Title	A holistic semantic similarity measure for viewports in interactive maps
Authors(s)	Ballatore, Andrea, Wilson, David C., Bertolotto, Michela
Publication date	2012-04
Publication information	Ballatore, Andrea, David C. Wilson, and Michela Bertolotto. “A Holistic Semantic Similarity Measure for Viewports in Interactive Maps.” Springer-Verlag, 2012.
Conference details	Web and Wireless Geographical Information Systems International Symposium (W2GIS 2012). April 12-13, 2012, Naples, Italy
Publisher	Springer-Verlag
Item record/more information	http://hdl.handle.net/10197/3740
Publisher's statement	The final publication is available at www.springerlink.com. http://www.springerlink.com/content/f27v127105406431/
Publisher's version (DOI)	10.1007/978-3-642-29247-7_12

A holistic semantic similarity measure for viewports in interactive maps*

Andrea Ballatore,[1] David C. Wilson,[2] and Michela Bertolotto[1]

[1] School of Computer Science and Informatics
University College Dublin, Belfield, Dublin 4, Ireland.
{andrea.ballatore,michela.bertolotto}@ucd.ie
[2] Department of Software and Information Systems
University of North Carolina
9201 University City Boulevard, Charlotte
NC 28223-0001, USA.
davils@uncc.edu

Abstract. In recent years, geographic information has entered the mainstream, deeply altering the pre-existing patterns of its production, distribution, and consumption. Through web mapping, millions of online users utilise spatial data in interactive digital maps. The typical unit of visualisation of geo-data is a viewport, defined as a bi-dimensional image of a map, fixed at a given scale, in a rectangular frame. In a viewport, the user performs analytical tasks, observing individual map features, or drawing high-level judgements about the objects in the viewport as a whole. Current geographic information retrieval (GIR) systems aim at facilitating analytical tasks, and little emphasis is put on the retrieval and indexing of visualised units, i.e. viewports. In this paper we outline a holistic, viewport-based GIR system, offering an alternative approach to feature-based GIR. Such a system indexes viewports, rather than individual map features, extracting descriptors of their high-level, overall semantics in a vector space model. This approach allows for efficient comparison, classification, clustering, and indexing of viewports. A case study describes in detail how our GIR system models viewports representing geographical locations in Ireland. The results indicate advantages and limitations of the viewport-based approach, which allows for a novel exploration of geographic data, using holistic semantics.

Keywords: Geographic Information Retrieval, Viewport, Holistic semantics, Geo-semantics, OpenStreetMap, Vector space model

1 Introduction

The term *neogeography* aptly describes the recent explosion of novel geographic practices, involving the mass production and consumption of geographic information over the Internet [31]. One of the most striking aspects of this nexus

* Research presented in this paper was funded by a Strategic Research Cluster grant (07/SRC/I1168) by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support.

of phenomena is the rapid diffusion of interactive web maps, enabled by enhancements in web technologies in the early 2000s [6]. In parallel, the amount of spatially-referenced information available online has kept increasing at an explosive rate, resulting in spatial information overload.

In order to satisfy the users' spatial information need, the development of effective geographic information retrieval techniques has become a core effort for both academia and industry. The discipline of text information retrieval emerged to find documents matching desired criteria in large collections [17]. Similarly, geographic information retrieval (GIR) aims at identifying relevant geographic objects in vast dataset, indexing locations, toponyms, and minimum bounding rectangles [12,22]. Several efforts have been made to enrich GIR systems with geographic knowledge, linking geographic data to ontologies [10,15,2,14]. However, as Leveling puts it, large-scale evaluations indicate that "more geographic knowledge typically had little or no effect on performance of GIR systems or that it even decreases performance compared to traditional (textual) information retrieval baselines" [12, p.29].

Popular GIR systems, such as Google Maps, Bing Maps, and Yahoo! Maps,[3] focus on the indexing of aspects of the geographic features, to allow efficient retrieval of individual features based on their textual meta-data, such as place name and street address [21]. The user is presented with a rectangular frame often called *viewport*, containing a pre-rendered image of a map that displays features based on their visibility at the current scale. Several actions can be applied on the web map, including text searches, panning - changing the bounding box location - and zooming - changing the map scale level in discrete, predefined steps. Beyond the details of each system, geographic data is most typically visualised and consumed in viewports, rendered at a specific map scale. In a trial-and-error process, users perform actions in order to satisfy their spatial information need [22].

Such web GIR systems can be utilised to examine aspects of objects, for example to observe the structure of a large building or a lake. This type of cognitive activity is generally considered to be an analytical process: complex objects are divided into their constituent parts and mutual relationships, in order to reach the desired piece of information. It has been argued that analytical thinking dominates Western culture [20,19]. This predominance notwithstanding, various forms of holism have emerged in psychology, cognitive science, and geography as an important mode of perception, learning, and thinking [33,25,1]. In particular, the field of landscape ecology strongly claims that landscape, as Antrop and Van Eetvelde put it, "should be considered a complex whole that is more than the sum of its composing parts" [1, p. 43].

