

**VORONOI CLASSIFIED AND CLUSTERED CONSTELLATION DATA  
STRUCTURE FOR THREE-DIMENSIONAL URBAN BUILDINGS**

**NOR SUHAIBAH BINTI AZRI**

**UNIVERSITI TEKNOLOGI MALAYSIA**

VORONOI CLASSIFIED AND CLUSTERED CONSTELLATION DATA  
STRUCTURE FOR THREE-DIMENSIONAL URBAN BUILDINGS

NOR SUHAIBAH BINTI AZRI

A thesis submitted in fulfilment of the  
requirements for the award of the degree of  
Doctor of Philosophy (Geoinformatics)

Faculty of Geoinformation and Real Estate  
Universiti Teknologi Malaysia

JANUARY 2017

*To my husband, son and ever-loving family...*

## ACKNOWLEDGEMENT

I would like to thank my academic supervisor, Professor Sr. Dr. Alias Abdul Rahman, for his great support throughout my PhD study. His high academic standards and great vision was always a source of inspiration to me. I would also like to thank Associate Professor Dr. Francois Anton from Geodesy Department, Technical University of Denmark (DTU) for his great supervision. This thesis wouldn't have been complete without the constant assistance and prodding from him. Despite the fact that we were located in different countries (Denmark and Malaysia) during the whole length of the project, he managed to assist and orchestrate everything in the best way. Special thanks to Associate Professor Dr. Darka Mioc from Geodesy Department, Technical University of Denmark (DTU) for her recommendations and advice in the learning process.

My very special thanks goes to my sponsors; Ministry of Higher Education Malaysia (MOHE) for the MyPhD, MyBrain15 scholarship and financial support during the whole period of my study. Thank you.

## ABSTRACT

In the past few years, the growth of urban area has been increasing and has resulted immense number of urban datasets. This situation contributes to the difficulties in handling and managing issues related to urban area. Huge and massive datasets can degrade the performance of data retrieval and information analysis. In addition, urban environments are very difficult to manage because they involved with various types of data, such as multiple types of zoning themes in urban mixed-use development. Thus, a special technique for efficient data handling and management is necessary. In this study, a new three-dimensional (3D) spatial access method, the Voronoi Classified and Clustered Data Constellation (VOR-CCDC) is introduced. The VOR-CCDC data structure operates on the basis of two filters, classification and clustering. To boost up the performance of data retrieval, VOR-CCDC offers a minimal percentage of overlap among nodes and a minimal coverage area in order to avoid repetitive data entry and multi-path queries. Besides that, VOR-CCDC data structure is supplemented with an extra element of nearest neighbour information. Encoded neighbouring information in the Voronoi diagram allows VOR-CCDC to optimally explore the data. There are three types of nearest neighbour queries that are presented in this study to verify the VOR-CCDC's ability in finding the nearest neighbour information. The queries are Single Search Nearest Neighbour query,  $k$  Nearest Neighbour ( $k$ NN) query and Reverse  $k$  Nearest Neighbour ( $Rk$ NN) query. Each query is tested with two types of 3D datasets; single layer and multi-layer. The test demonstrated that VOR-CCDC performs the least amount of input/output than their best competitor, the 3D R-Tree. Besides that, VOR-CCDC is also tested for performance evaluation. The results indicate that VOR-CCDC outperforms its competitor by responding 60 to 80 percent faster to the query operation. In the future, VOR-CCDC structure is expected to be expanded for temporal and dynamic objects. Besides that, VOR-CCDC structure can also be used in other applications such as brain cell database for analysing the spatial arrangement of neurons or analysing the protein chain reaction in bioinformatics applications.

## ABSTRAK

Dalam beberapa tahun kebelakangan ini, pertumbuhan kawasan bandar telah meningkat dan mengakibatkan terhasilnya jumlah set data bandar yang besar. Keadaan ini menyumbang kepada kesukaran dalam mengendalikan dan mengurus isu-isu berkaitan kawasan bandar. Set data yang besar dan banyak boleh menurunkan prestasi untuk mendapatkan semula data dan menganalisis maklumat. Di samping itu, persekitaran bandar sangat sukar untuk diurus kerana ia melibatkan pelbagai jenis data, seperti tema pengezonan berbilang jenis dalam pembangunan bandar bercampur. Oleh itu, teknik khusus diperlukan untuk mengendalikan dan mengurus data dengan lebih cekap. Melalui kajian ini, kaedah capaian spatial secara tiga dimensi iaitu kumpulan data klasifikasi dan pengelompokan *Voronoi* (VOR-CCDC) diperkenalkan. Struktur data VOR-CCDC beroperasi di atas dasar dua saringan iaitu klasifikasi dan pengelompokan. Untuk meningkatkan prestasi dapatan semula data, VOR-CCDC menawarkan peratusan pertindihan antara nod yang minimum dan kawasan liputan yang sedikit untuk mengelakkan kemasukan data yang berulang dan pertanyaan pelbagai laluan. Selain itu, struktur data VOR-CCDC dibekalkan dengan elemen tambahan maklumat jiran yang terdekat. Maklumat jiran yang telah dikodkan dalam gambar rajah *Voronoi* membolehkan VOR-CCDC menjelajah data secara optimum. Terdapat tiga jenis pertanyaan jiran yang terdekat dikemukakan dalam kajian ini untuk mengesahkan kemampuan VOR-CCDC dalam mencari maklumat jiran yang terdekat. Pertanyaan-pertanyaan tersebut adalah pertanyaan jiran terdekat tunggal, pertanyaan  $k$  jiran terdekat ( $k$ NN) dan pertanyaan  $k$  jiran terdekat secara songsang ( $Rk$ NN). Setiap pertanyaan diuji dengan dua jenis set data 3D, iaitu lapisan tunggal dan berbilang lapisan. Ujian telah menunjukkan bahawa VOR-CCDC menghasilkan amaun input/output yang paling sedikit berbanding dengan pesaingnya, iaitu 3D R-Tree. Disamping itu VOR-CCDC juga telah diuji untuk penilaian prestasi. Hasil ujian menunjukkan bahawa prestasi VOR-CCDC adalah melebihi pesaingnya dengan bertindak balas 60 hingga 80 peratus lebih cepat dalam operasi pertanyaan. Di masa hadapan, struktur VOR-CCDC dijangka akan dikembangkan untuk objek temporal dan dinamik. Selain daripada itu, struktur VOR-CCDC juga boleh digunakan untuk aplikasi lain seperti pangkalan data sel otak untuk menganalisis susunan spatial bagi neuron atau menganalisis reaksi rantaian protein dalam aplikasi bioinformatik.

## TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	<b>DECLARATION</b>	ii
	<b>DEDICATION</b>	iii
	<b>ACKNOWLEDGEMENT</b>	iv
	<b>ABSTRACT</b>	v
	<b>ABSTRAK</b>	vi
	<b>TABLE OF CONTENTS</b>	vii
	<b>LIST OF TABLES</b>	xi
	<b>LIST OF FIGURES</b>	xii
	<b>LIST OF ABBREVIATIONS</b>	xvi
	<b>LIST OF SYMBOLS</b>	xv
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Research Background	1
	1.2 Background to The Problem	3
	1.3 Problem Statement	6
	1.4 Research Questions	6
	1.5 Research Aim	7
	1.6 Research Objectives	8
	1.7 Research Design and Methodology	8
	1.8 Thesis Organization	13

<b>2</b>	<b>SPATIAL ACCESS METHOD FOR URBAN DATA MANAGEMENT</b>	<b>16</b>
2.1	Introduction	16
2.2	Urban Data Management in Three-Dimensional Space	17
2.3	Information Retrieval using Spatial Access Method	22
2.3.1	Quadtree	23
2.3.2	<i>k</i> -d Tree	24
2.3.3	Hilbert Space Filling Curve	26
2.3.4	R-Tree	29
2.4	Three Dimensional Spatial Access Method	31
2.5	Nearest Neighbour: Methods and Techniques	34
2.5.1	Linear Search	34
2.5.2	Space Partitioning	35
2.5.3	Voronoi Diagram	44
2.6	Clustering Algorithm	45
2.7	Proposed Structure of 3D Spatial Access Method for Urban Applications	46
2.8	Summary	47
<b>3</b>	<b>DESIGN AND DEVELOPMENT OF VORONOI CLASSIFIED AND CLUSTERED DATA CONSTELLATION (VOR-CCDC)</b>	<b>48</b>
3.1	Introduction	48
3.2	The Development Phase of Voronoi Classified and Clustered Data Constellation (VOR-CCDC)	49
3.2.1	First Phase: Classification	50
3.2.2	Second Phase : Clustering	58
3.2.3	Third Phase : Voronoi Partitioning	64
3.2.4	Fourth Phase: Hierarchical Tree Structure	66
3.3	Summary	69



