Privacy-Enabling Abstraction and Obfuscation Techniques for 3D City Models (Position Paper)

Martin Kada, Michael Peter, Dieter Fritsch
Institute for Photogrammetry, University of Stuttgart
Geschwister-Scholl-Straße 24D
70174 Stuttgart, Germany
++49 (0)711/685-83386
[firstname.lastname]@ifp.uni-stuttgart.de

Oliver Siemoneit, Christoph Hubig
Institute of Philosophy, University of Stuttgart
Seidenstraße 36
70174 Stuttgart, Germany
++49 (0)711/685-82491
[firstname.lastname]@philo.uni-stuttgart.de

ABSTRACT
Privacy and security issues within geospatial information systems are of growing public and scientific interest. Especially with the launch of Google Street View and Google Earth, geospatial data has come to the attention of the public, thereby not only raising support for these technologies, but also massive concerns. It is the duty of science to pick up today's uprising debates and to help structuring them, providing clarifications and different solutions. Thus, the aim of this paper is to contribute in form of an interdisciplinary discussion about privacy issues, both from a philosophical and an engineering point of view. Privacy and its importance are outlined as well as different privacy issues raised concerning the nowadays so popular 3D city models. In addition, technical solutions are shown which allow data providers to preserve privacy, but that won't interfere with the advancements of these technologies.

Categories and Subject Descriptors
K4.1 [Computers and Society]: Public Policy Issues - Privacy

General Terms
Human factors, legal aspects, algorithms

Keywords
Ethics, privacy, privacy-enhancing technologies, 3D city models, street views

1. INTRODUCTION
Technological developments usually come forth with a vast amount of new opportunities. But at the same time, technological innovations are also prone to novel, unknown problems and threats–for its latter users and/or society in general. It is the unavoidable, janus-faced nature of technology, which has also recently drawn a lot of attention to geospatial information systems and services. Especially in a lot of European countries–where privacy and data protection laws are far stricter than in the United States–some early kind of geospatial information system has proven to be highly controversial. The launch of Google Street View has caused many citizens to issue complaints to government officials about the project thereby claiming that it is a massive intrusion upon privacy and thus a violation of existing data privacy laws [1]. In this quite emotional, heated and sometimes even irrational debate, it is the job of the sciences to pick up the raised questions, to think about them, to analyze, and to restructure them in joint, interdisciplinary research and finally to think of adequate solutions. It is the main aim of this paper, to make a basic contribution to this debate and to shed some light on fundamental questions. Therefore, in section 2, it is first elaborated on what privacy is at all and which role privacy plays in western societies. In section 3, different privacy issues within geospatial information systems are identified and outlined. Section 4 presents techniques that offer the potential to better preserve privacy in 3D city models, but also discusses their limitations. Section 5 concludes our work, which we regard as a stepping stone to future research directions.

2. PRIVACY AND ITS IMPORTANCE
Privacy and the right of privacy is central for all liberal, egalitarian, and democratic societies [2][3]. Privacy assures personal freedom and autonomy, guarantees freedom from governmental interventions, or other societal institutions, parties or persons and thus allows a person to build his own individual scheme of life. It is privacy that establishes a sphere of non-intervention which is also crucial for self-fulfillment and the development of a personal identity. So-called *decisional privacy* concerns therefore basic decisions of a person about who he wants to be and how he wants to live. Decisional privacy is the core of one's personal freedom and the possibility to form one's own authentic identity. It is also the core of political freedom in the form of the absence of interferences with the sovereignty (negative freedom as "freedom from") and the assistance in fulfilling one's own potential (positive freedom as "freedom to").

On the contrary, so-called *informational privacy* deals with the fact that a person wants to be in control of personal information about intimacies of his life. In this clearly information-based or

knowledge-based conception of privacy, privacy intrusions are defined therefore as situations in which personal information is collected or disseminated without consent of the person who is topic of the information. Informational privacy is crucial for regulating personal relationships and establishing different social roles one plays in society: "If everyone knew everything about everyone else, differentiated relations and self-presentation would no longer be possible, nor would autonomy and the freedom to determine one's own life" [3].

However purely information-based conceptions of privacy are clearly flawed: There are also other privacy violations, which are not of a cognitive nature but of a physical one: *Local privacy* is the right of a person to restrict physical access of others to his body, his personal belongings and his home. Local privacy assures therefore a sphere of non-intervention, a protected, secured, private place or shelter. This definition of privacy corresponds well with the famous description of Samuel D. Warren and Louis D. Brandeis of the right of privacy as the right to be left alone [12]. Not only is it important to be left alone from the gaze and opinions of others, but also the right to control physical interference by others into one's private affairs.

