Positional Accuracy of Flickr and Panoramio Images in Europe

Hartwig H. HOCHMAIR and Dennis ZIELSTRA

Abstract

Flickr and Panoramio are fast growing photo sharing services that provide a rich resource of Volunteered Geographic Information through geotagged images. This study analyzes the positional accuracy of 211 image footprints for six cities in Europe by comparing the geotagged position of photos to the position from where the photos were most likely taken, based on the image content. The analysis reveals a better positional accuracy for Panoramio than for Flickr datasets. For Flickr data, some associations were observed between the category of photographed scene and positional accuracy, while this was not the case for Panoramio images. The accuracies identified in this study can be helpful when considering Flickr and Panoramio images as data source in future geo-applications and services.

1 Introduction

The Web 2.0 with its wide range of applications and services provides a continuously growing source of Volunteered Geographic Information (VGI) (GOODCHILD 2007), which complements geographic information that can be extracted from traditional sources, such as Earth imagery or census data. Flickr and Panoramio are prominent examples of Web 2.0 applications which facilitate the sharing of VGI in form of geotagged images. These photo sharing platforms allow a Web user to upload geotagged images to an application server, and to view the location of photo footprints on a world map. Photo sharing applications are not meant to build comprehensive map-like representations of the world, which is the goal of other Web 2.0 initiatives, such as OpenStreetMap (OSM)[1] or Wikimapia[2]. The images rather depict one's favorite locations that are, in the eyes of the photographer, worthwhile being photographed. Since Web 2.0 data from photo sharing applications are used in a growing number of geo-applications and spatial analysis tasks (GIRARDIN et al. 2008; SCHLIEDER & MATYAS 2009; HOCHMAIR 2010), there is need to assess the quality of image information, both in the spatial and semantic realm. Flickr and Panoramio images, like most Web 2.0 data, are provided by volunteer Web users with little formal training in measuring and mapping techniques, as opposed to how the collection and processing of geo-data is handled by authoritative agencies. Quality assessment for shared photos will help in deciding whether this comprehensive data source is accurate enough for consideration in future geo-applications, such as guided virtual tours that make use of shared images.

As a contribution into this direction, this paper assesses the positional accuracy of Flickr and Panoramio image footprints. More specifically, a Geographic Information System is used to measure the distance between the published footprint position (which is determined through geotagged latitude and longitude coordinates) and the position from where the

[1] http://www.openstreetmap.org
[2] http://wikimapia.org

image was presumably taken, based on the image content. This measurement was done for a total of 211 Flickr and Panoramio images distributed throughout six urban areas in Germany, Austria, and Italy. We assess the potential effect of data source (Flickr vs. Panoramio) and image content type (e.g., street building, natural panorama, or statue), on error distance. Further we quantify two types of blunders that can occur when a Flickr or Panoramio user manually places a photo on a map during the image upload process.

2 Previous Work

A growing number of research papers that analyze metadata and the spatial distribution of photos posted on publicly accessible photo sharing services shows the increasing interest of the research community in this source of VGI. GIRARDIN et al. (2008) used spatio-temporal information on Flickr images to analyze different aspects of mobility and travel in Rome. The authors reconstructed people's movement through the city, which was found to differ between Italian and foreign photographers. Time stamps were also used to identify temporal patterns of visiting activities for selected locations, such as the Colosseum, which indicated higher activities on weekends. SCHLIEDER & MATYAS (2009) analyzed the popularity of places based on the spatial distribution of Flickr images for four European cities. The study revealed that the popularity follows a power law with very few highly popular locations and a long tail of places in a city that are visited only occasionally. A recent study that analyzed user uploaded routes on trail sharing Web sites found that the density of Panoramio pictures is higher along scenic routes than along shortest routes (HOCHMAIR 2010). This indicates that photo footprints could be used as a complementary data source for trip planning in addition to typical street attributes.