<b>4</b>	<b>QUERY PROCESSING AND NEAREST NEIGHBOUR VERIFICATION</b>	<b>70</b>
4.1	Introduction	70
4.2	Semantically Identified Dataset	71
4.3	Spatially Identified Dataset	73
4.4	Nearest Neighbour Query Processing	75
4.4.1	Single Nearest Neighbour Search	75
	Example 1: Putrajaya Administrative Centre	77
	Example 2: Urban Mixed-Use Development	79
4.4.2	$k$ Nearest Neighbour Query ( $k$ NN)	80
	Example 1: Putrajaya Administrative Centre	82
	Example 2: Supply Chain Management	83
4.4.3	Reverse $k$ Nearest Neighbour Query ( $Rk$ NN)	86
	Example 1: Putrajaya Administrative Centre	88
	Example 2: Business and Market Analysis	89
4.4.4	Nearest Neighbour of VOR-CCDC and 3D R-Tree	92
4.5	Summary	95
<b>5</b>	<b>ANALYSIS AND PERFORMANCE EVALUATION</b>	<b>96</b>
5.1	Introduction	96
5.2	Coverage and Overlapping Node Analysis	97
5.2.1	Data Updating Analysis	99
	Case 1: Data Insertion	100
	Case 2: Data Deletion	103
	Case 3: Data Modification	105
5.3	Page Access or I/O Analysis	107
5.4	Query Response Time	110
5.5	Summary	115
<b>6</b>	<b>CONCLUSIONS</b>	<b>116</b>
6.1	Introduction	116
6.2	Discussion and Findings	117

6.3	Research Contributions	120
6.4	Recommendations for Future Research	121
6.5	Limitations of the Study	122
6.6	Conclusion	123
<b>REFERENCES</b>		<b>124</b>

## LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Comparison of 2D and 3D data size	20
3.1	The description of 3D Voronoi properties	67
4.1	$k$ NN from point $q$	83
4.2	Minheap $H$ List	85
4.3	$k$ NN for Supplier $q$	86
4.4	Minheap $H$ during the filtering process of $Rk$ NN	92
4.5	$k$ NN Query of VOR-CCDC and 3D R-Tree	93
5.1	Overlap percentage between nodes	88
5.2	Overlap percentage and number of nodes using different approaches.	104
5.3	Query Response time of VOR-CCDC and 3D R-Tree for $k$ number of objects	106
5.4	Query Response Time of $k$ NN Query	110
5.5	Query Response Time of $Rk$ NN Query	114

## LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1	Research design and methodology	9
1.2	Sample of single layer 3D datasets	10
1.3	Sample of multi-layer 3D datasets	11
2.1	Growth of megacities (National Geographic, 2009)	19
2.2	Densely populated area in Hong Kong. Images are courtesy of Michael Wolf (Booker, 2013)	21
2.3	Quad-tree data structure (Rigaux <i>et al.</i> , 2002)	23
2.4	$k$ -d tree data structures (Gomes <i>et al.</i> , 2009)	25
2.5	Adaptive $k$ -d Tree data structures (Bærentzen <i>et al.</i> , 2012)	26
2.6	First four iterations of Peano Curve	27
2.7	The first three stages in generating Hilbert Curve	28
2.8	R-Tree representation (Manolopoulos <i>et al.</i> , 2006)	30
	Based	
2.9	MBV in 3D R-Tree Structure (Gong, 2011)	32
2.10	Comparison of 3D R-Tree methods. (a) Original algorithm of 3D R-Tree and (b) Improved algorithm (Zhu <i>et al.</i> , 2007)	33
2.11	Improved 3D R-Tree algorithm over Original or Classic R- Tree and R*Tree (Gong <i>et al.</i> , 2009)	33
2.12	Nearest neighbour information retrieval using Linear Search	35

2.13	Nearest neighbour information retrieval using $k$ -d Tree (Danforth, 2015)	37
2.14	Nearest neighbour information retrieval using Depth First Algorithm (Tao et al., 2002)	39
2.15	Example of the MINDIST and MINMAXDIST metrics between a 2-dimensional query point $Q$ and two MBRs (Manolopoulos <i>et al.</i> , 2010)	41
2.16	Illustration of mindist; mindist( $p_1, r$ ) and mindist( $p_2, r$ )	42
2.17	Nearest neighbour information retrieval using Best First Search (Tao, 2015)	43
2.18	Voronoi diagram of eight points in two-dimensional space (Aurenhammer, 1991)	45
3.1	Four phases of VOR-CCDC	50
3.2	Conceptual models of horizontal and vertical mixed-use development	52
3.3	Classification result for urban mixed used development based on zoning themes	52
3.4	Bounded objects in parallelepiped	53
3.5	The process and concept of supply chain management	55
3.6	Data model diagram of the SCM dataset	56
3.7	Groups of classification for supply chain management in an urban area	57
3.8	Interior and exterior parallelepiped	58
3.9	Cluster centre $C_i$ using different approaches	61
3.10	Comparison of $k$ -means and $k$ -means++ crisp clustering	62
3.11	Clustered buildings based on classified residential group (Residential-R1)	63
3.12	Clustered customer's location based on classified customer group (Customer-C1)	64
3.13	3D Voronoi diagram (a) and Voronoi cell (b)	65
3.14	The VOR-CCDC structure	68
4.1	Retrieved records for completed status of residential units	72

4.2	Possible units for City District Office in Building 20	74
4.3	Putrajaya Administrative Centre	77
4.4	3D building of Putrajaya Administrative Centre	77
4.5	Nearest building to Point $Q$	78
4.6	Possible dine-in restaurants near Unit A	79
4.7	Voronoi diagram of potential restaurants from Unit A	80
4.8	Point $T$ as a point query $q$ of $k$ NN query	82
4.9	$k$ NN building from point $q$	83
4.10	(a) Location of customers and supplier $q$ and (b) Voronoi diagram of customer and supplier $q$ location	84
4.11	Point $S$ in Putrajaya Administrative Centre	88
4.12	$Rk$ NN query for Point query $S$ in Putrajaya Administrative centre	89
4.13	Location of Hot Pizzeria and other fast food franchisers in an urban area	91
4.14	Voronoi diagram of Hot Pizzeria and other franchisers locations	91
4.15	3D buildings in an urban area	93
4.16	VOR-CCDC of 3D buildings	94
4.17	MBV of 3D R-Tree for a set of buildings	95
5.1	Comparison of overlapping percentage among nodes	98
5.2	Percentage of overlapping nodes for varying numbers of objects	98
5.3	Comparison of coverage percentage using various data constellation methods	99
5.4	Clustered buildings using 3D geospatial data clustering with $M = 4$ and $m = 2$	101
5.5	Infilled building on the empty parcel in between MBV R and MBV Q	101
5.6	Overlap percentage among parallelepipeds	102
5.7	Percentage of overlap between nodes for the process of data insertion	102
5.8	Buildings $e, f, g$ and $h$ need to be demolished for new development plan in an urban area	103

5.9	Total overlap percentage during the process of data deletion	104
5.10	Buildings $i$ and $j$ are proposed for restructuring plan and renovation	105
5.11	Comparison of overlapping Parallelepiped $P$ , $Q$ and $R$ , after the renovation or expansion of buildings $i$ and $j$	106
5.12	Total overlap percentage during the process of data modification	107
5.13	I/O vs $k$ number of objects for $k$ NN query	108
5.14	I/O vs $k$ number of objects for $Rk$ NN query	109
5.15	Response time analysis for $k$ number of objects retrieval in search operation using (a) VOR-CCDC and (b) 3D R-Tree	111
5.16	Response time analysis for $k$ NN query operation	113
5.17	Response time analysis for $Rk$ NN query operation	114

