

A salience-based approach for the modeling of landscape descriptions

Jean-Marie Le Yaouanc
Naval Academy Research
Institute, BP 600
29240 Brest Naval, France
leyaouanc@ecole-
navale.fr

Éric Saux
Naval Academy Research
Institute, BP 600
29240 Brest Naval, France
saux@ecole-navale.fr

Christophe Claramunt
Naval Academy Research
Institute, BP 600
29240 Brest Naval, France
claramunt@ecole-
navale.fr

ABSTRACT

While conventional GIS maps have long been a privileged way for the integration and diffusion of geographical information, novel forms of representation and description of urban and natural environments are nowadays emerging. In particular, verbal and textual descriptions of landscapes are progressively considered as alternative modeling resources for GIS. The research presented in this paper introduces a salience-based approach whose objective is to identify the noticeable entities of a natural landscape description. The model is based on a structural analysis of a given description, where salient entities are identified at the linguistic and structural levels. Salient entities are also spatially qualified according to their relative location with respect to a given observer.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
I.2.7 [Artificial intelligence]: Natural Language Processing

General Terms

Human factors, Algorithms

Keywords

Spatial cognition, landscape modeling, salience

1. INTRODUCTION

Landscape perception encompasses cognitive principles that favor the memorization of the main properties of an environment, and potentially the communication of its salient properties to an external addressee using natural language [3]. We consider the case of an observer located in a landscape, perceiving its 360° surroundings, and who is asked to provide a description of its environment to an external addressee. The verbal description of such a natural environment should underline the salient entities that compose

space, the spatial relations relating them, and the structural properties of the environment [10]. In a previous work, an analysis of the spatial and structural properties that emerge from such descriptions was developed [5]. Entities in the environment are identified according to a semantic categorization, their proximity and orientation with respect to the observer. The research presented in the present paper extends this modeling approach, by identifying the salient concepts that result from a linguistic analysis of a verbal description, and a study of its spatial properties.

Salience can be informally defined as an emphasize of an element or a set of elements distinguishable from the whole [4]. Salience is related to naturalness, *i.e.*, the natural order of constituents in a sentence that usually reflects the natural order of the events, and to vividness, *i.e.*, the semantics or visual properties of the entities [7]. This point of view generally comes to consider as salient what is natural, simple and clear. The aim of our research is to contribute to the specification of the linguistic and spatial salience involved in scenery descriptions. The salience-based approach evaluates the authority of terms contained in sentences of a scenery description, but also the degree of spatial prominence of the related entities in the landscape. It is based on a mutually reinforced approach where the concept of salience is recursively specified. Overall, this approach provides a conceptual and spatial specification of a landscape description whose final aim is to facilitate the analysis of the described environment for its georeferencing.

The remainder of the paper is organized as follows. Section 2 presents the modeling background of our approach and a conceptual representation of an environmental scene. Section 3 develops our modeling approach of salience implemented by a mutual reinforcement algorithm, and the experimental results. Finally, section 4 draws the conclusions and outlines further work.

2. MODELING BACKGROUND

We have introduced a structural categorization of a landscape view based on panoramic photographs that act as a substitute of a given natural environment. Verbal descriptions of a landscape scene provide the modeling input of the approach [5]. The structural-based model identifies the spatial, relational and semantic constructs that emerge from



Figure 1: Natural landscape of a mountainous region - scenery 1

these descriptions. Concepts in the environment are qualified according to a semantic classification, their proximity and orientation to the observer, and the spatial relations that qualify them. The resulting model is schematized by a representation that constitutes a modeling support for the study of environmental scenes (Fig. 2).

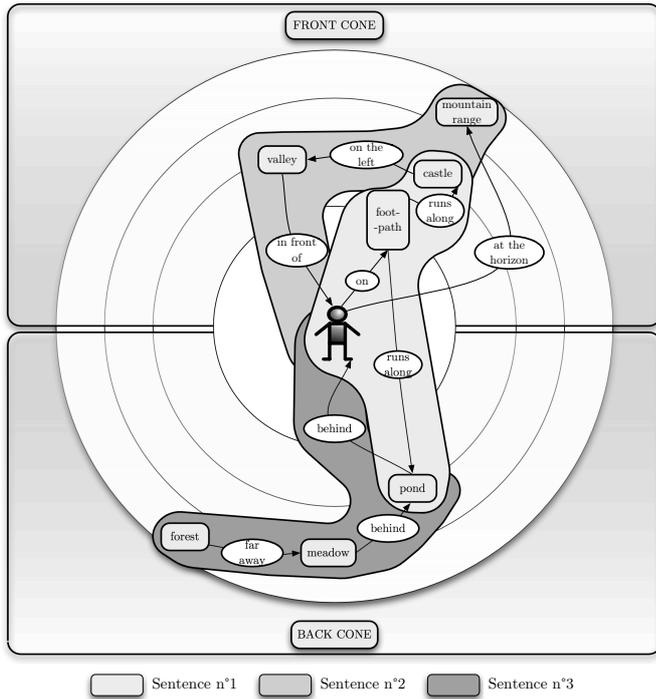


Figure 2: Conceptual map - scenery 1

Let us consider an example of description given by the landscape presented in figure 1: “I’m on a footpath that runs along a castle and a pond. In front of me, there is a valley with the castle on the left of it and at the horizon, I can distinguish a mountain range. Behind me, there is the pond with a large meadow behind and a forest far away”.

