Internal Version - Early Draft

Depth Cue of Occlusion Information as Criterion for the Quality of Annotation Placement in Perspective Views

Stefan Maass[1], Markus Jobst[2], Juergen Doellner[1]

[1]University of Potsdam, Hasso-Plattner-Institute, 14482 Potsdam, Germany
[2]Technical University of Vienna, Research Group Cartography, A-1020 Vienna, Austria
maass@hpi.uni-potsdam.de, markus@jobstmedia.at, doellner@hpi.uni-potsdam.de

Abstract. In cartography and computational geometry, concepts and techniques for automated label placement have been developed for two-dimensional maps. Less is known whether these methods can be applied to annotate geovirtual 3D environments. In this paper we discuss the application of these methods within geovirtual 3D environments and investigate the effects that can possibly harm the information transfer. To achieve high quality labeling readability, visibility, and the unambiguous correlation to the reference have to be ensured. Illustrated by examples, we show that perspective attributes inherently encoded in a depiction have to be considered as well. In particular, we focus on overriding occlusion information by added annotations and the impact on the complexity of the cognition process. A user test verifies our hypothesis that this disturbance is actually noticeable by users.

KEYWORDS: Annotation, Labeling, Depth Cue, GIS, Perspective Views, Geovirtual Environments

1 Introduction

The access to geo-information continuously becomes easier and ubiquitous (e.g., Google Earth, NASA Worldwind, Microsoft Virtual Earth) and effective user interfaces to geo-information are required for a growing number appliances and IT applications. Besides traditional 2D maps, *geovirtual 3D environments* establish themselves as flexible and effective platforms to communicate geo-information. Additionally, the wide range of available geo-data (e.g., satellite images, aerial photographs, cadastral data, and city models) and their increasing quality have created a market for geo-services. Applications can be found in the areas of urban and landscape planning (e.g., interactive presentation of planning alternatives, support for civil participation), environmental protection and nature conservation (e.g., simulation of noise emissions based on 3D city models), or location based services (e.g., virtual city maps, bus and underground timetables, recommendations for restaurant, hotel, and sights). Even non-expert users can configure and compose geo-information, e.g., to combine personal GPS-tagged data such as holiday pictures, cycling tours, and jogging paths with geodata over the Internet.

The wish to recapitulate own experiences in a spatial context, present them to friends, or to plan new activities leads to a wide approval of these services.

IT applications and systems progressively make use of interactive 3D geovisualization based on perspective views. There are several reasons for using 3D geodata, models, and visualization:

- 3D is required for a growing number of applications that represent 3D objects and 3D processes (e.g., radio-network planning, noise emission determination).
- Spatial 3D data becomes part of standard geo-data (e.g., 3D data for navigation systems or 3D city models).
- 3D technology and concepts benefit from an extensive body of scientific work in disciplines such as Virtual Reality, Computer Graphics, Gaming, and CAD/CAM. Furthermore, implementations can fall back on reliable, standardized software APIs and a broad variety of inexpensive 3D hardware.

The use of geovirtual 3D environments and perspective views can also be motivated by usability issues. By imitating a human observer's view, applications using interactive 3D geovisualization offer a more intuitive access, better comprehension of the content and improved virtual re-experience, e.g., in the case of recreational activities done in reality before.

Annotations play an essential role for geovirtual 3D environments: They provide meta information or detailed information about the presented objects by integrating texts or symbols in perspective views. This raises the question if techniques developed for automated label placement on cartographic maps can be used for perspective views without adaptations. To find an answer, we compare the kind of information inherently encoded in two-dimensional maps and perspective views (without annotations) and analyze whether existing annotation techniques consider the differences. Because we did not found techniques that explicitly use this as a criterion, we constructed a user test to validate the existence of an influence to the human perception. In this contribution we focus on overriding occlusion information by added annotations, the aspect for which we assume the strongest impact.

