reviewed paper

Exploring Crime Hotspots: Geospatial Analysis and 3D Mapping

Markus Wolff, Hartmut Asche

(University of Potsdam, Department of Geography, 3D Geoinformation Research Group; Karl-Liebknecht-Straße 24/25, 14476 Potsdam, Germany, Markus.Wolff@hpi.uni-potsdam.de, gislab@uni-potsdam.de)

1 ABSTRACT

This contribution presents a combined set of methods and techniques for geospatial analysis and 3D mapping of crime scenes. Geospatial clusters of robbery scenes are identified by applying methods of geoinformation science. Once hotspot areas are identified, relationships between robbery clusters and their spatial neighbourhood are analysed by including the urban context. For this purpose numerous geospatial data as well as a three-dimensional city model is included for analysis. To verify whether there exist any correlations between specific urban features and existent robbery clusters, statistical analyses are conducted. The results of these analyses are visualised within a three-dimensional geovirtual environment. At this point geospatial analysis is complemented with three-dimensional geovisualization techniques. This combination of crime mapping methods with innovative 3D geovisualization helps to facilitate an instant grasp of complex spatial phenomena in the field of crime mapping – for both, the public and responsible decision makers.

2 INTRODUCTION

Within the discipline of crime mapping geographic information systems (GIS) are widely used. Because crime scenes can virtually almost be localised in space – inside or outside of a building – a GIS can be considered as an adequate tool for managing and analysing crime data. Both, in academic research and in practical law enforcement GIS is applied for the analysis and the mapping of crime data (Murray et al. 2001). Digital analysis and mapping of crime offers a number of benefits, particularly in the following fields of applications: operational policing purposes, crime prevention, informing and interaction with the community, change monitoring in the distribution of crime over time and evaluation of efficiency of crime prevention initiatives (Hirschfield and Bowers, 2001).

Subsequent to geospatial analysis of crime scene datasets, the results have to be communicated to a broader audience. For this purpose thematic maps are created. Therefore, cartographic visualizations can be considered as fundamental to communicate the outcomes of crime analyses. However, those crime maps are predominantly presented in form of two-dimensional static maps. Frequently these maps show pattern- or feature distributions, for instance the spatial variation of crime hotspots related to certain offences.

The approach presented in this paper is twofold: at first robbery scenes are geospatially analysed (Section 2). This step contains the analysis of robbery scene patterns in order to discover regional clusters (Section 2.1). Once identified, in-depth analysis of the hotspot area is performed (Section 2.2). Positions of robbery scenes are statistically tested for spatial correlations with their particularly neighbourhood. For this purpose other geospatial data is included for analysis as, for instance, the road network and pedestrian frequencies.

The second part of the approach explores the potential of using interactive three-dimensional visualizations to communicate the findings of geospatial crime scene analysis to a broader audience – e.g. decision makers that might not be accustomed to map reading. Therefore a three-dimensional geovirtual environment of the study area is created. Into this environment the outcomes of crime scene analysis are integrated (Section 3).

While GIS is used for all kinds of spatial analysis, the interactive environment is modelled with a 3D visualization system. This process of linking GIS and 3D-VIS finally broadens the spectrum of geospatial crime scene analysis and crime scene mapping by facilitating an intuitive comprehension of complex geospatial phenomena – to both, decision makers in security agencies as well as for authorities related to urban planning.

3 DISCOVERING ROBBERY HOTSPOTS

This section deals with geospatial analysis of robbery crime scene data. Methods of geoinformation science are applied to process crime scene data and to reveal spatial clusters of crime. Furthermore, it is attempted to identify particularly spatial elements which might explain why a robbery hotspot exists in its given boundaries.

3.1 Determining the hotspots

Crime scene data for analysis is obtained from the police headquarters of the German city of Cologne. The dataset represents robbery crime scenes for the year 2007 whereas each robbery scene is represented as an individual point, geocoded by x and y co-ordinates. Beyond these co-ordinates each point has further attributes describing the time of the offence. To identify areas that are characterised by a higher crime density than other areas, hotspot analysis is conducted (Chainey and Ratcliffe 2005, McCullagh 2006, Ratcliffe 2004 cited in Boba 2005, Bowers et al. 2004).

