

SEMANTIC SEGMENTATION OF BUILDING ELEMENTS USING POINT CLOUD HASHING

M. Chizhova^{1,2,3,4,*}, A. Gurianov⁵, M. Hess¹, T. Luhmann², A. Brunn³, U. Stilla⁴

¹ Otto-Friedrich University of Bamberg, Bamberg, Germany - (mona.hess, maria.chizhova)@uni-bamberg.de

² Jade University of Applied Sciences, Oldenburg, Germany - (luhmann, maria.chizhova)@jade-hs.de

³ University of Applied Sciences Würzburg-Schweinfurt, Würzburg, Germany - ansgar.brunn@fhws.de

⁴ Technical University of Munich, Munich, Germany - stilla@tum.de

⁵ Ivanovo State University, Ivanovo, Russia - a.v.gur.2008@mail.ru

Commission II, WG 4

KEY WORDS: semantic segmentation, hashing, Orthodox church, point clouds

ABSTRACT:

For the interpretation of point clouds, the semantic definition of extracted segments from point clouds or images is a common problem. Usually, the semantic of geometrical pre-segmented point cloud elements are determined using probabilistic networks and scene databases. The proposed semantic segmentation method is based on the psychological human interpretation of geometric objects, especially on fundamental rules of primary comprehension. Starting from these rules the buildings could be quite well and simply classified by a human operator (e.g. architect) into different building types and structural elements (dome, nave, transept etc.), including particular building parts which are visually detected. The key part of the procedure is a novel method based on hashing where point cloud projections are transformed into binary pixel representations. A segmentation approach released on the example of classical Orthodox churches is suitable for other buildings and objects characterized through a particular typology in its construction (e.g. industrial objects in standardized environments with strict component design allowing clear semantic modelling).

1. INTRODUCTION

1.1 Motivation and problem definition

Point cloud segmentation is a common topic in point cloud processing. Many research projects deal with extraction of scene elements using geometrical and decision making methods for point cloud interpretation.

For the interpretation of point clouds the semantic definition of extracted segments from point clouds or images is a common problem. Its concept is outlined as semantic definition in many literature sources. Usually, the semantic of geometrical pre-segmented point cloud elements is determined using probabilistic networks and scene databases.

Current research focusses on semantic segmentation of building point clouds, which is based on human comprehension and understanding of building construction and knowledge about position and semantic (functional role) of particular building parts. In the context of this research, the expression *semantic segmentation* means the extraction in an distinguished church point cloud of sub-point clouds corresponding to particular building parts of this church with simultaneous definition of its semantics.

This work is carried out in the context of research for 3D reconstruction of destroyed Orthodox churches from incomplete point clouds. The reconstruction algorithm presented in (Chizhova et al., 2017b) includes some workflow steps; here wrong semantic segmentation leads to incorrect reconstruction results. In the context of Orthodox churches we speak about specific building canons and principles, on which Orthodox architecture relies.

A problem of church parts segmentation consists in geometrical complexity of such elements (usually a combination of different geometrical primitives and forms), that excludes pure geometrical segmentation.

The aim of this work is the development of an automatic semantic segmentation approach of church building parts in the 2. level of details - LoD2 (Figure 1), including human interpretation of such buildings. The key part of the procedure is a novel method based on hashing where point cloud projections are transformed into binary pixel representations. The proposed method can be extended for other architectural building types and objects with particular typology.

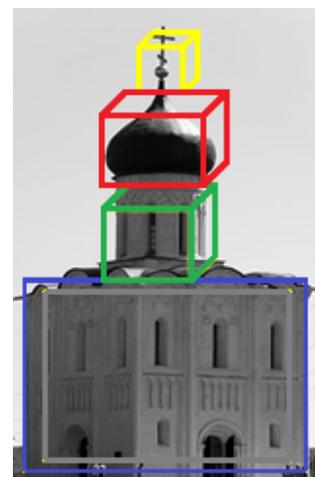


Figure 1. Semantic extraction of the church building parts

*Corresponding author

1.2 Research object

The specific feature of the Orthodox church architecture is a strict compliance with canonical religious requirements, which are based on the Orthodox dogmatics and architectural traditions.

Each Orthodox church building is canonically divided in three main spaces - altar, nave (or main space) and narthex (often with additional hall) - and has a typical set of building parts (Figure 2). Their position, relations, number and geometrical form fall within the strict compositional rules and can be semantically represented in a graph model (Chizhova et al., 2017b).

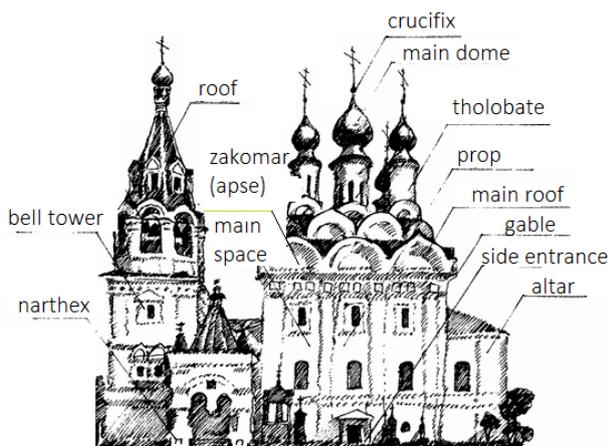
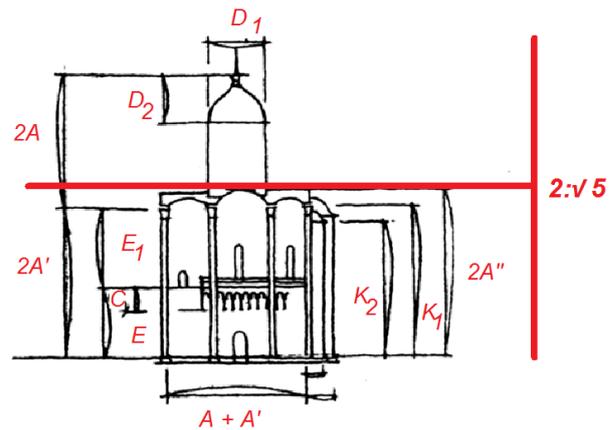


Figure 2. Typical building parts of the classical Orthodox Church

Structural completeness, significant proportionality and harmonic conformity of all building parts to one another and the whole were achieved through a specific joint building system. The term "golden ratio" in art and architecture is known from the Antiquity and defined later from Leonardo da Vinci. Proportions and relations, which are based on the "golden ratio", stand out due to high aesthetic properties corresponding to human and natural essence.