2 Related Work

2.1 3D Cartography

Various scientific disciplines offer theoretical work arguing that 3D can enhance the information transfer between systems and users. For example, in the scope of cartography, communication sciences generally focus on the understanding of content [12]. One indispensable requisite for a successful transmission of information is that both partners, the sender and the receiver, understand the content, thus the partners do have an intense overlapping of their knowledge base due to similar experiences, development environments or social influences [3]. For the field of cartography this fact supports the use of 3D, because the highly coded information of traditional maps, especially in case of topographic situations, calls for a massive cognitive load in the decoding process. Contrary, 3D presentations can deliver a more intuitive environment, in which the user has the possibility to behave like in the real world, which is the main environment of human development [5, 13].

The experience of the virtual world via its 3D presentation is highly related with the underlying human-computer interface, which enables the user to experience the surrounding – in the best case with the same impact as in the real world. Therefore several graduations of user-interface classes have to be defined, according to their support of intuitive information transfer. Although all kind of senses should be considered for the classification of user-interfaces, the main importance for this contribution lies in the transmission of geovirtual 3D environments via the visual mode. An interface with real 3D transmission uses all kinds of depth cues, psychological as well as physiological ones. Examples of this class of visual interfaces are holograms or light emitting volumes. Parallax 3D interfaces make use of autostereoscopic methods for an intuitive 3D impression. All psychological and a small selection of physiological depth cues are used for the transmission of the third dimension. Examples for Parallax 3D are lenticular lenses or autostereoscopic screens. The most widely spread interface-class for 3D environments is pseudo 3D, which is represented by computer displays. This class only uses psychological depth cues for information transmission, which can be synthetically generated in this 2D plane, the surface of the display, generally by using geometric rules and proportions [1]. The investigation shows that the impact for an intuitive spatial transmission of psychological depth cues can be damaged by the use of non-appropriate annotation techniques, which cause inhomogeneous transfer of depth. Therefore it becomes important that the role of occlusion information becomes classified in context with psychological depth cues.

2.2 Annotations

Independent of presentation media, depictions of geo-data become more informative by integrating annotations, such as text labels or symbols. Typical applications include:

- Naming of point, line, or area features for the mutual reference between positions in the virtual representation and the real world (e.g., cities, streets, or lakes).
- Support of orientation by providing names or symbolic depictions of land marks, e.g., churches, power supply lines, or rails.
- Spatial visualization of thematic attributes with symbols, e.g., amount of investments, occurrence of mineral resources, population density, election results, as a basis for policy makers.
- Collaboration support for geovirtual 3D environments (e.g., comments from civil participation processes).

Since amount, extensions, and positions of annotations in interactive applications can be influenced and configured by the user, there is a need of techniques supporting an automatic selection, scaling, and placement of annotations. Besides visibility of annotations, their readability and unambiguous correlation to their references has to be considered to achieve a visual aesthetic and understandable appearance. Taking all these criteria simultaneously into account is a complex task [17]; as result the annotation process is typically the last step in visualization applications.

2.3 Evaluation of Existing Annotation Techniques

A number of annotation techniques can be found in cartography [6,8,9], virtual reality environments [4,14,15,16], and virtual illustrations [2,10,11,19,20,21]. In general, two fundamental approaches are used: annotating in screen space and annotating in object space (Fig. 1).

The annotation in screen space is a classical approach used in cartography or book illustrations and is based on pictorial information as input. Here the placement can be understood as taking place in a separate layer lying over the depiction. Annotating perspective views of geovirtual 3D environments in this way can be seen as an annotation of an image of that view.

