TALP at MediaEval 2010 Placing Task: Geographical Focus Detection of Flickr Textual Annotations

Daniel Ferrés
TALP Research Center
Software Department
Universitat Politècnica de Catalunya
C. Jordi Girona Salgado, 1-3
08034 Barcelona, Spain
dferres@lsi.upc.edu

Horacio Rodríguez
TALP Research Center
Software Department
Universitat Politècnica de Catalunya
C. Jordi Girona Salgado, 1-3
08034 Barcelona, Spain
horacio@lsi.upc.edu

ABSTRACT
This paper describes our geographical text analysis and geotagging experiments in the context of the Multimedia Placing Task at MediaEval 2010 evaluation. The task consists of predicting the most probable coordinates of Flickr videos. We used a Natural Language Processing approach trying to match geographical place names in the Flickr users textual annotations. The resources employed to deal with this task were the Geonames geographical gazetteer, stopwords lists from several languages, and an electronic English dictionary. We used two geographical focus disambiguation strategies, one based on population heuristics and another that combines geographical knowledge and population heuristics. The second strategy does achieve the best results. Using stopwords lists and the English dictionary as a filter for ambiguous place names also improves the results.

Categories and Subject Descriptors
H.3 [**Information Search and Retrieval**]

General Terms
Design, Performance, Experimentation, Measurement

Keywords
Place Names Recognition, Geographical Tagging, Geographical Knowledge Bases, Natural Language Processing

1. INTRODUCTION
In this paper we present our system, experiments, and conclusions in the context of the MediaEval 2010 Placing task evaluation. The MediaEval 2010 Placing task requires that participants automatically assign geographical coordinates (latitude and longitude) to Flickr videos using one or more of: Flickr metadata, visual content, audio content, and social information. The data set are composed by 5125 and 5091 videos for the development and test sets respectively.

MediaEval 2010 Workshop, October 24, 2010, Pisa, Italy

Evaluation of results is done by calculating the distance from the actual point (assigned by a Flickr user) to the predicted point (assigned by a participant). Runs are evaluated finding how many videos were placed at least within some threshold distance: 1 km, 5km, 10km, 50km and 100km.

1.1 Motivation
In the context of GeoCLEF Geographical Information Retrieval (GIR) evaluations from 2005 to 2008 we developed several approaches for Geographical Information Retrieval in English with closed newspapers collections [1]. The MediaEval 2010 Placing task gives us an opportunity to deal with multilingual non grammatical language issues appearing in the users textual annotations and geographical disambiguation strategies. Our approach has the objective of maximising the precision of the predicted places at the expense of recall.

1.2 Issues in the Flickr textual annotations
After an analysis of the users textual annotations of the development test we detected some issues in the task of recognizing the textual use of the place names in the textual annotations: 1) joined place names (e.g. riodejaneiro, buenosaires) 2) acronyms (e.g. L.A., NY, MN), 3) parts of the place name (e.g. rio, paulo), 4) Place names with affixes (halloweenbrazil, brazilguides, inbraziltours), 5) multilingual place names (e.g. Cataratas de Iguaçu, iguazufalls, iguaçufalls), 6) place name plus a feature name (e.g. iguazufalls, newyorkcity), 7) orthographic errors (e.g. Rio da Janeiro, sao Paulo). We also detected some typical ambiguity problems: nouns that could be tagged as place names and viceversa (e.g. aurora (name), Aurora (place name)), and the referent ambiguity problems with place names (e.g. Barcelona, Spain or Barcelona, Colombia).

2. SYSTEM DESCRIPTION
The system can be divided in three phases: Place Names Recognition, Place Names Disambiguation and Geographical Focus Disambiguation. The Place Names Recognition phase uses the Geonames[1] Gazetteer for detecting the place names in the textual annotations. The Geonames gazetteer has been used in GIR and geo-tagging approaches[4] and allows us to deal with the recognition issues related with multilinguality, acronyms, lower and uppercase place names,

[1]**Geonames.** http://www.geonames.org

and joined place names. We use the following information from each Geonames place: country, state, and continent of the place, feature type, coordinates, and population. The Place Names Disambiguation phase tries to avoid ambiguity in Place Names Recognition due to the huge number of place names with meaning that could be a noun (e.g. aurora (noun), Aurora (city)). This phase uses stopwords lists[2] in several languages (including English) and an English Dictionary of 71.348 words from Freeling[3] software to filter names that could be erroneously tagged as place names.