## LIST OF ABBREVIATIONS

2D	-	Two-dimensional
3D	-	Three-dimensional
MBR	-	Minimum Bounding Rectangle
MBV	-	Minimum Bounding Volume
GIS	-	Geographical Information System
NN	-	Nearest Neighbor
TLS	-	Terrestrial Laser Scanning
LiDAR	-	Light Detection and Ranging
NW	-	North West
NE	-	North East
SW	-	South West
SE	-	South East
NASA	-	National Aeronautics and Space Administration
MODIS	-	Moderate Resolution Imaging Spectroradiometer
VOR-CCDC	-	Voronoi Classified and Clustered Data Constellation
MXD	-	Mixed-Use Development
SCM	-	Supply Chain Management
$k$ NN	-	$k$ Number of Nearest Neighbour
$Rk$ NN	-	Reverse $k$ Nearest Neighbour
VorID	-	Voronoi ID
VorNbr	-	Voronoi Neighbour
SQL	-	Structured Query Language



## LIST OF SYMBOLS

$k$	-	number of classes or number of candidate
$m$	-	minimum entry
$M$	-	maximum entry
$I$	-	$n$ -dimensional rectangle
$\mathbb{R}^3$	-	Three-dimensional space
$C_i$	-	Cluster Centre
$p$	-	points
$V_{Cell}$	-	Voronoi cell
$dist$	-	distance
$x_{max}$	-	Maximum value of x
$y_{max}$	-	Maximum value of y
$z_{max}$	-	Maximum value of z
$x_{min}$	-	Minimum value of x
$y_{min}$	-	Minimum value of y
$z_{min}$	-	Minimum value of z
$\emptyset$	-	Potential Function
$\sum$	-	Sum of
$\leq$	-	Less or equal to
$\geq$	-	Bigger or equal to
$dist(p',q)$	-	Distance from $p'$ to $q$
$\in$	-	Subset of
$S$	-	Set of Objects
$\forall$	-	For all
$H$	-	Minheap table

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 Research Background**

The number of megacities or large conurbations has been rapidly increasing over the last few decades. This phenomenon is known as the urbanisation process. 'Urbanisation' is a process of physical growth in urban areas as opposed to rural areas (Collier, 2006). According to a United Nations (UN) Report in 2012, there are nineteen megacities around the globe. This number is expected to increase to twenty seven by the year 2020. Managing the rapid process of urbanisation in a sustainable way is a big challenge. Rapid and uncontrolled urbanisation threaten communities with negative issues such as social problems, environmental pollution, deforestation, and global warming (Kennish, 2002). In recent years, the wave of urbanisation has grown, and it cannot currently be stopped. The only possible way to reduce its negative effects is to adequately control and manage urban areas.

Urban planners and geospatial professionals play a very significant role in the transformation of urban areas into beneficial living environments. Each urban object requires a well-planned and strategic location in the urban area. The location of urban objects such as buildings, road networks, communication towers, and bridges is very important in urban planning because it can affect the urban community in many ways, such as through lighting conditions, building air flow, the urban skyline and natural wind flow (Akdağ and Bostanci, 2013, van der Hoeven and Nijhuis, 2012). Umakoshi and Gonçalves (2011) suggest that the location of buildings in urban areas is very important in reducing the negative effects of urbanisation: a qualitative analysis of the effect of five tall buildings showed that the strategic location of buildings provides the surrounding area with good air ventilation and facilitates optimum sunlight for enjoyment during daytime. Studies such as these demonstrate that it is important to appropriately manage urban areas to provide positive and flourishing living environments.

Advanced systems such as geographical information systems (GIS) are used by urban planners and geospatial professionals for location analysis. The spatial and semantic data for urban areas is gathered in the system for further processing and examination, however, the process is not straightforward. Several factors need to be considered, such as data management, data constellation, data standardisation and data maintenance. The emphasis in this research is on managing the urban spatial data efficiently. Since urban applications are usually associated with large volume of data, an optimised data structure will be proposed in this research in order to maximise the efficiency of data and information retrieval.

This chapter is organised as follows. In Section 1.1, the background to the study is discussed. The research problems were investigated and discussed in Section 1.2. The research questions are presented in Section 1.3 and the research aims are presented in Section 1.4. Research objectives are presented in Section 1.5, research flow in Section 1.6 and research methodology in Section 1.7. This chapter ends with a description of the thesis structure in Section 1.8.

## 1.2 Background to the Problem

Two-dimensional Geographical Information System (2D GIS) has been used to assist planners and geospatial professionals in the handling of spatial objects and performance of spatial analyses such as land suitability analysis (Akıncı *et al.*, 2013), urban environment quality (Joseph *et al.*, 2014), and green infrastructure planning (Chang *et al.*, 2012). Urban environments are complex and complicated; however, and there are cases and situations that are beyond the limits of 2D GIS. For instance, object or location identification, such as grocery shops in multi-storey buildings, may be misidentified if there are other units, such as residential space or offices, above them. The limitations of 2D GIS when visualising and handling three-dimensional (3D) data were discussed many years ago (Stoter and Zlatanova, 2003; Zheng *et al.*, 2011; Zlatanova, 2000; Zlatanova *et al.*, 2002).

As a result of the above issues, the 3D Geographical Information System (3D GIS) is proposed as the best option for handling 3D data. The uses and applications of 3D GIS have rapidly expanded in the last several years, as a result of detailed information, realistic views of the environment and advanced spatial analysis such as noise prediction analysis (Pamanikabud and Tansatcha, 2009), urban water run-off analysis (Mohamad Yusoff *et al.*, 2010), and urban solar radiation analysis (Liang *et al.*, 2014). 3D GIS seems to be the best option for handling and analysing data from urban applications. Most of the objects in urban areas are in 3D, such as multi-storey buildings, the position of telecommunication towers, multilevel highways or roads, and underground roads or tunnels. The implementation of 3D GIS for urban applications could improve users' understanding towards the urban environment and improve the potential of urban analysis and its information.

Even though 3D applications could resolve the limitations of 2D applications, users might face problems when dealing with the voluminous of 3D data. 3D data is typically larger than 2D data due to its greater geometric dimensions. The size of 3D urban data for urban applications, is dependent on many factors, such as the Level of Detail (LoD) in a building, building density in urban areas and other urban objects

such as street furniture. According to Hsu (2014); Krawczyk *et al.* (2015); and Tian and Zhao (2015), the huge size of 2D datasets resulted in users experiencing difficulties associated with data installation and maintenance, fault tolerance, and low performance during data retrieval and analysis. The same issues were also found in 3D applications. According to Nguyen and Saupe (2001) and Reed (2005), the large volume of 3D data will mean that users are confronted with performance issues and this will affect its visualisation. Living in the era of ‘big data’ has made the situation even more crucial and complex. Vast amounts of big data need to be stored and maintained in the database. More issues involving geospatial big data, such as challenges in storing, managing, processing, analysing, visualising and verifying the quality of data can also be found in Li *et al.* (2015).

Spatial access methods are used in spatial databases to organise records and optimise data retrieval. Spatial access methods for 2D objects have been successfully developed for spatial databases such as *k-d* Tree, BSP Tree and space filling curves (Oosterom, 1999). Although only a few have been implemented within commercial databases, users still have several choices of 2D spatial access methods for use in their applications. The need for a 3D spatial access method for 3D data management was noted by scholars several years ago (Arens *et al.*, 2005; Deren *et al.*, 2004; Kofler and Gervautz, 2000; Wang and Guo, 2012; Zhu *et al.*, 2007 and Zlatanova, 2000). There have only been a few dedicated spatial access method structures designed for 3D objects until now, such as Octree and 3D R-Tree. Many issues can be raised regarding the current 3D spatial access method structures. One of the issues is the current framework which is based on 2D but was dedicated to 3D objects (Arens *et al.*, 2005). According to Gu *et al.* (2011), 3D data that was processed based on the expansion of the 2D spatial access method may affect data retrieval during queries. Along with the development of 3D GIS, 3D spatial access methods technology is attracting scholarly attention. For example, studies on Octree for 3D applications (Koes and Camacho, 2015, Xu *et al.*, 2015) and 3D R-Tree (Kofler and Gervautz, 2000, Lei, 2004) have been proposed. Based on the studies by (Kofler, 2000) retrieval efficiency is still low when using 3D R-Tree to render the visualization of 3D objects.