To conclude, we could say, with Ferdinand Schoeman, that "a person has privacy to the extent that others have limited access to information about him, limited access to the intimacies of his life, or limited access to his thoughts or his body" [4]. The right of privacy is then the right of a person to be protected against intrusions (negative form of privacy as being free from) and to be able to control cognitive or physical access to his personal things and affairs (positive form of privacy as being able to decide freely to). Thus, privacy allows for inner and outer freedom of an individual, helps building and assuring the personal integrity and autonomy, helps protecting his reputation, is enabling different forms of social self-representation in different social contexts.

If you have a look at the history of privacy, it becomes obvious, that what counts as "public" or "private" depends largely on the social tradition and varies from culture to culture [3]. Privacy is therefore of "conventional nature" only and subject to an on-going societal negotiation process. The personal right of privacy is delimited and overridden by other rights and competing moral principals so as to protect interests and rights of other parties or of society in general. (E.g. at the workplace, privacy is neither totally free from restrictions nor does a contract of employment nullify privacy claims at all. Or as Anders J. Persson and Sven Ove Hansson put it, taking another form of contractual relation as example: Having a rental contract will give the owner of the house the right to enter the house for certain purposes, but not to open closets and read private papers that are kept in there [5].)

All in all, from a theoretical point of view, this societal balancing and negotiation process could be best described by the concept of so-called reflective equilibrium [6]. On the one hand, certain "given" values or norms do restrict our social practices. On the other hand, looking at the practical effects of these norms, we do also change certain norms and values we consider as too restrict, inadequate or outdated. In our case this means: On the one hand personal privacy rights are weighed against other rights restricting/enlarging personal privacy. On the other hand existing rights and social norms are also changed, because we are not willing to accept them anymore (since they restrict privacy too much).

3. PRIVACY ISSUES IN GEOSPATIAL INFORMATION SYSTEMS

According to the directive 2007/2/EC of the European Parliament and Council, geospatial data is "any data with a direct or indirect reference to a specific location or geographical area" thereby often describing a spatial object which is further defined as an "abstract representation of a real-world phenomenon related to a specific location or geographical area" [7]. Geospatial data is therefore only object-related data and not initially subject to data protection laws [9]. However–under certain circumstances–geospatial data could become personal data [8][9]. This is the case, if 1) photos or photo-realistic views/models of spatial objects (i.e. a building or an estate) could be easily located by geo-coordinates and thus easily matched to its owner or residents and/or if 2) the data is–to put it more generally–able to describe personal or factual affairs [8][9]. In these cases, geospatial data is also subject to data protection laws. Then, the collection, storage, processing and dissemination of the data is only allowed, if the interests of the individuals, which are subject of the data, are not harmed and/or are not superseded by other rights and interests (such as homeland security) [8].

Picking up the above mentioned "dimensions" of privacy, privacy intrusions in the realm of geospatial data are of cognitive nature only and therefore mainly intrusions on informational privacy (with possible effects on decisional and local privacy in the future):

1) Geospatial data showing faces of people, license plates of cars (as Google Street View does) could be seen as problematic, since it conveys a lot of information on personal affairs, such as personal habits, preferences, circumstances [8][10]. Even obfuscating faces or license plates is often not sufficient, as a lot of things still remain recognizable because of other distinctive, individual characteristics [10].

2) The same is true for showing house numbers and detailed, photo-realistic images or representation of spatial objects, since it tells a lot about personal circumstances and thus could allow for geo-marketing or scoring of creditworthiness [8][10].

3) Especially Google Street View is criticized for having a "privileged view" on the spatial object: Pictures are taken at the height of 2.5m and not at the height of the eyes of a pedestrian, thus allow to look inside an estate or home–a per se secured, protected, intimate space [10].

However, data protection officials also agree on this: If the spatial object is obfuscated or the presentation of the spatial object is of abstract manner only, no interests of individuals are violated [8]. Therefore, in the following, different methods and techniques are to be presented and outlined, which meet these requirements and "remove" certain privacy issues. However, one has to bear in mind, that the core of data protection is not met by that: Especially in Germany, the data protection officials want to force Google not just to obfuscate images and grant individuals the right to get certain pictures removed from the database, but also to delete all non-obfuscated raw-data (so as not to be able to use the data anymore for e.g. commercial purposes in countries where data protection laws are not as strict as in Europe) [11]. And indeed this is the case: Google has collected data in Germany but has transferred all data for storing and further processing to the

US thereby not willing to delete the raw material and thus the data still being open for abuse and possible privacy violations [11].

