Identification of Practically Visible Spatial Objects in Natural Environments

Martin Tomko, Friedjoff Trautwein, Ross S. Purves

GIS Division, Department of Geography, University of Zurich-Irchel, Winterthurerstr. 190,
CH-8057 Zurich, Switzerland

{martin.tomko,trautwei,ross.purves}@geo.uzh.ch

Abstract. Image retrieval of landscape photographs requires accurate annotation using multi-faceted descriptions relating to the subject and content of the photograph. The subject of such photographs is dominantly the terrain and spatial objects visible from the photographer's viewpoint. While some spatial objects in the background may be obscured by foreground vegetation, other visible spatial objects beyond a certain distance may not present noteworthy elements of the captured scene (such as distant houses). Our aim is to assess approaches to improve the identification of practically visible spatial objects for image annotation. These approaches include the consideration of the apparent spatial object size and landcover information about occluding vegetation. These inputs are used to enhance viewshed analysis to accurately identify only spatial objects practically visible and therefore likely to be notable subjects of a photograph. The two approaches are evaluated in an experiment in a semi-rural area of Switzerland, whose results indicate that visual magnitude is key in accurate identification of visible spatial objects.

1 Introduction

Landscape photographs are records of the visible portion of the terrain and the objects and vegetation positioned on top of it. Current efforts in spatial image annotation, such as project TRIPOD (`http://tripod.shef.ac.uk/`) aim at accurate annotation and captioning of landscape photographs for image search and retrieval. Photographs can be annotated using multi-faceted descriptions relating to, among others, the subject of the photograph (Shatford, 1986). Therefore, the objects visible from a viewpoint contained within a photograph's viewport need to be reliably identified.

Consider a photograph of a rural landscape. Typically, objects in the middle distance or background are partially obscured by vegetation and other proximal objects. Furthermore, distant objects may be barely identifiable due to their small apparent size and reduced contrast from background as a consequence of atmospheric conditions. Hence, while visible, objects beyond a certain distance may not present noteworthy elements of the captured scene. Finally, photographs are printed or viewed on screen and the resolution of this visualization further reduces the number of noteworthy elements of the scene.

The aim of this paper is to assess approaches to improve the identification of *practically visible* objects for image annotation. Apparent object size and enhancement of the digital elevation model with information about vegetation occlusion need to be considered during the calculation of the viewshed in order to accurately identify the objects practically visible from the origin of the photograph and therefore likely to be the subject of the photograph. We test improvements brought about by limiting the computation to a distance beyond which the visual impact of objects is negligible and compare it to the improvements from DEM data enhanced by landcover information from global multispectral remote sensing imagery to infer the presence or absence of occluding vegetation.

As we wish to develop techniques which do not require detailed spatial data, since we wish to process photographs from a large area (such as Europe), only general purpose datasets with large-area coverage are practically usable for image annotation. Furthermore, parameters of the camera sensor and display system further impact on the visibility of an object in the photograph and its relevance to the captured scene.

This paper is structured as follows: in the next section, we review past research pertinent to visual impact analysis of landscapes from the perspective of image information retrieval. In Section 3 we present two methods that may improve the inference of practically visible spatial objects. We put these methods to a test in Section 4 and we present the results of the individual methods. In Section 5 the results are discussed and conclusions are drawn in Section 6, along with suggestions for further work.

2 Background

2.1 Information Retrieval

Accurately annotated documents improve the relevance of results during information search (Salton & Buckley, 1988; van Rijsbergen, 1979) and thus improve user experience. In recent years, the importance of the geographical scope of digital documents was widely recognized (Larson, 1996; Purves et al., 2007). Geographic Information Retrieval (GIR) emerged as a specific area of interest, where methods to infer and use the geographic scope of the documents – their footprint – are researched. Once a footprint is assigned to a document, spatial objects found within it can be used as source of highly contextual information for the annotation of the documents (Naaman et al., 2006; Purves et al., 2008). Such topical and accurate annotation is then used in retrieval to identify documents matching the query by geographic and thematic scope.

Digital photography is an emerging field of interest for GIR. Urban and rural landscape photographs have a clear geographic context provided by the photograph's origin (the location, focus and orientation of the camera) and the subject of the photograph. Photographers' annotations frequently reflect this geographic scope – consider Figure 1, with an example caption: "*A country house seen across an orchard, near Zurich, Switzerland*". A photographer might annotate this photograph with keywords such as *house*, *orchard*, *Zurich*, and *Switzerland*. One could also refer to the individual trees, grassy lawn, footpath in the foreground and forest in the background. These are, however, not prominent elements of the scene and inclusion in the annotation would reduce the precision of the search results by including this picture in the result sets for photographs of forests or footpaths.

To improve annotation of photographs, we focus on the determination of the *practically* visible portion of a rural landscape, to identify spatial objects of substantial visual impact contained in a photograph. This should lead to annotation accuracy superior to that resulting from the use of simple circular buffer regions around a photograph's origin, or viewsheds computed purely based on the terrain. Parallel research focusing on urban environments is being undertaken by De Boer et al. (2008), and related work identifying other qualities of the scene captured through multifaceted image descriptions is presented in (Edwardes & Purves, 2007).

Figure 1 A country house seen across an orchard, near Zurich, Switzerland (Photo and caption Martin Tomko)

2.2 Viewshed

The computation of a viewshed – the visible portion of a terrain and objects on top of it (De Floriani & Magillo, 2003), is a geographic analysis task applied to problems from urban planning to archaeology. Viewshed computation typically assumes that an object is visible to an observer if an unobstructed line of sight can be constructed between the observer's eye and the object. The computation is usually performed on an interpolated digital elevation model devoid of surface objects or vegetation (Fisher, 1996; Kaučič & Zalik, 2002; Maloy & Dean, 2001). Viewshed calculation can then be used to identify objects situated in the visible portions of the surface. The calculation of a viewshed can be limited to a specific direction and distance (by specifying, for instance, the maximum length of the line of sight).