Spatial access methods are also being used together with the execution of the query statement, which helps in improving query operations. There are many types of queries and analyses that could be performed for spatial data. An important and well-known queries and analyses is Nearest Neighbour (NN). According to (Taniar and Rahayu, 2013), NN is the most common query in spatial databases. An NN query finds and sorts a list of the nearest locations from a given point ( $P$ ). NN query is also used for urban applications. For example, Chatzimilioudis *et al.* (2012) used All  $k$  Nearest Neighbour, to query smartphone networks for enhancing public emergency services in urban areas, which allow users to send SOS (Save Our Souls) signals to the closest rescuers. In normal practice, spatial access methods are used together with the NN query to retrieve information from the database; however, currently successful algorithms such as R-Tree (Guttman, 1984) usually result in the unnecessary investigation of many nodes, meaning that none or only a small subset of the candidate list belongs to the actual result set of NN. This situation will result in excessive Input/Output (I/O) operations, which are inefficient for data retrieval. The situation will become more important when dealing with 3D objects due to the amount of volume and very detailed geometry of the objects.

The earlier discussions in this section and in the previous section have clearly shown that there are several issues in managing urban data in 3D space. Motivated by the above observation, this research proposes a new spatial access method for the applications of urban data in 3D space. The proposed structure is enable to tackle the issues on the management of urban data in 3D spatial databases and nearest neighbour identification. Three major cases or scenarios from urban areas are used as a case study in this research. The first case is urban mixed-use development, the second is supply chain management and the third case is business and market management. Further description and the issues involved in each case can be found in Chapters 3, 4 and 5.

### **1.3 Problem Statements**

The current trends in spatial technologies mean that urban data is best managed and analysed in 3D; however, 3D urban spatial data is typically large in volume and causes a deterioration in performance during data retrieval and analysis. Spatial access methods involve techniques to boost the performance of data retrieval. Even though it has been successfully developed for 2D data in various applications, no dedicated structure has been developed for 3D urban data and its applications. The only promising 3D spatial access method structure used in a commercial database is 3D R-Tree which still has several unresolved issues, such as overlapping among nodes, repetitive data entry and excessive Input/Output operations. Spatial access methods are also being used in combination with the execution of query statement. One of the most common queries for spatial data is the Nearest Neighbour (NN) query, which is also one of the most common queries for urban applications. There are several techniques used by spatial access method in retrieving the NN information; however, it is launched separately from the spatial access method structure and cause the deterioration of I/O-optimal processing. Current approach utilizes 3D R-Tree to address nearest neighbour query but still result unnecessary I/O operations. In this research, a dedicated 3D spatial access method that incorporates NN information is proposed and designed for the application of 3D urban data.

### **1.4 Research Questions**

The questions that are addressed in this research are:

1. How should 3D urban data be organised and managed in spatial databases?
2. How should information be retrieved and the neighbouring information about 3D urban data be preserved using 3D spatial access methods?

3. How can the data retrieval performance of 3D urban data in a spatial database be measured?

Based on these research questions, more specific research questions were designed as follows:

**Research Question 1:**

- What are the existing methods or approaches used for managing 3D urban data in spatial databases?
- What is a spatial access method?
- How can spatial access methods improve the efficiency of 3D urban data management?

**Research Question 2:**

- What are the existing methods or approaches used to retrieve neighbouring information in spatial databases?
- What techniques can be used to preserve the neighbouring information in urban data?
- How can information be retrieved using the 3D spatial access method?
- How can 3D neighbouring information be preserved in the 3D spatial access method?

**Research Question 3:**

- How can the data retrieval performance for urban data management be evaluated?
- How can data retrieval performance be described?



## **1.5 Research Aims**

The aim of this research is to propose and develop a three-dimensional (3D) spatial access method for managing 3D urban data. The spatial access method structure will be able to tackle several issues of data retrieval and query.

## **1.6 Research Objectives**

In this research, three main objectives are identified in order to answer all the research questions. The three main objectives in this research are:

1. To propose data structures that are spatial access methods for 3D urban data management.
2. To develop a new 3D spatial access method for 3D urban data management.
3. To test the performance of the proposed 3D access method structure with respect to existing approaches.

## **1.7 Research Design and Methodology**

There are four main processes to describe the design and methodology of this research. The process began with formulation phase. Then the next process is to identify data and tools used in this research and then designing and developing the new 3D spatial access method. The final process is the evaluation and analysis. Figure 1.1 shows the overall process of research design and methodology of this study. The detail description of these processes can be described as follows:

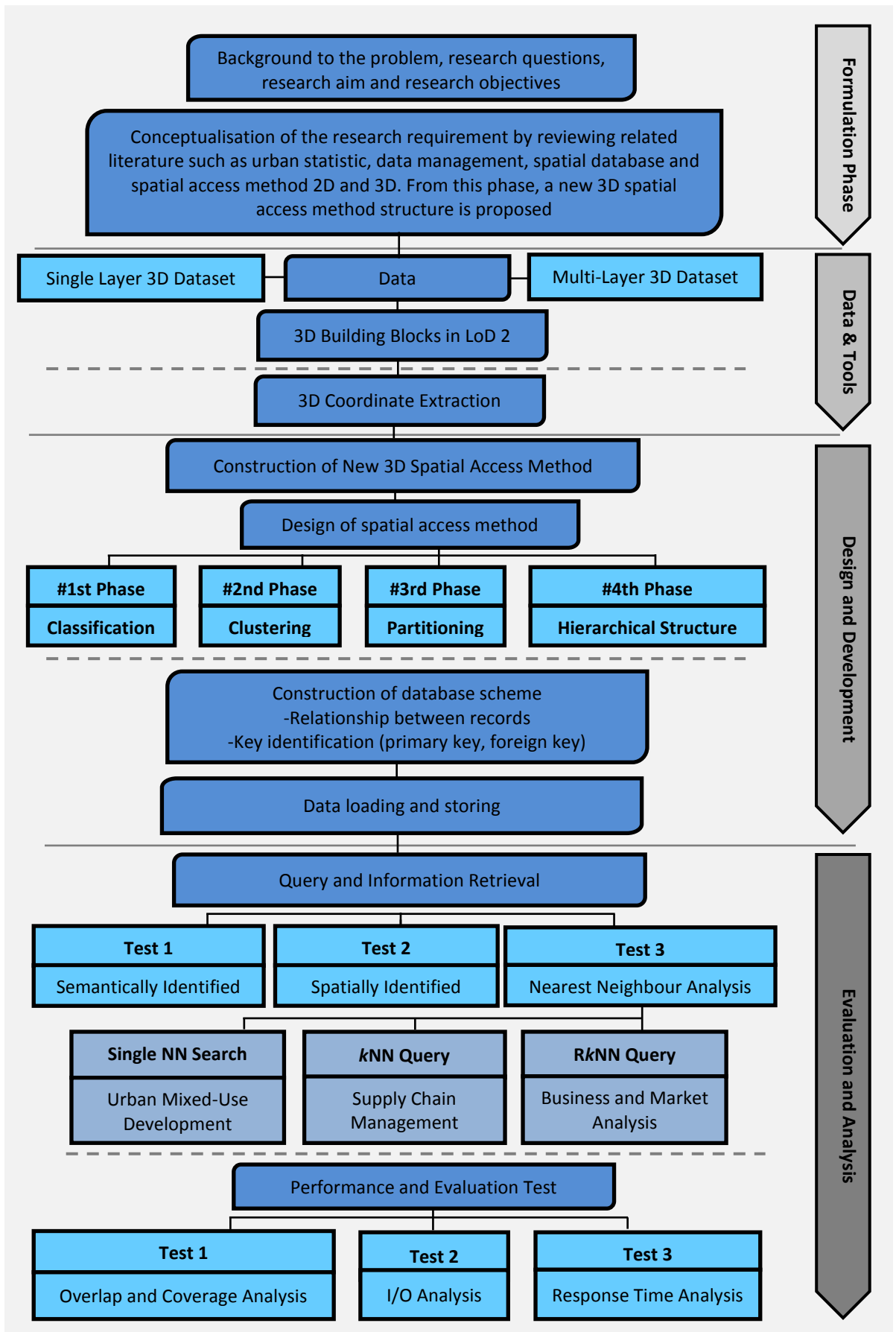


Figure 1.1: Research design and methodology

### (i) Formulation Phase

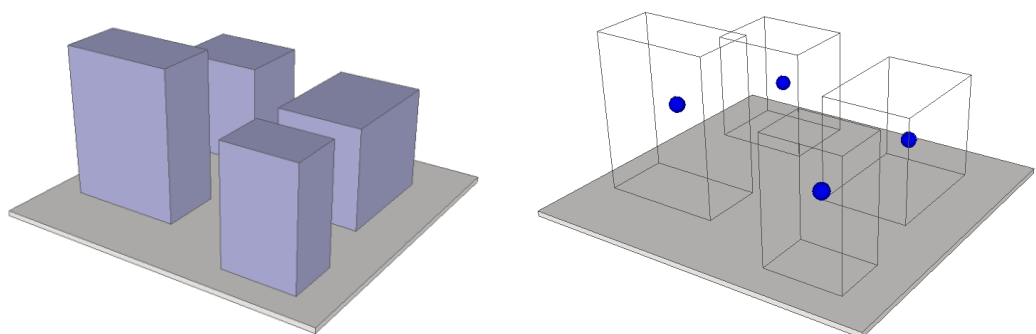
The process of this research is began with identifying the research formulations. Background to the problem is identified to formulate the problem statement. Based on the issues in the problem statement, research aims, research questions and research objectives are formulated. Related literature such as urban statistic, data management, spatial database and spatial access method 2D and 3D are thoroughly reviewed. With the current issues and related studies, a new spatial access method structure is proposed for the application of 3D urban data management.

### (ii) Data and Tools

Data such as 3D building blocks with Level of Detail (LoD) 2 were used as the main dataset in this research. Data is divided into two categories, single layer 3D datasets and multi-layer 3D datasets. The definitions of these categories are given in the following paragraphs.

- Single Layer 3D Datasets

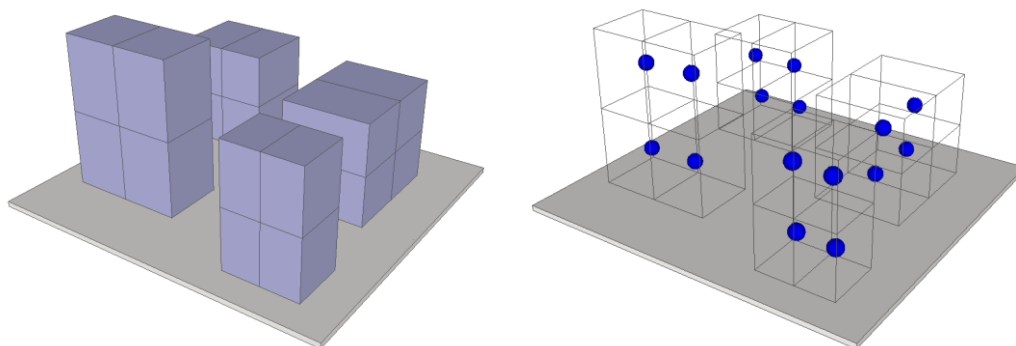
In this research, single layer datasets are 3D datasets with one layered point positions. For example, one 3D point location represents one unit of 3D building. Figure 1.3 shows 3D horizontal datasets for a set of buildings in an urban area.



**Figure 1.2:** Sample of single layer 3D datasets

- Multi-Layer 3D Datasets

Multi-layer datasets are defined as several 3D point locations that represent multiple units of space such as rooms or partitions in a building. 3D point locations are thus multi-layered. Figure 1.4 shows 3D vertical datasets for a set of buildings in an urban area.



**Figure 1.3:** Sample of multi-layer 3D datasets

From this datasets, coordinates are extracted and loaded into the spatial database based on database scheme.

### (iii) Design and Development

The construction of new 3D spatial access methods are based on four main phases; classification, clustering, partitioning and hierarchical structure. Each phase is significant in producing the spatial access methods structure. The search and query operation are depending on the first two phases which are classification and clustering. Detail description on each phases are described in the Chapter 3.

In the process of design and development, datasets are tabulated and mapped based on database scheme. This is to make sure the relationships among records are efficiently stored and loaded. Key identifications such as Primary Key and Foreign Key are also identified during this process. Then, data are loaded and stored based on scheme.

#### (iv) Evaluation and Analysis

The last process of this research is the evaluation and analysis. This process is important to test the developed 3D spatial access method structure. There are three main tests were performed in this research; semantically identified, spatially identified and nearest neighbour identification. The nearest neighbour analysis are tested with three different variants of nearest neighbour query which are, single search nearest neighbour,  $k$  Nearest Neighbour ( $k$ NN) and Reverse Nearest Neighbour ( $Rk$ NN). Each test is using single layer 3D dataset and multi-layer 3D dataset. Different case studies of urban applications are also implemented in these tests such as urban mixed-use development, supply chain management and business and market analysis.

To test the performance of proposed structure several analysis were performed such as overlap and coverage analysis, Input/Output or page analysis and response time analysis. The overlap and coverage analysis are performed to evaluate the percentage of overlap among nodes. Lower percentage mark the efficiency of the structure in terms of repetitive data entry and multi-path query. Page access analysis or Input/Output will mark the number of page accessed to the block memory during the query operation. The proposed structure is supposed to mark the lower page access. The last test is response time analysis where the time of data retrieval is measured. In this research the time is measured in millisecond (ms) and all of the tests are compared with the existing spatial access method in commercial software. The details description of this process is described in Chapter 4 and Chapter 5.

## 1.8 Thesis Organisation

This thesis is structured into six main chapters. A summary of each chapter is presented as follows:

- Chapter 1 – Introduction

This chapter introduces the background study and problems for research. It also clarifies the scope and aims of this research. A clear idea of the research is presented and the contributions of this research clarified.

- Chapter 2 – Spatial Access Methods for Urban Data Management

This chapter provides the basic knowledge of spatial access methods. Each method is reviewed and analysed to identify issues and limitations. Current researches on spatial access methods are also highlighted in this chapter. The gaps within existing methods were described in order to propose a new structure of spatial access method for the applications of urban data management.

- Chapter 3 – Design and Development of the Voronoi Classified and Clustered Data Constellation (VOR-CCDC)

This chapter focuses on the development of a new version of a 3D spatial access method for urban data management. The new structure was developed based on the issues raised in Chapter 1 and Chapter 2. The new structure of the spatial access method is called the Voronoi Classified and Clustered Data Constellation (VOR-CCDC). Several examples of applications are also explained in this chapter, such as urban mixed-use development, supply chain management and business and market analysis.

- Chapter 4 –Voronoi Classified and Clustered Data Constellation (VOR-CCDC) Data Retrieval and Nearest Neighbour Verification

In this chapter, the focus is on how to retrieve information from the database using the Voronoi Classified and Clustered Data Constellation (VOR-CCDC). Types of information retrieved from the database using VOR-CCDC structure are semantic information and spatial information. Other information that could be retrieved using VOR-CCDC is neighbouring information. There are three types of neighbouring information presented in this chapter: single search nearest neighbour,  $k$  number of nearest neighbour and reverse nearest neighbour. Different applications are tested for each type of neighbouring information, such as urban mixed-use development, supply chain management and business and market analysis applications. Several building blocks in the Putrajaya Administrative Centre were utilized for a nearest neighbour test for 3D single layer datasets.

- Chapter 5 - Voronoi Classified and Clustered Data Constellation (VOR-CCDC) Performance Evaluation and Verification

The Voronoi Classified and Clustered Data Constellation (VOR-CCDC) is tested through several evaluations to measure its efficiency in handling 3D urban data. The tests are to verify its efficiency in terms of repetitive data entries by performing overlap and coverage percentage analysis on the nodes. The structure was compared with other existing structures for page access analysis. Another test was to query its response time during data retrieval traversal and nearest neighbour analysis. The tests were compared with the existing access method structure in spatial databases. Based on the results presented in this chapter, the efficiency and practicality of VOR-CCDC for 3D urban data management can be verified.