A significant contribution to the investigations of the Orthodox church architecture was provided by the Russian architects Zhelohovzev, Zholtovsky, Shmelev and Rybakov.

The "golden ratio" in the context of Orthodox churches has been formulated from the outstanding architect J. Zholtovsky (Zholtovsky, 1955) and called as "Zholtovsky-function". The "Zholtovsky-function" determines a rectangle, in which the contours of church main space are inscribed with relations $1 \div \frac{2}{\sqrt{5}}$ to one another. A line going through the center of the main dome rectangle and connecting north and south portals divides a main space rectangle in proportion $1 \div \frac{2}{\sqrt{5}}$ (Figure 3). The same relation determines a rectangle, in which the main dome props are inscribed. The outside diameter of the main altar is equal to the diagonal line of underdome rectangle and diameter of sacrificial altars - a longer side of this rectangle. A church height with domes is equal to the double length of longer rectangle side that contains the whole main space in the ground plan. The main space height divides the church height in proportion of $\frac{2}{\sqrt{5}}$. A portal height is equal to the diameter of central zakomar and is derived through "Zholtovsky-function" from its diagonal.



$$ab:cd = ef:gh = ij:kl = 1:(2:\sqrt{5})$$

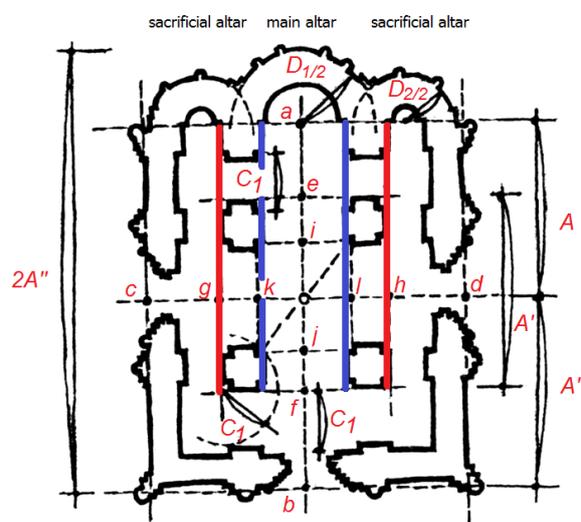


Figure 3. Proportioning system according to the "Zholtovsky-function"

According to another theory, the church building parts were constructed using native Russian units of measurements, which proportions represent geometrical multiplicity relations. (Rybakov, 1949) shows, that the most of native measurement units have been formulated through irrational relations that could be simply described using geometric constructing. This connection can be defined as proportion between square and its diagonal (Shevelev, 1973) in native Russian measurement units: "mernaya sazhen" 176,4 cm and "sazhen bez cheti" 197,2 cm. We can see that this relation $\frac{197,2}{176,4} \approx 1.11$ remains the "Zholtovsky-function" - $\frac{\sqrt{5}}{2} \approx 1.11$.

These facts demonstrate that the considered methods don't contradict each other. Their joint existence and application serve as template for further church building restoration works and investigation of Orthodox architecture.

1.3 State of the art

1.3.1 Semantic segmentation and classification of point clouds In the interpretation of 3D point clouds the most relevant problems are segmentation and semantic definition of segmented parts. Very often this process is defined as "semantical

segmentation". In this research and usually in the image processing community, it can be defined as semantics of geometrically segmented point cloud parts using probabilistic networks and scene databases.

(Weihaio and Yang, 2016) present the semantic segmentation of man-made objects (artefacts). This method has been developed for images where scene components have strong contextual connection to each other using fully connected CRFs (Conditional Random Fields) for its modelling. In this case, the idea of semantic segmentation consists of connecting information of low-level object features (on pixel-level between single points) with spatial scene information using datasets for semantic segmentation (e.g. eTRIMS and LabelMeFacade).

In (Niemeyer et al., 2016) a method of semantic segmentation has been developed for ALS point clouds (Aerial Laser Scanning). Input point clouds are hierarchically classified via CRF in two levels - point level and segment level. The process starts with a point-based classification, which serves as template/basis for scene classification. The segmentation process is considered as the detection context of connected components from the same class in the point cloud. The points are treated as CRF-vertices on the point-level (analog to segments as CRF-vertices on the segment-level). Extracted classes are recognized using analysis of 2D kNN-point neighborhoods and height differences between point segments using CEC (Conditional Euclidean Clustering).

(Roynard et al., 2016) present a classification approach for TLS (terrestrial laser scanning) point clouds, which allows ground extraction, segmentation of residual point clouds using connected components, extraction of geometrical features via histogram-based method, for example by a CVFH (Clustered Viewpoint Feature Histogram), GRSD (Global Radius-based Shape Descriptor), or ESF (Ensemble of Shape Functions), and point cloud classification using Random Forest. The principle of Random Forest insists in independent training of arbitrary number of decision-making trees (or each decision-making tree using random subset from a training dataset). In this path the point clouds of urban scenes will be classified.

(Riveiro et al., 2016) developed a method for automatic segmentation of arch bridges. The algorithm starts with the segmentation of vertical elements using azimuth-histogram estimation, which provides information about the presence of vertical walls in different orientations. Subsequently, a segmentation of non-vertical elements is carried out and during its processing particular object elements are extracted using logical operators based on construction of voxels and 3D images.