All in all, it should be highly appreciated, that privacy and geospatial data is discussed more and more on a broad public basis. It is the duty of society in general to decide, which technology to adopt (respectively how to adopt a technology): We need not do all the things we are (technically) capable to do. It is the job of philosophy and the engineering sciences to accompany, support and guide these debates, to clarify things and to outline and provide different solutions.

4. PRIVACY-ENABLING ABSTRACTION AND OBFUSCATION TECHNIQUES

As already mentioned, the obfuscation of faces and license plates still leaves a lot of information in the image, so that a person or one person's property could still be recognizable. The distinct characteristics could be very small or unusual, which makes it nearly impossible to automatically detect and remove them all by processing a single image at a time. Even if it would be possible, the resulting images would depict scenes where large parts are blurred or missing. As such images are not attractive to anyone this is obviously not a viable solution. The only alternative is therefore to use image sequences that show the same scene at different times and/or from different angles. Then the critical objects are hopefully gone or at least are located in a different part of the image and have cleared the view to the formerly occluded area. Multiple images allow for an image fusion to produce new ones without people and private objects.

Most comparable work has been done for the automatic generation of façade textures from terrestrial images, where occlusions from cars and pedestrians are avoided by a filtering of multiple images. Böhm [13] e.g. blends per-pixel registered images in a color clustering approach in order to synthesize occlusion-free texture images for building façades (see Figure 1 left). By only capturing a handful of images from multiple stations or a sequence of images from one point, both moving and static objects that are in front of the façade can be completely eliminated. However, each pixel must exactly point to the same planar part of the façade as the corresponding pixels of the other images. Such image correspondences can be reliably determined by the SIFT operator [14], which has also been implemented to run in real-time [23]. Although an automatic retouching of façade images is only possible if the underlying façade geometry is known, the necessary methods for a reconstruction from stereo imagery and laser scanning data at street level has long been shown (see e.g. [20][21][22]). And as recent reports have stated, the Google Street View vehicles of Google have been spotted with laser scanners mounted on the roof.

Until now, we have only regarded objects that are in front of the façade and not on the façade. This applies to house numbers, name plates and billboards. And although stores, firms and companies place them by the majority for advertising purposes, private persons and small firms might feel their privacy violated by this unwanted publicity. Such objects could be detectable by optical character recognition (OCR), which has reached a level where letters and numbers are reliably recognized. The question remains what to do with these areas? In contrast to persons and cars, a blurring of the characters would in most cases be sufficient to make them unrecognizable. Again, such an approach is not very appealing as it degrades the quality of the façade textures. Better would be to retouch these areas by copying similar parts of the façade image (see Figure 1 right, bottom).

Once the objects in front of the building have been eliminated and the façades been cleared, the next level of anonymization is to remove what can be seen of the interior of the building. The major intrusion into private homes can be expected coming from the windows. To counteract this, the glass parts could be grayed out and given a bright streak of reflected light to keep a realistic appearance. Another option would be to store the semantic information, so that a visualization application can adapt the window glass to better reflect the environment and weather conditions. However, before the relevant pixels can be altered, the

Figure 1: Left: Occlusion-free texture (bottom) by multiple image fusion. Right: Two abstraction levels (1st image, 2nd geometry) from a photo-realistic 3D building model (top).

locations and shapes of the windows must be detected. Several publications have addressed this problem. By using the Förstner operator [15], Mayer matches façade images to a database containing images of common window types. This enables the identification of the position and dimensions of the windows [16]. Ripperda and Brenner reconstruct the arrangement of doors and windows in a stochastic Reversible jump Markov Chain Monte Carlo process using formal grammars of façades [17]. Becker and Haala detect 3D edges in image pairs to do a hypothesis test on the existence of glazing bars and fanlights of windows and doors [18]. Also Wenzel et al. detects repetitive structures in facade images by using the SIFT operator in conjunction with a heuristic search method [19].

A photorealistic visualization might not always be necessary in all applications. Döllner and Kyprianidis e.g. present an automatic image abstraction approach that, applied to image sequences of 3D city models, results in a realistic, but cartoon-like presentation of virtual environments [24] (see Figure 1 right, center). On the one hand, such a presentation of real-life objects features enough details to recognize the spatial situation, but on the other hand changes enough to make re-identification of persons impossible and the judgment on people's living conditions inconclusive. In a last image abstraction step, the facades and roofs could be colored in a single color only.

The abstraction of (façade) images is only one aspect concerning the privacy of geospatial data. Another is the geometry, which can be regarded both by single buildings, but also by their arrangement into building blocks.

