Automatic Zoom Level Prediction for Informal Location Descriptions

Igor Tytyk and Timothy Baldwin
Department of Computing and Information Systems
The University of Melbourne

ihor.tytyk@gmail.com   tb@ldwin.net

ABSTRACT

This paper is concerned with the automatic prediction of the zoom level at which to present a map with results for an informal location description. We propose the use of identifiability (relative uniqueness) and zoom level of each component geospatial expression (GE) in the location description, as a means of predicting the appropriate zoom level for the overall description. We apply a simple classification approach to zoom level prediction, and compare results using gold-standard and automatically-inferred GE information. We find the approach to have strong promise, including relative to the zoom level used in results from Google Maps for our location descriptions dataset.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing

General Terms

Algorithms, Experimentation

Keywords

geospatial expression, zoom level, map, natural language processing

1. INTRODUCTION

While there is an extensive literature on geographic information retrieval (i.e. the prediction/ranking of relevant documents for a geospatial query [2, 1]), there is a relative sparsity of literature on the determination of the appropriate map zoom level to use in presenting map-based results. For example, if a user queries for:

(1)   corner of como parade east and parkers road, in the library building

it would seem appropriate to zoom the resulting digital map right down to the specific building (and immediate surrounds), whereas if they query for:

(2)   cafes in Carlton

the appropriate result is probably a map of the suburb of Carlton, with a point-based indication of all known cafes in that area. This is the primary task we target in this paper: the determination of the appropriate zoom level for a given spatial description. As part of this, we explore the interaction between properties of each geospatial expression in the description, and appropriate zoom level of the result set.

2. RELATED WORK

Map-based querying is the denizen of Geographic Information Systems (GIS: [4]) and assumes that location information is presented in the canonical form of a geoparsable address, which can be matched against a structured index of spatial items, based on which results can be rendered on a map or similar. Geographic information retrieval (GIR: [2, 1]) relaxes the assumption on the structure of both the query and the index, and attempts to map potentially unstructured queries with some spatial dimension against document collections, potentially rendering the results on a map via the disambiguated toponyms contained in the documents. This research is pitched between the two: we attempt to disambiguate unstructured text queries via a Google Maps interface, as part of which, we attempt to determine the appropriate zoom level for the resulting map.

This paper builds directly off the work of Tytyk and Baldwin [6], who manually annotated a selection of 1638 user-generated spatial descriptions sourced from the Tell Us Where spatial game [7]. This included: (1) manual identification of geospatial expressions (GEs); and (2) annotation of the "identifiability" and "zoom level". Identifiability was based on three values, as defined relative to the state of Victoria in Australia: identifiable non-ambiguous (uniquely identifying, e.g. *Flemington Road*), identifiable ambiguous (one of a small, bounded set of entities, e.g. *Canning St*, of which there are four in Victoria), and non-identifiable (one of a large or unbounded set, e.g. *the park*). The zoom level of the GE, as judged on a scale of 1–7, is based on the classification of Richter et al. [5], as detailed in Table 1. In addition to the GEs contained within a given location description, the

Zoom level	Description
(1) Furniture	Location within a room (e.g. *by my computer*)
(2) Room	Location within a building (*third floor*), or medium-sized vehicle
(3) Building	Location of a building, street no. or building name
(4) Street	Institution, public space or street level, larger than building and/or vaguer boundaries than building.
(5) District	Suburb, rural district or locality, or post code area
(6) City	Town or city level, and metropolitan areas
(7) Country	Everything beyond city level

Table 1: Zoom level class set

overall location description was tagged with a zoom level, determined based on the lowest zoom level that captures the set of point-based geospatial interpretations for the overall description. For example, *at home* would be tagged as zoom level 7, as the entire state of Victoria must be rendered in order to present the region which bounds the full result set for that description, whereas Example (1) would be zoom level 3. The primary focus of this paper is the interaction between these classifications and the prediction of zoom level for a given spatial description.