As noted by Ervin and Steinitz (Ervin & Steinitz, 2003), simple computation of viewsheds is not sufficient to assess the visual quality of a landscape. The way visual quality of a landscape impacts on a human observer is determined by a wide variety of factors, intrinsic to the landscape but also dependent on the observer's context (Litton, 1968).

2.3 Landscape Perception and Visual Impact

Typically, people are able to summarize the visual quality of a landscape in a few words. While some aspects of the visual quality are highly subjective and reflected in adjectives such as romantic, peaceful, serene, others are more tangible and relate to visible objects and landcover. These different facets of the landscape are similar to the facets of image descriptions, as studied in Shatford (1986).

The material aspects of landscape quality and its change (such as introduction of anthropogenic objects or landuse change) has been the focus of multiple studies (Bishop, 2003; Daniel, 2001; Gret-Regamey et al., 2007; Magill, 1990). These studies relied on the assessment of the visual impact of the introduced objects based on computer visualizations and digital photographs altered by computer animations (Bishop, 2002; Hadrian et al., 1988; Shang & Bishop, 2000) and are restricted to parameters that can be objectively determined, for example by measurement of physical qualities (Groß, 1991).

For an object to be notable in a scene, its apparent size must exceed a certain visual magnitude, also known as visual threshold (Iverson, 1985; Magill, 1990). Three different visual magnitudes derived from the parameters of human visual acuity (approximately 1') determine the thresholds for object detection, recognition (or identification) and visual impact (Shang & Bishop, 2000).

An object with a visual magnitude of 1' can just be detected by the retina (as a single dot, or pixel), but not recognized or have visual impact. Depending on the type of object and viewing conditions, a simple, well known object has to exceed a visual magnitude of approximately 5.5' in order to be recognized (Luebke et al., 2003). At this visual magnitude the most salient elements of the object's structure can be differentiated. This is reflected in common cartographic guidelines (for example (Spiess et al., 2005)) where map symbols are rendered as 5x5 pixels at least.

In natural landscapes, few objects have well defined familiar shapes. Furthermore, the viewer does not know *a priori* which objects will be visible (uninformed recognition). Studies performed on digital images of faces, outdoor and indoor objects and complex scenes showed that a natural object had to be rendered with a higher resolution to be recognized (Cai, 2004).

Visual thresholds based on visual magnitude can be used to limit the length of the line of sight during viewshed calculation. However, recognition of objects in natural settings is a much more complex task than the simple recognition of letters or symbols in controlled laboratory conditions. While it can be limited to the determination of visual magnitude for practical reasons, experience, personal objectives and atmospheric conditions play a strong role in recognition of objects (Pitchford & Malm, 1994). Furthermore, when the objects are to be detected or recognized in photographs as opposed to viewed in natural settings as such, the resolutions of the sensor, lens (optical) and display systems affect the visual thresholds as detailed in Section 3.2.

2.4 Visibility and Occlusion by Vegetation

Little research has directly addressed the influence of vegetation on the visibility of the surrounding space. Dean (1997) proposed a method to improve the prediction of object visibility in forests based on estimates of the vegetation's opacity, characterized by a visual permeability value. The study combined DEM data with extruded vegetation from detailed forest inventory data, including accurate tree heights. All evaluation was limited to lines of sight of 50 to 500m, with an orange air balloon as an artificial target.

Another method was proposed for object visibility prediction in paleoarcheology by Llobera (2007). It is based on principles derived from light attenuation by particles and relies on highly accurate data about spatial distribution of individual plants in the area studied. While plausible, the model has only been tested on a synthetic DEM using simulated vegetation coverage and relies on data of too high an accuracy for practical image annotation.

An attempt to use widely available, global coverage vegetation information of relatively high resolution for realistic visualization of terrain was proposed by Roettger (2007). Based on a classification of the well-known Normalized Difference Vegetation Index (NDVI) values, they infer the presence of vegetation at a particular location. Furthermore, they map NDVI values to vegetation height based on a linear interpolation between user defined maximum and minimum values. While not tested in a field experiment, the method could provide a simple and efficient way of estimating the distribution of vegetation over large areas at acceptable resolution and thus provide a viable basis for the consideration of vegetation occlusion in object visibility analysis.

3 Method

We propose two methods to improve the results of viewshed calculations. First, we determine a visual impact threshold for landscape images viewed on LCD displays. Second, we enhance the DEM used to calculate these viewsheds by adding extruded vegetation information.

3.1 Visual Impact Determination for Photographs

For the annotation of photographs, the impact of the sensor and display parameters to the determination of the visual impact threshold have to be considered. The acuity of human vision, as well as the resolution of consumer grade digital camera sensors is beyond the resolution of typical LCD displays. Photographs are displayed on displays at a fraction of their actual resolution. The display thus represents the effective limit to the identification of objects in photographs. The resampling *r* is equivalent to the ratio between the sizes of the sensor (*sensordim*) and the screen (*screendim*, in pixels)(Figure 2):