Humans tend to structure space using distance and bodily directions that relate the spatial entities to their own location. The way these distance and directional relations organize space implicitly generates a partition of space that can be modeled using proximity spaces and directional cones. We introduce a schematization approach that takes into account these principles, and whose objective is to facilitate the understanding of the spatial structures that emerge from the verbal descriptions of a natural landscape.

We define an *environmental scene* as the 360° environment, perceived by an observer from a static point of view. An environmental scene that provides an observer-centered reference for the location of entities, is ordered by the boundaries of four proximity spaces, and a cone-based partition whose number can vary from two (front-back or right-left) to four (front, back, right and left).

The conceptual map of figure 2 that results from the previous example clearly makes the difference between the space in front and behind the observer. This example shows evidence of a close relationship between the sentences and the location of the entities described.

Such a model qualifies and characterizes natural landscapes, and provides a framework for the analysis of the properties of verbal descriptions made by different observers, and cross-comparisons of different landscape descriptions. However, the salient entities of the scene are not always clearly revealed by the resulting conceptual map. This motivates an extension and analysis of the semantic and structural properties of the entities identified in the scenery.

3. SALIENCE COMPUTATION

The salience-based model integrates the properties of the linguistic description, and the structural characteristics of an environmental scene. Salience should reflect the particularities of the entities that result from a scenery description such as their linguistic properties, *i.e.*, the richness of information associated to each term, and their structural characteristics, *i.e.*, their degree of spatial isolation. We introduce a bipartite and oriented graph representation of a verbal description, where vertices represent sentences, entities, and cones, and edges the relations between them (Fig. 3).

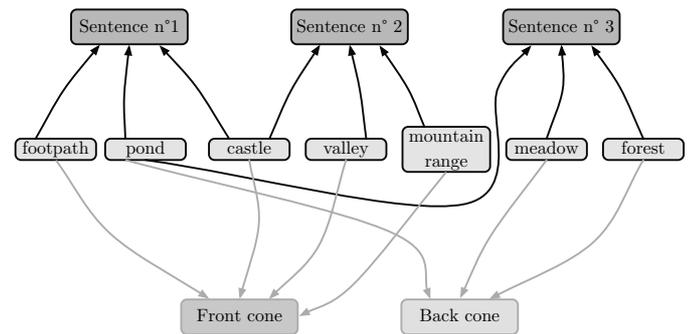


Figure 3: Graph representation - scenery 1

That graph supports a mutual reinforcement algorithm that

derives the linguistic and spatial salience scores associated to each entity, sentence and directional cone of the environmental scene.

3.1 Linguistic salience

Early research in Natural Language Processing, *i.e.*, automatic text retrieval and automatic summarization has largely raised the issues of salience related to the formal content and semantics of the data, focussing on recency and repetition [7, 2]. Recency is a relevant characteristic when scoring the importance of terms in documents, whether textual or verbal. Recency is based on the assumption that the more entities are recently quoted, the more they are salient [8]. This syntactic salience is related to the appearance order of the terms in each sentence. The number of occurrences of a term should also be considered as a relevant information for salience quantification. This assumption is based on the fact that a term mentioned more often than the others should receive a higher salience score [9].

Let \mathcal{S} be the set of sentences composing a verbal description, \mathcal{D} the set of verbal descriptions, \mathcal{U} the set of elementary units composing a sentence, \mathcal{E} the set of entities of an environmental scene, and \mathcal{R} the set of spatial relations including the null element \emptyset . A verbal description D is modeled as an ordered set of sentences $s_i \in \mathcal{S}$, *i.e.*, $D = [s_1, s_2, \dots, s_N]$ where $D \in \mathcal{D}$ and $N \geq 1$. A sentence s_i is an ordered set of elementary units $u_i \in \mathcal{U}$, *i.e.*, $\forall i \in [1, \dots, M]$ with $M \geq N$, $s_i = [u_1, u_2, \dots, u_M]$. An elementary unit u_i is a triplet such as $u_i = [e_j, r_k, e_l]$ with $e_j, e_l \in \mathcal{E}$, and $r_k \in \mathcal{R}$. The computational model of linguistic salience is valued by two functions, $ling_sentence(s_i, k)$ that values the score of a sentence s_i , and $ling_entity(e_j, k)$ that values the authority score of an entity e_j , with k the index of the iteration.