Research concerning the knowledge formalization approaches for describing heritage buildings has been deepened in the last ca 15 years. (De Luca et al., 2011) delivered an approach with an integrated platform containing the semantic description of architectural elements, based on morphological analyses of 3D elements including a description of hierarchical relations of semantic entities (building parts) and structural levels including a temporal component and hypothetical restitution (reconstruction). Hierarchical shape description of buildings and the application shape grammar in architecture for architecture modelling has been described by (Müller et al., 2006) and (Havemann and Fellner, 2011).

The process of semantic segmentation in (Hackel et al., 2016), based on extraction of geometrical 3D features in an approximated neighbourhood, is using eigenvalues and approximation

of 3D coordinates from context (covariance and histogram-based methods). In this case, the feature extraction has been accelerated using a scale pyramid in neighborhood approximation. For the classification of extracted features a provided learned Random Forest classifier is applied that computes conditional probabilities of different segments.

(Meidow et al., 2016) developed a method of semantic segmentation for aerial images based on CNN (Convolutional Neural Network). Each neuron from actual level receives input data with spatial localized aperture with a fixed size from previous level, and outputs a vector with different sums of values distributed through all neurons. Constructed networks have been trained using visual databases (e.g. ImageNet, Pascal VOC).

In (Yastikli and Cetin, 2016) a point-based classifier of different scene objects (buildings, vegetation etc.) in a point cloud is considered. The classification workflow is based on two approaches that apply particular feature types (e.g. spatial-based, echo-based or waveform-based). Hierarchical rules have been developed for each feature type.

As previously mentioned, the interpretation of point clouds is often similar to the human comprehension. A general study about human perception of building constructions based on geometrical building representation is presented in (Tutzauer et al., 2016).

1.3.2 Definition of building parts and architectural elements

In the description of architectural building parts and engineering structures, an agreed vocabulary for descriptive terms has been developed in many languages in order to communicate the appearance and geometrical design both with words and with drawings. For the visual products in particular there are drawing conventions to produce scaled drawings in two (plans, sections, facades) or three dimensions (axometry, isometry, perspective).

For the written or spoken description of architectural building parts, the Getty AATA (Art and Architecture Thesaurus) is an example of a semantic view and hierarchical view with definitions of terms in different languages with references (semantic view of the definition of *proportion* on Art and Architecture Thesaurus Full Record Display (Getty Research: Semantic view of the definition of "proportion" on Art & Architecture Thesaurus Full Record Display, 2004)).

There are many examples of visual descriptions and manuals of architectural building parts starting with Vitruvius (Vitruvius, 2001). The discipline of "Baustilkunde", German for the studies in stylistics of built environment and architecture, is a taught subject on the curriculum of architectural studies with illustrated textbook publications, for example by (Koch, 2006) and (Schwarz, 2017). For Russian Orthodox church buildings similar illustrated publications exist, such as (Hamilton, 1983) and (Kesler, 2003).

Buildings are described in their macro structure, and detailed down to their micro structure of the single construction element. The expression of the spatial and hierarchy of building parts has been used in many projects, be it new design or spatial analysis of existing buildings. An example for tree-like hierarchy of spatial and semantical description (ontologies) of buildings is the MonArch project in Germany (Digital Archive for monumental buildings) where an adaptable vocabulary of building parts are described in a so-called *partonomy*. This hierarchy is presented on a web-based user interface in connection with a relational

database for the storage and retrieval of documents and metadata of building and relational 2D maps and drawings highlighted in connection to the hierarchical nodes (Freitag and Schlieder, 2009) and (Stenzer et al., 2011).

1.3.3 Proportion in architecture and perception of architectural proportions First definitions of how buildings should be set up can be found in Vitruvius "De architecturae libri decem", where he describes symmetry and proportions as well as different types of cult and religious buildings (Vitruvius, 2001) and (Grütter, 2015).

The term "proportion" of a building is described by Getty's AATA as "Relation between respective parts or between parts and the whole, in a building or any work of art, whether considered purely visually or numerically" (Getty Research: Semantic view of the definition of "proportion" on Art & Architecture Thesaurus Full Record Display, 2004).

Especially the numerical consideration for the description of building parts will be used in this presented research project with the focus on creating automatic extractions by machine learning in conjunction with human perception.

In the modern architecture, the proportionality principles, which are applied in building projection, have prototypes in the past. One example is given by the elaboration of the prominent french architecture Le Corbusier. His proportionality system - the Modulor (Corbusier, 2004) - conveys aesthetic complete proportions and assembling diversity according to human dimensions. The Corbusier system has combined the experiences from modern and past West-European architecture. Such investigations have been started initially from the famous Italian mathematician Leonardo Fibonacci.

Studies of the perception of architectural proportions have been undertaken and described on the grounds of human visual perception on the one hand and from its psychology on the other hand (Grütter, 2015), (Jonak, 2015). This research, based on numeric programming for the classification of building elements, will be complemented with qualitative research through Human-Computer Interaction in Virtual Reality to support the numerical description of proportions and elements of Orthodox churches as future step by the authors.

It must be noted that an art historian, architect with an understanding of the structural and material composition of a building will have a trained and discerning look at the single architectural elements, and can thus perceive as well as discern (and draw and describe) the elements and the connected proportions.

2. METHODOLOGY

2.1 Algorithm outline

The developed semantic segmentation method is based on the psychological human interpretation of geometric objects, especially on fundamental rules of primary comprehension (Koffka, 1935), (Wertheimer, 1938), (Soegaard, 2012) of main object structures (e.g. long straight lines, simple geometries, vertical and horizontal orientation etc.). Primary comprehension of simple structures (without considering of smaller details, which is usually the second step of object interpretation) allows a human

operator (e.g. architect) to classify buildings into different building types and structural elements (dome, nave, transept etc.), including particular building parts, which are quite well visually detected.

In this case, the primary duty of semantical segmentation could be the extraction of the main building structures in a point cloud with simultaneous reduction of data volume. After the extraction, building structures will be compared with the construction of existing Orthodox churches, where position and semantics of particular church elements are already known. Hence, an input-church (ingested dataset of church point cloud) will be semantically segmented in analogy to a similar church found in the database. The steps of the proposed algorithms are:

1. Extraction of the main building (church) structures using point cloud hashing,
2. Church matching: search of a similar church in the prepared database of Orthodox church constructions with semantically pre-segmented elements,
3. Extraction of semantic segments in the form of bounding boxes analogical to the fitted church.

