



What3Words Geocoding Extensions

Wen Jiang¹ · Emmanuel Stefanakis¹

Published online: 15 February 2018
© Springer International Publishing AG, part of Springer Nature 2018

Abstract

With the advent of location-based services, the demand for location data has dramatically increased. Geocoded locations have become necessary in many GIS analysis, cartography, and decision-making workflows. A reliable geocoding system that can effectively return any location on earth with sufficient accuracy is desired. This study is motivated by a need for a geocoding system to support university campus applications. Address-based geocoding systems have been used for decades. However, they present limitations in address resources, address standardization, and address database maintenance. These limitations have recently sparked an interest in developing alternative geocoding systems that apply alphanumeric codes as a reference to locations, such as Geohash, Google's Open Location Code, and what3words to name a few. Comparing to other geocoders, what3words (w3w) has many advantages. It uses a simple format of code consisting of three words, it is less error-prone, codes are easier to memorize, and multiple languages are supported. However, its fixed resolution (consisting of 3 m by 3 m square cells) and lack of consideration of the third dimension may limit its applicability. To better support geographic applications with special requirements, the w3w geocoding system needs to be extended. This paper proposes extensions of w3w in two aspects: variable resolution and third dimension support. A geocoding processing tool that implements these extensions is being developed to support the need of a university campus' facility management, emergency evacuation and route navigation planning, student survey data management, and other location-based services.

Keywords Geocoding systems · what3words · Geocoding processing · Location-based services

Introduction

Methods to determine a location on a digital map or find the optimal paths to get there are becoming increasingly relevant to everyday life. Location encoding (also known as geocoding) is an approach that transforms an address, postal code, place name, or another geographic reference to geographic coordinates (Goldberg 2011; Lee 2009; Karimi et al. 2011). This allows spatial analysis, mapping, and other geolocation related processes to be performed in GIS software packages. As such, a variety of geocoding systems have been created to support specific applications including, but not limited to, marking locations on a map, route finding and

navigation, and local searching. A geocoding system consists of three components: (i) the input data—address, the name of a place, or code; (ii) the geocoder—the processing algorithm; and (iii) the output data—the location coordinates corresponding to the input data. Reverse geocoding functions perform the opposite task and convert coordinates into the address, name, or code of the location.

Address-based geocoding, which refers to “address matching” including a postal code and place's name, has been dominant for decades. Numerous efforts have been made to improve address-based geocoding. However, they were not able to completely restrain all limitations involved (Lee 2009). Alphanumeric code geocoding systems have emerged more recently. These systems provide alternative methods to overcome the problems with address-based geocoding.

An alphanumeric code geocoding uses codes as a reference to locations. It partitions the earth surface into arrays of cells, often in layers of variant resolution, and assigns each cell a unique alphanumeric code. Codes can then be converted into the centroid coordinates of the cells. Examples of these geocoding schemes include Geohash, MapCode, c-Squares, WMO squares, Open Post Code, Google's Open Location

✉ Wen Jiang
wjiang3@unb.ca

Emmanuel Stefanakis
estef@unb.ca

¹ Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada

Code, and what3words (Barr 2015; Stefanakis 2016). The latter one is of particular interest to this study.

what3words (w3w) divides the earth's surface into a grid of 3 m by 3 m squares; each assigned a unique code consisting of three dictionary words separated by periods, e.g., the entrance to the CN Tower in Toronto, Canada, is located at "select.threaten.shelters" (Barr 2015). This method allows for easier memorization of locations and is supported in multiple languages. Meanwhile, the three-word code is efficient to encode and decode. w3w provides a website, apps for iOS and Android, and an API that enables bidirectional conversion between the three dictionary words assigned to the grid cells and latitude/longitude coordinates (Barr 2015).

w3w has demonstrated some advantages over the traditional address-based geocoding. It is estimated that over 4 billion people on Earth are physically disconnected because of lack of a reliable street address (Barr 2015). But even when street addresses are available, they are very often unable to describe the location. For example, locations inside parks or large facilities (e.g., stadiums or hospitals with multiple entrances) may be hundreds of meters away from the nearest address. The use of directions (such as "behind the main building find a storehouse; deliver the package to the right door facing the park") instead of address has become a common practice. However, this is usually ambiguous as it relies on local knowledge, while it cannot be interpreted automatically (Stefanakis 2016).