Hotspot analysis is achieved by transforming the discrete point distribution of crime scenes to a continuous surface of crime scene density. For that purpose kernel density estimation (KDE) technique is applied (Smith et al. 2006, Williamson et al. 2001). Based on a given point dataset, this technique calculates a grid whose cell values represent density values related to a certain surface measure (for instance number of crimes scenes per square kilometre). For this purpose KDE-algorithms overlay a study area with a grid of user definable cell size. In a second step, density values for each cell are calculated – depending on the implemented kernel density function (cf. Smith et al. 2006). For the analysis presented in this paper the ArcGIS system, version 9.2, is used. Here KDE is implemented with a quadratic kernel density function:

$$g_j = \frac{3}{4}\left(1 - t^2\right), |t| \leq 1$$

with $t = d_{ij} / h$, h as bandwidth, i as robbery scene position

The value at each grid location gj with distance dij from each robbery scene i is calculated as the sum of all applications of the kernel function over all event points of the crime scene dataset. Therefore two parameters are crucial for every KDE-analysis and have to be specified: cell size and bandwidth. The cell size parameter defines the resolution of the resulting grid, the bandwidth describes the size of the search radius, i.e. how many crime scene locations (points) are used for analysis. A large bandwidth includes a larger area and therefore more points into analysis than a smaller bandwidth would include. Hence, a too large bandwidth might hamper the identification of smaller hotspots, while a too small bandwidth might result in many small clusters of crime. For KDE-analyses presented in this paper, a cell size of 20 meters and a bandwidth of 400 meters are considered as appropriate. However, the lack of rules and standards concerning reliable hotspot bandwidth parameterisation prompts Smith et al. (2006) to conclude that bandwidth selection "is often more an art than a science". The decision for the 400 meter search radius is taken predominantly as the result of experimental studies: compared with other settings, the 400 meter parametrisation produces the most reasonable output. The resulting hotspot grid reveals very clearly an inner-city hotspot-region while simultaneously preserving the overall representation of crime scenes (cf. Figure 1).

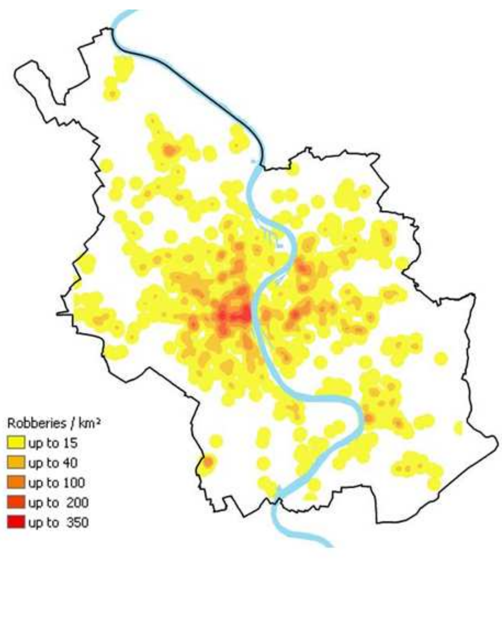

Fig. 1: 2007 hotspot grid with robbery densities defined as number of incidents per square kilometre.

Further hotspot analysis requires extracting hotspot boundaries from the KDE-grid. Since the grid represents a continuous surface of crime density the definition of discrete hotspot boundaries is not straightforward. However, to get a rough estimation of the boundary, focal neighbourhood statistic is applied. This method compares each pixel value of the grid to the values of its neighbours: each pixel of the KDE grid is compared to its 7 x 7 neighbourhood and the standard deviation of crime density is calculated. This results in a new grid, whose cell values represent standard deviation values of robbery scene densities. Using this method a gradient of crime scene density is represented. The higher the value, the higher is this gradient of an actual cell to its 7 x 7 neighbours. This value is finally used to detect the hotspot boundaries. Based on visual exploration, standard deviation equal to 15 is defined as the threshold value. Finally, the standard deviation grid is reclassified: a third grid is created where all cells with standard deviation < 15 become 0 while all cells > = 15 become 1. After vectorisation of this grid a simplification of the resulting polygons is proceeded. The result is polygonal boundaries of the three largest hotspots (cf. Figure 2).

Fig. 2: Simplified hotspot boundaries as identified with focal neighbourhood statistics.