2.2 Point cloud hashing

For the extraction of main church structures, a method is proposed which is based on a mathematical operation, namely *hashing*.

Hashing is a transformation of an input data array of arbitrary size to data of fixed size using a particular function for its realisation. Input information is called "hash-key" and results of its operation are called as *hash* (or hash code).

2D hashing and 3D (spatial) hashing is a quite simple procedure used in data structuring, particle simulations and all tasks where something should be quickly found. Some examples of spatial hashing application can be space partitioning tasks for relation recognition between objects (Teschner et al., 2003) or reorganisation of spatial data structure (e.g. point clouds) in most efficient and short form for its further processing (Han et al., 2012), (Lalonde et al., 2007). An interesting approach for point cloud processing using Z-order-based partitioning as locality-sensitive hashing is presented in (Alis et al., 2016).

Thereby, the volume of input data will be transformed in a bit string with predefined string length, that allows not only reduction of data volume (which is an important feature in the context of point clouds and big data problem, as considered in research of (Liu et al., 2016)), but extraction of main structures too (or low frequency components of building from a view of signal processing theory).

In detail, the workflow can be presented as follows:

1. The 3D point cloud of an architectural object (in our case an Orthodox church) will be projected, i.e. flattened, onto three planes (XY, XZ, YZ), which corresponds to frontal, horizontal and profile projections. These projections are a standard in mechanical and architectural drawings for complete representation of voluminous objects.

- Through the point cloud projection, the concentration of large building structures (e.g. walls, tambours etc.) perpendicular to the current projective plane is high, and can therefore be extracted through the filtering of the current 2D point cloud projection. A concentration intensity can be understood as pixel intensity normalized at range from 0 to 1 and then filtered with from user defined threshold.
- The hash of the recognized point cloud will be calculated using three projections of the filtered church point cloud, which has been previously normalized and voxelised of the complete volume of the input point cloud. These projections will be transformed into bitmap grids with 24×24 pixels on a continuous grayscale. In this case a recognized hash key will be presented as a one-dimensional array of real numbers ranging from 0 to 1 depending on the intensity of corresponding bitmap pixels (Figure 4).

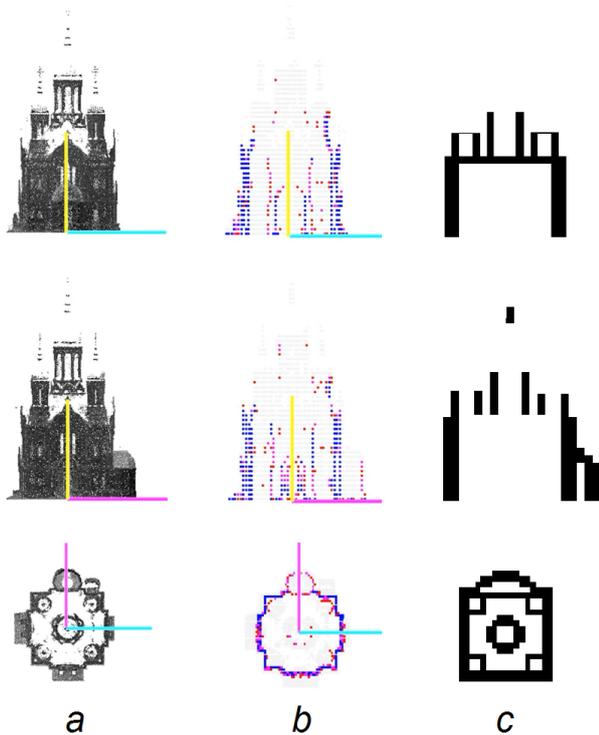


Figure 4. a - church point cloud, b - projection to be filtered, c - graphical hash (24×24 pixels)

Prior to the hashing the point cloud of an input church should be cleaned from extraneous objects and filtered for uniform point distribution (e.g. using VoxelGrid Filtering).

Analog to this principle, two additional (specifying) hashes are generated in the zones of altars and domes. Although, the resulting hash can be presented as a bit string resulting from the connection of one main and both additional hash-keys.

2.3 Creation of hierarchical database

As described below, filtered point cloud projections will be compared with the prepared database consisting information about localization and semantic of particular church elements. For the creation of database we have considered most common Orthodox church constructions described in (Kesler, 2003) (Figure 5).



Figure 5. Examples of typical church constructions: a - chetverik, b - ship, c - rotunda, d - cross-in-square (from: www.iamruss.ru)

These types of churches have been organized as hierarchical tree including some classification features (Figure 6).

1. Type of construction and nave

- Cross-in-square:
 - classical “crucifix” with symmetrical building parts in cardinal direction and
 - four-up “chetverik” (high-walled box, building parts are smaller and not symmetrical)
- Rotunda: classical rotunda, eight-sided build-up “vosmerik”
- Ship: the main property (feature) of this construction is that all elements (nave, entrance and bell tower) lie on the same axis.

- Number of altars: one (only main altar) or three (with two additional sacrificial altars),
- Number of domes: one (only main dome) or five (symbolising Christ and four Apostles).

Data in this structure are stored as hash-keys, which are binary (black and white) bitmap images resulting from three standard church projections. These bitmaps are convoluted in bit strings that will be connected in one string used as main hash-key for further church comparison.