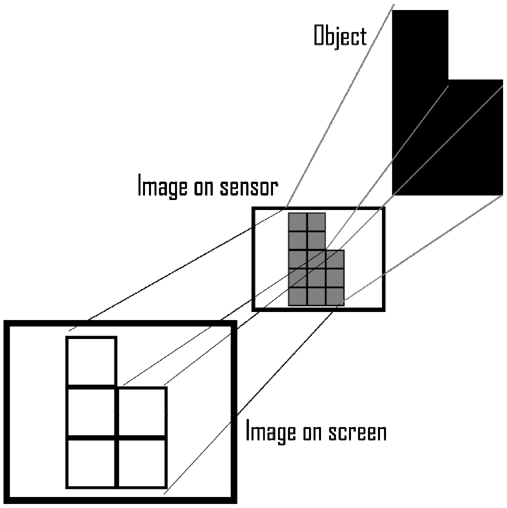

Figure 2 Resampling occurring in the object-sensor-display system.

$$r = \frac{screendim_{pix}}{sensordim_{pix}} \quad \textbf{(1)}$$

The angular field of view *afov* captured by a camera is characterized by the focal length *f* of the lens used and the physical size of the sensor, in *mm*:

$$afov = 2atan\left(\frac{sensordim_{mm}}{2f}\right) \tag{2}$$

Images of recognizable natural objects consist of at least 1024 pixels (32 x 32 pixels), compared to only 289 pixels(17x17 pixels) for familiar faces (Cai, 2004). If the object is to be recognized on screen, this is the size of the object's rendered image and not that image captured by the sensor. As the resolution of the screen is the limiting factor of the sensor-display system the image of the object has to be captured as a square of side *is* = *wr* (*is* – image size on sensor, *w* – image size on screen, in pixels).

The density of pixels on the sensor determines the angular resolution of the sensor. The angular resolution *ares* of the sensor – lens combination is the fraction of the angular field of view that is captured by one pixel of the sensor. The higher the sensor pixel density (or, the smaller the pixel size), the more pixels will capture the same extent of *afov*.

From the image size *is* and the angular resolution of the sensor – lens combination it is possible to determine the minimal angular field of view α occupied by an object of known size to exceed the visual impact threshold. The maximal distance *d* at which this magnitude is exceeded by the object of size *o* for a given sensor-lens-display combination is:

$$d = \frac{\frac{o}{2}}{tan(\frac{\alpha}{2})} \tag{3}$$

In Section 4.3, we use the approach outlined to compute the distance *d* for the combination of sensor, lens and display used in a set of field experiments. The value of *d* is then used to limit the computation of the viewsheds for observation points, in order to identify only practically visible objects for photographs of the given landscape scenes.

3.2 Occlusion by Vegetation

The second method explored aims at accurate inference of vegetation occlusion. This requires reliable information about the spatial distribution of vegetation and its height. In order to be practical for image annotation, the method should use general-purpose datasets of large-area coverage. Furthermore, accurate information about vegetation height is, usually, not available.

We build on the approach of Roettger (2007) using NDVI extracted from remote sensing imagery. NDVI values are computed from sampling the Earth's surface in the near infra-red (NIR) and visible red (VIS) bandwidth of the Landsat ETM+ sensor. The index is calculated as follows:

$$NDVI = (NIR — VIS)/(NIR + VIS) \tag{4}$$

The index gives an estimate of healthy vegetation land cover. While values beyond a given threshold are likely to relate to dense foliage and allow inference of the presence of forests or shrubs, it is impossible to directly relate the value of the index to the height of vegetation. We therefore chose a single threshold value to indicate the presence of dense vegetation, without relating the index values of the vegetated areas to vegetation height. The index value of 0.2 of Roettger (2007) was taken as a starting point and tested in 0.01 increments up to 0.3. Best matches between the vegetation layer derived from NDVI and thematic landcover datasets of the Swiss national mapping agency Swisstopo were achieved for values of 0.27 (Vector200 dataset) and 0.28 (Vector25 dataset) and confirmed by visual comparison with photogrammetric records of the area. The value of 0.28 was chosen for the extrusion of vegetation in the experiment due to its best match in the direct vicinity of the experiments' observation points.

As no detailed datasets of vegetation heights is associated with the vegetation layer derived from NDVI, and our motivation does not allow for specialized spatial datasets, we built on the knowledge of the forest types in the area of interest (mostly mixed beech and spruce forests), three tree heights were used to extrude the vegetation layer - 10, 20 and 30m (for more information on forest types, see `http://www.gis.zh.ch` and (BAFU, 2005)). The extruded vegetation was then added to the DEM of the area studied and viewshed were calculated. Results of the visibility analysis are reported in Section 4.4.

4 Experiment and Results

4.1 Overview

In two experiments we evaluated the possibility to identify visible objects for image annotation. Two approaches are tested - viewshed analysis enriched with heuristics about object's visual magnitude and viewshed analysis including consideration of occlusion by vegetation using an extruded layer of landcover information. The workflow of the two methods and their evaluation is outlined in Figure 3.

In the right strand, the workflow for experiment 1 is shown in parallel to experiment 2 (left strand). Joint data or analytical procedures overlap both strands.

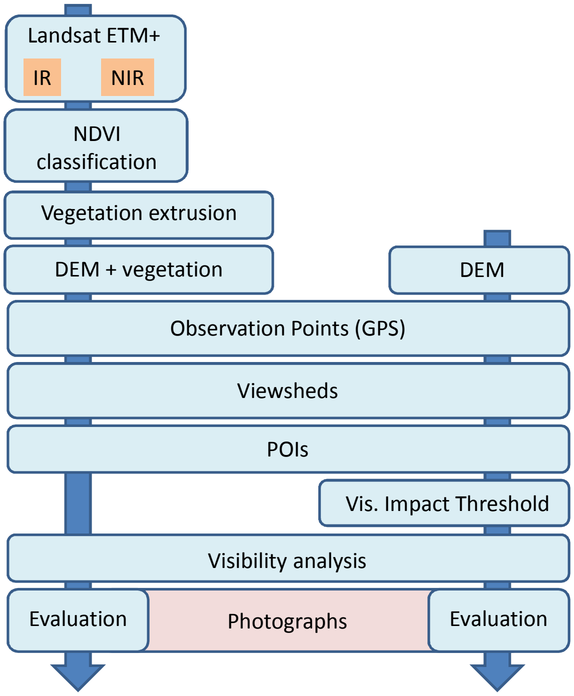

Figure 3 Workflow schema.