3.2 In-depth analysis of particular hotspot regions

Since previous analysis identified three large hotspots, this section covers specific analyses of robbery scenes inside these areas. In a first step some overall characteristics of these regions are identified by applying GIS methods. For that purpose the distribution of several facilities in the city of Cologne (schools, restaurants, clubs, sights, banks and many more) is analysed for each particular hotspot region (cf. Table 1, 2 and 3).

Feature	Number
restaurant, diner	94
bank	12
club	9
public transport stop	8
parking place, car park	7
hotels	7
book shop	7
pharmacy	7
supermarket	4
theatre, cabaret	3
cinema	2
sight	1
school	1
consulate	1
church	1
shopping centre	1

Table 1: Total number of specific facilities in hotspot region one.

Feature	Number
restaurant, diner	151
hotel	46
parking place, car park	24
bank	23
public transport stop	21
sight	16
pharmacy	14
museum	12
church	10
book shop	9
theatre, cabaret	7
shopping centre	6
supermarket	5
car dealership	3
club	3
stage, arena	3
railroad station	2
indoor swimming pool	2
bus terminal	1
tourist information	1
consulate	1
post	1
town-hall	1
school	1
petrol station	1

Table 2: Total number of specific facilities in hotspot region two.

Feature	Number
bank	7
restaurant, diner	6
supermarket	6
pharmacy	5
car dealership	5
public transport stop	5
book shop	3
hotel	2
shopping centre	1
stage, arena	1

parking place, car park	1
police department	1
theatre, cabaret	1

Table 3: Total number of specific facilities in hotspot region three.

This analysis reveals distinct differences between the three hotspot regions. Except for restaurants and diners that are frequently found in all three regions, hotspot number one with its numerous clubs shows evidence for a nightlife district. Similarly characterised is adjacent hotspot region two: here tourism plays a major role due to its high number of hotels, parking places, museums and sights. Unlike hotspot three: banks, supermarkets, pharmacies and car dealerships point rather to the direction of a housing area.

Subsequently one can conclude that robberies in hotspot regions one and two might be related to pickpocket predominantly, while hotspot region three seems to be a kind of social hotspot. Overlaying the hotspot boundaries with a city map (cf. Figure 3) reveals that hotspot number one covers an area famous for its nightlife (“Rudolfplatz”, “Friesenplatz”). Hotspot two is located in the very city centre of Cologne – an area that is highly frequented by tourists and for shopping. Hotspot three finally is located in the district of Cologne-Kalk which is a former industrial location with high unemployment rates.

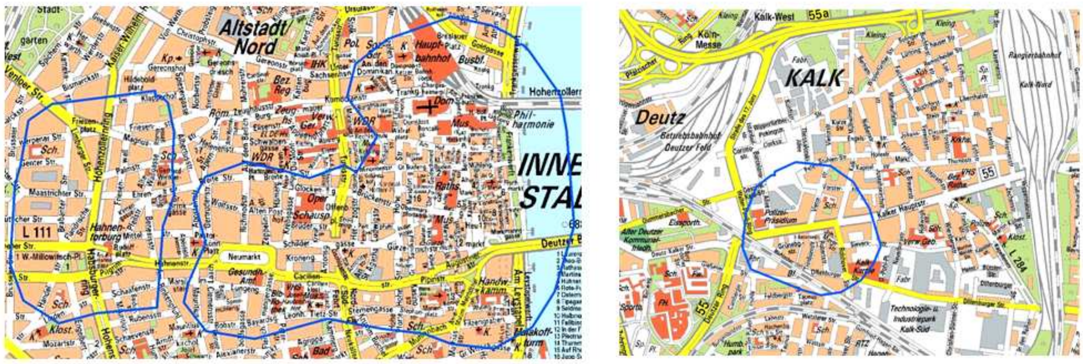

Fig. 3: Hotspot regions one, two (left image, blue outlines) and three (right image) on a city map.

Given the high pedestrian frequencies of inner-city pedestrian zones, position and frequency of robbery scenes are correlated with the number of pedestrians. It is expected, that many (few) robbery scenes can particularly be found near streets that are passed by many (few) people. To analyse this, a dataset is integrated that represents average pedestrian flows along every road segment within the city of Cologne as frequency values. This data is obtained from the FAW Frequency Atlas of the German Association for Outdoor Advertising (FAW). Frequencies are calculated as average values per hour on a working day basis for the years 1999 to 2005 (Data description FAW-frequency-Atlas 2006). Technically speaking, one FAW point exists with the corresponding frequency values for each road segment. Based on its geocoded coordinates this point-based FAW information is referred to the corresponding road segments via its unique segment ID. Subsequently each robbery scene is assigned to the closest road segment.

