Folkioneer: Efficient Browsing of Community Geotagged Images on a Worldwide Scale

Hatem Mousselly-Sergieh[1,2], Daniel Watzinger[1], Bastian Huber[1], Mario Döller[3], Elöd Egyed-Zsigmond[2], and Harald Kosch[1]

[1] Universität Passau, Innstr. 43, 94032 Passau, Germany
[2] Université Lyon, 20 Av. Albert Einstein, 69621 Villeurbanne, France
[3] FH Kufstein, Andreas Hofer-str. 7, 6330 Kufstein, Austria
firstname.lastname@uni-passau.de, firstname.lastname@insa-lyon.fr, mario.doeller@fh-kufstein.ac.at, {daniel.watzinger,huber.baste}@gmail.com

Abstract. In this paper, we introduce Folkioneer, a novel approach for browsing and exploring community-contributed geotagged images. Initially, images are clustered based on the embedded geographical information by applying an enhanced version of the CURE algorithm, and characteristic geodesic shapes are derived using Delaunay triangulation. Next, images of each geographical cluster are analyzed and grouped according to visual similarity using SURF and restricted homography estimation. At the same time, LDA is used to extract representative topics from the provided tags. Finally, the extracted information is visualized in an intuitive and user-friendly manner with the help of an interactive map.

1 Introduction

In the era of Web 2 users become able to contribute content to the web by themselves. Online image portals like Flickr[1] is witnessing an explosion in the number of images uploaded everyday. Browsing such huge amount of data is challenging. Therefore, efficient solutions are needed. Currently, an increasing number of the user uploaded images are geotagged beside being annotated by the users. In this paper, we present Folkioneer, a novel system for efficient browsing of community generated images in an intuitive manner on a worldwide scale. Folkioneer exploits geodesic information (longitude and latitude), visual similarity, and user tags to structure visual folksonomies.

The remainder of the paper is organized as follows. In the next section, individual steps of the processing pipeline are discussed briefly. Section 3 provides a description of the system through screenshots. We conclude and discuss future work in Section 4.

2 Folkioneer Architecture

Figure 1 depicts an abstract overview of the entire processing pipeline of Folkioneer. First, a world-scale sample is acquired from a visual folksonomy based on the

[1] http://www.flickr.com

C. Gurrin et al. (Eds.): MMM 2014, Part II, LNCS 8326, pp. 361–364, 2014.

small-world phenomenon. After that, the geodesic information of the crawled images is used to extract geographical clusters by applying Clustering Using Representatives (CURE) [1]. In a next step, characteristic polygons describing the shape and the physical extent of each geographical cluster are built by iteratively eroding Delaunay triangulation of the projected image locations [2]. The eroded triangulation is subsequently reused to approximate the area of geodesic polygons in order to prune the cluster hierarchy. Next, the geographical clusters are refined by analyzing and clustering respective images based on visual similarity. Furthermore, Folkioneer also exploits the provided user tags and mines representative tags for each geographical cluster and at different geographical granularity levels. Finally, the acquired information is presented by means of an interactive map.

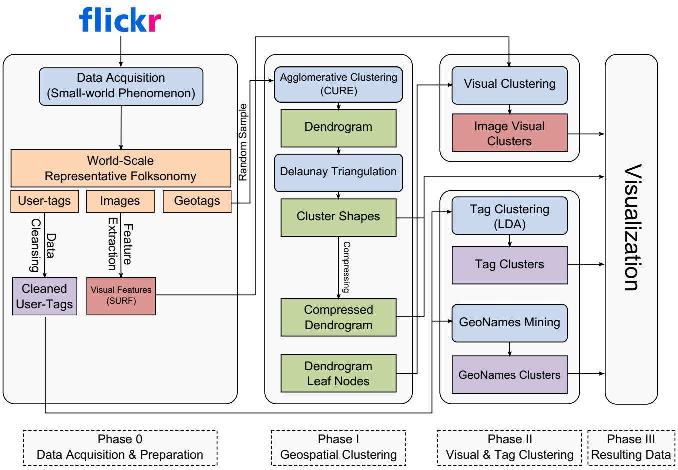


Fig. 1. The processing pipeline followed in Folkioneer

Data Acquisition. In a first step image metadata (identifiers, tags, titles and geodesic information) were obtained from Flickr. To ensure the representativeness of the collected data we applied a crawling strategy that exploits the small-world phenomenon. This was achieved by a breadth-first search on Flickr's friendship graph. The final dataset includes data of 14.1 million images which were uploaded by 211 thousands users in the period from 14/5/2000 to 01/04/2012. The images are annotated with 154 million tags with 6.3 million unique tags.