4.2 Data

We limit our analysis to datasets that are available at low costs and provide large area or global. For our experiments, the following datasets covering the region around Zurich, Switzerland, were used (all Swisstopo datasets in the Swiss *CH1903* national grid coordinate system):

- Orthorectified Landsat 7 ETM+ band 3 and 4 dataset (image p194r027_7), acquired on August 24th, 2001, referenced in WGS84 (transformed into CH1903), spatial resolution of 28.5m;
- A raster DEM raster dataset Swisstopo DHM25 with a spatial resolution of 25m. The height accuracy varies from 1.5m in flat lands to 3m in Alpine regions (Swisstopo, 2005);
- A dataset containing centroids of all named objects present on the 1:25000 Swisstopo maps (Swissnames);

While the Swissnames dataset is not an ideal source of point of interest (POI) data due to its explicit focus on cartographic content (it contains the centroids and labels of all toponyms on Swistopo maps), it is the best available dataset with comprehensive coverage in rural areas. The dataset was filtered to include only 29 categories of objects that can be considered point-like for the purpose of our assessment (excluding names of forests, meadows, hills etc.), with the exception of settlements and ponds, included due to their easy visual identification in photographs. Note that no information is available about the objects' size and height, and therefore their projective size cannot be computed.

Furthermore, the following data were collected:

- Coordinates of 12 points from which photographs of the surroundings were taken. These points served as centroids for the generation of viewshed and POI visibility analysis;
- 83 georeferenced photographs with directional information, taken from the 12 observation points, taken with an 8.13 Mpix Ricoh Caplio 500G digital camera (sensor size 3264 x 2448 pixels, physical sensor size 7.18 x 5.32 mm) with direct Bluetooth link to a GPS receiver. Image azimuths were measured with a handheld digital compass. All photographs were taken with a focal length of 5mm (wide angle) reported in EXIF data, equivalent to a field of view of 71^{o}. The 360^{o} panoramas for each of the observation points are shown in Figure 4**.** The photographs were viewed on an LCD display with resolution of 1280*1024 pixels (Philips Brilliance 200W) with a pixel size of approximately 0.294mm.

(a) Point 1

(b) Point 2

(c) Point 3

(d) Point 4

(e) Point 5

(f) Point 6

(g) Point 7

(h) Point 8

(i) Point 9

(j) Point 10

(k) Point 11

(l) Point 12

Figure 4 Views from the 12 test sites as panoramic collages of the photographs taken.

4.3 Experiment 1: Objects Exceeding the Visual Impact Threshold

The visibility of POI objects was analyzed by calculating a 360° viewshed on the DEM. For comparison of the results with Experiment 2, the location of each POI was rasterized to match the cells of the vegetation layer (spatial resolution of 28.5m). As no information about the real size of the spatial objects was available, this value was taken as input for the calculation of the visual impact threshold. We assert that 28.5m represent a reasonable size estimate for man-made spatial objects such as farm houses. The counts of POIs evaluated as visible in the viewshed analysis without distance limitation are shown in Table 1 (DEM).

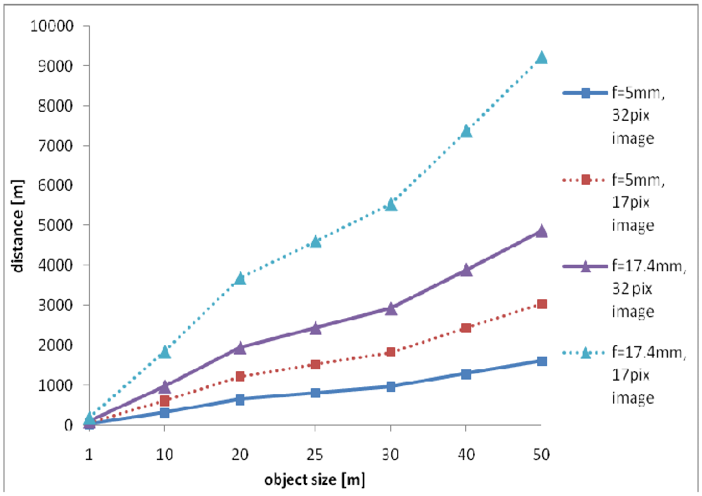

Figure 5 Dependence of minimum distance to object from object size, visual impact threshold and parameters of the sensor-lens system. For an object to be above visual impact threshold, it must be closer than the distance related to its size.

For comparison, the objects exceeding the visual impact threshold were identified. First, the distance at which the visual impact threshold for the POIs is exceeded was determined. An object of 28.5m occupies a screen space of 17x17pix to 32x32pix (approximately 0.5cm to 0.94cm on the screen used and 43x43 to 82x82 sensor pixels) when closer than 914m - 1730m, if photographed with *f*=5 mm lenses (wide angle lens). This is equivalent to an apparent visual magnitude of 0.94° to 1.78° for an object observed by naked eye. For the plot of dependencies between the focal length, object size and object distance to exceed the visual impact threshold see Figure 5. As shown, the visual impact threshold distance for the same object, but captured using a *f*=17.5mm lens is between 4 to 10km. A single value of 1km has been taken as a conservative substitute of the interval identified for *f*=5mm lens, allowing for degradation of visual impact due to, for example, contrast reduced by haze and unfamiliar object shapes. The counts of the objects exceeding the visual impact threshold are reported in Table 1.

