A semantic and language-based model of landscape scenes

Jean-Marie Le Yaouanc, Éric Saux, Christophe Claramunt

Naval Academy Research Institute,
BP 600, 29240, Brest Naval, France
{leyaouanc, saux, claramunt}@ecole-navale.fr

Abstract. The modelling of landscape environment is a cognitive activity, that still requires novel kinds of spatial representations. This paper introduces a structural categorisation of a landscape view based on panoramic photographs that act a substitute of a given natural environment. Verbal descriptions of a landscape scene provide the modelling input of our approach. The structural-based representation identifies the spatial, relational and semantic constructs that emerge from these descriptions. A landscape view is modelled according to the structure of its language description. Concepts in the environment are qualified according to an ontological classification, their proximity to the observer, and the relations that qualify them. The resulting model is schematised by a music score-based representation that constitutes a modelling support for the study of environmental descriptions.

Key words: Landscape perception, spatial cognition, verbal descriptions.

1 Introduction

Early developments of Geographical Information Systems have been widely influenced by quantitative representations of space. However, it has been recognised that quantitative models do not reflect the way humans perceive, conceptualise and describe their environment [1]. Over the past few years, novel approaches based on spatial cognition and qualitative perceptions of space have emerged as alternative solutions. In particular, the development of wayfinding and landscape representations have emphasised the potential and role of verbal descriptions in the modelling of large-scale spaces, including the way landmarks and spatial relations are formalised [2]. It has also been long recognised that natural language favours the understanding on how people conceptualise space [3].

The research presented in this paper addresses the conceptualisation and the formalisation of a panoramic natural landscape perceived by an observer, and its materialisation by a verbal description. Our study considers a verbal description of an environmental scene as the initial context of the modelling approach whose objective is to categorise the constructs and semantics exhibited. These environmental scenes consist of 360 ° panoramas materialised by photograph montages

presented to the observer with a dynamic interface. These photographs acts as a substitute of a given natural environment. The verbal descriptions considered are extracted from a panel of observers who are given the task of depicting these views in a way that allows recognition of the scene by an external addressee. The study focuses on the identification and categorisation of the spatial and structural constructs that emerge from these scenic perceptions. The modelling purpose of the research is to formalise the way a panoramic scenic environment is qualified by a verbal description. The forms that emerge from our modelling approach should help identify the structure, rhythm, patterns and properties of a scenic description. The schematisation results in a two-dimensional structural representation, based on the notions of proximity spaces, and a linear ordering of the verbal descriptions.

The remainder of the paper is organised as follows. Section 2 briefly introduces related work on the cognitive representations of natural landscapes. Section 3 presents our experimental study, where a panel of observers are given the task of producing verbal descriptions. Sections 4 and 5 provide a conceptualisation of these verbal descriptions. Finally, section 6 draws the conclusions and outlines further work.

2 Landscape perception

It has been shown that many important features in landscape descriptions cannot be observed by conventional cartographical representations. Photographs have been recognised as substitutes for *in-situ* landscape observations as they facilitate the acquisition of verbal descriptions [4]. Although photographs have limited capabilities for the identification of small or detailed elements due to the absence of depth, they provide a valuable support for a semantic interpretation of descriptive observations.

Cognitive studies usually distinguish two levels in the mechanisms and abstractions associated to mental interpretation. On the one hand, the internal level, focussing on the neuropsychological process, which generates a mental representation. On the other hand, the external level which reproduces the structural organisation of the observed phenomenon, for instance at the language level. Our research concerns the latter. As it has been shown in early studies, an observer acting in a given environment, organises space according to proximities and senses [5, 6]. The "proximate environment" is defined as the area perceived by the senses, while the "landscape area" is perceived by sight alone. The landscape area extends to the horizon separating the earth from the sky. The notion of mental spaces has been introduced and "schematised by eliminating detail and simplifying features around a framework consisting of elements and the relations among them" [7]. Tversky makes the difference between "the space of the body", "the space around the body", "the navigational space" and "the space of graphics" [8]. Similarly, Montello categorises the "figurative", "vista", "environmental" and "geographical" space [9]. These cognitive studies clearly show

that the perception of environment is organised at different levels, according to the perception and the behaviour of the observer.