Apparently, the w3w grid can meet the fundamental need of addressing people and places no matter the existence of a reliable street address. The opportunities, that w3w presents, have been widely recognized around the globe. Governments of at least five nations with poor addressing systems have already adopted w3w for their postal services, while w3w has been recently used as an addressing mechanism in disaster relief missions and emergency response (The Ethicalist 2017).

Besides the opportunities, w3w presents limitations that complicate its usefulness to certain geographic applications. Specifically, w3w has a fixed resolution of 3 m \times 3 m and provides no consideration of elevation. How can the what3words geocoding system be extended to address these limitations? How could these extensions be applied to support geographic applications such as a micro-scale area of a university campus? To answer these questions, this study has established three main objectives: (a) to design alternative w3w geocoding extension models, (b) to implement geocoding algorithms for the extended models, (c) to apply the extended models and associated algorithms in various applications relevant to a university campus. This paper focuses on the first objective.

The w3w geocoding system was extended towards two aspects: variable resolution and elevation. A potential application using a finer resolution extension of the w3w geocoding system is to locate a geographic entity (e.g., a desk) in a building room, as demonstrated in Fig. 1. The location of the desk can be represented by an extended model involving a

Fig. 1 Example application of a w3w extension that considers a finer resolution: **a** level E of Head Hall building and the w3w grid of 3 m \times 3 m squares, **b** one w3w square within room E5 sub-divided into cells of 33 cm \times 33 cm, and **c** the same cells in larger scale and the encoding of two cells covering a desk

code consisting of four words with a resolution of 33 cm by 33 cm. On the other hand, a coarser resolution extension would be required for an evacuation planning application. As shown in Fig. 2, an open space in a university campus covered by square cells of a size that exceeds the 3 m \times 3 m could be represented by an extended model involving the central 3 m \times 3 m cells accompanied by a number describing the size of these cells.

An elevation extension of the w3w geocoding system can be beneficial for locating entities in three-dimensional space. In the example of Fig. 1, the desk is located on the fifth floor of a building. Its relative height (above the ground) is 13 m (i.e., each floor is 3 m high, and the desk is 1 m high). Hence, the location of the desk can be described by three words supplemented by a code that provides its relative height, e.g., "psychic.rolling.recital.tan.RH7". A reverse geocoding process would turn the above code into the location of the desk expressed in (x, y, z) coordinates, i.e., (-66.64174, 45.950082, 13).

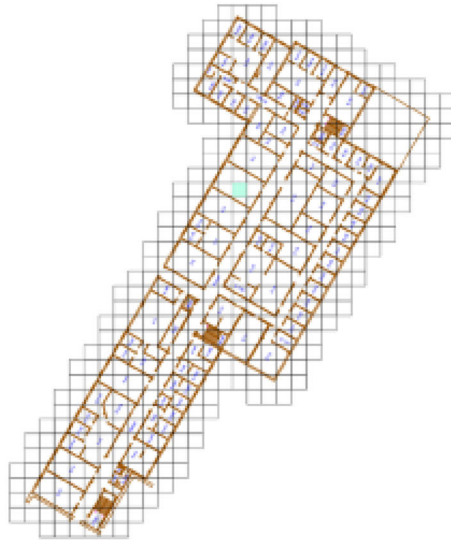
It is believed that the development of the extensions above could facilitate various geographic applications in a university campus such as campus navigation, emergency evacuation, facility management, and student survey data management.

The paper content is organized into five Sections. In "Geocoding Systems and what3words" section provides the state-of-the-art of geocoding systems including a discussion on the advantages and disadvantages of these systems. In "Extensions of what3words" section introduces the models to extend the w3w geocoding system to support variable resolution and elevation. In "Applications" section discusses some potential applications and example scenarios of the extended w3w geocoding system. In "Forward and Reverse Geocoding Transformations" section presents the forward and reverse transformations between extended w3w codes and the corresponding geographic coordinates. In "Conclusion" section summarizes the proposed extensions and presents some future developments.