Afterwards the road network carries two new thematic attributes: the average frequency of pedestrians passing this segment per hour and the total number of robberies closest to it. Finally, robbery scenes and pedestrian frequencies are tested for correlation.

For the whole city of Cologne a weak but significant positive correlation between the number of offences and the number of pedestrians (Spearman's rank correlation coefficient = 0.202, significant for $p=0.01$) can be detected. Only little robbery is registered near to segments passed by a few pedestrians. However, by far the most robbery scenes are not located close to segments passed by the highest number of pedestrians. Instead, most robberies (as analysed for the whole city of Cologne) are committed close to segments passed by up to 45 pedestrians per hour (cf. Figure 4).

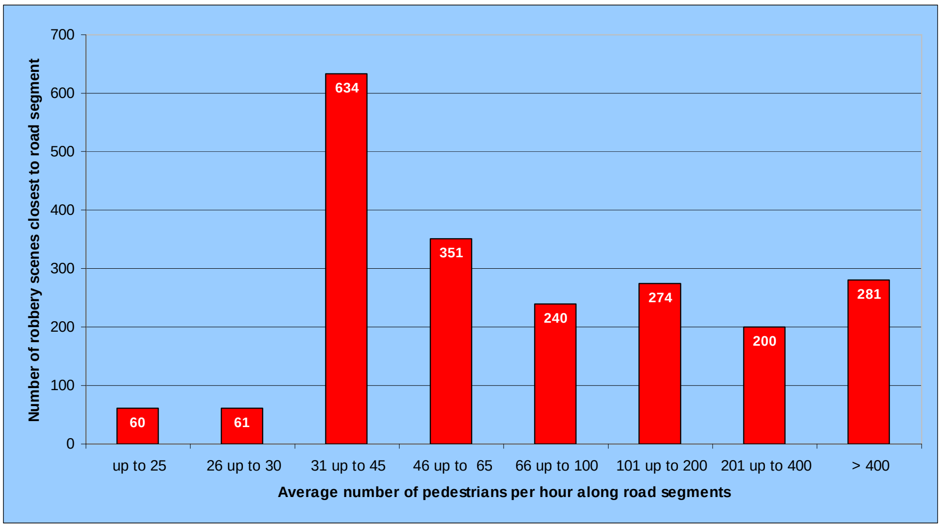

Fig. 4: Number of robbery offences compared to pedestrian frequencies along road segments for the whole city of Cologne.

Afterwards the same analysis is conducted for the subset of robbery scenes and road segments encircled by the hotspot boundaries. Compared to the results presented in Figure 4, different relationships between robbery positions and pedestrian frequencies are observed. Figure 5 shows that in most cases the closest street segment to a robbery scene is frequented by 400 and more pedestrians per hour. Roughly the half of all road segments with pedestrian frequencies equal to 400 and more pedestrians are located within the hotspot boundaries (296 from 588 segments).

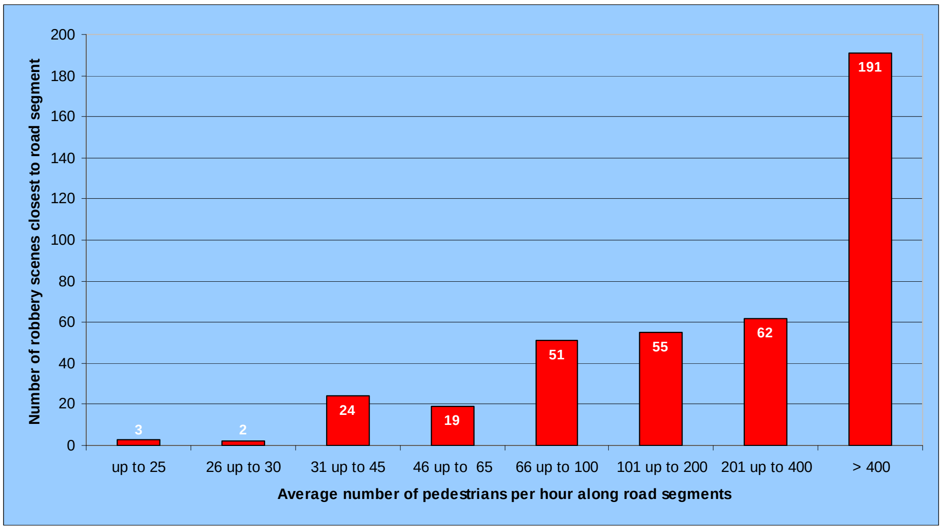

Fig. 5: Number of robbery offences compared to pedestrian frequencies along road segments for areas inside the hotpot boundaries.

