

Did it change? Learning to Detect Point-of-Interest Changes for Proactive Map Updates

Björn Bebensee (2019–21343)

Topics in Artificial Intelligence

September 10, 2019

In our modern society maps are of increasing importance as many people utilize the maps provided by their smartphones to navigate in every day life. However, these maps are still largely updated manually which is a very time-consuming process. Revaud, Heo, Rezende, You, Jeong [1] propose a method to automatically detect changes in points of interests (POIs) from geo-tagged street-view images which can allow maps to be updated much more proactively as a step towards fully-automated map maintenance. The task at hand is challenging as images of the same POI may vary widely due to different viewpoints, noise and lighting and even if the POI does not change it is possible that its appearance changes.

As part of their work they release a dataset of geo-localized images of POIs captured in indoor shopping malls. The dataset was collected in two in-door shopping malls. It includes many different viewpoints for the 578 POIs and each geo-localized image additionally contains 6-DoF camera pose information. There are two main tasks in POI change detection. First, one has to determine a function that can predict the similarity between two images $d \in D^t$ and $d' \in D^{t'}$ where D and D' are geo-localized datasets and where the time $t' > t$. Once these POIs are identified, one wants to determine a POI change scoring function which subsequently scores images of the same POI to predict whether it has changed (for any reason).

To learn the similarity function Revaud et al. learn an embedding function $f(I) = x \in \mathcal{X} \subset \mathbb{R}^N$ and define the ground truth such that two images are only similar when they show the same POI (at least partially). The POI scoring function is implemented by max-pooling pairwise similarity scores so that a POI change is predicted when no pair of images of the same location is similar. In order to predict POI changes one only needs to learn the embedding function now.

This approach is based on metric learning where a distance metric is learned. Ideally this metric should be invariant to noise, lighting and viewpoint changes in the image. Revaud et al. evaluate their approach on scoring similarities between image pairs as well as end-to-end on the geographical level. They find that the metric learning approaches work much better than ImageNet and SIFT and conclude that triplet loss works best in this problem setting.

However, just as the authors state in their conclusion, this problem needs further research as a lot of important points are not covered. Not only can an image contain more than one POI but the proposed approach focuses on in-door shopping malls alone whereas most of the interesting and important map changes in a real-world setting happen outside in streets.

References

- [1] Revaud, Jerome, Minhyeok Heo, Rafael S. Rezende, Chanmi You, and Seong-Gyun Jeong. "Did It Change? Learning to Detect Point-Of-Interest Changes for Proactive Map Updates." *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4086-4095. 2019.