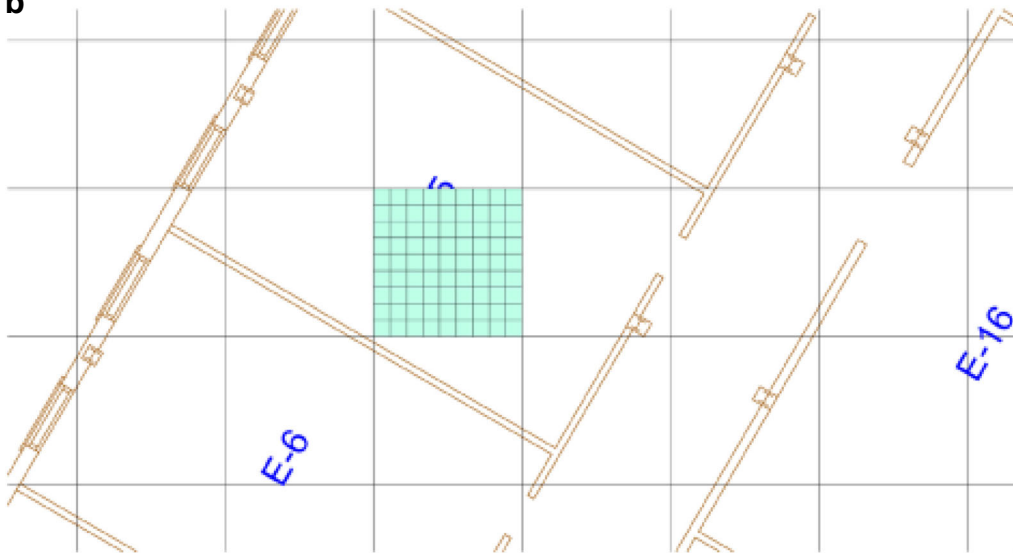
Geocoding Systems and what3words

Geocoding is one of the basic geospatial operations that convert addresses, postal codes, place names, or other geographic references to geographic coordinates (Goldberg 2011; Lee 2009; Karimi et al. 2011). It plays a vital role in the spatial analysis as geocoding technology has been utilized in many application areas such as epidemiology, environmental