Geospatial Clustering. In this phase, the crawled images are clustered according to their geographical coordinates. To efficiently handle large amounts of geographical coordinates, we applied a clustering solution based on CURE algorithm for agglomerative clustering [1]. Since spherical coordinates are not suitable for spatial clustering, the geodesic coordinates are projected into the 3-dimensional Cartesian reference system ECEF (Earth-Centered Earth-Fixed). To facilitate nearest-neighbor search CURE is implemented by indexing the 3-dimensional data points using K-D Tree data structure. Furthermore, we used

Fibonacci min-heap to preserve and effectively reuse dissimilarity information across iterations of the clustering algorithm. The output of this a phase is a dendrogram of geographical coordinates. The geospacial clusters are determined by pruning the dendrogram according to a minimum area threshold. Approximation of the geodesic area of geographical clusters is achieved by using Delaunay triangulation [2].

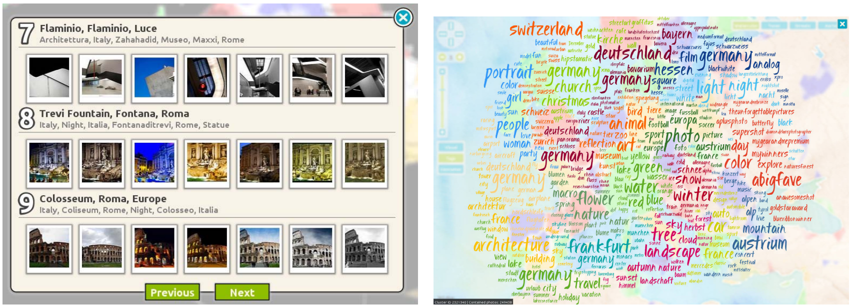
Visual Image Clustering. The geographical clusters are refined by grouping the images of each cluster according to the visual similarity. The main objective is to cluster images representing popular landmarks and events and to determine representative images for each cluster. Visually similar images are identified using the Speeded Up Robust Features (SURF) [3]. SURF is an efficient algorithm for the extraction and matching of local image features. We improved SURF by extending the algorithm with a homography estimation step using LO-RANSAC [4]. Since images depicting the same scene are usually taken from a restricted set of canonical views, the instantiation of new hypotheses for RANSAC is accelerated by restricting the homographies to affine transformations. Homography estimation results in a final number of inliers that obey the generated model. Accordingly, any monotonically decreasing function provides a possible dissimilarity metric. The final visual clusters are then determined by building a complete graph from the pairwise image dissimilarities. To identify the final visual clusters, the graph is fed into a clustering algorithm based on Maximum Standard Deviation Reduction approach (MSDR) [5].

Topic Extraction. After the cleansing step, respective tags of each geographical cluster are analyzed in order to extract representative terms. For this purpose, we used the method of LDA (Latent Dirchlet Allocation) [6] to extract representative topics from larger clusters. For this purpose, tags corresponding to a particular image are regarded as a single document. Subsequently, the document corpus consisting of the respective tags of a geographical cluster is analyzed using LDA to generate possible topics.

3 Demonstration Details

Folkioneer's visualization is based on an interactive map with well-known features like changing the zoom level and panning. In order to visualize contents of the map, the Mercator projection is used to provide a pleasant and familiar viewing experience to the users.

The system is able to manually adjust the granularity of displayed geodesic clusters to meet user preferences by using a simple slider control. Retrieving additional information of particular clusters is initiated by clicking on the respective outlines. Visual clusters are displayed and ordered according to their importance. Additionally, a collection of representative images are displayed with suitable captions for each cluster (Figure 2.a). Furthermore, it is possible to browse through the complete set of images of a given visual clusters as well as



(a) (b)

Fig. 2. Visual clusters and tag clouds

to view the images in full-size. Finally, Folkioneer supports the visualization of representative tag sets in form of a cloud of tag clouds (Figure 2.b).

4 Conclusion

In this paper, we presented Folkioneer, a system that originally combines different kinds of state-of-the-art algorithms to provide a novel and intuitive browsing experience of community images. In future work, we will consider using the newly gathered information to assist applications that exploit folksonomies such as automatic image annotation.

References

1. Guha, S., Rastogi, R., Shim, K.: Cure: an efficient clustering algorithm for large databases. In: Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, SIGMOD 1998, pp. 73–84. ACM, New York (1998)
2. Duckham, M., Kulik, L., Worboys, M., Galton, A.: Efficient generation of simple polygons for characterizing the shape of a set of points in the plane. Pattern Recogn. 41(10), 3224–3236 (2008)
3. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). Comput. Vis. Image Underst. 110(3), 346–359 (2008)
4. Chum, O., Matas, J., Kittler, J.: Locally optimized RANSAC. In: Michaelis, B., Krell, G. (eds.) DAGM 2003. LNCS, vol. 2781, pp. 236–243. Springer, Heidelberg (2003)
5. Grygorash, O., Zhou, Y., Jorgensen, Z.: Minimum spanning tree based clustering algorithms. In: Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2006, pp. 73–81. IEEE Computer Society, Washington, DC (2006)
6. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. The Journal of Machine Learning Research 3, 993–1022 (2003)