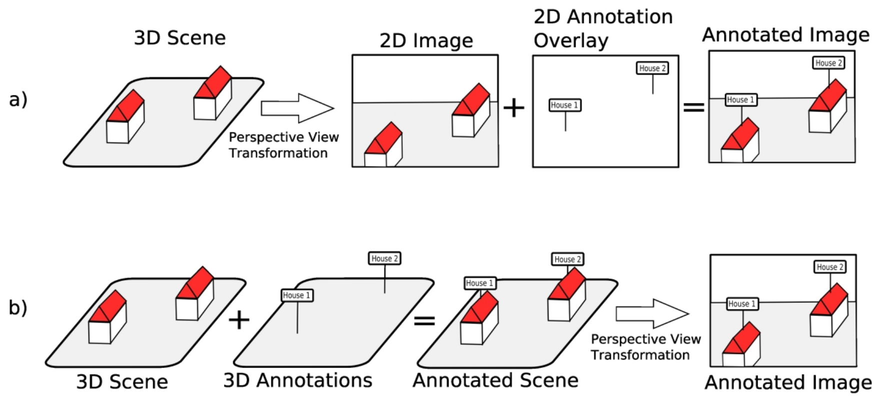

Fig. 1. Two different approaches to annotate virtual scenes with symbols or text: a) screen space annotation and b) object space annotation.

Annotation in object space means that the annotations are modeled as elements of the scene and become a part of it. This approach rather suits the expectations from users that are familiar with the modeling process of virtual environments. For them, it is more intuitive to place labels as banners on poles in a scene undergoing the perspective view transformation, than thinking of annotating a depiction for a specific viewpoint.

Both approaches aim at the identical objective: enabling the communication of information in a complete, fast and correct manner. Screen space placement techniques support the visibility criterion by concept. Here, visibility conflicts among annotations or among annotations and important parts of the depiction can be solved effectively, because the whole problem is reduced to the two dimensional space. Additionally, readability is supported by the fact that annotations are oriented parallel to the screen, whereby character deformation is avoided.

However, in the following sections we point out why screen space techniques applied natively to perspective views evoke visual conflicts that can harm the information communication. Being aware of this possible influence could help to extend known annotation techniques or to develop new, specialized ones.

3. 3D Information in Perspective Depictions

Generating a 2D image for a perspective view of a 3D scene implies a loss of information, which represents a characteristic property of pseudo-3D depictions compared to 2D representations. This disadvantage results primarily from occlusion, which impedes information transfer due to hidden objects and complicates the comparison of geometric objects from one part of the image with the same object in another part of the image. Additionally, perspective views combine an infinite quantity of scales within a single image, which demands for generalization.

Despite these characteristics, perspective views offer information that allows users to mentally reconstruct depicted geovirtual 3D environments. These attributes are known as *psychological depth cues*. The main components include occlusion, shadowing, linear perspective, texture gradients and aerial perspective [7, 18,22].

- **Occlusion** represents the most primitive psychological depth cue. Objects are hidden by objects that are closer to the viewer. The effects of occlusions are counteracted by the process of completion performed for partly hidden objects and represents an integral part of the human perception.
- **Shadowing/Shading** allows observers to estimate the direction of light, the vertical shape of objects, and distances between objects casting and receiving shadows. Additionally these attributes carry further information about the surface structure of receiving objects, e.g., polished, corrugated, or bumpy surface. Shadowing results in the combined perceptional processing of all listed aspects, which help to perceive object forms that are not explicitly visible.
- Due to **linear perspective** objects of constant size seem to grow with decreasing distance. The human perception process does generally not result in the size change of the objects but combines the perspective influence with distance dependency. The perspective influence is defined by the geometric situation that parallel lines join at one or more vanishing points.
- **Texture gradients** support the perception of surface characteristics. Texture gradients can be in form of size, density, and brightness gradients. For instance the brightness gradient enhances the perception of surface characteristics with the help of light-to-dark transitions depending on directional light sources.

- The effect of **Arial (or atmospheric) perspective** refers to differences in clarity, colour, and contrast of objects seen from a point far away. These differences are caused by a growing number of scattering light particles in the atmosphere. As classical example, this effect can often be studied on photographs, where tree-covered mountains show a bluish tint with an increasing distance.

Psychological depth cues play an important role for the intuitive transmission and perception of virtual 3D environments. The impact on effectiveness of the interface depends on the particular depth cue parameter as well as all possible combinations of these. Therefore it is important to ask for a possible influence on perspective information by adding annotations in various ways.