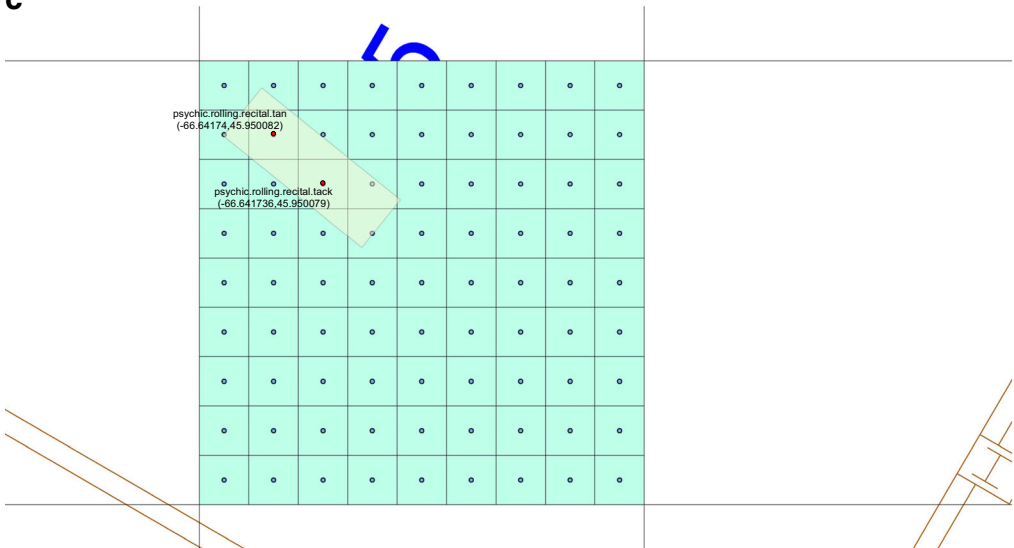
a



b



c



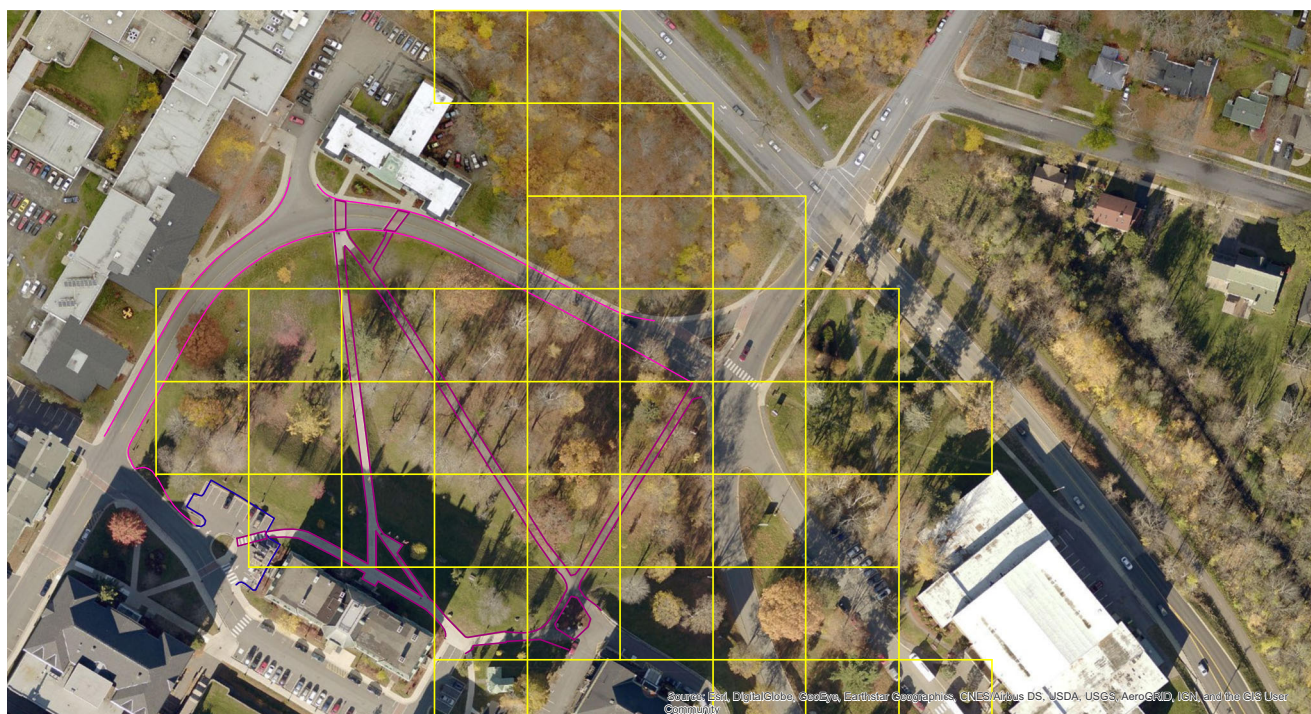


Fig. 2 An example application of a w3w extension that considers a coarser resolution (cell size, 30 m × 30 m)

science, emergency management, marketing, planning, and location-based services. These applications involve a broad range of disciplines including, but not limited to, geography, geographic information science, computer science, digital libraries, history, and economics (Goldberg 2011; Karimi et al. 2011). Geocoded data provides a basis for subsequent spatial analysis and mapping. Errors associated with the geocoded data are likely to propagate through subsequent processing, analysis, modeling, and decision-making (Goldberg 2011; Karimi et al. 2011). Therefore, it is important to obtain accurate locations from geocoding processes.

There are two categories of geocoding schemes adopted by geocoding systems: address-based and alphanumeric code-based. The address-based geocoding scheme makes use of two main models: street network geocoding and rooftop geocoding (also known as address-point geocoding), which have been widely used for decades (Lee 2009; Karimi et al. 2011; Zandbergen 2008).

