invoked in response to the headings ‘parts’, ‘activities’ and, ‘attributes’. Smith and Mark [13] also looked at the issue of how geographic concepts are encoded in everyday language, drawing on the methodology of Battig and Montague [2] for eliciting category norms.

### 2. EXPERIMENT DESIGN

Previous research has demonstrated the value of eliciting terms from participants in empirical studies. Hence, a set of experiments have been designed aimed at obtaining representative descriptors and the dimensions that might structure these. They each employ stimuli consisting of photo sets of natural landscapes from different locations in Europe. Three types of experiment have been developed, a semi-structured free description experiment, a similarity sorting experiment and, a controlled vocabulary description experiment, of which the first two are described here.

The aim of the free description experiment is to build a database of descriptive terms associated with particular photographs. In the experiment the participant is shown randomly selected photographs, and for each asked to give terms in response to four headings; overall (they are asked for a short phrase to describe the scene in its entirety, e.g. a relaxing beach), activities (they are asked to describe things in the picture), elements and qualities that might be done at the locations shown), and qualities (they are asked to give adjectives describing the feelings and ideas invoked by the picture) (Figure 1).

The aim of the sorting experiments is to identify images that are alike and dimensions that people use to evaluate similarity. In this experiment participants are presented with randomly selected photographs and asked to sort them into groups using whatever criteria they feel is appropriate. The participant is then asked to describe the particular characteristics of each of their groups (Figure 2).

### 3. IMAGE DESCRIPTION ANALYSIS

The final method uses a dataset obtained from Geograph (www.geograph.org.uk). Geograph is a project with the aim of collecting “geographically representative photographs and information for every square kilometre of the UK and the Republic of Ireland”. Since each photograph has an associated description given by the photographer we contend that the dataset may well be a digital representation of place.

Experiments are being conducted with this dataset to attempt to identify characteristics (e.g. components, activities, and qualities) of basic level scene categories using co-occurrence of terms taken from seed word lists. For example, Table 1 describes the most frequently used qualities of the basic level scenes ‘beach’, ‘village’, ‘hill’ and ‘mountain’ using the adjective list of [4].

---

**Figure 1** Interface for the free description experiment

**Figure 2** The describe task of the sorting experiment
Table 1. Most commonly co-occuring adjectives for four basic level scene categories.

<table>
<thead>
<tr>
<th>Beach</th>
<th>Village</th>
<th>Hill</th>
<th>Mountain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy</td>
<td>Deserted</td>
<td>Steep</td>
<td>Distant</td>
</tr>
<tr>
<td>Deserted</td>
<td>Pretty</td>
<td>Distant</td>
<td>Black</td>
</tr>
<tr>
<td>Eroded</td>
<td>Green</td>
<td>Wooded</td>
<td>Remote</td>
</tr>
<tr>
<td>Soft</td>
<td>Quiet</td>
<td>Black</td>
<td>Rocky</td>
</tr>
<tr>
<td>Rocky</td>
<td>Lovely</td>
<td>Rough</td>
<td>Grassy</td>
</tr>
<tr>
<td>Warm</td>
<td>Pleasant</td>
<td>Grassy</td>
<td>Steep</td>
</tr>
<tr>
<td>Glacial</td>
<td>Beautiful</td>
<td>Round</td>
<td>Natural</td>
</tr>
<tr>
<td>Low</td>
<td>Remote</td>
<td>Big</td>
<td>Dark</td>
</tr>
<tr>
<td>Beautiful</td>
<td>Unusual</td>
<td>White</td>
<td>Broad</td>
</tr>
<tr>
<td>Lovely</td>
<td>Large</td>
<td>Broad</td>
<td>Running</td>
</tr>
</tbody>
</table>

4. CONCLUSIONS
This paper reports on the early stages of on going research to gather terms and their relationships for use in a geographic concept ontology to support image retrieval based on the location of the images themselves. The methodology applied includes a theoretical grounding based on the semantics of place, together with the design and implementation of a set of experiments based on empirical experiments where users are asked to describe and group images and abductive approaches where a large database of geographically grounded image descriptions has been analysed to explore terms and their co-occurrences.

Future work will include analysis of the results with regard to the need for developing the ontology itself. One possible technique for achieving this that is currently being explored is Latent Semantic Indexing [6].

5. ACKNOWLEDGMENTS
We would like to acknowledge contributors to Geograph British Isles, see http://www.geograph.org.uk/credits/2007-02-24, whose work is made available under the following Creative Commons Attribution-ShareAlike 2.5 Licence (http://creativecommons.org/licenses/by-sa/2.5/). This research reported in this paper is part of the project TRIPOD supported by the European Commission under contract No. 045335.

6. REFERENCES