3.1 Perspective Information Impairment by 2D Annotation

The deliberate disturbance of perspective information is a widely used technique in art and design. For example, M.C. Escher applied these effects to create his famous depictions of impossible buildings and objects. Historical depictions of townscapes often communicate information about the town in a pleasing manner instead of presenting a correct view of a human observer. Furthermore, commercials use this kind of confusion to attract consumer's attention.

Human designers or illustrators intuitively avoid a placement that strongly impairs the perspective impression when they add annotations by hand. For automated annotation we cannot find a screen space technique considering the properties of perspective view transformation. In contrast, we discovered some examples, where an automated placement disturbs that impression. One is conceptually illustrated in Fig. 2.

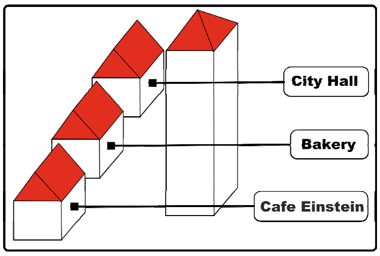

Fig. 2. Impairment of the perspective impression caused by annotations infringing occlusion information.

Here, the lines that connect the labels with the reference violate occlusion information. The conflict increases the complexity of the cognition process with an impact at the quality and speed of the information transfer. Other examples of annotations impairing the perspective impression include:

- Integration of annotations with different (font-) sizes, inconsistent to the scaling of the perspective view transformation.

- Simulation of object-embedded annotations that violate occlusion information e.g., embedded street names that overlay buildings or vegetation elements, self occluding the street within the annotation area.
- Adding information in a (3D) speech balloon that cast a fake shadow at a 2D top view of a map.

4 User Test

4.1 Experimental Setup

Our user test verifies whether annotations placed in the screen space can affect human cognition of perspective views. Provided with the evidence of an influence, additional tests can study the quantitative aspects.

The test described in this paper focuses on overriding occlusion information by annotations that overlay a perspective depiction. To avoid the influence of other depth cues, we use depictions of a virtual 3D city model, containing simple buildings without textures and rendered with approximated global illumination.

We first experimented with tests where users have to rank the single or overall placement quality of an annotated depiction or have to compare different possible placements against each other. Unfortunately, more attributes than expected can affect a user judgment about the quality of annotations. For example, some people rate annotations more important than others if they reference a larger area of the virtual scene (e.g., a place) and associate this fact with a high screen presence of the annotation. Furthermore, on depictions containing multiple annotations the overall layout, e.g., in form of symmetry or alignment, has an effect on the aesthetic integration of each single annotation.

For this reason, we can hardly create depictions that allow us an isolate test of the effect we want to examine. To solve this conflict we changed the way of testing to a setting that is effective for validation although unusual for annotations: We placed two external annotations on a depiction referencing two different points. The function of the first annotation is to provoke the effect we are looking for. For this, the line connecting this annotation and its reference point crosses an object closer to the observer. The second annotation is placed to create a depth reference. Its line crosses the line of the first annotation. Because occlusion information allows observers only an interpretation at an ordinal scale [22], the referencing lines were designed in a way that the participants can decide which line overdraws the other. We use this setup with the following hypotheses:

- **H1:** If both annotations and their referencing lines do not overlay a scene object that is closer than their anchor points, the line referencing a position closer to the observer will be rated as the front line (Fig. 3, left).
- **H2:** If the line of an annotation crosses an scene object that is closer than its anchor point, this line will be interpreted as coming closer to the observer because the depth information is continuously interpreted along this line. As a

consequence it could be rated as being in front of an annotation with an anchor point closer to the observer. (Fig. 3, middle)

- **H3:** Assume the situation described in hypothesis H2: If the annotation line that breaks the perspective is partially overlaid by the closer object that is crossed, the effect described in hypothesis H1 becomes valid again. (Fig. 3, right)