The International Organization for Standards (ISO) codes 19113 (quality principles) describe several aspects of quality for geographical information, including completeness, logical consistency, positional accuracy, temporal accuracy, and thematic accuracy (International Organization for Standardization 2009), where positional accuracy is significant in the evaluation of fitness-for-use of data not created by professionals and without data-collection standards (HAKLAY 2010). However, only few papers tackle the problem of positional accuracy of VGI. HOLLENSTEIN & PURVES (2010) analyzed for 9775 Flickr images around Hyde Park in London whether tagged geographic terms match the geotagged position, revealing an accuracy of about 86%. However, no point-by-point comparison, as presented in the current study, was made. The authors further classified errors of misplacement, including object mismatch during the georeferencing process (e.g., an image placed in the wrong park), bulk upload, and the tendency to place images towards the town center. The authors conclude that the precision and accuracy of user generated data appear to be high enough to describe city neighborhoods, and that people appear to be better at describing what they have seen than they are at assigning an accurate georeference. In the realm of road networks, HAKLAY (2010) shows that OSM road data are on average within 6 meters of the position recorded by Ordnance Survey. Several other studies compare network completeness between OSM and proprietary datasets (NEIS et al. 2011; ZIELSTRA & HOCHMAIR 2011), but less with a focus on positional accuracy.

3 Data and Methods

3.1 Data Sources

Flickr and Panoramio, two of the most prominent photo sharing Web sites, were used as primary data sources for this study. Flickr hosted over 6 billion images in August 2011 (KREMERSKOTHEN 2011) with about 3-4% being geotagged (FLICKR 2009). For Panoramio, where all images are geotagged, current numbers are not published, but back in January 2008 6.8 million geotagged images were submitted (TAYLOR 2008). Both data sources provide similar methods for users to upload images and add location information. Most newer photo cameras use a built-in GPS receiver which obtains the photographer's position at the time of taking the photograph. This location information, together with other image and camera metadata, is commonly stored in the exchangeable image file (EXIF) format, and can be uploaded to the Flickr and Panoramio portals together with the image. Position information can also be added manually, e.g., when obtained from an external GPS device. Both portals also allow to manually drag an image to the corresponding position on a background map. Each of these methods can cause position errors of picture footprints due to different causes. While GPS handheld receivers that have visibility to enough satellites provide an accuracy level in the magnitude of a few meters, canopy and nearby buildings will deteriorate positioning accuracy. Also, a certain initial time is needed to get a lock on by the GPS receiver after the receiver starts. Manually placing an image over a map requires the photographer to be somewhat familiar with the area, which may not be the case when visiting areas, e.g. as a tourist. Also, confusing similar locations, e.g. bridges, statues, or parks, can lead to misplacement of the image, even by a more advanced user.

Flickr and Panoramio provide a public API that enables anyone to download photos that were declared public by the person uploading the photos. In Flickr one can assign uploaded photos to one or several thematic groups to more clearly specify the content of the image, besides providing tag information. Flickr is not restricted to outdoor scenes only, but contains also a large percentage of images showing indoor or close up images of objects or persons that do not well reflect the outdoor environment. Since for our analysis we are interested in images of outdoor scenes, we applied a query filter on the API that retrieves Flickr images only from selected groups, such as “Architecture” or “Travel Photography”. In Panoramio all photos show outdoor scenes and therefore no filtering was necessary.

3.2 Measurement Procedure

The positional accuracy assessment is based on distance measurements. More specifically, all image footprints obtained through the APIs were added to ArcMap as point features based on their tagged, i.e., published, geographic coordinates. Each point feature also contained an attribute with the URL linking to the photo on the Flickr or Panoramio Web site. This way the image content could be viewed in a browser. Based on the image content the analyst estimated the photographer's approximate position (which was estimated to be accurate within a few meters). For each image, the distance between the imported footprint feature point and the assumed photographer's position was measured in ArcMap, where ellipsoidal distances on the reference ellipsoid of the map projection were used to avoid distortion of distances caused by the map projection.