4 THREE-DIMENSIONAL MAPPING OF CRIME SCENE ANALYSIS

This section deals with the presentation of geospatial crime scene analysis. To facilitate an instant grasp of these complex geospatial phenomena, the results analysis are visualised with a three-dimensional urban environment. To provide a basis for subsequent urban crime data visualization, a three-dimensional geovirtual environment is created for the city of Cologne. This geovirtual environment consists of a digital terrain model, a 3D city model, high resolution aerial photography (25 cm/pixel), digital cadastral map and further vector-based datasets including rivers, administrative boundaries and others (cf. Figure 6). Using GIS, all datasets are processed for 3D visualization. Afterwards the datasets are integrated into the LandXplorer software, an appropriate system for interactive three-dimensional visualizations (Döllner et al. 2006).

Fig. 6: Virtual three-dimensional environment of the city of Cologne.

For visual analysis of hotspot areas the hotspot grid is integrated as a three-dimensional surface into the 3D geovirtual environment. This thematic relief facilitates an intuitive exploration and interactive visual analysis of crime scene densities. In addition the surface can be overlaid with various geocoded textures – for instance with (classified) choropleth or isopleth maps of the hotspot grid or with topographic maps (cf. Figure 7). This multiple feature coding of crime scene densities can be considered as an effective visualization method to single out certain hotspot regions.

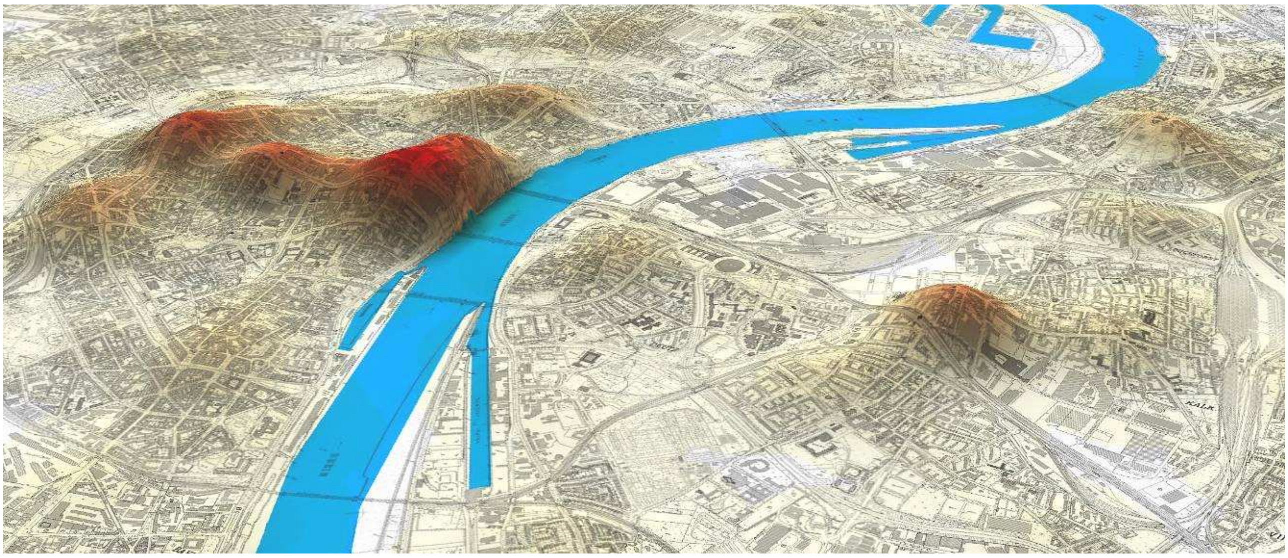

Fig. 7: 3D visualization of a classified KDE surface.