In a holistic process, the emphasis is not on individual objects but on the set of objects considered as a unified whole. In GIR, users often judge the *overall semantic content* of a large area to evaluate it against their information need. Holistic judgements on geographic areas are done in several instances. For ex-

[3] http://maps.google.ie, http://www.bing.com/maps, http://maps.yahoo.com (accessed on 24/1/2012)

ample, a user might want to evaluate possible areas when looking for houses on sale. If the user is interested in seaside towns, they want to retrieve areas matching an overall semantic content, such as a small urban settlement located on the coast, with a high density of amenities, beaches, etc, without focusing on specific, individual features. Similarly, a geographer might compare several viewports to study large-scale phenomena affecting the landscape. As in these use cases the focus is on entire viewports rather than on specific features, users can benefit from a holistic semantic query tool.

For all the aforementioned reasons, we think that a computable measure of semantic similarity *between viewports*, taken as holistic units of geographic information, and not between individual features, offers a different approach to GIR. However, this approach does not aim at superseding text-based retrieval, but rather at offering an additional tool that can be integrated with existing GIR methods. In this paper we describe a novel technique to extract holistic semantic descriptions of viewports, and the computation of their similarity, as a foundation of a holistic, viewport-based GIR system.

The remainder of this paper is organised as follows: Section 2 surveys related work in the area of GIR, viewports, and holistic cognition. Section 3 reports the core of the proposed approach, while Section 4 illustrates a case study walk-through in the computation of holistic semantic descriptors in a typical web map. Finally section 5 presents concluding remarks, and outlines directions for future research.

2 Related work

Web mapping is one of driving technologies that brought geographic information into the mainstream, enabling the explosion of neogeography since 2005 [31]. Haklay et al give an account of the recent developments in Internet web mapping, including map mash-ups, crowdsourcing, geostack, and folksonomies, under the umbrella-term 'Web Mapping 2.0' [6]. A major innovation is that map users are not only consumers of geographic information, but also producers of the so-called 'Volunteered Geographic Information' [5].

In such web mapping services, geographic data is distributed through a viewport, a rectangular viewing frames that represent a geographic area at a given scale. Typically, users can zoom and pan, updating the viewport. The concept of viewport is inscribed in a long-standing representational tradition of the *screen*. Manovich traces a compelling genealogy of the screen, seen as a flat, rectangular surface "acting as a window into another space" [16, p. 115].

While interacting with map viewports, users aim at fulfilling their spatial information need. This process is often focused on specific individual map features, decomposing the represented landscape analytically. However, the field of landscape ecology strongly argues that landscape is perceived holistically, as a complex whole. Antrop states that the holistic approach was stimulated by aerial photography, which represents the landscape in its holistic complexity [1]. Naveh, in his broad discussion on landscape ecology and system theory, identifies holism as "perceiving all parts in their full context", and criticises analytical, reductionist approaches "focused on single, isolated parts of the system" [18, p .13]. Moreover, in cognitive science and psychology, holistic cognition is believed to play a major role in perception [29,20].

To be interpreted by humans, the geographic information represented in viewports has to convey some intelligible meaning. The semantics of geographic data has been discussed extensively by Kuhn, who points out the difficulties of grounding meaning in symbolic systems [11]. It is a tautology to state that meaning is crucial in geographic information retrieval (GIR), which aims at identifying relevant features in large datasets [22]. To date, most GIR systems focus on individual map features, with particular emphasis on text-based retrieval [4].

In order to compare, classify, index and cluster geographic objects by their semantics, several analytical approaches have been devised [23,8]. Schwering surveys and classifies the similarity measures for geographic data [26]. In these approaches, the similarity of individual semantic geographic concepts are compared based on their commonalities, differences, positions in taxonomies, and so on. While such models can compute plausible similarities between specific features or feature types, they do not consider the computation of holistic similarity of the map fragments that are, ultimately, displayed to and manipulated by the users in rectangular viewports.

When presented to users, viewports are bi-dimensional images. For this reason, our approach can be seen as analogous to techniques used in image processing systems [30] to compare raster images. However, while such techniques are based on the analysis of low-level image features, such as colour, to compute the similarity of raster viewports, our focus is on the semantics of specific objects. Therefore, our GIR system focuses exclusively on vector data, regardless of the particular visual display of the rendered viewport.

Moreover, viewports show geo-information at a specific map scale. In a typical web map, a viewport is associated with a scale and includes different types of features depending on specific visibility rules. None of the traditional semantic similarity measures discussed above take scale into account, as they focus on abstract psychological classes rather than on viewports. Our system, on the other hand, captures the scale of a viewport by building a semantic descriptor that includes only features that are present (i.e. represented) at the viewport scale – but independently of how they are represented.

To the best of our knowledge, no GIR system focuses on the holistic semantics of map viewports. To explore this concept, the next Section outlines a holistic information retrieval system, based on viewport semantic descriptors.