3. ZOOM LEVEL PREDICTION USING GOLD STANDARD DATA

In our first set of experiments, we investigate the utility of the identifiability and zoom level of the component GEs in a given location description in determining the zoom level for the description. In this, we use only the identifiability and zoom level data (and no geocode, lexical or other features) and attempt to predict the zoom level in a standard supervised classification setup, using a linear-kernel support vector machine (SVM).[1]

The set of Boolean features we use to represent a given location description relative to its component GEs are:

- **zoom**: the zoom level of every GE

- **unambiguous**: the zoom level of each unambiguous GE

- **ambiguous**: the zoom level of every ambiguous GE

- **non-identifiable**: the zoom level of every non-identifiable GE

- **double**: zoom levels that occur two or more times in GEs within the description

- **minimum zoom**: the smallest zoom level value for all GEs

We run our experiments based on 10-fold stratified cross-validation over the 1638 location descriptions, and evaluate in terms of: (a) classification accuracy (the proportion of

---

[1]We also experimented with regression models, but achieved the best results with a classifier.

Feature set	RMSE	Accuracy
Minimum zoom	1.915	0.593
Most frequent zoom	1.779	0.322
All	0.932	0.838
−zoom	0.956	0.825
−unambiguous	1.067	0.821
−ambiguous	0.936	0.833
−non-identifiable	0.950	0.832
−double	0.934	0.840
−min. zoom	0.911	0.838
−(min. zoom, double, ambiguous)	**0.901**	**0.841**

Table 2: Predicting the zoom level of location descriptions with gold-standard GE annotations

instances where the model generates a correct zoom level prediction); and (b) root mean squared error (RMSE), between the predicted zoom and actual levels (as integers). As baselines for the task, we use:

1. **minimum zoom**: assign the zoom level of the GE in the description with the lowest zoom level

2. **most frequent zoom**: assign the most frequent class (zoom level) in the training data to all test instances; for all partitions, the most frequent zoom was level 4 (Street level).

The results are presented in Table 2, including sub-results where we ablate a given feature from the feature set. The classifiers (including under feature ablation) outperform both baselines in all instances, with the best-performing feature set being the combination of zoom, unambiguous, and non-identifiable, at RMSE = 0.901 and accuracy = 0.841. That is, knowledge of the zoom levels of all component GEs, in addition to the zoom levels of the subset of unambiguous and non-identifiable GEs, leads to the highest-quality predictions.

4. ZOOM LEVEL PREDICTION USING AUTOMATICALLY PREDICTED DATA

The results in Section 3 are informative in shedding light on the relative impact of the zoom level and identifiability of each GE contained within it. However, they are based on the unrealistic assumption of perfect knowledge of these properties of each component GE. In this section, we consider the more realistic setting of needing to predict the zoom level and identifiability of each GE as part of the prediction of the zoom level of the overall location description. Note that we still assume knowledge of the GEs in the description, although this is a more realistic assumption, with the best-published results for chunk-level GE identification over TELLUSWHERE being around an F-score of 80% [3].

4.1 Predicting the Zoom Level and Identifiability of Component GEs

First, we need to predict the zoom level and identifiability of each GE in a given location description. To do this, we part-of-speech tag and full-text chunk parse the location

| Feature set | Zoom level | | Identifiability |
	RMSE	Accuracy	Accuracy
All	**1.025**	**0.564**	0.597
−preposition	1.028	0.556	**0.609**

Table 3: Results for GE-level zoom level and identifiability prediction

Feature set	RMSE	Accuracy
Minimum zoom	1.762	0.428
Most frequent zoom	1.779	0.322
SVM (zoom, unambiguous, non-id)	**1.760**	**0.457**

Table 4: Predicting the zoom level of location descriptions with automatically-predicted GE zoom level and identifiability

description using OpenNLP. Next, we extract a feature representation of each GE as follows, based on the composition and context of use of the GE:

- **following GEs**: the number of GEs within the description that follow the current expression

- **non-noun count**: the number of words within the GE that are not nouns

- **noun count**: the number of words within the GE that are nouns

- **capitalised**: are all of the words within the GE capitalised?