To allow for further analysis this virtual environment is extended by a 3D city model. The analytical and geovisual potential of 3D city models can be instrumental for decision makers working in security agencies concenring an instant comprehension of complex spatial phenomena related to urban security issues. In this study a city model is used that consists of approximately 22,000 buildings.

To facilitate geovisual analysis in terms of comparing single buildings with the robbery hotspots, the city model is combined with the KDE-hotspot grid. Figure 8 shows hotspot area number two with the corresponding 3D city model.

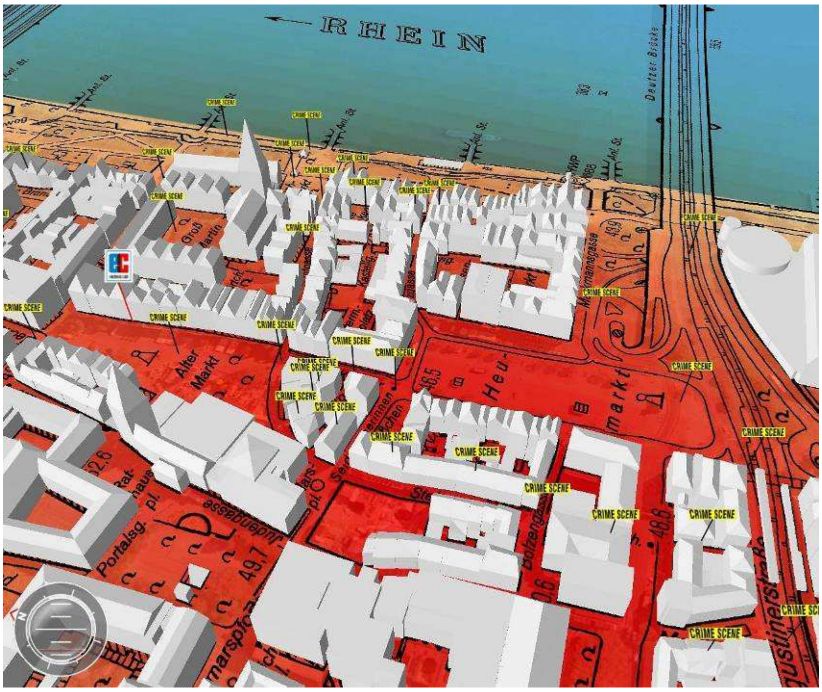

Fig. 8: 3D city model with additional hotspot texture and crime scene positions.

To broaden this visual approach and to facilitate further analysis, the minimum distance to the closest robbery scene is calculated for each building. Based on the crime scene dataset an Euclidean-distances-grid with a cell size of two meters is calculated. Each pixel of this grid represents the distance to the closest crime scene. This grid is combined with the city model: for each of the 22,000 buildings those pixels are detected that are included by the respective building footprint. From this set of pixels that one with the lowest value is determined – which is the minimum distance of the building to the closest crime site. This value is added to the building database as a new attribute.

Afterwards, the building dataset is classified and coloured according to these minimum distance values. The subsequent 3D visualization allows for exploring particular buildings of urban districts affected by a high number of robberies in their neighbourhood (cf. Figure 9). This visualization facilitates an intuitive geo-communication of each building's distances from the closest crime scene. Since the distance values are stored in the buildings database, further selections of buildings for analysis is supported.