3 A viewport-based, holistic GIR system

In this Section we detail our proposal of a viewport-based GIR system, by examining the structure of a typical web map, and constructing vector-based semantic descriptors for viewports. In a session in our GIR system, the user can retrieve viewports that are similar to a query viewport, indicated as fulfilling the users'

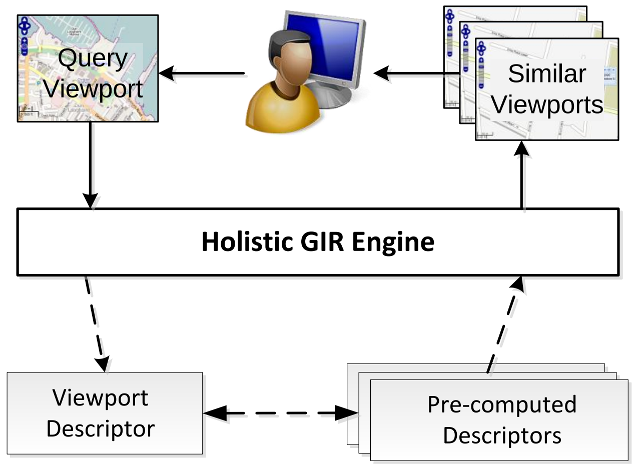

Fig. 1. The architecture of a viewport-based, holistic GIR system. The user submits a query viewport to the system, and the system retrieves the most similar viewports from a set of precomputed semantic descriptors.

spatial information need. The system compares the query viewport with pre-computed viewports, and returns to the user the most similar viewports it has found. This GIR architecture is schematised in Figure 1.

3.1 Viewports

When using GIR systems, users are presented with geographic data displayed in *viewports*, defined as a rectangular, bi-dimensional images rendered on a screen. To capture the semantic content of a viewport, it is useful to start from the visualisation structure of a typical web map. In order to test our approach, we utilised the OpenStreetMap vector dataset, released under a Creative Commons license [7]. As a representative of typical web mapping, we selected the Cloud-Made service, which renders OpenStreetMap data as interactive online maps.[4] As opposed to other commercial geo-services, this service enables exploration of the internal structure of a viewport, and its underlying geographic content.

In the CloudMade maps, the scale can be set to 19 discrete zoom levels, ranging from scale 1:446M (zoom level 0) to 1:1700 (zoom level 18). The map scale is controlled by Equation 1, where y is either the distance in meters or the map scale, and z is the zoom level. The constant C is $78,271$ in the case of meters per pixels, and $223 \cdot 10^6$ in the case of map scale. This equation allows the conversion between map scale, screen pixels, and zoom levels:

$$y = C\ 2^{1-z} \qquad z \in [0, 18] \tag{1}$$

[4] http://maps.cloudmade.com (accessed on 24/1/2012)

Zoom Level	Meters per pixel	Scale	Visible Types	Description
1	78271	$1 : 223M$	2	Region
3	19568	$1 : 55M$	2	–
5	4892	$1 : 14M$	3	Country
7	1123	$1 : 3.5M$	5	–
9	306	$1 : 871K$	10	County
11	76	$1 : 217K$	23	–
13	19	$1 : 54K$	44	Neighbourhood
15	5	$1 : 13K$	64	–
17	1	$1 : 3400$	93	Building

Table 1. Overview of the zoom levels of a CloudMade web map. Total number of feature types: 101.

At each zoom level, the map displays certain types of features, e.g. at the region level, only countries and seas are shown. For the purpose of our study, the geographic dataset has been subdivided into 101 feature types, closely modelled on the visualisation rules of the CloudMadeMap.[5] For example, types include *restaurant*, *stadium*, and *prison*. The visibility of each type per zoom level is defined as a range, e.g. restaurants are visible when $z \in [16, 18]$, while stadiums, being generally larger objects, in range $z \in [14, 18]$. For the sake of clarity, all the notations used in this paper are displayed in Table 2. Intuitively, the number of visible types increases as the scale decreases. The characteristics of each zoom level are summarised in Table 1.

In this context, a viewport $v_{bb,z}$ is defined by a bounding box bb, specified by the latitude/longitude coordinates of its bottom-left and top-right corners, and a zoom level $z \in [0, 18]$ as defined in Table 1. As stated in Section 1, a viewport can be seen as the visualisation unit of geographic data in a web map. A user session on a web map consists of a sequence of manipulative actions on the map, such as panning and zooming, resulting in the visualisation of a sequence of viewports $\{v_1 \dots v_n\}$. In our GIR system, the user can select a viewport by drawing a bounding box on the map, and the selected viewport v_s is used as a query, described in the next Section.

3.2 Holistic viewport descriptors

In order to retrieve semantically similar viewports, our GIR system constructs a holistic semantic descriptor for each viewport v in a vector space model. To extract the overall semantic content from a viewport, the system performs spatial queries on the OpenStreetMap dataset. Given the input viewport v, the system will perform spatial queries to retrieve F_v, all the visible features in that viewport (Equation 2):

$$\forall t \in S_z, \quad q(bb, z, t) \to F_t, \qquad F_v = \{F_{t_1} \dots F_{t_{|S_z|}}\} \tag{2}$$