Over the last decade, quite some work has been dedicated to the simplification of 3D building models for cartographic purposes (e.g. [25][26][27][28]). In contrast to surface simplification algorithms known from the field of computer graphics, these algorithms are specifically designed for buildings. They strictly maintain global symmetries and enforce geometric properties like the co-planarity, parallelism and rectangularity of façade walls with the purpose to avoid that they become tilted in the simplification process.

Because not every application needs highly detailed models, as long as the final outcome is the same, we suggest using such a geometric abstraction as a way to protect peoples' privacy. During route guidance, e.g., only those details that assure a high recognition rate of the landmarks and that are necessary for a particular route should be presented to the user. Objects that are of no interest in the application's context must not be exposed to the public in every detail (see Figure 1 right, bottom), because a highly detailed representation could reveal private information about someone's living condition. For the driver of a vehicle, the single colored flat façade of a building has the same informative value as a detailed façade with windows, dormers, doors, etc. However, if highly detailed façade models are a requirement, their appearance could be obfuscated by a generalization that changes the number and arrangement of façade elements [26].

The above mentioned generalization algorithms try to maintain the object's shape characteristics as best as possible. This is especially useful for landmarks and other buildings with unusual architectures. For residential buildings, which are generally of higher concern regarding privacy issues, an even stronger shape simplification can be achieved by the use of 3D building symbols or standard roof shapes (see e.g. [29][30]).

At this point, we want to leave single buildings behind us and take a look at spatial situations with several buildings. There are two alternatives for an abstraction: typification and aggregation. Typification is a generalization technique where the spatial situation is analyzed to detect similar objects and their arrangement. Then the number of objects is reduced while maintaining the global appearance. For example, two buildings in a row of five similar looking houses could be removed and the remaining three re-located and increased in size to fill the idle space. Traditionally, such a technique is used to make room in a map when its scale changes and all objects won't fit anymore in the same space. This technique, however, could well be used to obfuscate a spatial situation or even hide buildings that are at risk concerning their security. As typification of 3D city models is really only a 2D problem, an algorithm like the one described in [31] could be used.

While such an abstraction approach still results in models that comprise of several entities, the aggregation operation replaces all buildings with a building block. Anders shows e.g. an approach that works on 3D building models [32]. Glander and Döllner aggregate building blocks while highlighting landmarks [33]. Such approaches could be context-sensitive, thus presenting only the detailed information that is vital to the task. The remaining objects are simplified to protect the privacy of residents, owners, public and private facilities.

5. CONCLUSION

In this paper, we gave an in-depth discussion on privacy in general and privacy issues in special within geospatial information systems. Furthermore, different abstraction and obfuscation techniques have been presented helping to circumvent possible (re-) identification of people and their living conditions and shield institutions that are at risk from prying eyes, which otherwise would be possible from highly accurate and detailed 3D city models.

6. REFERENCES

[1] Spiegel Online International. 2008. Privacy Concerns: German Towns Saying 'Nein' to Google Street View. URL=http://www.spiegel.de/international/germany/0,1518,581177,00.html.

[2] Rössler, B. 2004. The Value of Privacy. Wiley.

[3] Rössler, B. 2006. New Ways of Thinking about Privacy. In Phillips, A., Honig, B. and Dryzek, J. (eds.) The Oxford Handbook of Political Theory. Oxford University Press. pp. 694-712.

[4] Schoeman, F. 1984. Philosophical Dimensions of Privacy. Cambridge University Press.

[5] Persson, A.J., and Hansson, S.O. 2003. Privacy at Work: Ethical Criteria. In Journal of Business Ethics, Vol. 42., pp. 59-70.

[6] Norman, D. 2008. Reflective Equilibrium. In Zalta, E. N. (ed.) The Stanford Encyclopedia of Philosophy (Fall 2008 Edition). URL=http://plato.stanford.edu/archives/fall2008/entries/reflective-equilibrium/.

[7] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

2007. URL=http://www.emwis.net/documents/fol962872/euro_legislation/stl078838.

[8] Resolution of the German top data protection authority "Düsseldorfer Kreis" on the provision of digital street views especially in the internet (German). 2008. URL=http://www.bfdi.bund.de/cln_118/SharedDocs/Publikationen/Entschliessungssammlung/DuesseldorferKreis/141108DigitaleStrassenansichten.html?nn=409242.

[9] Forgó, N., Krügel, T., and Reiners, N. 2008. Expert's report on geospatial data and data protection (German). URL=http://www.iri.uni-hannover.de/tl_files/pdf/Gutachten%20GEODAT.pdf.

[10] Privacy International. 2009. PI files complaint about Google Street View. URL=http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-564039.