Studies on the perception of distances show evidence that proximity to concepts composing an environment is particularly influenced by salient features. Concepts act as references when they are salient and permanent in the environment [6]. The term *concept* includes regions, parcels of land and water bodies, structural features such as mountains, hills or valleys, and human-made features such as roads or buildings. Although they are not always precisely located in the environment, concepts often serve as frames of reference. This is largely due to the nature of concept boundaries and the distinction made between "fiat boundaries", *i.e.*, the one defined by human demarcation, and "bona fide boundaries", *i.e.* the ones that refer to physical discontinuities, such as a mountain, valley or meadow [10].

3 Experimentation

This study has been carried out in the context of a semi-natural landscape located in the Alpine region of France. The experimentation was conducted using a panel of 22 participants, 17 males and 5 females, mainly non-experts in spatial cognition studies. After a quick overview of the panoramic photograph using a dynamic interface[1], they were asked to spontaneously describe the photograph for an external addressee whose goal was to recognise the scene. Verbal descriptions were recorded using a data storage device and were not to exceed 5 minutes.

Fig. 1. Experimental panorama of a semi-natural environment

The 360 ° panoramic images were displayed on a computer screen. The photographs were presented interactively and non interactively. When the user chose a non-interactive method, the panorama was displayed with a constant rotational speed. The panoramic view was also presented in an interactive manner: the observer was able to explore the scene by rotating the view, just as he would have done it by rotating his body in a natural environment. Photographs were displayed in a large size in order to limit the restriction due to the difference of size between an environment perceived *in situ*, and the one perceived through a photograph.

An example of description that emerges from the landscape presented in figure 1 is as follows: *"I'm in a mountainous region. In front of me, there is a house with a little garden. 500 meters behind the house, a footpath crosses a*

[1] The experimental setup is available at http://experimentation.yaou.org/

large meadow. Far away, behind the meadow, I can see a huge town in a valley. On the horizon, there is a mountain range. Behind me, there is a large road that runs around a hill and a footpath winding around it. There is a little cottage along the road, and crossroads on the right of the cottage."

The verbal descriptions generated by the panellists have led the following results. Most of the participants describe the panoramic photograph from left to right, taking a reference point to begin their description. Most verbal descriptions are implicitly organised with a hierarchy. 60 % of the participants first describe the scene as a whole with sentences like "I am in a mountainous region". The salient concepts structuring the scene are first mentioned but not localised. Next, participants describe the content of the landscape, and the way these concepts are related to each others. This shows evidence of a hierarchical perception of space where the landscape is first perceived and described as a whole, before being detailed. This confirms previous studies and evidence of hierarchies in the perception of spatial information [11]. When observing the environment, humans perceive distinct concepts that are part of the landscape. Most descriptions attempt to identify this variety of concepts which are directly related to the cultural and ontological background, and common sense of the observer. The scene description is largely based on salient concepts, as their role is prominent in the structure and organisation of the environment [12]. Concepts are qualitatively associated to others using spatial relations, preferably with landmarks as frames of reference. With respect to the considered environment, participants used man-made (50% of the identified entities), relief (30%) and vegetation (15%) concepts, where landmarks are assimilated to punctual entities, and forms to extended entities. The salience and legibility of these concepts largely influence their role in the perception process.

The observers interpret photographs of an environmental scene by using spatial relations to qualify the position of the concepts such as "behind the mountain", "in front of the house", "in the background", "in the foreground", "in the long distance", proximity adjectives such as "near", "close to", "far from", "further", directional relations such as "to the right of" and a few tridimensional constructs such as "above" or "below". These concepts refer to "image schematas" used in the perception of large-scale spaces [13]. These terms are either egocentric constructs and vary depending on the observer's location, or relative constructs associated to the location of another concept identified in the landscape. This shows the complementary roles played by egocentric and relative frames of reference in scene perception. It also appears that the roles played by the frames of reference selected by the observer, are determined by their contribution to the hierarchical organisation of the environment. This confirms Tversky's intuition on the organisation of environment that impacts the structure of a verbal description and precedes its linear structure [14]. The fact that different frames of reference are used within a description may also reflect the fact that speakers tend to use a greater variety of terms and rhythms which, in turn, generate the richness of the description. The location of these concepts depends on their proximity to the observer. Proximate spaces which emerge from

these descriptions, vary according to the terms used and the spatial extent of the concepts.