[5] The visibility rules are defined at http://maps.cloudmade.com/editor (accessed on 24/1/2012)

z	Zoom level $\in [0, 18]$ (see Table 1 for details).
bb	Bounding box, specified by the latitude/longitude coordinates of its bottom-left and top-right corners.
v	A viewport on bounding box bb at zoom level z. h_v and w_v is the viewport height and width in screen pixels.
t	A type of map feature (e.g. *restaurant*, *prison*, etc). In this work 101 types were defined.
$\sigma(z)$	Function mapping the visibility of feature types at zoom level z. $S_z \leftarrow \sigma(z)$.
S_z	Set of visible t at zoom level z. $S_z = \{t_0 \dots t_n\}$, where $\forall t$ is visible at zoom level z.
D_v	Semantic descriptor of viewport v.
$q(bb, z, t)$	Spatial query on bounding box bb, zoom level z, and feature type t. $q(bb, z, t) \rightarrow F_t$
F	Set of all existing map features.
F_t	Set of features of type t, $F_t = \{f_0 \dots f_n\}$
F_v	Set of features visible in viewport v. $F_v = \{F_{t_1} \dots F_{t_n}\}$
$I(t)$	Self-information of type t, assuming a random distribution of types in the map.
$a(f)$	Area of feature f.
V_g	Set of viewports extracted from a geographic area g.

Table 2. Notations

The service can now compute a holistic descriptor D_v, combining all the visible types S_z in a multidimensional vector as in Equation 3, where n is the cardinality $|S_z|$, $t \in S_z$, and w are non-negative normalised weights.

$$D_v = w_1 t_1 + w_2 t_2 + \ldots + w_{n-1} t_{n-1} + w_n t_n, \qquad w \in [0, 1], \ \sum_{i=1}^{n} w_i = 1 \quad (3)$$

In order to characterise D_v, we propose four ways to compute the weights w: linear, logarithmic, information-theoretic, and surface-based approaches.

(a) Linear weights. The simplest approach consists of assigning them proportionally to the cardinality of sets $F_t \in F_v$, using the normalised cardinality:

$$w_i = \frac{|F_{t_i}|}{\sum_{j=1}^{n} |F_{t_j}|} \quad (4)$$

The main limitation of this approach lies in the fact that the statistical distribution of types t is heavily skewed in favour of very frequent features, such as *road*. In a viewport on an urban area, the number of features *road* is often greater than other types by several orders of magnitude, such as *restaurants*, which matches our intuition on the fact that roads are very common map objects, while restaurants are less frequent. For example, it is uncommon to find 2 restaurants, and 200 roads in a viewport. In this case, the weighting function in Equation 4 would assign a very low weight to the type *restaurant*, and an extremely high weight to *secondary*.

(b) Logarithmic weights. A variant that focuses on magnitude rather than number of features is the following, where δ is a positive quantity that prevents the nullification of a term if $|F_{t_i}| = 1$:

$$w_i = \frac{log(|F_{t_i}| + \delta)}{\sum_{j=1}^{n} log(|F_{t_j}| + \delta)}, \qquad \delta = 1 \tag{5}$$

This logarithmic version is less sensitive to small changes in the statistical distribution of types t, and tends to preserve the importance of infrequent features.

(c) Information theoretical weights. A second variant to compute weights w_i taking into account the statistical occurrence of feature types, is based on the information theoretical approach [27]. Given a set of spatial features F, the probability p and the self-information I of feature type t randomly from the dataset are defined as in Equation 6. The self-information of type t can then be used to weight its importance in the vector, by combining it with the number of features:

$$p(t) = \frac{|F_t|}{|F|} \qquad I(t) = -log(p(t)) \qquad w_i = \frac{I(t_i)|F_{t_i}|}{\sum_{j=1}^{n} I(t_j)|F_{t_j}|} \tag{6}$$

In this case, the importance of a type t in the descriptor D_v is increased or reduced depending on its frequency in the dataset. Therefore features of type *secondary* carry low self-information, while features *restaurant* are emphasised. Although this weighting approach intuitively seems the most promising among the three we have presented (Equations 4, 5, and 6), it can assign very high weights to rare features. While in some cases this might be a desirable behaviour (for example to detect landmarks), in general it risks skewing the descriptor towards unusual features, regardless of their actual semantic weight in the viewport.

(d) Area weights. With features modelled as a polygon, it is possible to attribute a weight proportionally to the feature area, on the assumption that large features should have higher semantic importance in the descriptor. Defining the feature area as $a(f)$, and $a(F_t)$ as the sum of all the areas of the features in the set, the surface weights are computed as:

$$w_i = \frac{a(F_{t_i})}{\sum_{j=1}^{n} a(F_{t_j})} \qquad a(F_t) = \sum a(f), \quad \forall f \in F_t \tag{7}$$

In this approach, the feature area is weighted against the area of the other features, and not of the viewport. Thus, the weight can account for overlapping features.

The effectiveness of these four approaches to weight the semantic types in the viewport descriptor, summarised in Table 3, largely depends on the specific

Approach	Key Parameter	Description
(a) Linear	Number of features	When types have different magnitude, large magnitudes take a large section of the descriptor.
(b) Logarithmic	Log of number of features	Represent the magnitude of types. Low sensitivity when types have the same magnitude.
(c) Self-Information	Self-Information of feature type	Common feature have low weight, while unusual feature types have very high weight.
(d) Area	Sum of feature areas	Large features have high weight. Not computable when feature is not a polygon (e.g. for points of interest)

Table 3. Weighting approaches in semantic viewport descriptor D_v

application context. As each approach captures different aspects of the holistic semantics of a viewport, and presents specific limitations, the weights can be computed by averaging different approaches. Without doubt, one of the main advantages of such vector-based viewport semantic descriptors is the wide range of techniques to compare, cluster, and classify them, discussed in the next Section.