2.4 Definition of the semantic segments in database

Classic Orthodox churches were built according to particular architectural canons, which define position and meaning of church elements. Moreover, sacral buildings have specific proportions and relations between elements. (Kesler, 2003) give detailed description of Orthodox architectural principles (typical building

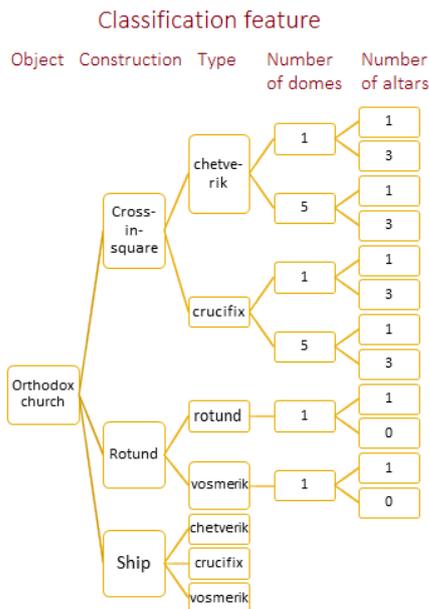


Figure 6. Example part of hierarchical church type tree

part forms, number and position of elements, proportions, common constructions etc.).

According to these principles we put into the database the region coordinates (or coordinates of bounding box in 3D) of the most probable position of building components (considering its proportion) and its semantic role. Thus, the database is pre-segmented and serves as template for further semantic segmentation of the input church.

In (Chizhova et al., 2017a) a method for detection of nave forms has been developed using wavelet analysis. Based on this approach, a nave position has been extracted in a graphical hash and its parameters (length, width, height) serve for proportional calculation and appointment of all probable church parts (Figure 7).

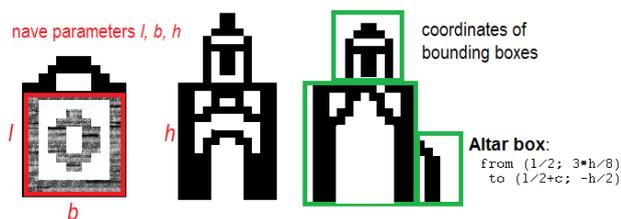


Figure 7. Coordinates definition of segment bounding boxes calculated from nave parameters

Appointed segments in a graphical hash will be transferred into the hash of the input church for its semantic segmentation.

2.5 Comparison workflow of hashed point clouds

Because of the search of fitting to an input-church in a database we need to compare the array of real numbers of corresponding to recognized patterns with a bit string of church hash-key from a database. Most obviously in this case would be an application of a particular threshold for the transformation of real number to integer and further calculation of Hamming distance.

The value (extent) of hash coincidence (or matching) is estimated through a “matching index” with its calculation based on the Hamming distance. The Hamming distance is an integer positive number that will be 0 (zero) only in the case of whole coincidence of compared bit strings of hashed point cloud projections and 1 (one) in the case of a complete mismatch. Such a relation between an absolute value of the Hamming distance and the string size could hinder the decision making procedure. Therefore, a calculated Hamming distance will be normalized through its division to the string length fitting its value in the range from 0 to 1. For better demonstration the index of strings matching I is applied instead of the normalized Hamming distance. This index is calculated by means of subtraction of Hamming distance from 1 and is equal to 0 in the case of whole hash mismatch, conversely - in the cases of whole coincidence (equation 1).

$$I = 1 - \left(\frac{HD}{Len} \right) \quad (1)$$

with

I - the index of strings matching;

HD - Hamming distance;

Len - length of bit strings.

During the upload of patterns, three hashes will be computed to each of them:

1. Main (join) hash - to all pixels in all three projections of a pattern;
2. “Altar” hash - to pixels included in the probable altar zone of the church in all three projections;
3. “Dome” hash - to pixels included in the probable zone of domes defined by architecture.

It allows to simultaneously calculate three matching indexes: joint index (I_t), altar index (I_a) and dome or cupola index (I_c). Since all indexes are normalized, it is quite simple to compute an integral index taking in account all church zones. In our program two versions of integral index are proposed:

1. Multiplicative integral index:

$$I_{mult} = I_t \times I_a \times I_c \quad (2)$$

2. Additive integral index:

$$I_{ad} = \frac{I_t + I_a + I_c}{3} \quad (3)$$

Both integral indexes remain in the range from 0 to 1, which allows a user to define a threshold of its recognition (decision making about its matching or mismatch) through a relative value from the same range. Moreover, a user can bring a weighted coefficient for each of the input indexes that allows to change their influence to an integral index in a particular extent. In general, considering the range of index values, a multiplicative index can be more sensitive to the difference of hashes in comparison to an additive index.

In preparation for the construction of a point cloud hash it is necessary to provide good conditions for a successful church query and comparison with hash-patterns in the database. In this case, a main deviation/error source could be variations in spatial church position or orientation as well as scaling and proportions of its projections.

For the minimization of such errors a procedure is required, which is based on the selection of the optimal object position, scale and proportions before the Hamming distance calculation.

Hash-keys stored in a database are discrete and, therefore, are of minor usability for continuous transformations. For this reason, we propose to vary point cloud parameters (e.g. spatial orientation) before its discretization. For numeric reduction of varied parameters we propose the following: since Orthodox churches are built according to the cardinal directions facing East and a deviation of its point cloud spatial position and orientation will be slight, a deviation correction can be made manual using a graphical interface before the point cloud discretization.

Visual control will be applied using the roof plan projection, on which a nave centre (e.g. projection of main dome) will be matched with coordinate origin, and then the nave wall will be oriented parallel to a coordinate axis using rotation about coordinate origin.

At the beginning, approximate values are appointed for point cloud hashing.

3. TESTS AND RESULTS

The algorithm has been tested on three classical Orthodox churches.

Figure 8 shows three Orthodox churches located in Kyiv and Wiesbaden. The Cathedral of the Transfiguration of the Saviour in Kyiv (Figure 8, a) was built from 2005 to 2010 and is an example for a modern but less complex Orthodox church. It was scanned by a Leica ScanStation C10 with 15mm resolution in object space which resulted in about 4 GB of 19.5 millions 3D points (see Figure 9). The Transfiguration Church in Kyiv (Figure 10) was scanned by a Faro Focus scanner with 6 GB data of 23.9 millions 3D points in a resolution of about 10mm. Finally, the St. Elizabeths Church in Wiesbaden (Figure 11) has been scanned with 7.6 mm resolution and 108.3 millions points.