The experiment has also highlighted the prominent role played by direction relations (50%) which are twice as often used as proximity (30%) and topological constructs (20%). This stresses the fact that direction relations are among the most appropriate in structuring a panoramic view. It is also worth noting the low role played by quantitative measures in the descriptions although two participants used metric data in their description to specify a proximity. Moreover, these quantitative measures are associated to a fuzzy linguistic term such as "about 500 meters from my position" or "around 400 meters from the house".

4 Modelling approach

We define an *environmental scene* as the 360 ° environment perceived and described by an observer from a static point of view. An environmental scene is composed of a combination of concepts associated by using spatial relations. These concepts are perceived as forms in the landscape *i.e.*, mountains, hills, vegetal and cultivated areas, or as landmarks, *i.e.* buildings and all other man-made features, salient in the environment. A landscape scene is qualitatively divided into proximity spaces which are defined according to their distance to the observer. These proximity spaces correspond to the ones visually accessible to the observer. Our classification also considers the space on the horizon in order to mark the boundary between the environment accessible to the observer, and the geographical space beyond the line of sight. These proximity spaces are defined as follows:

- The *space around the body* as introduced by Tversky [8] that is described by terms such as "near me".
- The *experienced space* that can be easily experienced through locomotion. It is described by terms such as "not so far".
- The *distant space* is the environment, between the experienced space and the space on the horizon. It is is described by terms such as "far away".
- The *space on the horizon* is the scene in the far distance, and made of silhouettes that constitute landform boundaries. It is described by terms such as "in the background" or "on the horizon".

A verbal description is made of a corpus of sentences. A sentence contains several concepts associated using relations that create the structure of the scene. We characterise a sentence by the concepts, *i.e.*, forms and landmarks, modelled as entities, and the relations that relate them. More formally, a verbal description D is formalised as an ordered set of sentences $s_i \in \mathbb{S}$, the set of sentences, *i.e.*, $D = [s_1, s_2, \ldots, s_n]$. A sentence is an ordered set of entities and relations where every relation has an entity as a predecessor, and an entity as a successor. Let $\mathbb{E}$ be the set of entities and $\mathbb{R}$ the set of relations, a sentence s_i is defined as follows:

– $s_i = [t_1, t_2, \ldots, t_n]$ with $t_1, t_2, \ldots, t_n \in \mathbb{E} \cup \mathbb{R}$ such as
– $\forall t_j \in s_i / t_j \in \mathbb{R}$ then $\exists t_{j-1}, t_{j+1} \in s_i / t_{j-1}, t_{j+1} \in \mathbb{E}$

Let us consider the sentence "500 meters behind the house, a footpath crosses a large meadow" taken from the example introduced in section 3. This sentence is formalised as $s_i = [t_1 =$"500 meters", $t_2 =$"behind", $t_3 =$"house", $t_4 =$"footpath", $t_5 =$"crosses", $t_6 =$"large meadow"]. Entities and relations are instanced, labelled by attributes and associated to a semantics (figure 2).

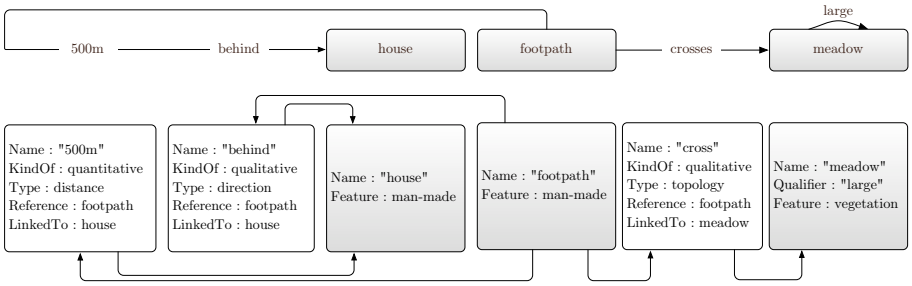


Fig. 2. Sentence description principle

5 Schematisation of a verbal description

This modelling approach constitutes a preliminary logical and structural representation of the verbal description of a landscape scene. The next objective is its schematisation in order to facilitate the understanding of the concepts and structures that appear in the description. The main notions that play an important role are the:

– hierarchical and linguistic properties that emerge from the sentences,
– proximity spaces,
– concepts, *i.e.*, forms and landmarks,
– relations, *i.e.*, direction, proximity and three-dimensional constructs,
– ontology references that classify the identified forms and landmarks.

We consider that a verbal description is ordered by two orthogonal dimensions: the boundaries and rhythm given by the sentences, and the proximity spaces. This leads us to retain a graphical representation designed by analogy to scores used in music, and where notes rhythm the time line, and where the pitch and duration characterise every note. Within our representation, the rhythm is given by the sentences, forms, landmarks and relations. Proximity spaces model the pitch of every concept in the landscape, its extent and its duration on the music score-based notation (figure 3). The following subsections illustrate the instantiation of the ontological and music score-based schematisation.

5.1 Ontological view

The ontological view is a semantic schematisation linked to the properties of the forms and landmarks that appear in the verbal description. Concepts and relations are represented by their linguistic terms associated to a semantic derived from the ontology of foot-orienteering. This ontological view combines a conceptual diagram with a proximity-based structural representation, and a phrasing structure that reproduces the rhythm of the verbal description (figure 3). This example shows that different frames of reference, located in different proximity spaces, are used in the scenic description (e.g, the observer, the house, the meadow, the valley and the road).

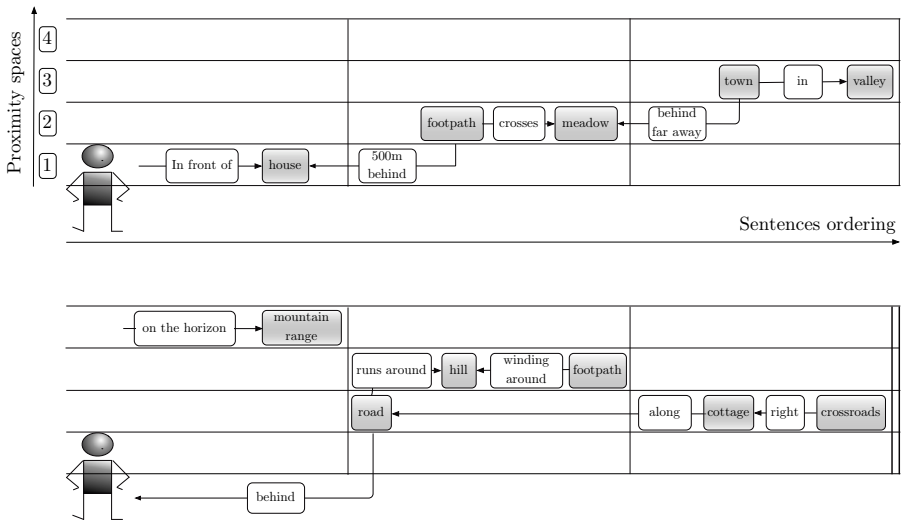


Fig. 3. Ontological view of a verbal description with the proximity spaces. 1. is the space around the body, 2. is the experienced space, 3. is the distant space and 4. is the space on the horizon

5.2 Music score-based view

The music score-based view is an interpretation of the ontological view where concepts are preferably represented by their extent in space, by similarity to the way notes are time-stamped on a music score (figures 5 and 6). With respect to the different levels of proximity spaces, concept symbols are appropriately located and indicate their distance to the observer. Note values show the spatial extent of the concepts rather than the relative duration as it is the case with a musical notation. For example, a one dimensional concept is represented by a sixteenth note as it is punctual in space. Similarly, the half note refers to a valley and the whole note to a chain of mountains. As the boundaries of proximity spaces are not always well-defined, the concepts identified are likely to extend over several proximity spaces.