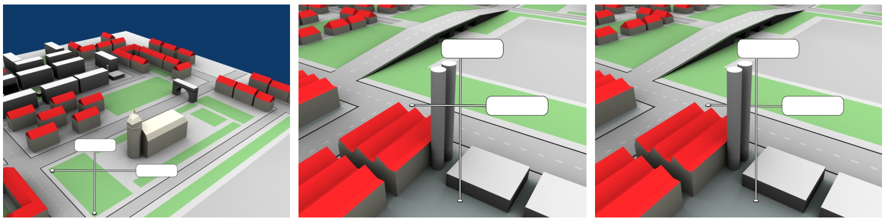

Fig. 2. User test examples for H1 (left), H2 (middle), and H3 (right)

4.2 Procedure

For each of these annotated depictions we created a variant that only differs in the way which line overdraws the other. We created a slide presentation showing these variants side by side and ask our participants with which depiction they feel more comfortable. Valid answers were: left, right, or both.
For each hypothesis we created four tests resulting in 24 annotated depictions (three hypotheses, four test depictions, two variants). To avoid the influence of learning effects they where presented in randomized order.

4.3 Participants

We had a group of 58 participants for our user test, 39 male and 19 female. The majority were students or staff members experienced with virtual environments, such as scientific 3D visualizations or computer games. All participants had normal or corrected to normal vision.

4.4 Results and Discussion

Table 1 shows the results for our first hypothesis, stating that an annotation with a closer reference point is judged nearer that an annotation with a reference point far away. The columns contain the values for each depiction and their average value. As expected a majority agrees with us in this point.

	Test 1	Test 2	Test 3	Test 4	Average
Agreements	56,9%	60,3%	82,8%	67,2%	66,8%
Undecided	13,8%	6,9%	5,2%	19,0%	11,2%
Rejections	29,3%	32,8%	12,0%	13,8%	22,0%

Table 1. Results for hypothesis H1.

The following table shows that our participants do not follow our hypothesis H2. In most cases they choose again the annotation with the anchor point closer to the observer as the one in front. However, an interesting point here is the difference to the results in table 1. The evidence for hypothesis H1 here is not as strong as in the tests where the reference line does not cross a scene object in front. Additionally the number of people who were undecided increased in the tests for H2. This could be interpreted as the effect we wanted to verify, but with a smaller influence than expected.

	Test 1	Test 2	Test 3	Test 4	Average
Agreements	15,5%	48,3%	20,7%	32,8%	29,3%
Undecided	24,1%	13,8%	20,7%	27,6%	21,6%
Rejections	60,3%	37,9%	58,6%	39,7%	49,1%

Table 2. Results for hypothesis H2.

Table 3 shows the results for the hypothesis H3 tests. As expected the line referencing a point farer away is voted as the back line. The number of participants that were undecided or reject this hypothesis decreased compared to the test for H1. A reason could be the strengthening of the depth cue of the line that is disconnected for a perspective more pleasing integration.

	Test 1	Test 2	Test 3	Test 4	Average
Agreements	84,5%	75,9%	84,5%	89,7%	83,6%
Undecided	3,4%	8,6%	10,3%	6,9%	7,3%
Rejections	12,1%	15,5%	5,2%	3,4%	9,1%

Table 3. Results for hypothesis H3.

5 Summary and Conclusions

In this paper we work out how 2D annotation techniques, operating only in screen space, can influence the human cognition if they are applied to a perspective view. We exposed different kinds of depth cues that are commonly not considered in the most 2D annotation techniques. For one of them, the occlusion information, we developed a user test to verify a possible influence to the interpretation of a human observer.

Even if our tests gave only a small hint for such an influence, we are still convinced that considering depth information can improve the quality of annotation placement techniques. We take our motivation for this thesis from the depictions below. Fig. 4a demonstrates the ability of humans to correct their interpretation of an annotated image until it fits the accustomed rules of perception and become plausible. Despite the lines of the right and lower annotation, overdrawing closer scene objects, most people are still able to interpret this as a 3D scene.

On the other side, Fig. 4b shows that annotations can strongly harm the perspective impression of a picture. Here another placement for the annotations resolving this conflict will be suggested by the most observers. Again, changing the annotation positions go along with varying a high number of parameters like the length of the reference lines, the selection of areas overlaid by the annotations, or the overall symmetry of the picture. This makes it hard to construct user tests proving the influence of only one attribute on the overall quality of the labeling.