[11] Spiegel Online International. 2009. Protecting Privacy: Hamburg reaches Deal with Google on Street View. URL=http://www.spiegel.de/international/zeitgeist/0,1518,626075,00.html.

[12] Warren, S.D., and Brandeis, L.D. 1890. The Right of Privacy. In Harvard Law Review, Vol. 4, No. 5, pp. 193-220.

[13] Böhm, J. 2004. Multi-Image Fusion for Occlusion-Free Façade Texturing. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXV, Part B, Istanbul, Turkey.

[14] Lowe, D.G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. In International Journal of Computer Vision, 60-2, 91-110.

[15] Förstner, W., and Gülch, E. 1987. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, 281-305.

[16] Reznik, S., and Mayer, H. 2007. Implicit Shape Models, Model Selection and Plane Sweeping for 3D Façade Interpretation. In the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, 3/W49A, 173-178.

[17] Ripperda, N., and Brenner, C. 2009. Application of a Formal Grammar to Façade Reconstruction in Semiautomatic and Automatic Environments. In Proceedings of the 12th AGILE Conference on GIScience, Hanover, Germany.

[18] Becker, S., and Haala, N. 2007. Refinement of Building Facades by Integrated Processing of LIDAR and Image Data. In Proceedings of Photogrammetric Image Analysis (PIA07), Munich, Germany, 7-12.

[19] Wenzel, S., Drauschke, M., and Förstner, W. 2008. Detection of Repeated Structures in Facade Images. In Pattern Recognition and Image Analysis, Vol. 18, No. 3, 406-411.

[20] Früh, C., and Zakhor, A. 2004. An Automated Method for Large-Scale, Ground-Based City Model Acquisition. In International Journal of Computer Vision, 60 (1), 5-24.

[21] Cornelis, N., Cornelis, K., and Van Gool, L. 2006. Fast Compact City Modeling for Navigation Pre-Visualization. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Vol. 2, 1339-1344.

[22] Pollefeys, M., Nistér, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Kim, S.-J., Merell, P., Salmi, C., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewénius, H., Yang, R., Welch, G., and Towles, H. 2008. Detailed Real-Time Urban 3D Reconstruction from Video. In International Journal of Computer Vision, Vol. 78 (2-3), 143-167.

[23] Sinha, S.N., Frahm, J.-M., Pollefeys, M., and Genc, Y. 2006. GPU-Based Video Feature Tracking and Matching. In Proceedings of EDGE 2006, Workshop on Edge Computing using New Commodity Architectures, Chapel Hill, USA.

[24] Döllner, J., and Kyprianidis, J.E. 2009. Approaches to Image Abstraction for Photorealistic Depictions of Virtual 3D Models. Proceedings of the First ICA Symposium for Central and Eastern Europe, 371-385.

[25] Forberg, A. 2004. Generalization of 3D Building Data based on a Scale-Space Approach. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, Vol. XXXV, Part B.

[26] Thiemann, F., and Sester, M. 2004. Segmentation of Buildings for 3D-Generalisation. In Working Paper of the ICA Workshop on Generalisation and Multiple Representation, Leicester, UK.

[27] Poupeau, B., and Ruas, A. 2007. A Crystallographics Approach to Simplify 3D Building. In Proceedings of the 23rd XXIII International Cartographic Conference, Moscow, Russia.

[28] Kada, M. 2005. 3D Building Generalisation. In Proceedings of the 22th International Cartographic Conference, La Coruna, Spain.

[29] Thiemann, F., and Sester, M. 2006. 3D-Symbolization using Adaptive Templates. In Proceedings of the GICON 2006, Vienna.

[30] Kada, M. 2007. Scale-Dependent Simplification of 3D Building Models Based on Cell Decomposition and Primitive Instancing. In Spatial Information Theory: Proceedings of the 8th International Conference, COSIT 2007, 222-237.

[31] Sester, M. 2000. Maßstabsabhängige Darstellung in digitalen räumlichen Datenbeständen (German) (Postdoctoral thesis). Deutsche Geodätische Komission, Reihe C, Heft 544.

[32] Anders, K.-H. 2005. Level of Detail Generation of 3D Building Groups by Aggregation and Typification. In Proceedings of the 22th International Cartographic Conference, La Coruna, Spain.

[33] Glander, T., and Döllner, J. 2008. Automated Cell Based Generalization of Virtual 3D City Models with Dynamic Landmark Highlighting. In Proceedings of the 11th ICA Workshop on Generalization and Multiple Representation, Monpellier, France.

This work has been developed within the NEXUS project (Collaborative Research Centre 627 "Spatial World Models for Context-Aware Applications"), funded by the German Research Foundation (DFG).