Table 1 Counts of visible POIs based on viewshed analysis without distance limitation and with a distance limitation of 1km based on visual impact threshold (DEM without vegetation). Image – POI visible in the photograph. DEM – POI evaluated visible using the DEM. 1km buffer– POI within 1km of the observation point. 1km buffer + DEM – POI predicted to be visible using the DEM within 1km of the observation point.

Observation Point	Image	DEM	1km buffer	1km buffer + DEM
p1	1	33	3	1
p2	0	36	2	0
p3	0	44	0	0
p4	1	91	4	1
p5	0	92	3	0
p6	1	91	5	1
p7	4	261	5	4
p8	0	82	3	1
p9	1	46	5	1
p10	0	57	6	0
p11	1	70	9	1
p12	0	34	7	1

Each object that was evaluated as visible in either of the two viewshed analyses was searched for in the corresponding photograph and marked as visible or invisible. Only objects considered large enough to be of visual impact to the subject of the image were identified as visible (executed as an image labeling exercise similar to that from Russell et al. (2008), Figure 6). The counts of the visible objects are reported in Table 1 (Image).

The results reported can be interpreted using the standard measures to assess the quality of remote sensing classifications through contingency tables. As none of the points visible in the photograph were reported as invisible in the DEM or not present in the 1km buffer region, the full contingency table can be reconstructed by the interested reader. As shown, the results of viewshed analysis neglecting vegetation information greatly exaggerate the number of visible POIs in all cases. The limitation

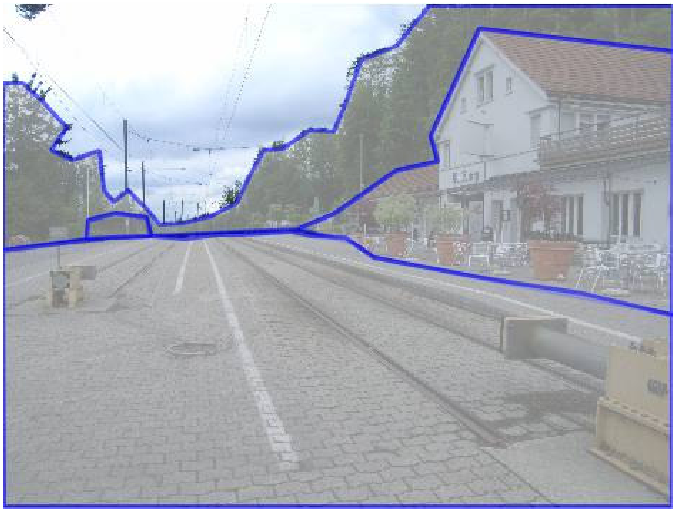

Figure 6 Example of detection of visible objects in a labeled photograph. The train station building (image right) is contained in the POI database.

of the visibility analysis to the distance at which the objects exceed the visual impact threshold achieves significantly higher precision of detection. Only in two out of 12 cases, an extra POI has been reported for a given image.

The apparent size of the smallest object considered of significant visual impact found in the labeled photographs is approximately 170 sensor pixels. The object has an apparent height of approximately 5.8mm on the screen, or approximately 20 screen pixels. This size is slightly inferior to the theoretical visual impact threshold used in this study. The corresponding object is a barely visible radio tower and hence it has a particular, familiar elongated shape and it is positioned on a prominent hill on the horizon. Radio antennas are prominent spatial objects frequently used as landmarks due to their good visibility and their high figure-ground contrast.

4.4 Experiment 2: Visibility Analysis Simulating Occlusion by Vegetation

The dataset based on NDVI classification provides information about presence or absence of vegetation. A threshold NDVI value of 0.28 was selected for vegetated areas and the extruded vegetation was added to the DEM and used for viewshed computation. The pixel incident with the observation point used for the calculation of a viewshed was kept at the original altitude of the DEM (the observation points were all on the ground or man–made structures).

The results indicate that the consideration of vegetation does not perform as well as the simple combination of DEM with a visual magnitude threshold consideration (Table 2). Only for seven out of 12 observation points the counts of visible POIs are accurate, in all cases where there were no visible objects in the photograph and hence the effect of over-filtering cannot be detected. No significant dependence was found

Table 2 Visibility analysis of POIs with vegetation occlusion. Horizontal reading: counts of spatial objects identified as visible (V) or not visible (NV) on a DEM with extruded vegetation of 10m, 20m and 30m, without distance limitation are shown for each point and vegetation combination. Reading by column: corresponding counts of the same spatial objects visible or not visible in photographs.

p1		image		p1		image		p1		image	
		V	NV			V	NV			V	NV
10m	V	0	3	20m	V	0	0	30m	V	0	0
veg	NV	1	NA	veg	NV	1	NA	veg	NV	1	NA
p2		image		p2		image		p2		image	
		V	NV			V	NV			V	NV
10m	V	0	0	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p3		image		p3		image		p3		image	
		V	NV			V	NV			V	NV
10m	V	0	8	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p4		image		p4		image		p4		image	
		V	NV			V	NV			V	NV
10m	V	1	17	20m	V	1	5	30m	V	0	1
veg	NV	0	NA	veg	NV	0	NA	veg	NV	1	NA
p5		image		p5		image		p5		image	
		V	NV			V	NV			V	NV
10m	V	0	12	20m	V	0	3	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p6		image		p6		image		p6		image	
		V	NV			V	NV			V	NV
10m	V	0	6	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p7		image		p7		image		p7		image	
		V	NV			V	NV			V	NV
10m	V	3	103	20m	V	0	43	30m	V	0	7
veg	NV	1	NA	veg	NV	4	NA	veg	NV	4	NA
p8		image		p8		image		p8		image	
		V	NV			V	NV			V	NV
10m	V	0	21	20m	V	0	13	30m	V	0	9
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p9		image		p9		image		p9		image	
		V	NV			V	NV			V	NV
10m	V	0	45	20m	V	0	0	30m	V	0	0
veg	NV	1	NA	veg	NV	1	NA	veg	NV	1	NA
p10		image		p10		image		p10		image	
		V	NV			V	NV			V	NV
10m	V	0	0	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p11		image		p11		image		p11		image	
		V	NV			V	NV			V	NV
10m	V	0	4	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA
p12		image		p12		image		p12		image	
		V	NV			V	NV			V	NV
10m	V	0	2	20m	V	0	0	30m	V	0	0
veg	NV	0	NA	veg	NV	0	NA	veg	NV	0	NA