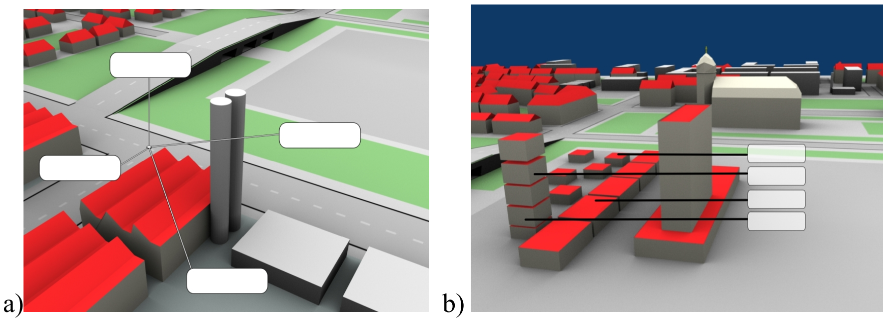

Fig. 3. Examples of annotations infringing occlusion information.

As a conclusion, we suggest the introduction of another parameter for the annotation of perspective views that includes all psychological depth cues. This parameter should be defined by the degree of perceivable impairment of the perspective attributes that are encoded in a perspective view of a 3D scene. Annotation placement techniques can additionally consider this parameter to make a decision among different positions candidates or to optimize the selection of visible annotations. Thereby situations as shown in Fig. 4b can be avoided or improved.

6 Future Work

In this contribution we focused on how annotations impair the perspective impression by violating occlusion information. For our future work we plan to extend our work to other perspective attributes and to study the influence of annotations on shading information, shadows, and linear perspective of a depiction.

Additionally, measurements for a perceivable perspective disturbance have to be developed and approved for different applications. One of our assumptions here is that the influence of harmed occlusion information to the overall information transfer depends on the depth complexity of the scene. This means it has a higher impact for geovirtual 3D environments such as wide area city models than for the labeling of single objects, e.g., a single 3D object centered in the screen. After we are able to weight this parameter to others, we plan to design new annotation strategies that additional minimize perceivable perspective disturbance in the optimization strategy, try to avoid it at all, or use it for a controlled communication of information.

Another field of interest is raised by the possibility of interaction in geovirtual 3D environments. Instead of handling the annotation process as an extra step, i.e., labels and symbols fade in shortly after the interaction has stopped, they can be dynamically placed for each frame. Here, we are planning to verify our hypothesis that this would increase the influence of a perceptible disturbing.

Acknowledgements

This work has been funded by the German Federal Ministry of Education and Research (BMBF) as part of the InnoProfile research group '3D Geoinformation' (www.3dgi.de). We would like to thank Johannes Bohnet and Marc Nienhaus for their comments on this work and all the participants of our user test for their collaboration.

References

1 Albertz, J. Die dritte Dimension – Elemente der räumlichen Wahrnehmung; in Wahrnehmung und Wirklichkeit; Verlag der Freien Akademie; Berlin (1997).

2 Ali, K., Hartmann, K., and Strothotte, T. Label Layout for Interactive 3D Illustrations. Journal of the WSCG, 13 (1): (2005), pp. 1–8.

3 Badura, A. Social Cognitive Theory of Mass Communication. In Bryant, J./Zillmann, D. (eds.): Media Effects. Hillsdale, N.J., pp. 121-153 (2002).

4 Bell, B., Feiner, S., and Höllerer, T. View Management for Virtual and Augmented Reality. In Proceedings of the 14th ACM Symposium on User Interface Software and Technology (UIST). ACM Press, 2001. pp. 101–110.

5 Buchroithner, M. Autostereoskopische kartografische 3D-Visualisierung; Kartographische Schriften, Band 6, eds. Deutsche Gesellschaft für Kartographie e. V., Kirschbaum Verlag, Bonn (2002).