- **word count**: number of words within the GE

- **preceding preposition**: is there a preposition immediately preceding the GE?

- **head cap**: is the head noun capitalised?

- **determiner**: is the GE specified by an article (*a* or *the*), and if so, what is the article?

- **preposition**: does one of the following ten prepositions precede the GE: *on, off, behind, in front of, inside, opposite, outside, near, at, in*

Based on this feature representation, we once again use a linear-kernel SVM to classify each GE for identifiability and zoom level. The results over the 3061 GEs in the dataset, based on 10-fold cross-validation with stratification at the description level, are presented in Table 3 for each of zoom level and identifiability prediction.

In the case of zoom level, the best results were achieved with all the features, but for identifiability, the omission of the preposition feature boosted results slightly. Note that we do not present the RMSE for identifiability prediction, as there is no natural numeric decomposition of the class set. While the accuracy of our best models are somewhat lacklustre, it is important to bear in mind that a given location description can contain multiple GEs, meaning that there is the possibility of smoothing across the noisy predictions of multiple component GEs.

4.2 Predicting the Zoom Level of the Location Description

Having obtained predicted values of the zoom level and identifiability for all 3061 GEs, we now return to the task of predicting the zoom level of the containing location descriptions, using the same classifier setup as in Section 3,
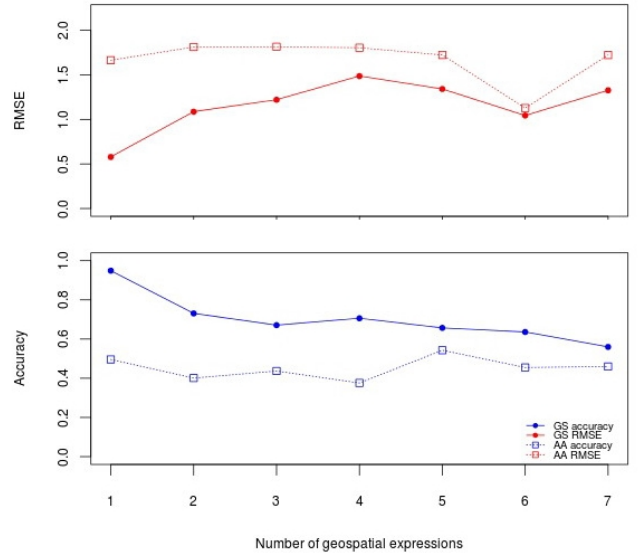

Figure 1: RMSE and accuracy split by the number of GEs per description

focusing on what we found to be the most predictive feature set, namely the combination of the zoom, unambiguous and non-identifiable features. Our baselines are as before, except that "minimum zoom" is now based on the predicted zoom levels of the GEs; most frequent zoom is unchanged (as it does not rely on GE features). The results are presented in Table 4.

When run over the automatic GE data, our SVM classifier is only slightly better than the two baselines, with the accuracy roughly halving and the RMSE roughly doubling, due to the propagation of errors from the GE classifiers.

To better understand these results, we broke down the results for the description-level classifiers based on both gold-standard and automatic GE data, according to the number of GEs contained in the location description. The results are presented in Figure 1 for RMSE and accuracy, where "GS" and "AA" stand for gold-standard and automatic, respectively.

In both cases, the performance tends to fall (accuracy drops; RMSE rises) as the number of GEs increases, due to the difficulty of composing the semantics of the component GEs to arrive at a holistic interpretation of the location description. More surprisingly, the lines in the two graphs

Method	Accuracy
Gold-standard GE	0.841
Automatic GE	0.457
Google Maps	0.528

Table 5: Comparison of our method for zoom level prediction (using gold-standard or automatic GE properties) with Google Maps

appear to be converging as the number of GEs increases. That is, the more GEs the description contains, the smaller the negative influence of noise in the GE zoom level and identifiability predictions.

