TALP at MediaEval 2011 Placing Task: Georeferencing Flickr Videos with Geographical Knowledge and Information Retrieval

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 Georeferencing approaches, experiments, and results at the MediaEval 2011 Placing Task evaluation. The task consists of predicting the most probable geographical coordinates of Flickr videos. Our approaches used only Flickr users textual annotations and tagsets to predict. We used three approaches for this task: 1) a Geographical Knowledge approach, 2) an Information Retrieval based approach with Re-Ranking, and 3) a combination of both (GeoFusion). The GeoFusion approach achieved the best results within the margin of errors from 10km to 10000km.

Categories and Subject Descriptors

H.3 [**Information Search and Retrieval**]

General Terms

Design, Performance, Experimentation, Measurement

Keywords

Georeferencing, Toponym Disambiguation, Geographical Knowledge Bases, Information Retrieval

1. INTRODUCTION

The MediaEval 2011 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 (see [1] for more details about this evaluation). 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 distances.

2. SYSTEM DESCRIPTION

We used three approaches for the MediaEval 2011 Placing Task (see more details about this approaches in [4]):

MediaEval 2011 Workshop, September 1-2, 2011, Pisa, Italy

1) Geographical Knowledge (GeoKB). This approach was used for MediaEval 2010 Placing Task [3] and then was improved (see [4]). The GeoKB approach uses the Geonames[1] Gazetteer for detecting the place names, stopwords lists, and an English Dictionary. The system uses the following rules from Toponym Disambiguation techniques [4] to get the geographical focus of the video: 1) select the most populated place that is not a state, country or continent and has its state apearing in the text, 2) otherwise select the most populated place that is not a state, country or continent and has its country apearing in the text, 3) otherwise select the most populated state that has its country apearing in the text 4) otherwise apply population heuristics.

2) Information Retrieval with Re-Ranking. This approach is similar to the one presented by [6]. It uses the Terrier[2] IR software (version 3.0) with the Hiemstra Language Modelling (HLM) weighting model [5]. The HLM default lambda (λ) parameter value in Terrier (0.15) was used. See in equation *1* the Terrier implementation of the HLM Weighting model (version 1 [5]) score of a term t in document d ; where $tf_{t,d}$ is the term frequency in the document, cf_t is the collection frequency of the term, $\sum_i cf_i$ is the number of tokens in the collection, and $\sum_i tf_{i,d}$ is the document length.

$$Score(t,d) = \log\left(1 + \frac{\lambda * tf_{t,d} * \sum_i cf_i}{(1-\lambda) * cf_t * \sum_i tf_{i,d}}\right) \qquad (eq.1)$$

The indexing of the metadata subsets were done with the coordinates as a document number and their associated tagsets as the document text. We indexed with filtering using the multilingual stopwords list and without stemming. The following metadata fields (lowercased) from the videos were used for the query: Keywords (tags), Title and Description. A Re-Ranking (RR) process is applied after the IR process. For each topic their first 1000 retrieved coordinates pairs from the IR software are used. From them we selected the subset of coordinates pairs with a weight equal or greater than the two-thirds (66.66%) of the weight of the coordinates pair ranked in first position. Then for each geographical coordinates pair of the subset we sum its associated weight (provided by the IR software) and the weight of their neighbours at a threshold distance (e.g. 100km). Then we select the one with the maximum weighted sum.

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

3) Combination of GeoKB and Information Retrieval with Re-Ranking (GeoFusion). The GeoFusion approach is applied by combining the results of the GeoKB approach and the IR approach with Re-Ranking. From the GeoKB system are selected the predicted coordinates that come only from the Geographical Knowledge heuristics 1, 2 and 3 (avoiding predictions from the population heuristics rules). When the GeoKB rules (applied in priority order: 1, 2, and 3) do not match then the predictions are selected from the IR approach with Re-Ranking.

We used two corpora for training the IR system for MediaEval 2011: 1) the MediaEval 2011 Flickr corpus (3,185,258 photos) and 2) the union of the MediaEval corpus with the CoPhIR[3] image collection [2] (106 million processed images). From the MediaEval corpus we filtered and extracted 1,026,993 coordinates (accuracies between 6 and 16 zoom levels) with their associated tagsets. From CoPhIR we selected the photos with geographical referencing with accuracies between 6 and 16 zoom levels (8,428,065 photos). Then we filtered repeated content and null content (7,601,117 photos). The union of the extracted data from CoPhIR and MediEval gives a total of 2,488,965 unique coordinates with associated tagsets.