The zoom level used while taking the image may affect the estimated photographer’s position within the analysis. For example, a high zoom level makes the uploaded image appear to be taken from closer proximity to the scene than this was actually the case. The EXIF file, if available, provides some indirect information relating to the used zoom level, including the focal length. A larger focal length results in a narrower angle of view and thus a higher magnification. However, the focal length value varies by photo camera models and is different for interchangeable lenses, which makes it difficult to derive a zoom level from a given focal length value in the image metadata. When analyzing the geometry of scene position, geotagged and assumed photographer’s position it was found that these three locations were rarely located along a straight line. This shows that the role of varying zoom levels on point accuracy in the analyzed examples is negligible and that other reasons lead to incorrectly placed images.

To be able to determine the photographer’s position based on the scene shown in an image, the analyst needed to have good local knowledge of the area. This measurement task was performed by the two authors and a co-worker, which lead to the selection of six regions in three European countries. The regions and corresponding numbers of analyzed images are listed in Table 1. As an example, Fig. 1 shows the locations of Flickr and Panoramio footprints in Cologne used in this study that were obtained through the Flickr and Panoramio APIs.

Table 1: Number of Flickr and Panoramio images analyzed for different regions

	City/Town	Flickr	Panoramio
Germany	Bonn	23	25
	Cologne	13	20
Austria	Vienna	22	20
	Zell am See	19	17
Italy	Florence	15	10
	Rome	14	13
TOTAL		**106**	**105**

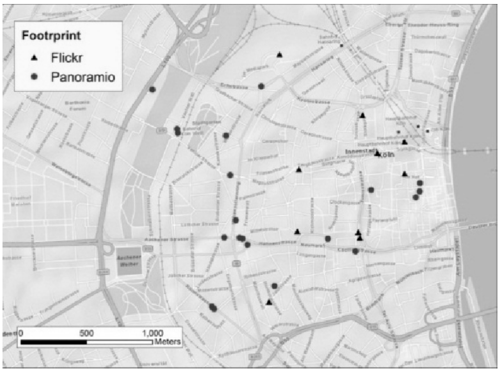

Fig. 1: Flickr and Panoramio photo footprints used from Cologne

Before the analysis, photos provided through bulk-upload were removed. Further, only images located near streets and footpaths (within 60 meters) were analyzed, since these could be most readily used for routing and navigation related geo-applications, as opposed to photos that were taken from open water or from the top of a building.

The chosen approach on distance measurements is based on the assumption that the photographer's position can be correctly mapped and is more accurate than what is provided through the geotagged photo information. As HAKLAY (2010) points out, a basic problem in desk based quality assessment of any spatial dataset is the selection of the reference dataset. In our case, the reference dataset is a background map provided through various Web mapping services in ArcGIS 10.0 used to mark the photographer's most likely position based on image content. Although we assume that because of the analyst's local knowledge the photographer's position can be accurately marked on the background map, the accuracy of the background map itself plays a role, since a shift of the ArcGIS background map relative to the map which was used by a user to place a footprint (or a shift between the geodetic datum used for retrieving GPS coordinates and the one used for the ArcGIS background map without corresponding datum transformation), could affect the distance measurements. In our study, the position of the photographer was determined based on satellite image layers in ArcGIS with different accuracies for different regions. While the accuracy of the different layers is not published, we compared the shift between the two provided background layers in ArcGIS. Some of the tested areas, such as Rome, Florence and Vienna, showed a shift of seven to ten meters between the Bing Maps Aerials in comparison to the World Imagery content, while other areas, such as Bonn and Cologne, did not show this shift. For Zell am See distances needed to be determined in Google Earth since none of the satellite image layers in ArcGIS provided a sufficient resolution to map the photographer's position accurately. While Google Earth may show shifts of up to 20 meters in certain areas, other areas, such as Zell am See, appear to be highly accurate due to background layers that use orthophotography images with sub-meter resolution.