4.3 Comparison with Google Maps

For a better, more objective, evaluation of the obtained results, we compare the zoom level predictions for the descriptions with the results when the location description is issued as a query to the Australian instance of Google Maps.[2] There is one large caveat associated with this comparison: we are predicting only the zoom level, while Google Maps needs to both geocode the location description and determine the appropriate zoom level to present the results to the user. To ameliorate this difference, we: (a) ignore all location descriptions which do not return results from Google Maps; and (b) restrict the query by appending *Victoria, Australia* to each location description. Predictably, the mapping service did not have any problems with more formal location descriptions that resemble addresses (e.g. *At home, 43 Campbell St, Kew*. It also returned results for shorter geometric descriptions such as *corner of arden st and drybrugh st*. Where it struggled was with longer descriptions that contained more complex information (e.g. *In my apartment in the block of flats on the corner of Park Street and Kings Way. The apartments take up the whole block of Park Street, Eastern Road, Napier Street and Kings Way ...*) and descriptions which contained vernacular or abbreviated placenames (e.g. *p.i.* as an abbreviation for *Phillip Island*, *richy* as the vernacular form of *Richmond*).

Table 5 compares our method, based either on gold-standard or automatic GE properties, with Google Maps. We present only accuracy numbers, as there is no natural mapping from the Google Maps zoom levels onto our formal zoom level class set. Instead, we used a somewhat informal interpretation of our zoom level class set for Google Maps. Additionally, there were instances where the query returned no results whatsoever in Google Maps, which we simply count as incorrect.

The accuracy for Google Maps is substantially lower than that for our method when we assume full gold-standard GE information, but moderately higher than that for our method using automatic GE information. This points to the potential utility of GE-based zoom level and identifiability features in determining the appropriate zoom level for map-based results. Given the relatively modest GE-based results achieved by our method, a natural avenue for future research is to improve the base classifier results, or alter-

natively explore the possibility of joint GE and description classification.

5. SUMMARY

We have presented a framework for predicting the zoom level of informal place descriptions, based on analysis of the zoom level and identifiability of its component geospatial expressions. First, we experimented with gold-standard GE information and found that we were able to predict the zoom level of the description with remarkable success. Next, we looked to automatically predict the zoom level and identifiability of each GE, and then repeated the original experiment over description levels, and found that the results drop appreciably, although they were still above baseline. Finally, we compared our method with Google Maps, by issuing the location descriptions as queries (with state-level location specification, to disambiguate), and found that our method was more accurate than Google Maps when it was provided with gold-standard GE information, but slightly less accurate when it only had access to automatic GE information. Overall, our experiments point to the potential utility of zoom level and identifiability information of component GEs in the determination of the appropriate zoom level for map-based results when presented with informal location descriptions.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] C. B. Jones and R. S. Purves. Geographical information retrieval. *International Journal of Geographical Information Science*, 22, 2009.

[2] R. R. Larson. Geographic information retrieval and spatial browsing. In L. Smith and M. Gluck, editors, *GIS and Libraries: Patrons, Maps and Spatial Information*, pages 81–124. University of Illinois, 1996.

[3] F. Liu. Automatic identification of locative expressions from informal text. Master's thesis, The University of Melbourne, Melbourne, Australia, 2013.

[4] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind. *Geographic Information Systems and Science*. John Wiley & Sons, Hoboken, USA, second edition, 2005.

[5] D. Richter, M. Vasardani, L. Stirling, K.-F. Richter, and S. Winter. Zooming in–zooming out hierarchies in place descriptions. In D. Cartwright, G. Gartner, and M. Peterson, editors, *Progress in Location-Based Services*, pages 339–355. Springer, 2013.

[6] I. Tytyk and T. Baldwin. Component-wise annotation and analysis of informal placename descriptions. In *Proceedings of the International Workshop on Place-Related Knowledge Acquisition Research (P-KAR 2012)*, Kloster Seeon, Germany, 2012.

[7] S. Winter, K.-F. Richter, T. Baldwin, L. Cavedon, L. Stirling, A. Kealy, M. Duckham, and A. Rajabifard. Location-based mobile games for spatial knowledge acquisition. In *Location-Based Mobile Games for Spatial Knowledge Acquisition*, Belfast, USA, 2011.

---

[2]`http://maps.google.com.au`