3. EXPERIMENTS AND RESULTS

We designed a set of four experiments (see Table 1) for the MediaEval 2011 Placing Task test set of 5347 Flickr videos. The experiment *TALP1* used the IR approach with Re-Ranking up to 100 km and the MediaEval 2011 photos corpu as a training data. The experiment *TALP2* used the GeoKB approach. The experiment *TALP3* used the GeoFusion approach with the MediaEval training corpora. The experiment *TALP5* used the GeoFusion approach with the MediaEval and the CoPhIR corpora of photos for training. The results are shown in Figure 1 and Table 2.

Table 1: MediaEval 2011 Placing task Experiments.

run	Approach	Training Corpus
TALP1	IR Re-Rank (100km)	MediaEval
TALP2	GeoKB	-
TALP3	GeoKB + IR Re-Rank (100km)	MediaEval
TALP5	GeoKB + IR Re-Rank (100km)	MediaEval+ CoPhIR

Table 2: Results at the Placing Task (5347 videos)

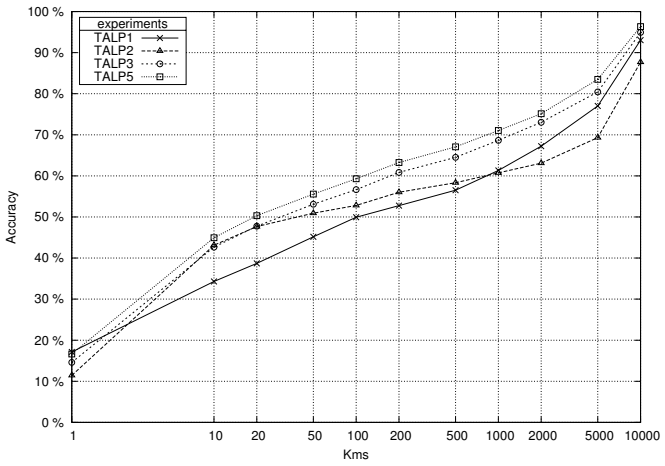
Margin	TALP1	TALP2	TALP3	TALP5
1km	**916**	611	781	890
10km	1834	2306	2281	**2403**
20km	2070	2549	2553	**2690**
50km	2415	2723	2840	**2971**
100km	2670	2823	3029	**3171**
200km	2821	2995	3253	**3382**
500km	3022	3119	3450	**3587**
1000km	3278	3247	3670	**3799**
2000km	3594	3374	3906	**4017**
5000km	4119	3706	4301	**4465**
10000km	4975	4688	5076	**5151**

4. CONCLUSIONS

We used three approaches at MediaEval 2011 Placing Task. The GeoFusion approach achieved the best results in the experiments clearly outperforming the other approaches. This approach achieves the best results because combines high precision rules based on Toponym Disambiguation heuristics and predictions that come from a data driven IR Re-Ranking approach. The GeoKB rules used in the GeoFusion approach achieved 80.18% of accuracy (1789 of 2231 videos) predicting up to 100km. As a further work we plan to improve the accuracy of the GeoKB rules.

[3]**CoPhIR**. http://cophir.isti.cnr.it

Figure 1: Accuracy against margin of error in kms



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] Adam Rae and Vannesa Murdock and Pavel Serdyukov and Pascal Kelm. Working Notes for the Placing Task at MediaEval 2011. In *Working Notes of the MediaEval 2011 Workshop*, Pisa, Italy, September 2011.

[2] Paolo Bolettieri, Andrea Esuli, Fabrizio Falchi, Claudio Lucchese, Raffaele Perego, Tommaso Piccioli, and Fausto Rabitti. CoPhIR: a Test Collection for Content-Based Image Retrieval. *CoRR*, abs/0905.4627v2, 2009.

[3] Daniel Ferrés and Horacio Rodríguez. TALP at MediaEval 2010 Placing Task: Geographical Focus Detection of Flickr Textual Annotations. In *Working Notes of the MediaEval 2010 Workshop*, Pisa, Italy, October 2010.

[4] Daniel Ferrés and Horacio Rodríguez. Georeferencing Textual Annotations and Tagsets with Geographical Knowledge and Language Models. In *Actas de la SEPLN 2011*, Huelva, Spain, September 2011.

[5] Djoerd Hiemstra. *Using Language Models for Information Retrieval*. PhD thesis, Enschede, Netherlands, January 2001.

[6] 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.