10m, lower than the mean height of the typical vegetation in the area.

The results indicate that the method is prone to over-filtering – the elimination of objects that are actually visible and can be identified in photographs (see values for Image[V]/model[NV] in Table 2). This is mostly due to the binary classification of the terrain surface as vegetated and not vegetated. As a result, sparse vegetation is extruded as an opaque cell (Figure 7). Thus, while the vegetation classification may be spatially correct, a simple extrusion of the vegetation layer may not present the most appropriate method for vegetation modeling. It also appears that positional accuracy of the vegetation dataset has higher impact on the results than accurate information about vegetation height.

While the results are often over-filtered, they also contain frequent false matches. POIs are reported as visible while they are not visible. This is likely due to occlusion by objects in the foreground, close to the observer. Hence, we conclude that the method is extremely sensitive and highly dependent on accurate vegetation information, as well as requiring complex data processing. As such, it is not suited for automated annotation of images for GIR.

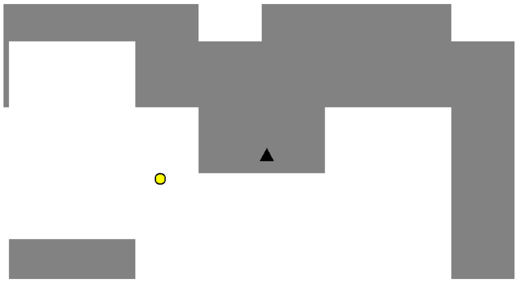

Figure 7 Visibility of an object (circle) obstructed by vegetation (adjacent pixel). In reality, this is an orchard and the vegetation is visually permeable (Dean, 1997). The observation point P11 is shown as triangle. The photograph of the scene is shown in Figure 1.

5 Case Study and Discussion

5.1 Case Study

In order to verify our findings indicating that visual magnitude thresholds and viewshed analysis based on DEM data (without vegetation) provide sufficient inputs for the inference of practically visible objects, we tested 4 arbitrarily selected georeferenced landscape photographs from different authors, similar to those available from photo-sharing sites such as Flickr. The photographs were selected from the area covered by identical datasets to those used earlier. All photographs were acquired within the last 2 years for the project TRIPOD. The photographs did not contain directional information and this information was therefore computed by relating the edges of the photographs with available spatial data and consequent computation of azimuths.

Table 3 Number of POIs evaluated as visible using five combinations of viewsheds (calculated on DEM), distance thresholds and the actual photograph's field of view (FOV). The values of POI Image indicate the number of POIs actually visible in the photographs.

Image	POI Viewshed	POI buffer 1km	POI Viewshed 1km	POI within 1km in FOV	POI FOV viewshed	POI Image
A	484	11	3	1	1	1
B	130	6	2	4	2	2
C	90	9	2	2	0	1
D	96	1	0	0	0	0

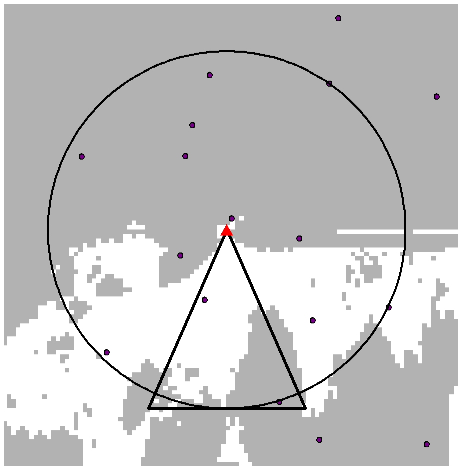

Figure 8 Viewshed of point A overlaid with the 1km buffer and the visual field of view of the image. POIs are represented as points. Visible cells of the DEM are white, invisible cells are grey.

For each photograph, the viewshed, 1km buffer and its directional field of view were calculated (Figure 8). The results, shown in Table 3 confirm that the combination of visibility calculation based on DEM (without vegetation), combined with a visual impact threshold value (expressed as 1km buffer) and field of view information provide together a reliable means to identify objects captured in a photograph. Note that alone, neither the viewshed analysis, nor the distance limitation within the available field of view yield optimal results. Their combination, however, allows for reliable identification of visible objects. The result for image C (containing one visible object but resulting in a prediction of no objects) points to method's dependence on accurate estimate of the objects size – image C contains a distant airport, with a building exceeding the size of 25m. As such, the airport is a significant element in the photograph.

5.2 Discussion

Two experiments were performed in a semi-rural environment with abundant vegetation and sparse man-made objects – POIs. In the experiments, a substantial reduction in the counts of spatial objects incorrectly classified as visible was achieved by limiting the visibility calculation to a distance at which an object has a visual magnitude above a visual impact threshold. Such a limitation based on a simple heuristic determination of the visibility impact threshold allows the elimination of objects that do not present a significant element of the observed and photographed scene if rendered on a computer screen.