As mentioned above, the result of the church point cloud semantic segmentation are the bounding boxes of particular church elements (Figs. 9, 10, 11). The segments of the input church have been extracted using pre-segmented graphical hashes of best fitted church, each segment has information about its semantic separately in a text file.

Figure 12 displays a part of the console output of the implemented program. It shows a matrix of Hamming distances between each of the different object elements, hence a value of 1.0 appear in the main diagonal of the matrix indicating the similarity with an object itself. In order to separate the extracted elements from each other, the Hamming distances should have a significant difference. As an example, the highlighted boxes in Figure 12 (CH-1-1-1 vs. CH-1-1-3 and K-1-1 vs. K-5-1) represent results where the differences of Hamming distances are too small for a confident decision. We can state that the implementation of two

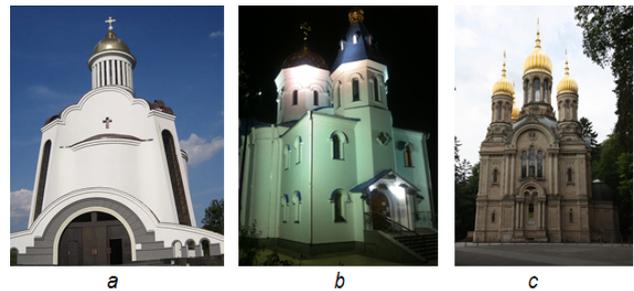


Figure 8. Three Orthodox churches used as test objects: *a* - Cathedral of the Transfiguration of the Saviour, Kyiv, Ukraine; *b* - The Transfiguration Church, Kyiv, Ukraine; *c* - St. Elizabeths Church, Wiesbaden, Germany.

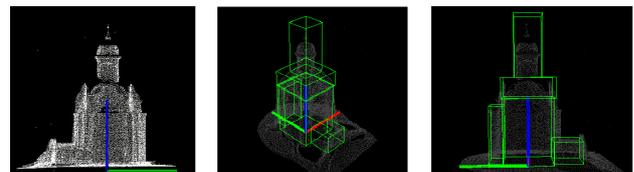


Figure 9. Segmentation results of church 1 (side view point cloud, side view with bounding box, axonometric view).

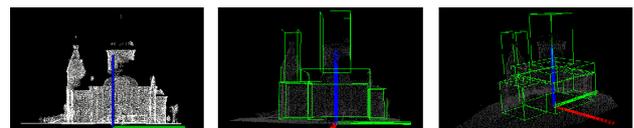


Figure 10. Segmentation results of church 2 (side view point cloud, side view with bounding box, axonometric view).

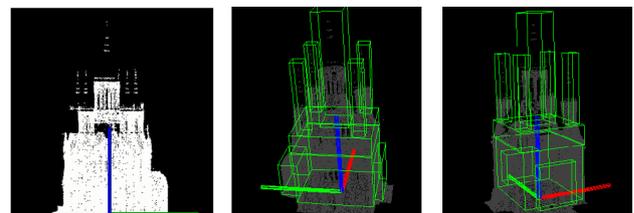


Figure 11. Segmentation results of church 3 (side view point cloud, side view with bounding box, axonometric view).

additional hashes in altar and dome zones will deliver better results in comparison to the application of only a main (join) hash: less difference between not similar and similar churches (Figure 12). Optimisation was carried out through precise localisation of probable hash distinction.

Extracted bounding boxes can differ slightly from a real building part size, but this not critical for the reconstruction process: smaller boxes will have enough parameters for element reconstruction; in bigger or overlapping boxes, points from other elements cause extraneous points (not corresponding to element, to be detected) to be eliminated as outliers. In a case of insufficient number of parameters needed for element recognition, a bounding box size will be optimized using manual manipulation and correction in HCI (Human-Computer-Interaction) in an interactive VR (Virtual Reality) environment (in development).

Hamming distance test:

	CH-1-1-1	CH-1-3	CH-5-1	CH-5-3	K-1-1
CH-1-1-1	1,000				
CH-1-3	0,863	1,000			
CH-5-1	0,649	0,640	1,000		
CH-5-3	0,712	0,719	0,670	1,000	
K-1-1	0,656	0,696	0,690	0,655	1,000
K-5-1	0,670	0,661	0,697	0,672	0,803
KCH-1-1	0,741	0,751	0,645	0,674	0,674
KCH-1-3	0,734	0,744	0,639	0,667	0,667
KCH-5-1	0,738	0,748	0,641	0,667	0,676
KCH-5-3	0,732	0,742	0,634	0,661	0,670
KK-1-1	0,652	0,667	0,630	0,655	0,590
KK-1-3	0,652	0,668	0,633	0,658	0,599
KK-5-1	0,652	0,668	0,650	0,619	0,613
KK-5-3	0,649	0,665	0,648	0,616	0,616
KW-1-0	0,748	0,752	0,650	0,686	0,650
KW-1-1	0,756	0,760	0,645	0,686	0,645
RO-1-0	0,752	0,764	0,650	0,714	0,718
RO-1-1	0,780	0,793	0,677	0,709	0,685
RR-1-0	0,744	0,764	0,692	0,704	0,684
RR-1-1	0,719	0,726	0,688	0,705	0,674

Figure 12. Console output: examples of low difference between church with 1 (CH-1-1-1) and 3 (CH-1-3) altars, 1 (K-1-1) and 5 (K-5-1) domes

In this test we applied large bounding boxes to segment some building blocks like the dome or the bell tower without precise extraction of smaller details. For example, on Figure 9 a dome block includes a crucifix, what is not suitable for the separated detection of a particular crucifix type. In Figs. 10, 11 the bounding boxes of church roofs include some parts of dome tambours and its props. A problematic situation, when the bounding box of a particular building part includes a completely another building part, is research in progress.