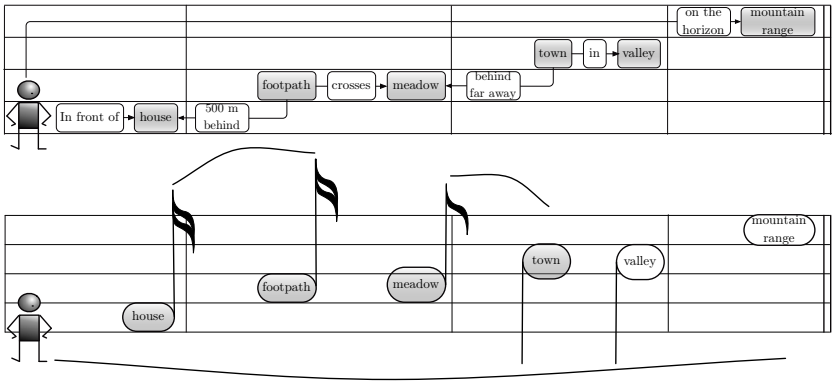
Not surprisingly, the analogy with a music score is not complete. With a musical notation, a bar is a segment of time defined as a given number of beats of a given duration. This constraint is not respected by the music score-based view as sentences do not contain a regular quantity of information. This is due to the fact that the number of landmarks and forms contained in a two-dimensional space is practically not constraint, as it is in a one-dimensional space. The rhythm that emerges from this representation reflects the richness of the sentences, and overall, of the verbal description. Figures 4, 5, and 6 provide two scenes, with examples of ontological and music score-based views derived from a linguistic description produced by an observer. It appears that this observer first describes nearby, and then distant features. We also observe that the rhythm diversity, and the number of sentences exhibited by scene 1 is higher than the one of scene 2. This reflects the fact that the observer has perceived scene 1 as semantically richer than scene 2.

(a) Scene 1 (b) Scene 2

Fig. 4. Studied environmental scenes

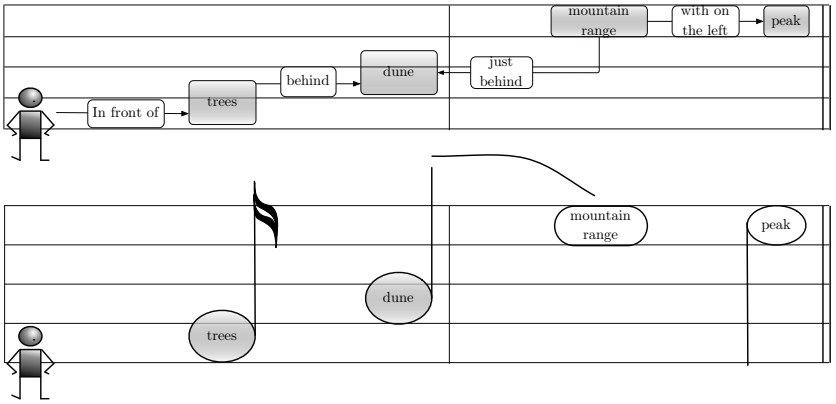
The ontological and music score-based representations can be combined in order to provide a global schematisation of an environmental scene. The approach provides a mean for the analysis of differences and similarities in the linguistic descriptions of landscape scenes.

The formalism is adapted to the identification of the rhythm given by the structure of the sentences, and the depth of the landscape as reflected by the properties of identified concepts. The formalism and its graphical expression can be used for analysing similarities and differences across either a same or different landscape scenes. This gives a methodological support for ethnophysiography studies where the objective is to characterise linguistic and cultural patterns used in landscape perception [15]. It also provides a preliminary resource for the development of a mapping between a verbal description and a computational representation. This formalism constitutes a complementary representation to conceptual and spatial models oriented to the semantics of geographical data.



In front of me, there is an house. 500 meters behind the house, a footpath crosses a large meadow. Far away, behind the meadow, I can see a town in a valley. On the horizon, there is a mountain range.