3.3 Image Classification

All images were grouped into nine different categories based on the image content (Table 2) since we hypothesized that the type of image content would affect the positional accuracy of a photo footprint.

Table 2: Scene types and their definitions

Scene type	Description
Street building	building close to streets (e.g. store)
Landmark	building with special character or meaning (e.g. city hall) taken from close vicinity
Distal landmark	landmark visible from far away (e.g. light house)
Street view	intersections, view along street
Natural panorama	panoramic view of natural landscapes (e.g. bay)
Urban panorama	panoramic view of urban landscapes (e.g. city skyline)
Square	squares and plazas (e.g. St. Peter's Square)
Statue	nearby statues, sculptures, monuments and signs
Bridge	bridge for any transportation mode

3.4 Error Classification

Some footprints were misplaced due to two types of blunders. A *footprint mismatch error* occurs if the Flickr or Panoramio user places the image to the location of the scene that was being photographed, instead of dragging the image to the photographer's position. Figure 2 illustrates an example. The right part shows the image content from the Flickr Web site. This picture was clearly taken from the bridge to the south-west of the castle, as indicated by the lower arrow in the left part of the figure. However, the photo footprint is (incorrectly) placed on top of the castle (upper arrow).

Fig. 2: Footprint mismatch error

A *similar object error* occurs if a Flickr or Panoramio user confuses a similar feature, such as a building or a bridge, with the correct location of the photographer. An example is shown in Fig. 3 where two bridges were confused when placing the footprint on the map.

Fig. 3: Similar object error

4 Analysis

4.1 Distance Errors

The first four rows for Flickr and Panoramio data sources in Table 3 present the sample characteristics of measured error distances for the nine scene types. Since error distances are skewed to the right, median distances are reported in the table. The rows labeled "#F" show the number of footprint mismatch errors for each scene type. The relative error frequency (%F) is computed as #F/Nx100, where N is the total number of images analyzed for a given category. The sample size is largest for the first two scene types, i.e., street building and landmark. The table suggests that median error distances are larger for Flickr than for Panoramio footprints for all scene types observed in both datasets, except for distal landmark (which has only sample size 1 for Panoramio) and street view, which is tied. Considering distance values from all scene types, a Mann-Whitney test indicated that the error distance was larger for Flickr (*Mdn*=58.5) than for Panoramio (*Mdn*=0.0), $U = 3582$, $Z = -4.69$, $p < .001$. To control for the potential effect of scene type (and the different sample sizes associated with each type) on error distance, we further compared error distances for each scene type separately. Distances for street buildings were larger for Flickr (*Mdn*=49.0) than for Panoramio (*Mdn*=0.0), $U = 171$, $Z = -2.68$, $p = .007$. Distances for landmarks were also larger for Flickr (*Mdn*=55.0) than for Panoramio (*Mdn*=0.0), $U = 326.5$, $Z = -2.20$, $p = .028$. No significant differences were observed for the other scene types, which can most likely be explained by the smaller sample sizes and in consequence the lower power of the test for these scene types.

Table 3: Descriptive characteristics of error distances (in meters) for Flickr and Panoramio

		Street building	Landmark	Distal landmark	Street view	Natural panorama	Urban panorama	Square	Statue	Bridge	**ALL**
Flickr	N	22	28	6	4	13	1	10	17	5	106
	Mdn	**49**	**55**	**25.5**	**0**	**250**	**1491**	**68**	**35**	**315**	**58.5**
	Min	0	0	0	0	0	1491	0	0	108	0
	Max	1410	2227	4800	252	48000	1491	606	476	2619	48000
	# F	3	4	0	0	0	0	3	8	1	19
	% F	14	14	0	0	0	0	30	47	29	
Pano-ramio	N	27	34	1	9	10	4	9	11	0	105
	Mdn	**0**	**0**	**40**	**0**	**91.5**	**0**	**0**	**0**	**-**	**0**
	Min	0	0	40	0	0	0	0	0	-	0
	Max	430	4870	40	78	2270	0	103	508	-	4870
	# F	1	4	0	0	0	0	1	3	0	9
	% F	4	12	0	0	0	0	11	27	-	