- Chapter 6 – Conclusions, Recommendations and Further Research

This chapter concludes the thesis by summarising the advantages of the spatial access method developed for 3D urban data management. The contributions of this research are also described in this chapter. Recommendations for further research are given.



- Green building or sustainable building refers to construction or processes that are environmentally efficient in urban areas. It is one of the recent concerns in urban planning and management; however, data and information about green building is still minimal and for some countries implementation is still in the early stages. Analysis of urban data for green building standards will become very important in the near future. It is important to manage these datasets efficiently in the spatial database management system. It would be interesting for further research to expand the structure of VOR-CCDC to fit with the green building standard which falls under the ISO/TS 21931:2006 standard.

#### **6.4 Limitations of the Study**

For future reference and further research, the limitations of this research are described in the following list.

- There are two types of data used in this research: real datasets and synthetic datasets. To test the performance of VOR-CCDC, a huge amount of 3D data is required; however, due to limited sources of real datasets in 3D, a synthetic dataset was generated to verify the performance of VOR-CCDC.
- Distance values used in this research are based on Euclidean distance. Thus, the distance values used in this research are rough calculations between two points. For general urban application purposes, the use of Euclidean distance is reliable.

## 6.5 Conclusion

The increasing number of metropolitan and urban areas has affected the surrounding environments in many ways. Planners and geospatial professionals are obliged to help the local authorities in maintaining sustainable environment. Various technologies are being used to help planners and geospatial professionals to analyse urban areas, such as 3D geospatial applications, however, a special technique is required to store and manage urban data in 3D databases. The management of urban data is important to produce accurate, efficient, fast and reliable results of analyses. The proposed spatial access method in this research demonstrates an efficient way of handling 3D urban data in spatial databases.

The efficiency of the VOR-CCDC spatial access method structure in retrieving objects in terms of time and page access was demonstrated. Various cases of urban applications were successfully demonstrated. It has been demonstrated that VOR-CCDC is practical for handling various categories of urban data. Another advantage of VOR-CCDC is that it is supplemented by neighbouring information and the analyses of 3D urban data can thus be optimised.

## REFERENCES

- Aggarwal, A., Deshpande, A. and Kannan, R. (2009). Adaptive Sampling for  $k$ -Means Clustering. Dinur, I., Jansen, K., Naor, J. And Rolim, J. (eds.). *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Springer Berlin Heidelberg. Page 15-28.
- Akdağ, S. G. and Bostanci, S. H. (2013). The Impacts of Prestige Projects on The Skyline of Istanbul. *International Journal of Architectural Research (Archnet-IJAR)*, Volume 7, Page 269-281.
- Akhbari, M., Zare Mehrjerdi, Y., Khademi Zare, H. and Makui, A. (2014). A Novel Continuous  $k$ NN Prediction Algorithm to Improve Manufacturing Policies in a VMI Supply Chain. *International Journal of Engineering, Transactions B: Applications*, 27, 1681-1690.
- Akıncı, H., Özalp, A. Y. and Turgut, B. (2013). Agricultural Land Use Suitability Analysis using GIS and AHP Technique. *Computers and Electronics in Agriculture*, 97, 71-82.
- Aloise, D., Deshpande, A., Hansen, P. and Popat, P. (2009). NP-hardness of Euclidean Sum-of-Squares Clustering. *Machine Learning*, 75, 245-248.
- Anandhakumar, P., Priyadarshini, J., Monisha, C., Sugirtha, K. and Raghavan, S. (2010). Location Based Hybrid Indexing Structure - R  $k$ -d Tree. *First International Conference on Integrated Intelligent Computing (ICIIC)*, 5-7 August. Page 140-145.
- Anders, S. (2009). Visualization of Genomic Data with the Hilbert Curve. *Bioinformatics Conference*, Page 1231-1235.
- Ang, C. H. and Tan, T. C. (1997). New Linear Node Splitting Algorithm for R-trees. Scholl, M. and Voisard, A. (eds.). *Advances in Spatial Databases*. Springer Berlin Heidelberg. Series 1262, Page337-349.

- Arens, C., Stoter, J. and van Oosterom, P. (2005). Modelling 3D spatial Objects in a Geo-DBMS using a 3D primitive. *Computers & Geosciences*, 31, 165-177.
- Arthur, D. and Vassilvitskii, S. (2007). *k*-means++: The Advantages of Careful Seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete algorithms*, 2007. Society for Industrial and Applied Mathematics, Page 1027-1035.
- Aurenhammer, F. (1991). Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 23, 345-405.
- Bærentzen, J., Gravesen, J., Anton, F. and Aanæs, H. (2012). Spatial Data Indexing and Point Location. *Guide to Computational Geometry Processing*. Springer London. Page 213-225.
- Bahmani, B., Moseley, B., Vattani, A., Kumar, R. and Vassilvitskii, S. (2012). Scalable *k*-Means+. *Proceedings of the VLDB Endowment*, Volume 5.
- Balasubramanian, L. and Sugumaran, M. (2012). A State-of-Art in R-Tree Variants for Spatial Indexing. *International Journal of Computer Applications*, Page 42.
- Beckmann, N., Kriegel, H.-P., Schneider, R. and Seeger, B. (1990). The R\*-tree: An Efficient and Robust Access Method for Points And Rectangles. *Special Interest Group on Management of Data.*, Volume 19, Page 322-331.
- Bellman, R. (1962). Dynamic Programming Treatment of the Travelling Salesman Problem. *Journal of Association of Computing Machinery*, Volume 9, Page 61-63.
- Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Association of Computing Machinery*, Volume 18, Page 509-517.
- Boccaro, N. (2007). Lindenmayer Systems. *Essentials of Mathematica*. Springer New York. Page 407-416.
- Booker, M. (2013). *These Unbelievable Photos Make Hong Kong Look Like Abstract Art* [Online]. Available: <https://newrepublic.com/article/115209/michael-wolf-photography-hong-kongs-architecture-density> [Accessed on 10 May 2016].
- Brinkhoff, T., Kriegel, H.-P. and Seeger, B. (1993). Efficient processing of spatial joins using R-trees. *Special Interest Group on Management of Data*, Volume 22, Page 237-246.

- Castro, J. and Burns, S. (2007). Online Data Visualization of Multidimensional Databases using the Hilbert Space-Filling Curve. Lévy, P., Grand, B., Poulet, F., Soto, M., Darago, L., Toubiana, L. and Vibert, J.-F. (eds.). *Pixelization Paradigm*. Springer Berlin Heidelberg. Series 4370, Page 92-109.
- Chang, Q., Li, X., Huang, X. and Wu, J. (2012). A GIS-based Green Infrastructure Planning for Sustainable Urban Land Use and Spatial Development. *Procedia Environmental Sciences*, Volume 12, Part A, Page 491-498.
- Chatzimilioudis, G., Zeinalipour-Yazti, D., Lee, W.-C. and Dikaiakos, M. D. (2012). Continuous All  $k$ -Nearest-Neighbor Querying in Smartphone Networks. *IEEE 13th International Conference on Mobile Data Management (MDM)*, Page 79-88.
- Chung, J., Jung, K. H., Jung, S. Y., Kang, S. W. and Gil, J. M. (2011). Nearest-Neighbour Query Processing with Non-Spatial Predicates For Service Allocation In Smart Space Environment. *Institution of Engineering and Technology Communications*, Volume 5, Page 2470-2481.
- Chung, K.-L., Huang, Y.-L. and Liu, Y.-W. (2007). Efficient Algorithms for Coding Hilbert Curve of Arbitrary-Sized Image and Application to Window Query. *Information Sciences*, Volume 177, Page 2130-2151.
- Collier, C. (2006). The Impact of Urban Areas on Weather. *Quarterly Journal of the Royal Meteorological Society*, Page 132, 1-25.
- Danforth, A. (2015). Nearest Neighbor Search Problem: What's the Nearest Restaurant to My Hotel? <http://slideplayer.com/slide/3427837/> [Accessed 5 February 2016].
- Deren, L., Qing, Z., Qiang, L. and Peng, X. (2004). From 2D and 3D GIS for CyberCity. *Geo-spatial Information Science*, Volume 7, Page 1-5.
- Deren, L., Qing, Z., Qiang, L. and Peng, X. (2004). From 2D and 3D GIS for CyberCity. *Geo-spatial Information Science*, Volume 7, Page 1-5.
- Dinu Popa, M. D. (2014). Supply Chain Management-Key Factors. *Network Intelligence Studies*, Page 32-37.
- Emporis (2016a). *Building Directory* [Online]. Available: <http://www.emporis.com/continent/100004/asia> [Accessed 5 February 2016 2016].
- Emporis (2016b). *Building Directory - New York City* [Online]. Available: <https://www.emporis.com/city/101028/new-york-city-ny-usa> [Accessed 7 December 2016 2016].