3.3 Sampling the Viewport Space

In order to describe the semantics of a web map viewport, we have defined a vector-based descriptor D_v, in four variants (linear, algorithmic, self-information, and surface). Given a web map covering a certain geographic area g, e.g. Ireland, we aim at extracting a number of descriptors that represent its semantics. The viewport space is the set of all possible viewports in g at zoom level z. The area g can be sampled in a number of viewports $V_g = \{v_1 \dots v_n\}$, where n is the desired number of viewports. The theoretical number of viewports that can be extracted in a geographic area delimited by a bounding box bb_g at zoom level z, where h_g, w_g are the height and width of the geographic area converted into pixels with Equation 1. h_v and w_v are the height and width of the viewport in pixels:

$$|V_g| = (h_g - h_v)(w_g - w_v) \tag{8}$$

For example, a geographic area g delimited by a bounding box of size $\approx 300 \times 230\ km^2$ corresponds at $z = 9$ (county level) to a screen of 4096×3072 pixels. Sampling g with 1024×768 pixel viewports, a common resolution for web maps, the possible viewports amount to ≈ 7.6 million. With higher zoom levels, the number of possible viewports increase rapidly following the power law in Equation 1. It is therefore evident that a sampling technique has to be utilised to extract a computable number of viewports, in particular for high zoom levels.

The most intuitive way of choosing viewports is based on user interests. Viewports in which user activity is performed are automatically included in the sample. However, to overcome the cold start problem that arises in this case, a general sampling mechanism is necessary to index g. A possible technique is that of random sampling, based on the law of large numbers. Even though it is difficult to compute all the possible viewports, it is possible to extract a sufficient number of random viewports to represent accurately the whole set of viewports, setting the sample size to a limit β, as shown in Equation 9. In order to determine β, a confidence level and a confidence interval have to be chosen.

$$|V_g|_\beta = \frac{\beta}{|V_g|}(h_g - h_v)(w_g - w_v) \qquad 0 < \beta < |V_g| \tag{9}$$

Similarly, it is possible to sample g by defining an arbitrary grid γ of pixels h_γ and w_γ, which reduces the number of viewports as follows:

$$|V_g|_\gamma = \frac{(h_g - h_v)(w_g - w_v)}{h_\gamma w_\gamma} \qquad 0 < h_\gamma < h_g, \;\; 0 < w_\gamma < w_g \tag{10}$$

A third possibility is a combination of grid and random sampling. The geographic area g is divided into an arbitrary number of grid cells, and each cell is sampled randomly. The choice of the sampling technique (random, grid-based or both) has to be done on an empirical basis, depending on the specific application context. Once a sample V_g has been selected, the corresponding descriptors D_v can be computed. Subsequently, the system can compare, cluster, and retrieve viewports (see Figure 1). The next Section describes comparison techniques for the semantic descriptors.

3.4 Comparing Viewport Descriptors

A geographic area g can sampled as a set of viewports V_g, with the techniques described in the previous Section. The corresponding descriptors D_v can then be pre-computed through one of the approaches presented in Section 3.2. The similarity of two viewports is therefore the similarity of their descriptors (Equation 11).

$$s(v_a, v_b) = s(D_{v_a}, D_{v_b}) \qquad s(v_a, v_b) = s(v_b, v_a) \qquad 0 \le s(v_a, v_b) \le 1 \tag{11}$$

Given that these descriptors are multidimensional vectors, encoding semantic aspects of the viewports, it is possible to compare them with well-known techniques in a vector space [24,9,32]. In the viewport semantic vector space, every viewport can be modelled as a row in a multidimensional matrix $|V_g| \times |t|$, having a column for each feature type t. Vector similarity is traditionally computed using linear algebra techniques, such as the Euclidean, cosine, Chebyshev, and Manhattan distances [3,13].

In a semantic information retrieval system in which the user submits a viewport v_q as a query, the most similar k viewports must be retrieved and displayed to them. To do so, these distance measures can be efficiently computed between the descriptor of the query viewport D_{v_q} and all the pre-computed descriptors D_v, where $v \in V_g$. Additionally, the similarity computation can be constrained in several ways to match specific user information needs. Among others, common constraints are the maximum or minimum distance from the query viewport v_q, a zoom level range ($z \in [z_{min}, z_{max}]$), and a weight constraint on a feature type ($w_{min} < w_t < w_{max}$).

The next Section presents a case study, in which the descriptors are used to capture the holistic semantics of viewports taken from the Dublin area in Ireland, and are used to retrieve semantically similar viewports.

4 A case study walkthrough

To capture a holistic impression of semantics in a web map, we have defined semantic descriptors as vectors D_v that represent the overall semantic content of the viewport v in which the map is displayed. This Section illustrates an interaction with our viewport information retrieval system (see Figure 1), suggesting possible applications, strengths and weaknesses of the approach.