We were hypothesizing that photos of scenes that comprise wider spaces would lead to a larger distance error, since Web users may not be able to identify their position as accurately on a map that shows an open space compared to urban spaces with more distinct map features. However, comparison of error distances between scene types, analyzed separately for Flickr and Panoramio, revealed a relatively homogenous footprint accuracy. The only exception are bridges for Flickr (*Mdn* = 315.0), which showed a significantly higher error distance than type street building (p = .010), landmark (p = .022), distal landmark (p = .082), street view (p = .063), square (p = .019), and statue (p = .015) (no mathematical correction was made for multiple comparison in reported p-values). The same procedure did not reveal any significant distance differences between scene types in the Panoramio dataset.

Panoramio provides possibilities for peer correction of photograph positions and commenting on an image. However, the API does not provide the history of coordinates, in case they were changed, or indicate whether image coordinates have been changed from their original value. Nevertheless, some other information can be obtained when directly viewing an image on the Panoramio and Flickr Web site. That is, the Web site displays the number of views for an image, and one can manually count the number of comments posted for an image. We hypothesized that a larger number of views or comments for an image would be associated with a better positional accuracy, but this could not be statistically confirmed.

The overall results of these analyses indicate that Flickr footprints provide less accurate positional information than Panoramio. The latter data source with an overall median error distance of 0 meters, provides a more reliable and consistent way to access imagery for a given location than Flickr. The difference in accuracy between the two data sources could be explained by the types of images both services host. While Flickr allows users to upload images without specifying their geographic location, Panoramio reminds the user of geotagging during the upload process. Panoramio users could also be motivated by the fact that some Panoramio images are featured in Google Earth. However, to be considered for Google Earth and to participate in a monthly contest on Panoramio itself, the images need to be accurately geotagged and of good quality. Thus, Panoramio seems to attract a community of users that is more spatially aware and interested in mapping than this is the case for the Flickr community.

4.2 Error Types

For both Flickr and Panoramio the highest frequency of footprint mismatch errors was found with scene type statue at 47.1% (8/17) for Flickr and 27.3% (3/11) for Panoramio. Some other categories, in particular those with a small sample size, did not show a single footprint mismatch error (see Table 3). Flickr had a total of 19 footprint mismatch errors (out of 106 point features), while Panoramio had only 9 (out of 105). This difference is partially caused by the higher number of images of scene type statue in the Flickr dataset (17 images) compared to Panoramio (11 images). For statistical testing of the association between scene type and frequency of footprint mismatch error we only compared scene types of sample size five or larger to obtain a reasonable power of the test, which led for Flickr data to 7 scene types being compared. A two-tailed Fisher's exact test was used, with p-values reported in Table 4. Parentheses indicate a smaller error rate for the scene type listed in the row header to the left compared to the scene type in the column header, while absence of parentheses indicates the reverse. If no mismatch error was found in both scene types, the p-value is printed in italics. Results indicate a significantly higher rate of footprint mismatch errors for scene type statue, compared to four other types (at different levels of significance). No differences at a 5% level of significance were observed between other categories. Interestingly, this observed footprint mismatch problem for scene type statue does not seem to affect footprint accuracy for Flickr statue images, as was shown in connection with median distance values (Table 3) before. The reason may be that statues, sculptures, and signs, are usually photographed from within close distance, so that the positional error caused by a footprint mismatch error is small. For Panoramio, no significant association between scene type and relative error frequency was observed.