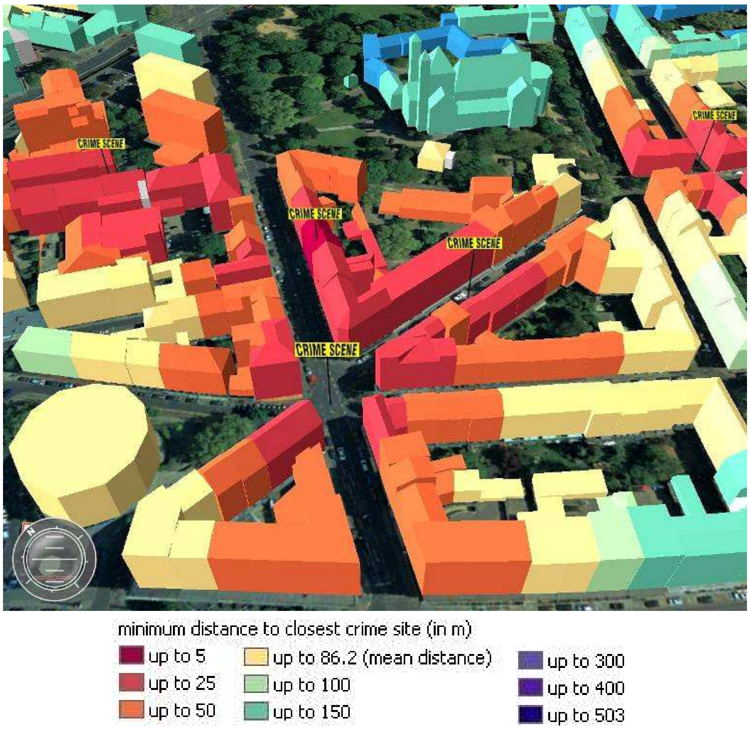

Fig 9: Minimum distances of each building to the closest robbery crime scene.

Finally, another 3D visualisation is produced to described that depicts the number of robbery scenes as assigned to the closest road segment (cf. Figure 10). In this Figure road segments are extruded according to the number of robbery scenes closest to it. The higher the segments, the more robberies are committed close to it. The colour of the segments represents pedestrian frequency.

Fig. 10: Robbery scenes aggregated to nearest road segment.

5 CONCLUSION

This paper presented an approach for exploring robbery hotspots by coupling geospatial crime scene analysis with 3D mapping methods. For this purpose robbery scenes were analysed for spatial clustering by applying KDE techniques. This led to hotspot identification. Subsequently a method was presented to determine the boundaries of hotspots. Based on these boundaries in-depth-analysis of certain hotspot regions was conducted. Against this background methods of geoinformation science were applied to typify these regions by analysing the distribution of certain urban facilities. However, for further studies the 4th dimension should be included in further analysis. Therefore next steps in this project will comprehend time related analysis of hotspot patterns.

6 ACKNOWLEDGEMENTS

Funding of this study by the German Federal Ministry of Education and Research (BMBF) within the framework of the InnoProfile research group ‘3D Geoinformation’ (www.3dgi.de) is gratefully acknowledged. The author also like to thank the police headquarters of the city of Cologne for providing extensive crime datasets. Furthermore the author like to thank Virtual City Systems GmbH for providing the 3D city model and 3D Geo GmbH (now: Autodesk, Inc.) for supplying the LandXplorer system and the German Association for Outdoor Advertising (FAW) for providing frequency atlas data.

7 REFERENCES

Boba, R. L. (2005), Crime Analysis and Crime Mapping. Thousand Oaks, California, Sage Publications

Bowers, K. J., Johnson, S. D., Pease, K., (2004), Prospective Hot-Spotting The Future of Crime Mapping? The British Journal of Criminology 44(5): 641-658

Chainey, S., Ratcliffe, J. (2005), GIS and Crime Mapping. Chichester, John Wiley & Sons Inc

Craglia, M., Haining, R., Wiles, P. (2000), A Comparative Evaluation of Approaches to Ur-ban Crime Pattern Analysis, Urban Studies 37(4): 711-729

Döllner, J., Baumann, K., Buchholz, H. (2006), Virtual 3D City Models as Foundation of Complex Urban Information Spaces. CORP, Vienna

Hirschfield, A., Bowers, K. (2001), Mapping and analysing crime data. London and New York, Taylor & Francis

McCullagh, M. J. (2006), Detecting Hotspots in Time and Space. ISG06

Murray, A. T., McGuffog, I., Western, J. S., Mullins, P. (2001), Exploratory Spatial Data Analysis Techniques for Examining Urban Crime Implications for Evaluating Treatment, British Journal of Criminology 41(2): 309-329

Ratcliffe, J. H. (2004), The Hotspot Matrix: A Framework for the Spatio-Temporal Targeting of Crime Reduction, Routledge, 5-23

Smith, de M.J., Goodchild, M. F., Longley, P.A. (2006), Geospatial Analysis, Troubador Publishing

Williamson, D., S. McLafferty, P. McGuire, T. Ross, J. Mollenkopf, V. Goldsmith and S. Quinn (2001). Tools in the spatial analysis of crime. Mapping and analysing crime data. In: A. Hirschfield and K. Bowers. London and New York, Taylor & Francis. 1: 187.