Fig. 5. Ontological and music score-based views - scene 1



In front of me, I can see a few trees and an huge dune behind them. Just behind the dune, there is a mountain range with a peak on the left.

Fig. 6. Ontological and music score-based views - scene 2

6 Conclusion

This paper introduces a language-based and cognitive approach that models a verbal description of a landscape scene. A verbal description is modelled by a structural and proximity-based representation that reflects its semantics. The structural properties of a verbal description are linked to a music score-based schematisation, where sentences are characterised by staves, and proximity spaces by their distance from the observer. Ontological-based and music score-based shematisations provide a symbolic view of a landscape scene.

This approach relies on a semantic-based analysis of verbal descriptions as used in landscape perception. It potentially allows cross-comparison of different

types of descriptions and environments, comparing their structures, the patterns that emerge, and the different terms and relations used. Further work concerns an extension of the ontological background of the model, and additional experimentations in different natural environments and user contexts.

References

1. Egenhofer, M., Mark, D.: Naive Geography. In Frankand, A.U., Kuhnand, W., eds.: COSIT'95: Conference on Spatial Information Theory. Volume 988 of LNCS., Springer (1995) 1–15
2. Brosset, D., Claramunt, C., Saux, E.: A location and action-based model for route descriptions. In Fonseca, F., Rodrígues, M.A., Levashkin, S., eds.: Proceedings of the 2nd International Conference on Geospatial Semantics. Volume 4853 of LNCS., Springer Verlag (2007) 146–159
3. Mark, D., Franck, A.U.: Cognitive and Linguistic Aspects of Geographical Space. Volume 63 of Behavioural and Social Sciences. Springer (November 1991)
4. Shuttleworth, S.: The use of photographs as an environment presentation medium in landscape studies. Journal of Environmental Management **11**(1) (1980) 61–76
5. Granö, J.: Reine Geographie. Volume 2 of 202. Acta Geographica (1929) Tr. as Pure Geography, reprinted in English from Johns Hopkins Press in 1997.
6. Tversky, B., Lee, P.U.: How space structure language. Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge **1404** (1998) 157–175
7. Mainwaring, S., Tversky, B., Ohgishi, M., Schiano, D.: Descriptions of simple spatial scenes in English and Japanese. Spatial Cognition and Computation **3**(1) (2003) 3–42
8. Tversky, B.: Cognitive maps, cognitive collages and spatial mental models. In Frank, A.U., Campari, I., eds.: COSIT'93: Conference on Spatial Information Theory. Volume 716 of LNCS., Springer-Verlag (1993) 14–24
9. Montello, D.: Scale and multiple psychologies of space. In Frank, A.U., Campari, I., eds.: Spatial Information Theory: A Theoretical Basis for GIS. Number 716 in LNCS, International Conference COSIT '93, Marciana Marina, Elba Island, Italy, September 19-22, 1993, Springer-Verlag (September 1993) 312–321
10. Smith, B., Varzi, A.: Fiat and bona fide boundaries: towards an ontology of spatially extended objects. In Hirtle, S.C., Frank, A.U., eds.: COSIT'97: Conference on Spatial Information Theory. Volume 1329 of LNCS., Springer (1997) 103–119
11. Hirtle, S., Jonides, J.: Evidence of hierarchies in cognitive maps. Memory & Cognition **13**(3) (1985) 208–217
12. Couclelis, H.: Worlds of information: the geographic metaphor in the visualization of complex information. Cartography and Geographic Information Systems **25**(4) (1998) 209–220
13. Franck, A.U., Raubal, M.: Formal specification of image schemata – A step towards interoperability in geographic information systems. Spatial Cognition and Computation **1**(1) (1998) 67–101
14. Tversky, B.: Narratives of space, time, and life. Mind & Language **19**(4) (September 2004) 380–392
15. Mark, D.M., Turk, A.G.: Landscape categories in Yindjibarndi: ontology, environment, and language. In Heidelberg, S.B., ed.: COSIT'03: Conference on Spatial Information Theory. Volume 2825 of LNCS. (2003) 28–45