Table 4: p-values for Fisher's exact test analyzing footprint mismatch errors for Flickr

	Street building	Landmark	Distal landmark	Natural panorama	Square	Statue	Bridge
Street building	-	(1.000)	1.000	0.279	(0.346)	(0.033)*	(1.000)
Landmark		-	1.000	1.000	(0.351)	(0.034)*	(1.000)
Distal landmark			-	*1.000*	(0.250)	(0.058)+	(1.000)
Natural panorama				-	(0.068)+	(0.004)**	(0.278)
Square					-	(0.448)	1.000
Statue						-	(0.360)
Bridge							-

* $p < .05$, ** $p < .01$, + $p < .10$ (no mathematical correction made for multiple comparison)

The number of observed similar object errors (Fig. 3) was smaller than that of footprint mismatch errors. This type of error was only found for Flickr data, with two occurrences for buildings (Rome, Vienna), two for bridges (Rome), and one for type statue (Vienna).

5 Summary and Outlook

This empirical study provides new insight into the positional accuracy of Flickr and Panoramio images. It showed that Panoramio images with a median error distance of 0 meters are more accurate than Flickr images with a median error distance of 58.5 meters. Within the Flickr data set, it was observed that the type of scene that was being photographed partially affected both positional accuracy and percentage of footprint mismatch errors. The Panoramio dataset did not show these effects. It can therefore be considered homogeneous in spatial accuracy across all content types and more reliable than the Flickr data. Since geotagged images from both data sources are freely available, they provide an attractive potential resource for a variety of geo- and mapping applications. The findings of this study should provide some help in deciding which of the analyzed data sources would meet the required point accuracy for future geo-applications, such as automated image annotations to route direction services. While this study analyzed positional accuracy for Europe, we plan to expand the accuracy assessment of shared photographs through these services to other continents, including Asia, and North- and South America.

References

FLICKR (2009), code.flickr. http://code.flickr.com/blog/2009/02/04/100000000-geotagged-photos-plus/, accessed 2/1/2012.

GIRARDIN, F., BLAT, J., CALABRESE, F., FIORE, F. D. & RATTI, C. (2008), Digital Footprinting: Uncovering Tourists with User-Generated Content. Pervasive Computing, 10-11/2008, 36-43.

GOODCHILD, M. F. (2007), Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0 (Editorial). International Journal of Spatial Data Infrastructures Research (IJSDIR), 2, 24-32.

HAKLAY, M. (2010), How good is Volunteered Geographical Information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B, Planning and Design, 37 (4), 682-703.

HOCHMAIR, H. H. (2010), Spatial Association of Geotagged Photos with Scenic Locations. In: CAR, A., GRIESEBNER, G. & STROBL, J. (Eds.), Proceedings of the Geoinformatics Forum Salzburg. Berlin/Offenbach, Wichmann, 91-100.

HOLLENSTEIN, L. & PURVES, R. S. (2010), Exploring place through user-generated content: Using Flickr to describe city cores. Journal of Spatial Information Science, 1, 21-48.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2009), Standards Guide: ISO/TC 211 Geographic Information/Geomatics. http://www.isotc211.org/Outreach/ISO_TC_211_Standards_Guide.pdf, accesseed 2/1/2012.

KREMERSKOTHEN, K. (2011), flickr Blog – 6,000,000,000. http://blog.flickr.net/en/2011/08/04/6000000000/, accessed 2/1/2012.

NEIS, P., ZIELSTRA, D. & ZIPF, A. (2011), The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007-2011. Future Internet, 4 (1), 1-21.

SCHLIEDER, C. & MATYAS, C. (2009), Photographing a City: An Analysis of Place Concepts Based on Spatial Choices. Spatial Cognition and Computation, 9 (3), 212-228.

TAYLOR, F. (2008), Google Earth Blog. http://www.gearthblog.com/blog/archives/2008/01/panoramio_layer_adds_2_million_phot.html, accessed 2/1/2012.

ZIELSTRA, D. & HOCHMAIR, H. H. (2011), A Comparative Study of Pedestrian Accessibility to Transit Stations Using Free and Proprietary Network Data. Transportation Research Record: Journal of the Transportation Research Board, 2217, 145-152.