We consider that the informative value of a short sentence is higher than the one of a long sentence that should have the default of drowning information. A sentence containing few entities should then be more salient than the others. The algorithm is defined as “a salient entity is an entity that occurs in a lot of sentences and a salient sentence is a sentence that contains few, but salient entities”. Consequently, the more entities there are in the sentence, the lesser the score of each entity. Since the entity with the highest authority usually influences and structures the sentence, the score of each sentence is weighted by the highest entity authority. Let N be the number of sentences of the verbal description, and n_j the number of entities in sentence s_j , then

$$\begin{aligned} \forall e_j, ling_entity(e_j, 0) &= 1, \\ \forall s_i, ling_sentence(s_i, 0) &= 1. \\ \forall e_i, k \geq 1, ling_entity(e_i, k) &= \\ &\sum_{\substack{j=1 \\ s_j \ni e_i}}^N ling_sentence(s_j, k-1). \\ \forall s_i, k \geq 1, ling_sentence(s_i, k) &= \\ &\frac{1}{\sum_{\substack{j=1 \\ e_j \in s_i}}^{n_i} ling_entity(t_j, k)} * Max(ling_entity(e_j, k)). \end{aligned}$$

Considering that the recency score of each entity e_i exponentially decreases with the total number of entities of the description, the recency of each term is given by

$$Recency(t_i) = exp^{-\lambda * \frac{Card(\mathcal{E}) - Index\ of\ t_i}{\mathcal{E}}}, \lambda \in \mathbb{R}.$$

The resulting linguistic score $ling_entity$ of each entity e_i is then given by

$$\begin{aligned} \forall e_i, ling_entity(e_i, k) &= \\ &\sum_{\substack{j=1 \\ s_j \ni e_i}}^N ling_sentence(s_j, k-1) * Recency(e_i). \end{aligned}$$

3.2 Spatial salience

Salient entities constitute a reference for the orientation and the description of an environmental scene [6]. The study of visual salience generally begins by physiological considerations, resulting from the perception of specific entities that are identified by their physical characteristics, ontological categories, and spatial location, *i.e.*, their isolation or membership to a group that makes them salient [11].

The salience of an entity is intrinsically linked to the environmental context in which it is embedded in [1]. An entity is particularly considered as salient when it is spatially isolated from other entities. The spatial isolation of an entity is calculated on the basis of the directional cones that structure an environmental scene. We consider that a salient entity is an entity that occurs in a lot of salient cones, and a salient cone is made up of few, but salient entities. Consequently, the more entities there are in the cone, the lesser the score of each entity. Since the term with the highest authority usually structures the cone, the score of each cone is weighted by the highest term authority.

The computational model of spatial salience is valued by two functions, $struc_cone(c_i, k)$ that values the score of a cone c_i , and $struc_entity(e_j, k)$ that values the score of an entity e_j , with k the index of the iteration. Let N_c be the number of cones of the conceptual map, and n_j the number of entities in cone c_j , then

$$\begin{aligned} \forall c_i, struc_cone(c_i, 0) &= 1, \\ \forall e_j, struc_entity(e_j, 0) &= 1. \\ \forall e_i, k \geq 1, struc_entity(e_i, k) &= \\ &\sum_{\substack{j=1 \\ c_j \ni e_i}}^{N_c} struc_cone(c_j, k-1). \\ \forall c_i, k \geq 1, struc_cone(c_i, k) &= \\ &\frac{1}{\sum_{\substack{j=1 \\ e_j \in c_i}}^{n_i} struc_entity(e_j, k)} * Max(struc_entity(e_j, k)). \end{aligned}$$

The spatial isolation of an entity can also be calculated relatively to the proximity spaces. The algorithm is then enriched by the fact that “a salient proximity space is a space that contains few, but salient entities”.

3.3 Experimental results

Let us consider the example of description introduced in section 2. Figure 4 shows the computation of the algorithm that evaluates the linguistic salience of the entities, $ling_{entity}$. Entities “pond” and “castle” that are the only ones quoted twice in the verbal description are the most salient. As a result, their significant score affects the sentences in which they are, and particularly the first sentence that is composed of both entities. Consequently, sentence one that obtains the lowest score influences the score of the last entity of the sentence, *i.e.* “footpath”. Finally, sentences two and three are both composed of one salient entity, respectively “castle” and “pond”, and two other entities that are only quoted one time in the verbal description. These two last sentences get the same structure, and do not really influence entities “valley”, “mountain range”, “meadow”, and “forest” that obtain a middle salience score.