We noted certain tasks that users perform on web maps are not only analytical, i.e. focused on the decomposition of large objects into simpler parts, but are also holistic, treating a geographic area as a unified entity. Analytical tasks involve the examination of specific target objects, and so on. On the other hands, examples of spatial holistic tasks are the classification of an urban area versus a rural area, in which the user needs not to focus on individual objects, but classify the area as a whole. These holistic tasks should not be considered in opposition to analytical tasks, but they are intertwined in the complex cognitive interplay that occur in the interaction with information retrieval systems.

In a holistic geographic information retrieval, the user can retrieve, instead of specific geographic features, viewports that represent visually geographic areas rendered at a given zoom level. In this case, the user's information need is not a specific spatial information, e.g. where is the target object, or what is the area of the target object, etc, but is an implicit semantic judgement on viewports displayed on the screen. A viewport representing a geographic area that fulfills the user's information need, e.g. a seaside town or a commercial port, is used as a query viewport to retrieve viewports conveying similar semantic content. To achieve this, the system has to be able to model this *implicit judgement* on geographic content displayed in a viewport v.

Sample viewports. In this case study we consider a small set of viewports selected from a bounding box bb, corresponding to the surroundings of Dublin. This geographic area g contains a total of $\approx 20,000$ features. Five sample viewports were extracted at zoom level $z = 15$, including a Dublin suburb, a park, a port, and two seaside towns. Five semantic descriptors D_v were then computed for each of the six viewports v, linear, logarithmic, information-theoretic, surface, and a mean of the first four, limiting for the sake of illustration the number of feature type to 10 out of the 64 visible types. The self-information of each

Viewport	Feat Type	Linear	Log	Area	Self-Info	Mean
v_1	building	.495	.321	.162	.388	.341
	coastline	–	–	–	–	–
Milltown: suburb of Dublin, with hospital, college, and residential estates.	commercial	.029	.113	.034	.058	.058
	hospital	.01	.056	.111	.021	.05
	industrial	–	–	–	–	–
	park	.01	.056	.043	.025	.033
	port	–	–	–	–	–
	road	.427	.309	–	.462	.3
	town	.01	.056	.59	.021	.169
	wood	.019	.089	.06	.025	.048
v_2	building	.229	.272	.035	.156	.173
	coastline	–	–	–	–	–
Phoenix Park: Large urban park with zoo, polo grounds, and American embassy.	commercial	–	–	–	–	–
	hospital	–	–	–	–	–
	industrial	–	–	–	–	–
	park	.029	.086	.789	.064	.242
	port	–	–	–	–	–
	road	.257	.285	–	.242	.196
	town	–	–	–	–	–
	wood	.486	.358	.175	.539	.389
v_3	building	.087	.16	.089	.047	.096
	coastline	.442	.268	–	.542	.313
	commercial	–	–	–	–	–
Howth: seaside town with tourist attractions, cliffs, and trekking trails.	hospital	–	–	–	–	–
	industrial	–	–	–	–	–
	park	.029	.096	.276	.052	.113
	port	–	–	–	–	–
	road	.308	.243	–	.233	.196
	town	.01	.048	.309	.015	.095
	wood	.125	.184	.325	.111	.186
v_4	building	.296	.207	.158	.175	.209
	coastline	.194	.183	–	.257	.158
Dun Laoghaire: seaside town with a small port, a private school, and a hospital.	commercial	.143	.165	.175	.214	.174
	hospital	.01	.042	.088	.017	.039
	industrial	.02	.067	.026	.028	.035
	park	.01	.042	.018	.02	.022
	port	.01	.042	.184	.021	.064
	road	.306	.209	–	.251	.191
	town	.01	.042	.351	.017	.105
	wood	–	–	–	–	–
v_5	building	.17	.19	.12	.08	.14
	coastline	.136	.176	–	.144	.114
Dublin Port: docks of the Dublin port, where large ships load and unload containers.	commercial	.386	.244	.326	.461	.354
	hospital	–	–	–	–	–
	industrial	.227	.209	.174	.251	.215
	park	–	–	–	–	–
	port	.011	.048	.38	.019	.115
	road	.068	.133	–	.044	.062
	town	–	–	–	–	–
	wood	–	–	–	–	–

Table 4. Viewport semantic descriptors D_v for 5 sample viewports, with weights computed using four approaches, and their mean. Symbol '–' corresponds to 0.

Viewport Name	$\star$	Cosine v_1	v_2	v_3	v_4	v_5	Euclidean v_1	v_2	v_3	v_4	v_5
Milltown	v_1	–	–	–	–	–	–	–	–	–	–
Phoenix Park	v_2	.55	–	–	–	–	.52	–	–	–	–
Howth	v_3	.54	.65	–	–	–	.55	.59	–	–	–
Dun Laoghaire	v_4	.82	.38	.68	–	–	.72	.48	.66	–	–
Dublin Port	v_5	.37	.15	.29	.73	–	.46	.35	.45	.68	–

Table 5. Similarity of sample viewports. The matrices are symmetrical, and their diagonal values are equal to 1. The euclidean similarity has been computed as $1 - d$, where d is the euclidean distance.

feature type was computed on g, and not on the entire OpenStreetMap dataset. The viewports and corresponding descriptors are reported in Table 4.