Similarly, the visibility analysis including landcover information shows a reduction in the number of spatial objects visible compared to viewsheds calculated on pure DEM. The consideration of vegetation should allow the elimination of objects occluded by foreground vegetation and leads to more realistic results of the visibility analysis. The vegetation in the foreground has high impact on the results compared to background vegetation, as objects in the foreground occlude a larger proportion of the visual field. It seems therefore that the accuracy of the data about the presence or absence of vegetation is more important than the exact knowledge of the vegetation's height. The variation of the vegetation height has had little impact on the results. The results obtained from the experiment performed, however, indicate that the consideration of vegetation is much more sensitive to the data available and the results obtained do not justify the computationally intensive process. In a follow-up case-study, we have shown that DEM data, combined with the simple visual impact threshold allows us to infer the objects actually visible in arbitrary photographs.

7 Conclusions and Future Work

Limiting the visibility analysis to objects appearing larger than the visual impact threshold is an efficient and effective method to reduce the computation of viewsheds and at the same time identify spatial objects relevant to image annotation. The visual magnitude of photographed objects is significantly influenced by the display on which the photographs are viewed, and the consideration of the resampling between the sensor and the display influence the estimate of the visual magnitude of the photographed object. It is important to note that the object's shape and the observer's position in relation to the object alter the visual impact of the observed object. If an object is viewed from a familiar perspective (also known as canonical perspective) (Palmer et al., 1981), its recognition is better and its visual impact is greater than when observed from an unfamiliar perspective. It is, however, difficult to infer whether an object is viewed from a canonical perspective if information about the object's shape and additional contextual information about viewpoints selected by other photographers in the region is not available. The latter point is currently addressed in research on geographic recommenders (Schlieder, 2007) researching amongst other the context of the photograph as defined by the past photographic activity of the photographer or their peers.

We further presented a simple method to enhance the estimate of the visual impact of an object with information about occlusion by foreground vegetation. The consideration of vegetation information may, in some cases, further improve the veracity of the visibility analysis, but care has to be taken not to over-filter the visible objects. Further research on vegetation visual permeability could lead to improved results, as suggested by Dean (1997). Note, however, that such approaches seem to be less reliable and more data expensive than a simple heuristic about the visual magnitude of the photographed objects.

The visual impact of an object can be further deteriorated by external factors altering the contrast of the object from the background, such as atmospheric conditions and the surface properties of the object. The consideration of atmospheric influences on visual threshold may be more practical than that of vegetation and could further improve the results. Meteorological services broadcast weather information including visibility range and haze information (for instance, METAR (OFCM, 2005)) that could be included in the threshold determination similar to (Pitchford & Malm, 1994). Heuristics allowing for accurate inference of the objects' size will, however, provide the greatest improvement. Such heuristics could be based, for instance, on the analysis of the category of spatial objects and the use of a mean size value per category.

Image annotation is an important step for the organization and management of searchable image libraries. Images annotated only with keywords related to the image content of practical visual impact allow for better image search relevance. Previously, Tomko and Purves (2008) focused on the analysis of the spatial distribution of POI in a given region as a means to infer an object's relevance for the annotation of the region. The identification of only practically visible spatial objects is a necessary requirement for such a classification method, providing inputs for multifaceted image descriptions (Edwardes & Purves, 2007).

Acknowledgments

The research reported in this paper is part of the project TRIPOD supported by the European commission under contract No. 045335.

References

BAFU. (2005). *Waldtypen der Schweiz* Bern, Switzerland: Bundesamt für Umwelt BAFU.

Bishop, I. (2002). Determination of Thresholds of Visual Impact: the Case of Wind Turbines. *Environment and Planning B: Planning and Design, 29*, 707-718.

Bishop, I. (2003). Assessment of Visual Qualities, Impacts, and Behaviours, in the Landscape, by Using Measures of Visibility. *Environment and Planning B: Planning and Design, 30*(5), 677-688.

Cai, Y. (2004). Minimalism Context-Aware Displays. *CyberPsychology and Behavior, 7*(6), 635-644.

Daniel, T. C. (2001). Whither Scenic Beauty? Visual Landscape Quality Assessment in the 21st Century. *Landscape and Urban Planning, 54*, 267-281.

De Boer, A., Dias, E., & Verbree, E. (2008). Processing 3D Geo-Information for Augmenting Georeferenced and Oriented Photographs with Text Labels. In A. Ruas & C. Gold (Eds.), *Headway in Spatial Data Handling* (pp. 351-365). Berlin, Heidelberg: Springer-Verlag.

De Floriani, L., & Magillo, P. (2003). Algorithms for Visibility Computation on Terrains: a Survey. *Environment and Planning B: Planning and Design, 30*(5), 709-728.

Dean, D. J. (1997). Improving the Accuracy of Forest Viewsheds Using Triangulated Networks and the Visual Permeability Method. *Canadian Journal of Forest Research, 27*, 969-977.

Edwardes, A. J., & Purves, R. S. (2007). *Eliciting Concepts of Place for Text-based Image Retrieval.* Paper presented at the 4th ACM Workshop On Geographic Information Retrieval, GIR 2007, Lisbon, Portugal.

Ervin, S., & Steinitz, C. (2003). Landscape Visibility Computation: Necessary, but not Sufficient. *Environment and Planning B: Planning and Design, 30*(5), 757-766.

Fisher, P. F. (1996). Extending the Applicability of Viewsheds in Landscape Planning. *Photogrammetric Engineering & Remote Sensing, 62*(11), 1297-1302.