The Geographical Focus Disambiguation phase uses some of the Toponym Resolution strategies presented in the GIR literature [3]. We assume the one reference per discourse hypothesis: one geographical place/coordinates per video. This phase has been designed with some heuristics described in [2] and [3]. Using the information of all possible referents of all the place names detected by the Geonames (phase 1) and not filtered (phase 2) we apply the following heuristics:

- **Population heuristics**. Only using population information to disambiguate between all the possible places. We used the following rules: 1) if exists a place select the most populated place that is not a country, state (administrative division type 1) or a continent, 2) otherwise if exists a state, select the most populated one, 3) otherwise select the most populated country, 4) otherwise select the most populated continent.

- **Geographical Knowledge and Population heuristics**. From the set of different places appearing in the text apply the following rules in this priority order to select the scope (focus) of the text: 1) select the most populated place that is not a state, country or continent and has its country or its state apearing in the text, 2) otherwise select the most populated state that has his country apearing in the text 3) otherwise apply the population heuristics presented above.

3. EXPERIMENTS AND RESULTS

For the MediaEval 2010 Placing task evaluation we designed a set of experiments that consist in tagging the test set and applying different baseline configurations (see Table 1). We used only textual content from the following metadata fields from the Flickr videos to perfom the task: Title , Description, and Keywords. We allowed the Gazetteer the recognition of place names of a maximum of five tokens (e.g. Sierra Nevada de Santa Marta). Our best results (see Table 2) are achieved by the TALP_2 run which has country/state heuristics and population heuristics combined with the use of stopwords and the English dictionary. The number of videos in which their geographical focus could not be predicted were 918 videos for runs TALP_1 and TALP_2, 454 for runs TALP_3 and TALP_4, and 410 for the run TALP_5. In those videos the latitude and longitude were set to 0.0 0.0 because there were no place names detected.

4. CONCLUSIONS

[2]Lingua::StopWords 0.09. http://search.cpan.org/dist/Lingua-StopWords

[3]Freeling 2.1. http://nlp.lsi.upc.edu/freeling/

Table 1: MediaEval 2010 Placing task Experiments.

run	Parameters		
	Disambiguation	StopWords	Dictionary
TALP_1	population	yes	yes
TALP_2	knowledge+population	yes	yes
TALP_3	knowledge+population	yes	no
TALP_4	population	yes	no
TALP_5	knowledge+population	no	no

Table 2: Results at the MediaEval 2010 Placing Task with Test Data (5091 videos).

run	#videos_correctly_predicted				
	1km	5km	10km	50km	100km
TALP_1	441	1417	1811	2227	2271
TALP_2	**536**	**1665**	**2153**	**2635**	**2740**
TALP_3	510	1604	2052	2526	2635
TALP_4	413	1315	1698	2092	2126
TALP_5	497	1587	2035	2507	2615

We used an approach based on Geographical Knowledge Bases to deal with the MediaEval 2010 Placing Task. The strategy that combines geographical knowledge and population heuristics for geographical focus detection achieves the best results in the experiments. Our experiments show that stopwords lists and controlled dictionaries can help the disambiguation of placing names and the focus detection. Future work, includes the study of the use of controlled dictionaries and disambiguation techniques in order to reduce the impact of geographical ambiguity problems.

Acknowledgments

This work has been supported by the Spanish Research Dept. (KNOW 2, TIN2009-14715-C04-04). Daniel Ferrés is supported by the EBW II Project, which is financed by the European Commission within the framework of the Erasmus Mundus Programme. TALP Research Center is recognized as a Quality Research Group (2001 SGR 00254) by DURSI, the Research Department of the Catalan Government.

5. REFERENCES

[1] Daniel Ferrés and Horacio Rodríguez. TALP at GeoCLEF 2007: Using Terrier with Geographical Knowledge Filtering. In *Working Notes for the CLEF 2007 Workshop*, September 2007.

[2] Alexander G. Hauptmann and Andreas M. Olligschlaeger. Using Location Information From Speech Recognition Of Television News Broadcasts. In *Proceedings of the ESCA ETRW Workshop on Accessing Information in Spoken Audio*, pages 102–106, Cambridge, England, 1999. University of Cambridge.

[3] J. L. Leidner. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names.* PhD Thesis, University of Edinburgh, 2007.

[4] Pavel Serdyukov, Vanessa Murdock, and Roelof van Zwol. Placing flickr photos on a map. In James Allan, Javed A. Aslam, Mark Sanderson, ChengXiang Zhai, and Justin Zobel, editors, *SIGIR*, pages 484–491, 2009.