As mentioned above, this work is carried out in the context of destroyed Orthodox churches reconstruction. In this case, the success of building part extraction depends essentially on the destruction level of the whole building. Due to a low number of real church point clouds collected will model the most common destruction cases which have particular influence on classification process:

1. Abandoned churches that undergo a natural erosion due to "devastation" and absence of preservation or restoration. The characteristic of these constructions is a partwise or complete destruction of dome blocks and roofs. Thereby, an input church with completely destroyed dome blocks could have either 1 or 5 domes.
2. Blown up churches (explosion) with essential destruction of the main structures (nave or main space and altar walls, bell towers). In this case one extraction problem could arise due to an unknown number of altars and undetected type of construction, if the total building length have been essentially reduced.

The significant feature of previous church type classification is the presence of a bell tower (min on *one* hashed projection) and an object length on the ground plan projection that clearly defines a "ship" construction (or not). In similar constructions ("ship"-constructions and "not-ship" constructions) the building part blocks will be similar, too. This rule helps to reduce the number of further calculations by factor 2. Only in the case of blown up churches with a reduction of whole object length with a remaining proportion of approx. 1.7 between church length and width, a construction type of input church will not be defined with a high probability (by the proportion less than 1.7 an input church will be classified as "not ship"). Thereby, the probable cases could be:

1. undetected construction type ("ship" or not);
2. undetected number of domes (1 or 5);
3. undetected number of altars (1 or 3).

The number of altars or domes and construction type are defined as "undetected", if no detection feature has been found on three hashed projections. Coming from a particular case or a combination of these cases, the probability of a correct church type detection and its parts extraction can be estimated using the Bayes theorem (4) and statistic information from the hierarchical database with 20 churches.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (4)$$

with

$P(A|B)$ - ikelihood of correct church type definition considering already detected details (e.g. number of domes or altars etc.);

B - information about detected details;

$P(B)$ - probability (frequency) of such details in the database (e.g. 5-domes churches in the whole database);

$P(A)$ - probability (frequency) of details to be distinguished;

$P(A|B)$ - ikelihood of event B occurring given that A is true.

The proposed method needs detailed empirical sensitivity settings of such transformation on a quite big learning sample. Moreover, some problems can exist through classification of partly or very roughly scanned buildings, because a Hamming distance reflects all differences of compared images resulting from measurements errors and the absence or certain church elements.

For this reason, a matching index will be calculated using a partially normalized dot product of a real number array of pixel intensity from the recognized object, with a bit string of pattern in a database.

In this case, bitmap pixels corresponding to absent church elements of recognized patterns will have an intensity value close to 0. This automatically excludes relevant pattern pixels from the comparison procedure reducing general index value only due to the normalization procedure.

The algorithm for the index calculation (based on Hamming distance as well) is given as described in equation 5.

$$I_{new} = \frac{\sum_{j=1}^{P_s} X_j E t_j}{\sum_{j=1}^{P_s} E t_j} \quad (5)$$

There are two patterns to be compared: the recognized pattern of the input church X and the pattern from the database $E t$. In the whole pattern range P_s we calculate a sum of element-wise products (multiplications) and divide it with the number of nonzero bits in a pattern from the database.

This procedure makes our matching index non-symmetrical relative to patterns transposition. To save the symmetry property, normalization could be made on the product of nonzero bits number in both patterns, but in that case a matching index value will be too small.

For the moment, the method described above is released for the search using the main hash-key. Obviously, it is not possible to reach the absolute match of positions and scales of both patterns and thus full convergence of its hashes. Integral part of the recognition process is a method searching for the best matching between input pattern and a pattern from the database. This process should be done repeatedly varying patterns position and scale. For the purpose of increasing the reliability of matching results, an iterative comparison procedure for best hash fitting matching, based on preliminary variation of point cloud position and its scaling is carried out for each key in a database.

In such cases, the application of scanning methods is quite resource-intensive, because the number of varying parameters is big. That is why we propose a random search method (Monte-Carlo-Method) with limitation of tests number to about some thousands and the variation range of each parameter to about $\pm 16\%$ from its nominal value (it consists of four pixels in the 24×24 bitmap net) and random scaling in range from 0.5 – 1.2. It allows to apply a random search method for the definition of optimal pattern matching. The number of iterations was limited to 10000. In each test, matching indexes are calculated. Parameter values, by which the maximal index values will be reached, are stored for its further application in the segmentation process. Generally, the proposed segmentation method could be optimized using human-computer-interaction.

4. CONCLUSION AND FURTHER WORKS

As mentioned above, the proposed semantic segmentation approach is suitable not only for Orthodox churches but, in principle, for other buildings and objects characterized through particular typology in their construction. In this case, the method could be applied using a different hash presentation, pattern database and adapted to specific conditions (like additional hash matching of altar and dome zones).

In the context of destroyed objects a segmentation has particular limitations depending on the specific object. In case of Orthodox churches, a necessary condition could be the presence of bearing structures (walls, apsis) presented at least on roof plan projection (it means 50% of probable height building destruction). The reliability will be 30% less if a dome block is missed. In this case a cellular automata for reconstruction of the most probable church would be applied as presented in (Chizhova et al., 2017b).

The method is relatively sensitive to database construction; that is why some notes should be taken into account: For the construction of a hash-key in a database, it is necessary to consider not only the presence of the main structure but also its extension. For example, if the length of a nave wall is nearly equal to the building length, then its parallel wall at the exterior will be several orders shorter and therefore have less "contribution" in hash. Thereby, it is necessary to delete or "blur" such objects in a hash due to present a difference modulus for the calculation of the Hamming distance: if a "small" object with "intensity" of 0.1 will be graphically represented with black points in hash, by the comparison of similar churches their hashes will differ in 0.9

(and 0.1 if we delete such object from the graphical hash); this leads to wrong matching results.

For the definition of point cloud parameters, a random search method has been applied, which could be optimized through application of more efficient simulations methods such as unscented transformation (Julier, 2002) or bootstrap simulation (Efron and Tibshirani, 1993).

Generally, the proposed algorithm shows good results necessary for the semantic extraction of LoD2 building parts. However, some boxes have not exactly fitted a building part form, which should be corrected through optimization of semantic database or HCI support.