- Gomes, A. P., Voiculescu, I., Jorge, J., Wyvill, B. and Galbraith, C. (2009). Spatial Data Structures. Gomes, A. P., Voiculescu, I., Jorge, J., Wyvill, B. and Galbraith, C. (eds.). *Implicit Curves and Surfaces: Mathematics, Data Structures and Algorithms*. Springer London. Page 41-62.
- Gong, J., Ke, S., Li, X. and Qi, S. (2009). A Hybrid 3D Spatial Access Method Based On Quadtrees and R-Trees For Globe Data. *International Symposium on Spatial Analysis, Spatial-temporal Data Modeling, and Data Mining*. International Society for Optics and Photonics, Page 74920R-74920R.
- Gong, J., Zhu Qing, Zhang Hanwu, Li Xiaoming, Zhou Dongbo (2011). An Adaptive Control Method of LODs for 3D Scene Based on R-tree Index. *Acta Geodaetica et Cartographica Sinica*, Page 531-534.
- Gu, W., Wang, J., Shi, H. and Liu, Y. (2011). Research on a Hybrid Spatial Index Structure. *Journal of Computational Information Systems*, Volume 7, Page 3972-3978.
- Gunasekaran, A. and Ngai, E. W. T. (2004). Information Systems In Supply Chain Integration and Management. *European Journal of Operational Research*, Volume 159, Page 269-295.
- Guttman, A. (1984). R-Trees: A Dynamic Index Structure for Spatial Searching. *Special Interest Group on Management of Data*, Volume 14, 47-57.
- Hill, J. S. (2008). *International Business: Managing Globalization*. SAGE Publications, Inc.
- Hoppenbrouwer, E. and Louw, E. (2005). Mixed-use development: Theory and practice in Amsterdam's Eastern Docklands. *European Planning Studies*, Volume 13, Page 967-983.
- Hsu, C.-H. (2014). Intelligent Big Data Processing. *Future Generation Computer Systems*, Volume 36, Page 16-18.
- Joseph, M., Wang, F. and Wang, L. (2014). GIS-Based Assessment of Urban Environmental Quality in Port-au-Prince, Haiti. *Habitat International*, Volume 41, Page 33-40.
- Jun, G. and Shengnan, K. (2011). 3D Spatial Query Implementation Method Based on R-tree. *International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE)*, Page 2828-2831.

- Kamel, I. and Faloutsos, C. (1994). Hilbert R-tree: An Improved R-tree using Fractals. *Proceedings of the 20th International Conference on Very Large Data Bases*, Morgan Kaufmann Publishers Inc., Page 500-509.
- Kennish, M. J. (2002). Environmental Threats and Environmental Future of Estuaries. *Environmental conservation*, Volume 29, page 78-107.
- Khan, F. (2012). An Initial Seed Selection Algorithm For  $k$ -Means Clustering of Georeferenced Data to Improve Replicability of Cluster Assignments for Mapping Application. *Applied Soft Computing*, Volume 12, Page 3698-3700.
- Kiekintveld, C., Miller, J., Jordan, P. R. and Wellman, M. P. (2011). Forecasting Market Prices in a Supply Chain Game. *Proceedings of the International Conference on Autonomous Agents*, Page 1323-1330.
- Koes, D. R. and Camacho, C. J. (2015). Indexing Volumetric Shapes with Matching and Packing. *Knowledge and information systems*, Volume 43, Page 157-180.
- Kofler, M. and Gervautz, M. (2000). R-trees for Organizing and Visualizing 3D GIS Databases. *The Journal of Visualization and Computer Animation*, Volume 11, Page 129-143.
- Korotkov, A. (2012). A New Double Sorting-Based Node Splitting Algorithm for R-tree. *Programming and Computer Software*, Volume 38, Page 109-118.
- Kovács, F., Legány, C. and Babos, A. (2005). Cluster Validity Measurement Techniques. *Proceedings of the fifth WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases*. Madrid, Spain, Page 388-393.
- Krawczyk, B., Stefanowski, J. and Wozniak, M. (2015). Data Stream Classification and Big Data Analytics. *Neurocomputing*, Volume 150, Part A, Page 238-239.
- Lancioni, R. A., Smith, M. F. and Oliva, T. A. (2000). The Role of the Internet in Supply Chain Management. *Industrial Marketing Management*, Volume 29, Page 45-56.
- Lee, D.-H. and Park, D.-J. (2005). An Efficient Incremental Nearest Neighbor Algorithm for Processing  $k$ -Nearest Neighbor Queries with Visual and Semantic Predicates in Multimedia Information Retrieval System. Lee, G., Yamada, A., Meng, H. and Myaeng, S. (eds.). *Information Retrieval Technology*. Springer Berlin Heidelberg. Series 3689, Page 653-658.

- Lee, D.-T. and Wong, C. (1977). Worst-Case Analysis for Region and Partial Region Searches In Multidimensional Binary Search Trees and Balanced Quad Trees. *Acta Informatica*, Volume 9, Page 23-29.
- Lei, S. Y. P. M. Z. (2004). Study of R-Tree Spatial Access Method in Three Dimensional GIS. *Journal of Computer Engineering and Applications*, Volume 14, Page 003.
- Li, M.-H., Hong, B.-R., Cai, Z.-S., Piao, S.-H. and Huang, Q.-C. (2008). Novel Indoor Mobile Robot Navigation using Monocular Vision. *Engineering Applications of Artificial Intelligence*, Volume 21, Page 485-497.
- Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A. and Cheng, T. (2015). Geospatial Big Data Handling Theory and Methods: A Review and Research Challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 115, Page 119-133.
- Liang, J., Gong, J., Li, W. and Ibrahim, A. N. (2014). A Visualization-Oriented 3D Method for Efficient Computation of Urban Solar Radiation Based on 3D-2D Surface Mapping. *International Journal of Geographical Information Science*, Volume 28, Page 780-798.
- Liu, Y., Fang, J. and Han, C. (2009). A New R-Tree Node Splitting Algorithm using MBR Partition Policy. *17th International Conference on Geoinformatics*. IEEE, Page 1-6.
- MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. California, USA, Page 14.
- Mamoulis, N. (2011). Spatial Data Management. *Synthesis Lectures on Data Management*, Volume 3, Page 1-149.
- Manolopoulos, Y., Nanopoulos, A., Papadopoulos, A. N. and Theodoridis, Y. (2010). *R-Trees: Theory and Applications*. Springer Science & Business Media. Page 3-13.
- Manewongvatana, S. and Mount, D. M. (2001). On The Efficiency of Nearest Neighbor Searching with Data Clustered in Lower Dimensions. *International Conference on Computational Science*, 2001. Springer, Page 842-851.
- Mark, D. M. and Lauzon, J. P. (2011). Approaches for Quadtree-Based Geographic Information Systems at Continental or Global Scales. *Proceedings of Auto-Carto*, 1985. Page 355-365.