Gret-Regamey, A., Bishop, I. D., & Bebi, P. (2007). Predicting the Scenic Beauty Value of Mapped Landscape Changes in a Mountainous Region Through the use of GIS. *Environment and Planning B: Planning and Design, 34*(1), 50-67.

Groß, M. (1991). The Analysis of Visibility—Environmental Interactions Between Computer Graphics, Physics, and Physiology *Computers & Graphics, 15*(3), 407-415.

Hadrian, D. R., Bishop, I. D., & Mitcheltree, R. (1988). Automated Mapping of Visual Impacts in Utility Corridors. *Landscape and Urban Planning*(3), 261-282.

Iverson, W. D. (1985). And that's about the Size of it: Visual Magnitude as a Measurement of the Physical Landscape. *Landscape Journal 4*(1), 14-22.

Kaučič, B., & Zalik, B. (2002). *Comparison of Viewshed Algorithms on Regular Spaced Points.* Paper presented at the 18th Spring Conference on Computer Graphics, Budmerice, Slovakia

Larson, R. R. (1996). *Geographic Information Retrieval and Spatial Browsing.* Paper presented at the Geographic information systems and libraries: patrons, maps, and spatial information. 1995 Clinic on Library Applications of Data Processing, April 10-12, 1995.

Litton, R. B., Jr. (1968). *Forest Landscape Description and Inventories - a Basis for Landplanning and Design*. Berkeley, CA: Pacific Southwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture.

Llobera, M. (2007). Modeling Visibility Through Vegetation. *International Journal for Geographical Information Science, 21*(7), 799-810.

Luebke, D., Reddy, M., Cohen, J. D., Varshney, A., Watson, B., & Huebner, R. (2003). *Level of Detail for 3D Graphics*. Amsterdam, NL: Morgan Kaufmann Publishers.

Magill, A. W. (1990). *Assessing Public Concern for Landscape Quality: A Potential Model to Identify Visual Thresholds*. Berkeley, CA: Pacific Southwest Research Station, Forest Service, U.S. Depratment of Agriculture.

Maloy, M. A., & Dean, D. J. (2001). An Accuracy Assessment of Various GIS_Based Viewshed Delineation Techniques. *Photogrammetric Engineering & Remote Sensing, 67*(11), 1293-1298.

Naaman, M., Songa, Y. J., Paepckea, A., & Garcia-Molina, H. (2006). Assigning Textual Names to Sets of Geographic Coordinates. *Computers, Environment and Urban Systems, 30*(4), 418-435

OFCM. (2005). *Federal Meteorological Handbook No. 1: Surface Weather Observations and Reports*. Washington, D.C., USA: Federal Coordinator for Meteorological Services and Supporting Research, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical Perspective and the Perception of Objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (pp. 135-153). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pitchford, M. L., & Malm, W. C. (1994). Development and Applications of a Standard Visual Index. *Atmospheric Environment, 28*(5), 1049-1054.

Purves, R., Clough, P., Jones, C. B., Arampatzis, A., Bucher, B., Finch, D., et al. (2007). The Design and Implementation of SPIRIT: a Spatially Aware Search Engine for Information Retrieval on the Internet. *International Journal for Geographical Information Science, 21*(7), 717-745.

Purves, R., Edwardes, A. J., & Sanderson, M. (2008). *Describing the Where – Improving Image Annotation and Search Through Geography.* Paper presented at the 1st Intl. Workshop on Metadata Mining for Image Understanding (MMIU 2008), Funchal, Madeira – Portugal.

Roettger, S. (2007). *NDVI-based Vegetation Rendering.* Paper presented at the Computer Graphics and Imaging CGIM '07, Innsbruck, Austria.

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: a Database and Web-based Tool for Image Annotation *International Journal of Computer Vision, 77*(1-3), 157-173.

Salton, G., & Buckley, C. (1988). Term-Weighting Approaches in Automatic Text Retrieval. *Information Processing & Management, 24*(5), 513-523.

Schlieder, C. (2007). Modeling Collaborative Semantics with a Geographic Recommender. In J.-L. Hainaut, E. A. Rundensteiner, M. Kirchberg, M. Bertolotto, M. Brochhausen, Y.-P. P. Chen, S. S.-S. Cherfi, M. Doerr, H. Han, S. Hartmann, J. Parsons, G. Poels, C. Rolland, J. Trujillo, E. S. K. Yu & E. Zimányi (Eds.), *Advances in Conceptual Modeling – Foundations and Applications. Workshop on Semantic and Conceptual Issues in Geographic Information Systems, Auckland, New Zealand, November 5-9, 2007. Proceedings* (Vol. 4802, pp. 338-347). Berlin: Springer-Verlag.

Shang, H., & Bishop, I. D. (2000). Visual Thresholds for Detection , Recognition and Visual Impact in Landscape Settings. *Journal of Environmental Psychology, 20*, 125-140.

Shatford, S. (1986). Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging and Classification Quarterly, 6*(3), 39-62.

Spiess, E., Baumgartner, U., Arn, S., & Vez, C. (2005). *Topographic Maps. Map Graphics and Generalization*. Wabern, Switzerland: Swiss Society of Cartography.

Swisstopo. (2005). *DHM25. Das digitale Hohenmodell der Schweiz. Produktinformation.* Bern, Switzerland: Swisstopo, Bundesamt fur Landestopografie.

Tomko, M., & Purves, R. S. (2008). *Categorical Prominence and the Characteristic Description of Regions* Paper presented at the Semantic Web meets Geospatial Applications, held in conjunction with AGILE 2008, Girona, Spain.

van Rijsbergen, C. J. (1979). *Information Retrieval*: Butterworth.