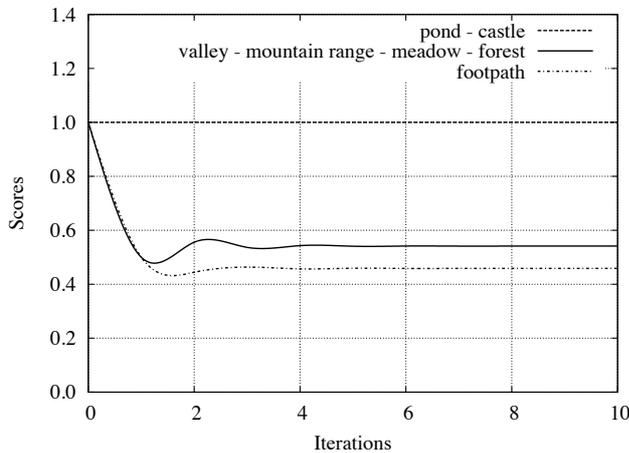


Figure 4: Linguistic salience of entities

Figure 5 illustrates the spatial salience of entities. The back cone is composed of few entities compared with the front one, and gets the highest score. Since no entities are located in several directional cones, the scores of the entities are directly related to the ones of the directional cones, and entities of the back cone obtain the highest score while entities composing the front cone obtain the lowest score.

4. CONCLUSION

The preliminary research presented in this paper introduces a salience-based approach of a verbal description of a natural landscape. The model is qualified by the linguistic and spatial measures of salience that characterize both verbal descriptions and natural landscapes, and provide a framework for the analysis of their properties. This approach reflects the main structure of a natural scenery perceived by an observer possibly lost in a natural environment, but also its

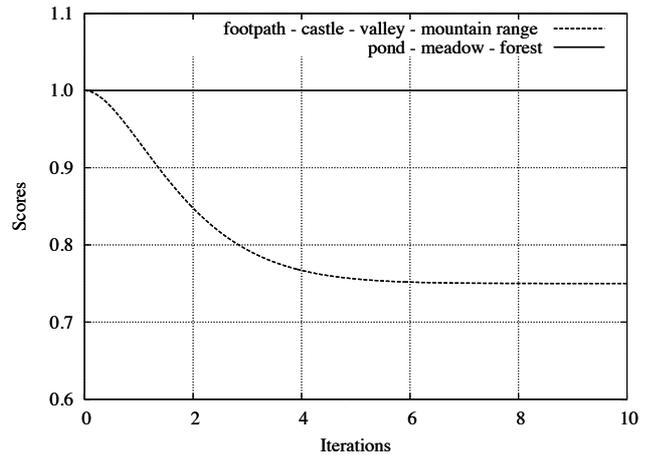


Figure 5: Spatial salience of entities

salient constituents.

Further work will be oriented toward an extension of the approach using ontological attributes for the characterization of entities, and a mapping of these conceptual maps towards a GIS representation in order to locate a given observer. The objective is to bridge the gap between the resulting conceptual map that identifies the salient entities and spatial relations, and locations within a scenery.

5. REFERENCES

- [1] P. Edmonds. A computational model of collaboration on reference in direction-giving dialogues. Master’s thesis, University of Toronto, oct 1993.
- [2] E. Hajicová. *Issues of sentence structure and discourse patterns*, volume 2 of *Theoretical and Computational Linguistics*, chapter 2. Charles University Press, Prague, Czech Republic, 1993.
- [3] A. Herskovits. *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Cambridge University Press, 1986.
- [4] F. Landragin. Taking situational factors into account when resolving anaphora: an approach based on events and salience. In A. Branco, T. McEnery, R. Mitkov, and F. Silva, editors, *Sixth Discourse Anaphora and Anaphor Resolution Colloquium*, pages 71–76, Lagos, Portugal, 2007.
- [5] J.-M. Le Yaouanc, E. Saux, and C. Claramunt. A semantic and language based model of landscape scenes. In M. P. Il-Yeol Song, editor, *ER 2008 Workshops*, volume 5232 of *LNCIS*, pages 334–343. Springer-Verlag, 2008.
- [6] K. Lynch. *The Image of the City*. MIT Press, Cambridge, 1960.
- [7] T. Pattabhiraman. *Aspects of salience in natural language generation*. PhD thesis, Simon Fraser University, Burnaby, BC, Canada, 1992.
- [8] G. Salton and C. Buckley. Term weighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5):513 – 523, 1987.

- [9] K. Spärck. A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–21, 1972.
- [10] B. Tversky and P. U. Lee. How space structures language. *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, 1404:157–175, 1998.
- [11] S. Winter. Route adaptive selection of salient feature. In W. Kuhn, M.F. Worboys, and S. Timpf, editors, *Proceedings of the International Conference on Spatial Information Theory*, volume 2825, pages 349–361. Springer Heidelberg, 2003.