- Miocevic, D. and Crnjak-Karanovic, B. (2012). The Mediating Role of Key Supplier Relationship Management Practices on Supply Chain Orientation-The Organizational Buying Effectiveness Link. *Industrial Marketing Management*, Volume 41, Page 115-124.
- Mohamad Yusoff, I., Uznir Ujang, M. and Abdul Rahman, A. (2010). 3D Volumetric Soft Geo-objects for Dynamic Urban Runoff Modeling. Neutens, T. and Maeyer, P. (eds.). *Developments in 3D Geo-Information Sciences*. Springer Berlin Heidelberg. Page 200-219.
- Moore, A. W. (1991). An Introductory Tutorial on KD Trees. Technical Report No. 209. *Carnegie Mellon University*. Page 1-20.
- Murray, C. (2009). Oracle Spatial Developer's Guide, 11g Release 1 (11.1). *Oracle Spatial User's Guide and Reference*. Page 133-145.
- National Geographic (2009). *Megacities* [Online]. Available: <http://ngm.nationalgeographic.com/static-legacy/ngm/0211/feature3/> [Accessed 13 January 2016].
- Nene, S. A. and Nayar, S. K. (1997). A Simple Algorithm for Nearest Neighbor Search in High Dimensions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Volume 19, Page 989-1003.
- Nguyen, K. G. and Saupe, D. (2001). Rapid High Quality Compression of Volume Data for Visualization. *Computer Graphics Forum*. Wiley Online Library, Page 49-57.
- Oosterom, P. V. (1999). Spatial Access Methods edited by Longley, Goodchild, Maguire en Rhind, John Wiley *Geographical information systems*, Volume 1, Page 385-400.
- Pamanikabud, P. and Tansatcha, M. (2009). Geoinformatic Prediction of Motorway Noise on Buildings In 3D GIS. *Transportation Research Part D: Transport and Environment*, Volume 14, Page 367-372.
- Park, D.-J. and Kim, H.-J. (2003). An Enhanced Technique for  $k$ -Nearest Neighbor Queries with Non-Spatial Selection Predicates. *Multimedia Tools and Applications*, 19, 79-103.
- Peuquet, D. J. (1984). Data Structures for A Knowledge-Based Geographic Information System. *Proceeding of International Symposium on Spatial Data Handling*. Zurich, Switzerland, Volume 2, Page 372-391.

- Power, D. (2005). Supply Chain Management Integration and Implementation: A Literature Review. *Supply Chain Management: An International Journal*, Volume 10, Page 252-263.
- Rabianski, J. S., Gibler, K. M., Alan Tidwell, O. and Sherwood Clements III, J. (2009). Mixed-Use Development: A Call for Research. *Journal of Real Estate Literature*, Volume 17, Page 205-230.
- Ravada, S., Kazar, B. and Kothuri, R. (2009). Query Processing in 3D Spatial Databases: Experiences with Oracle Spatial 11g. Lee, J. and Zlatanova, S. (eds.). *3D Geo-Information Sciences*. Springer Berlin Heidelberg. Page 153-173.
- Reed, C. (2005). Data Integration and Interoperability: OGC Standards for Geo-Information. Sisi Zlatanova, D. P. (ed.) *Large-scale 3D Data Integration: Challenges and Opportunities*. Springer. Page 163-166.
- Rigaux, P., Scholl, M. and Voisard, A. (2002). Spatial Access Methods. *Spatial Databases*. San Francisco: Morgan Kaufmann. Page 201-266.
- Rodger, J. A. (2014). A Fuzzy Nearest Neighbor Neural Network Statistical Model for Predicting Demand for Natural Gas and Energy Cost Savings in Public Buildings. *Expert Syst. Appl.*, Volume 41, Page 1813-1829.
- Samet, H. (1984). The quadtree and related hierarchical data structures. *ACM Computing Surveys (CSUR)*, Volume 16, Page 187-260.
- Samet, H., Rosenfeld, A., Shaffer, C. A. and Webber, R. E. (1984). A geographic information system using quadtrees. *Pattern Recognition*, Volume 17, Page 647-656.
- Samet, H. (2003). Depth-First  $k$  Nearest Neighbor Finding using the MaxNearestDist estimator. *International Conference on, Image Analysis and Processing*, Page 486-491.
- Sample, N., Haines, M., Arnold, M. and Purcell, T. (2001). Optimizing Search Strategies in kd Trees. *Fifth WSES/IEEE World Multiconference on Circuits, Systems, Communications & Computers*.
- Sellis, T. K., Roussopoulos, N. and Faloutsos, C. (1987). The R+-Tree: A Dynamic Index for Multi-Dimensional Objects. *Proceedings of the 13th International Conference on Very Large Data Bases*, Morgan Kaufmann Publishers Inc., Page 507-518.

- Sharkawi, K., Ujang MU, Abdul-Rahman A, MU, U. and A, A.-R. (2008). 3D Navigation System for Virtual Reality Based on 3D Game Engine. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXVII, Page 513–518.
- Stoter, J. and Zlatanova, S. (2003). 3D GIS, Where Are We Standing. *Proceedings ISPRS Workshop on spatial, temporal and multi-dimensional data modelling and analysis, Quebec, Canada*. Page 8-16.
- Taniar, D. and Rahayu, W. (2013). A Taxonomy for Nearest Neighbour Queries In Spatial Databases. *Journal of Computer and System Sciences*, Volume 79, Page 1017-1039.
- Tao, Y., Papadias, D. and Shen, Q. (2007). Continuous Nearest Neighbor Search. *Proceedings of the 28th international conference on Very Large Databases. VLDB Endowment*, page 287-298.
- Tao, Y. (2015). Nearest Neighbor Queries using R-trees. <http://slideplayer.com/slide/3427837/> [Accessed 5 February 2016].
- Tian, W. and Zhao, Y. (2015). 2 - Big Data Technologies and Cloud Computing. Tian, W. and Zhao, Y. (eds.). *Optimized Cloud Resource Management and Scheduling*. Boston: Morgan Kaufmann. Page 17-49.
- Umakoshi, E. M. and Gonçalves, J. C. (2011). A Qualitative Analysis of Five Tall Buildings Presented As Sustainable. *International Conference on Passive and Low Energy Architecture and Sustainable Development*, Page 447-452.
- Van der Hoeven, F. and Nijhuis, S. (2012). Developing Rotterdam's skyline. *Council on Tall Buildings and Urban Habitat Journal*, Page 32-37.
- Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les paralléloèdres primitifs. *Journal für die reine und angewandte Mathematik*, Volume 134, Page 198-287.
- Wang, Y. and Guo, M. (2012). An Integrated Spatial Indexing of Huge Point Image Model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXIX-B4 Page 397-401.
- Weber, R., Schek, H.-J. and Blott, S. (1998). A Quantitative Analysis And Performance Study For Similarity-Search Methods In High-Dimensional Spaces. Volume 98, Page 194-205.

- Xu, H., Cheng, L., Li, M., Chen, Y. and Zhong, L. (2015). Using Octrees to Detect Changes to Buildings and Trees in The Urban Environment from Airborne LiDAR Data. *Remote Sensing*, Volume 7, Page 9682-9704.
- Yan, D.-M., Wang, W., Lévy, B. and Liu, Y. (2010). Efficient Computation of 3D Clipped Voronoi Diagram. *International Conference on Geometric Modeling and Processing*, Springer Berlin Page 269-282.
- Yuan, J. and Liu, X. (2013). Transform Residual  $k$ -means trees For Scalable Clustering. *Proceedings - IEEE 13th International Conference on Data Mining Workshops, ICDMW* Page 489-496.
- Zhang, J., You, S. and Gruenwald, L. (2011). Parallel quadtree coding of large-scale raster geospatial data on GPGPUs. *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Page 457-460.
- Zhang, P., Cheng, R., Mamoulis, N., Renz, M., Züfle, A., Tang, Y. and Emrich, T. (2013). Voronoi-Based Nearest Neighbor Search for Multi-Dimensional Uncertain Databases. *International Conference on Data Engineering (ICDE)*, Page 158-169.
- Zheng, K., Liu, X.-g. and Yang, H. (2005). Research on Spatial Index Structure of LOD-OR Tree in 3D GIS. *Bulletin of Surveying and Mapping*, Volume 5, Page 27-29.
- Zheng, L., Yang, C. and Hu, Y. (2011). The Research of Digitized Substation Fault Diagnosis Based on 3D GIS. *International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, Page 1374-1377.
- Zhu, Q., Gong, J. and Zhang, Y. (2007). An Efficient 3D R-Tree Spatial Index Method for Virtual Geographic Environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 62, Page 217-224.
- Zlatanova, S. (2000). *3D GIS for urban development*. International Inst. for Aerospace Survey and Earth Sciences (ITC).
- Zlatanova, S., Rahman, A. and Pilouk, M. (2002). 3D GIS: Current Status And Perspectives. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, Volume 34, Page 66-71.