Looking at the weights of the descriptors, it is possible to trace behaviour, pros and cons of each of the four weighting mechanism proposed in Section 3.2. The linear weights reflect the number of features visible in the viewport. For this reason, important feature such as the Phoenix Park in v_2 rank very low, and roads rank very high in most viewports. This is because roads are represented in numerous small chunks, while large objects such as a park consist of one large polygon, and the linear weights do not take this aspect into account.

This problem is partly addressed by the logarithmic weights, which smooth the results by increasing the importance of types with few features and by decreasing that of types with many occurrences in the viewport. In v_3, the logarithmic weight of *building* has been doubled, while that of *coastline*, another type of feature modelled in small chunks, has been reduced. However, despite the smoothing, the resulting weights are still strongly biased towards types such as *road*, and *building*, while large and infrequent features are squeezed into small weights.

The area-based technique tends to correct this bias. In v_2, the type *park*, which is almost ignored by the previous weighting mechanisms, is the most important in the viewport. Thanks to its large area, this type gains a lot of influence in the descriptor. As it is possible to notice by the frequent 0 values in the area column, only a subset of features are polygons and can be included in this descriptor, resulting in a limited information problem.

The behaviour of the self-information weights is more difficult to interpret. For types that occur very frequently in g, such as *road* and *building*, $I(t)$ is low (3.51 and 4.85), while less frequent types have higher $I(t)$ (12.21 for *port*, and 11.56 for *park*). This is correct, but when the self-information values are multiplied by the number of features, they smooth the results to a limited extent. In the case of viewport v_2, according to the self information weights, buildings are more important than parks, maintaining the bias of the linear and logarithmic weights. As it is expected, the mean weights are heavily smoothed, but maintain some of the bias of the linear, logarithmic, and self-information weights.

Viewport similarity. The semantic similarity of the viewports can be computed via the cosine distance between their descriptors, as shown in Table 5. In this case, the two vector similarity measures rank the pairs in the same way. The pairs with highest similarities are $\langle v_1, v_4 \rangle$, and $\langle v_4, v_5 \rangle$. Viewports v_1 and v_4 are semantically very similar, and this result is satisfactory. The pair $\langle v_4, v_5 \rangle$ is surprising, because a tourist seaside town appears very different to a commercial port. This result is easily explained with the omission of feature types that would increase the distance between the two viewports, such as restaurants, tourist attractions, amenities, which are strongly present in v_4 but not in v_5. The two viewports share the fact of including the seaside, the presence of a port, many buildings and commercial activities, all aspects that are captured by the 10 feature types considered in this case study.

On the other hand, the least similar pairs are $\langle v_2, v_5 \rangle$, and $\langle v_3, v_5 \rangle$. This seems to be a valid result, as these pairs represent very different areas, sharing very few feature types. Based on this case study, it can be concluded that the holistic semantic descriptors proposed in this paper are a promising approach to compute semantic similarity of viewports.

5 Conclusions and future work

In this paper, we have proposed a technique to extract semantic descriptors for viewports, which can be used in a viewport-based, holistic GIR system (see Section 3). Instead of focusing on specific geographic features as in traditional GIR systems, our system aims at capturing the overall semantic content in a viewport. Thus, viewports are treated in a manner similar to documents in text-based IR. Based on the work presented in this paper, the following conclusions can be drawn:

- The vector-based semantic descriptors capture the overall semantic content of a viewport in a holistic mode, without focusing on specific individual features. The user retrieves viewports that present similar characteristics to the query viewport that is submitted to the system.
- The viewports display a map at a given zoom level. The corresponding descriptor captures the semantic content displayed at the specific zoom level. This enables the semantic analysis of the viewports, taking the map scale into account.
- Four weighting mechanisms are proposed to extract semantic content from a viewport. Such weights can be combined to achieve different representation of the same viewport. Being based on the well-known vector space model, our approach can benefit from a wide range of techniques to index, compare, classify, and cluster large sets of vectors. Descriptors can be used to perform collaborative filtering, and user profiling.
- Our holistic GIR system offers an additional technique to retrieve relevant geographic information from a large dataset. It is not conceived as antagonistic to traditional GIR, but rather as a complementary approach that can be combined with existing analytical techniques to enable retrieval through holistic semantics.

The case study outlined in Section 4 indicates that the system is able to capture the holistic semantic of a viewport. However, further work is necessary to assess its accuracy and recall on a big spatial dataset, in the context of realistic information needs. The limitations of the presented approach can be overcome by incorporating more sophisticated semantic techniques for vector space models, such as latent semantic analysis [32,28]. Besides, other holistic metrics be added to the descriptors, such as heterogeneity, entropy, and fractal dimension [1].

The holistic GIR system presented in this paper provides a different approach to traditional geographic information retrieval, modelling the overall semantic content of a viewport in a vector space model. Treating viewports as documents enables the exploration of digital maps from a holistic perspective, stressing the need for reconsidering the undisputed centrality of analytic approaches.