REFERENCES

- Alis, C., Boehm, J. and Liu, K., 2016. Parallel processing of big point clouds using z-order-based partitioning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B2*: 71-77.
- Chizhova, M., Gurianov, A., Brunn, A., Stilla, U. and Luhmann, T., 2017a. Wavelet-basierte erkennung architektonischer formen aus präzisen punktwolken. In: *Luhmann T / Schuhmacher C (Hrsg) Beiträge der Oldenburger 3D-Tage 2017*, 129 - 141.
- Chizhova, M., Korovin, D., Brunn, A., Stilla, U. and Luhmann, T., 2017b. Probabilistic reconstruction of 3d buildings using cellular automata. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, XLII(2/W3)*: 187-194.
- Corbusier, L., 2004. *The Modulor and Modulor 2: A Harmonious Measure to the Human Scale, Universally Applicable to Architecture and Mechanics*. Vol. 579, Birkhäuser, Basel.
- De Luca, L., Busayarat, C., Stefani, C., Veron, P. and Florenzano, M., 2011. A semantic-based platform for the digital analysis of architectural heritage. *Comput. Graph., Virtual Reality in Brazil 35*: 227-241.
- Efron, B. and Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 456 p. Macmillan Publishers Limited.
- Freitag, B. and Schlieder, C., 2009. MonArch: Digital archives for monumental buildings. *Kuenstliche Intelligenz, 4(2009)*.
- Getty Research: Semantic view of the definition of "proportion" on Art & Architecture Thesaurus Full Record Display, 2004.
- Grütter, J., 2015. *Grundlagen der Architektur-Wahrnehmung*. Springer Fachmedien, Wiesbaden.
- Hackel, T., Wegner, J. D. and Schindler, K., 2016. Fast semantic segmentation of 3d point clouds with strongly varying density. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, III(3)*: 177 - 184.
- Hamilton, G., 1983. *The Art and Architecture of Russia*. Yale University Press.
- Han, S., Jang, J., Kim, S. and Heo, J., 2012. Development of a hashing-based data structure for the fast retrieval of 3d terrestrial laser scanned data. *Computers & Geosciences, 39*.
- Havemann, S. and Fellner, D., 2011. *Rainbow of Computer Science, Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, chapter Towards a New Shape Description Paradigm Using the Generative Modeling Language, pp. 200–214.

- Jonak, U., 2015. *Architekturwahrnehmung*. Springer Fachmedien, Wiesbaden.
- Julier, S. J., 2002. The scaled unscented transformation. In: *Proceedings of the American Control Conference in Anchorage*, pp. 4555-4559.
- Kesler, M., 2003. *Pravoslavnye hramy i komplexy (Handbuch für Kirchenbau und -projektion)*. MDS 31-9.2003.
- Koch, W., 2006. *Baustilkunde*. Wissen Media Verlag.
- Koffka, K., 1935. *Principles of Gestalt Psychology*. Lund Humphries, London.
- Lalonde, J., Vandapel, N. and Hebert, M., 2007. Data structures for efficient dynamic processing in 3-d. *The international Journal of robotics research*, 26(8): 777-796. Research supported by ARL, Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-209912.
- Liu, K., Boehm, J. and Alis, C., 2016. Change detection of mobile lidar data using cloud computing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41: 309-313.
- Meidow, J., Hammer, H., Pohl, M. and Bulatov, D., 2016. Enhancement of generic models by recognition and enforcement of geometric constraints. *International Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III(3).
- Müller, P., Wonka, P., Haegler, S., Ulmer, A. and Gool, L. V., 2006. Procedural modeling of buildings. In: *ACM SIGGRAPH 2006 Papers, SIGGRAPH 06*. ACM, New York, NY, USA: 614 - 623.
- Niemeyer, J., Rottensteiner, F., Soergel, U. and Heipke, C., 2016. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI(B3): 655 - 662.
- Riveiro, B., DeJong, M. and Conde, B., 2016. An automatic method for geometric segmentation of masonry arch bridge for structural engineering purposes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI(B5): 719 - 724.
- Roynard, X., Deschaud, J.-E. and Goulette, F., 2016. Fast and robust segmentation and classification for change detection in urban point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI(B3): 693 - 699.
- Rybakov, B., 1949. Russische Längenmaße der xi-xv jhs. *Sowjetische Ethnographie*, 1: 69-71.
- Schwarz, T., 2017. *Wörterbuch der Architektur*. Reclam.
- Shevelev, I., 1973. *Logik architektonischer Harmonie*. Strojizdat, Moskau.
- Soegaard, M., 2012. *Gestalt Principles of form Perception*. Interaction design foundation.
- Stenzer, A., Woller, C. and Freitag, B., 2011. Monarch: digital archives for cultural heritage. In: *Conference: iiWAS - The 13th International Conference on Information Integration and Web-based Applications and Services, 5-7 December 2011, Ho Chi Minh City, Vietnam*.
- Teschner, M., Heidelberg, B., Müller, M., Pomeranets, D. and Gross, M., 2003. Optimized spatial hashing for collision detection of deformable objects. *Computer Graphics Laboratory, ETH Zurich*.
- Tutzauer, P., Becker, S., Fritsch, D., Niese, T. and Deussen, O., 2016. A study of the human comprehension of building categories based on different 3d building representations. *Photogrammetrie - Fernerkundung - Geoinformation*, 5-6: 319-333. - ISSN 1432-8364. - eISSN 2363-7145.
- Vitruvius, 2001. *Vitruvius: Ten Books on Architecture*. Cambridge University Press.
- Weihao, L. and Yang, M. Y., 2016. Efficient semantic segmentation of man-made scenes using fully-connected conditional random field. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI(B3): 633 - 640.
- Wertheimer, M., 1938. *A Source Book Of Gestalt Psychology*. Routledge And Kegan Paul Ltd., chapter Laws of organization in perceptual forms, p. 7194.
- Yastikli, N. and Cetin, Z., 2016. Classification of lidar data with point based classification methods. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI(B3): 441 - 445.
- Zholtofsky, I., 1955. *Projekte und Gebäude*. Staatlicher Verlag der Literatur für Bauwesen und Architektur, Moskau.