6 Christensen, J., Marks, J., and Shieber, S. An Empirical Study of Algorithms for Point-Feature Label Placement. ACM Transactions on Graphics, 14 (3):(1995), pp. 203–232.

7 Drascic, D., Milgram, P. Perceptual Issues in Augmented Reality. In Proceedings of SPIE 2653; Stereoscopic Displays and Virtual Reality Systems III, pp. 123-134 (1996).

8 Ebner, D., Klau, W.K., Weiskircher, R. Force-Based Label Number Maximization, Technical Report TR-186-1-03-02, Technical University Vienna, June

2003, available at http://www.apm.tuwien.ac.at/publications/bib/pdf/ebner-03.pdf.
9 Edmondson, S., Christensen, J., Marks, J., Shieber, S. M. A General Cartographic Labeling Algorithm. TR1996-004, Cartographica 33, 4, 13-23, 1996, available at http://www.merl.com/pub lications/TR1996-004/.
10 Hartmann, K., Ali, K., and Strothotte, T. Floating Labels: Applying Dynamic Potential Fields for Label Layout. In Smart Graphics: 4th International Symposium (SG 2004), volume 3031. Springer-Verlag, pp. 101–113 (2004).
11 Hartmann, K., Götzelman, T., Ali, K., and Strothotte, T. Metrics for Functional and Aesthetic Label Layouts. In A. Butz, B. Fisher, A. Krüger, and P. Olivier, editors, Smart Graphics: 5th International Symposium, SG, volume 3638. Springer, Frauenwoerth Cloister, Germany, pp. 115–126 (2005).
12 Jobst M., Brunner-Friedrich, B., Radoczky, V. User Centered Cartography – Preparations for Ubiquitous Cultural Access, Studies in Communication Sciences 5/1, Universitá della Svizzera Italiana, Lugano, pp. 93-110 (2005).
13 Kirschenbauer, S. Empirisch-kartographische Analyse einer echt-dreidimensionalen Darstellung am Beispiel einer topographischen Hochgebirgskarte, Dissertation, Technische Universität Dresden, Mensch und Buch Verlag, Berlin (2003).
14 Kolbe, T. H. Augmented Videos and Panoramas for Pedestrian Navigation. In G. Gartner, editor, Proceedings of the 2nd Symposium on Location Based Services and TeleCartography, (2004).
15 Maass, S. and Döllner, J. Dynamic Annotation of Interactive Environments using Object-Integrated Billboards. In J. Jorge and V. Skala, editors, 14-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, WSCG'2006. Plzen, Czech Republic, pp. 327–334 (2006).
16 Maass, S. and Döllner, J. Efficient View Management for Dynamic Annotation Placement in Virtual Landscapes. 6th Int. Symposium on Smart Graphics 2006, Vancouver, Canada, July, pp. 1-12 (2006).
17 Marks, J. and Shieber, S. The Computational Complexity of Cartographic Label Placement. Technical Report TR-05-91, Harvard University, (1991).
18 Palmer, S. E., Vision Science, Photons to Phenomology, Cambridge, Massachusetts: MIT Press, ISBN: 0-262-16183-4, (1999).
19 Preim, B., Raab, A., Strothotte, T. Coherent Zooming of Illustrations with 3D-Graphics and Textual Labels. Proceedings. of Graphics Interface, pp. 105-113 (1997).
20 Ritter, F., Sonnet, H., Hartmann, K., and Strothotte, T. Illustrative Shadows: Integrating 3D and 2D Information Displays. In IUI'03: Proceedings of the 8th International Conference on Intelligent User Interfaces. ACM Press, pp. 166–173 (2003).
21 Sonnet, H., Carpendale, S., and Strothotte, T. Integrating Expanding Annotations with a 3D Explosion Probe. In AVI '04: Proceedings of the Working Conference on Advanced Visual Interfaces. ACM Press, pp. 63–70 (2004).
22 Yonas, A. How the Eye Measures Reality and Virtual Reality. Behavior Research Methods, Instruments and Computers, 29(1), pp. 27-36.