References

1. M. Antrop and V. Van Eetvelde. Holistic aspects of suburban landscapes: visual image interpretation and landscape metrics. *Landscape and urban planning*, 50(1-3):43–58, 2000.
2. A. Ballatore and M. Bertolotto. Semantically Enriching VGI in Support of Implicit Feedback Analysis. In *Proceedings of the Web and Wireless Geographical Information Systems International Symposium (W2GIS 2011)*, volume 6574/2011 of *Lecture Notes in Computer Science*, pages 78–93. Springer, 2011.
3. M.W. Berry, Z. Drmac, and E.R. Jessup. Matrices, vector spaces, and information retrieval. *SIAM review*, pages 335–362, 1999.
4. F. Gey, R. Larson, M. Sanderson, H. Joho, P. Clough, and V. Petras. Geoclef: the clef 2005 cross-language geographic information retrieval track overview. In *Accessing Multilingual Information Repositories: 6th Workshop of the Cross-Language Evalution Forum, CLEF 2005, Vienna, Austria, 21-23 September, 2005.*, number 4022, pages 908–919. Springer, 2006.
5. M.F. Goodchild. Citizens as Sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.
6. M. Haklay, A. Singleton, and C. Parker. Web Mapping 2.0: The Neogeography of the GeoWeb. *Geography Compass*, 2(6):2011–2039, 2008.
7. M. Haklay and P. Weber. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing*, 7(4):12–18, 2008.
8. K. Janowicz, C. Ke, I. Panov, M. Wilkes, M. Espeter, and M. Schwarz. A study on the cognitive plausibility of SIM-DL similarity rankings for geographic feature types. In Sara Fabrikant and Monica Wachowicz, editors, *The European Information Society*, pages 115–134. Springer, 2008.
9. L. Jing, M.K. Ng, and J.Z. Huang. Knowledge-based vector space model for text clustering. *Knowledge and Information Systems*, pages 1–21, 2010.
10. C. Jones, H. Alani, and D. Tudhope. Geographical Information Retrieval with Ontologies of Place. In *Spatial Information Theory*, volume 2205/2001 of *Lecture Notes in Computer Science*, pages 322–335. Springer, 2001.
11. W. Kuhn. Geospatial Semantics: Why, of What, and How? In *Journal of Data Semantics III*, volume 3534 of *Lecture Notes in Computer Science*, pages 1–24. Springer, 2005.

12. J. Leveling. Challenges for Indexing in GIR. *SIGSPATIAL Special*, 3(2):29–32, 2011.
13. H. Li, R. Shi, W. Chen, and I.F. Shen. Image tangent space for image retrieval. *Pattern Recognition*, 2:1126–1130, 2006.
14. Wei Liu, Hehe Gu, Chunmin Peng, and Dayu Cheng. Ontology-based retrieval of geographic information. In *2010 18th International Conference on Geoinformatics*, pages 1–6, june 2010.
15. M. Lutz and E. Klien. Ontology-based retrieval of geographic information. *International Journal of Geographical Information Science*, 20(3):233–260, 2006.
16. L. Manovich. *The Language of New Media*. MIT Press, Cambridge, Massachusetts, 2001.
17. C.T. Meadow, B.R. Boyce, and D.H. Kraft. *Text information retrieval systems*. Academic Press, 2007.
18. Z. Naveh. What is holistic landscape ecology? A conceptual introduction. *Landscape and Urban Planning*, 50(1-3):7–26, 2000.
19. R.E. Nisbett and Y. Miyamoto. The influence of culture: holistic versus analytic perception. *Trends in Cognitive Sciences*, 9(10):467–473, 2005.
20. R.E. Nisbett, K. Peng, I. Choi, and A. Norenzayan. Culture and systems of thought: Holistic versus analytic cognition. *Psychological review*, 108(2):291, 2001.
21. A.M. Nivala, S. Brewster, and L.T. Sarjakoski. Usability Evaluation of Web Mapping Sites. *The Cartographic Journal*, 45(2):129–138, 2008.
22. R. Purves and C. Jones. Geographic information retrieval. *SIGSPATIAL Special*, 3(2):2–4, 2011.
23. M.A. Rodríguez and M. Egenhofer. Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. *International Journal of Geographical Information Science*, 18(3):229–256, 2004.
24. G. Salton, A. Wong, and C.S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.
25. G. Schwarzer, S. Huber, and T. Dümmler. Gaze behavior in analytical and holistic face processing. *Memory & Cognition*, 33(2):344–354, 2005.
26. A. Schwering. Approaches to Semantic Similarity Measurement for Geo-Spatial Data: A Survey. *Transactions in GIS*, 12(1):5–29, 2008.
27. C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, 1948.
28. S. Sizov. GeoFolk: latent spatial semantics in web 2.0 social media. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 281–290. ACM, 2010.
29. J.D. Smith and J.H. Shapiro. The occurrence of holistic categorization. *Journal of Memory and Language*, 28(4):386–399, 1989.
30. H. Tan, P. Yu, X. Li, and Y. Yang. Digital image similarity metrics and their performances. In *Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pages 3922–3925. IEEE, 2011.
31. A. Turner. *Introduction to Neogeography*. O'Reilly Media, Inc., Sebastopol, CA, USA, 2006.
32. P.D. Turney, P. Pantel, et al. From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1):141–188, 2010.
33. T.B. Ward and J. Scott. Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition*, 15(1):42–54, 1987.