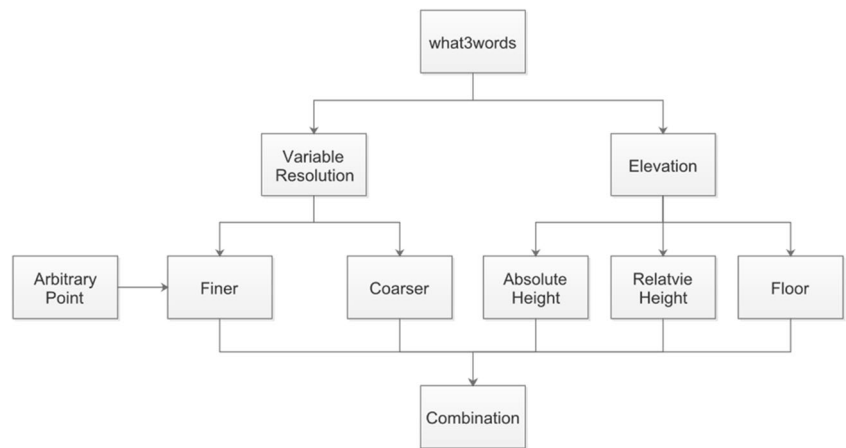
Address-based geocoding systems have many constraints including coverage, standardization, maintenance, and precision. The coverage issues occur because these systems are unable to geocode locations with no official address (Barr 2015). It is estimated that, worldwide, over two billion people live at locations with no official street name or house number (Geelen 2015). Standardization is an issue because address-based systems require properly formatted input, whereas address formats vary from location to location (Lee 2009). Also, these systems require high maintenance as address databases must be updated regularly to reflect real-world changes for the

entire coverage area (Lee 2009). Last but not least, precision can be a major concern. Geocoding in rural areas is often offset-prone (Kellison 2012). Even in urban areas, where geocoders are typically more precise, they use as a reference the centroid of structures or parcels. Entire university buildings, business parks, and farms are abstracted to single points that do not carry sufficient precision for many applications (Chen et al. 2016; Goldberg et al. 2007; Karimi et al. 2011).

Alphanumeric code geocoding scheme provides an alternative way for describing a geographic location. Multiple alphanumeric code-based geocoding systems are already available to convert between alphanumeric codes and geographic coordinates: Geohash, MapCode, c-Squares, WMO squares, Open Post Code, Google's Open Location Code, and what3words (Barr 2015; Stefanakis 2016). These systems assign systematic alphanumeric labels to locations (polygons) over the earth, which are converted to geographic coordinates using mapping formulas, instead of graphs and an address database (Barr 2015). Alphanumeric code geocoding systems share several advantages: (i) every cell is assigned a unique and static code, (ii) codes are simple to encode and decode, and (iii) codes are efficient for communication (Chen et al. 2016).

w3w, compared to other alphanumeric code geocoding systems, has three significant advantages. First, it is the cleanliness of the coding scheme as the use of dictionary words is less error-prone than a code mixing letters and numbers. Secondly, it is easier to remember (Barr 2015). Finally, it supports multiple languages. Other alphanumeric code geocoding systems

Fig. 3 Alternative w3w geocoding extension models



provide a code combining random Latin characters with numbers, which can be as long as ten characters for a high resolution. The codes provided by these systems are hard to remember, while they are not language- or culture-independent (Rhind 2015). On the other hand, as opposed to other alphanumeric geocoding systems, w3w has a fixed resolution of 3 m × 3 m. This may impede an efficient modeling in applications, such as in indoor environments where a finer resolution is required, or in outdoor environments where a coarser resolution is preferable.

None of those mentioned above geocoding systems considers the third dimension, i.e., elevation (Stefanakis 2016). 3D geocoding and 3D reverse geocoding services are still a challenge (Verbree and Zlatanova 2007). An address-based

3D geocoding system for the indoor environment was proposed by Lee (2009). However, this method inherits the limitations of address-based geocoding systems and does not offer an appropriate solution for 3D location-based services.

Overall, alphanumeric code geocoding systems may demonstrate advantages over address-based geocoding ones in many aspects. Furthermore, a comparison between the alphanumeric code geocoding systems reveals that w3w has many significant advantages that can be utilized to better support specific geographic applications (Barr 2015). However, w3w still has limitations, such as a fixed resolution and no consideration of elevation. This study introduces a series of extensions of the w3w geocoding system to overcome the above limitations.

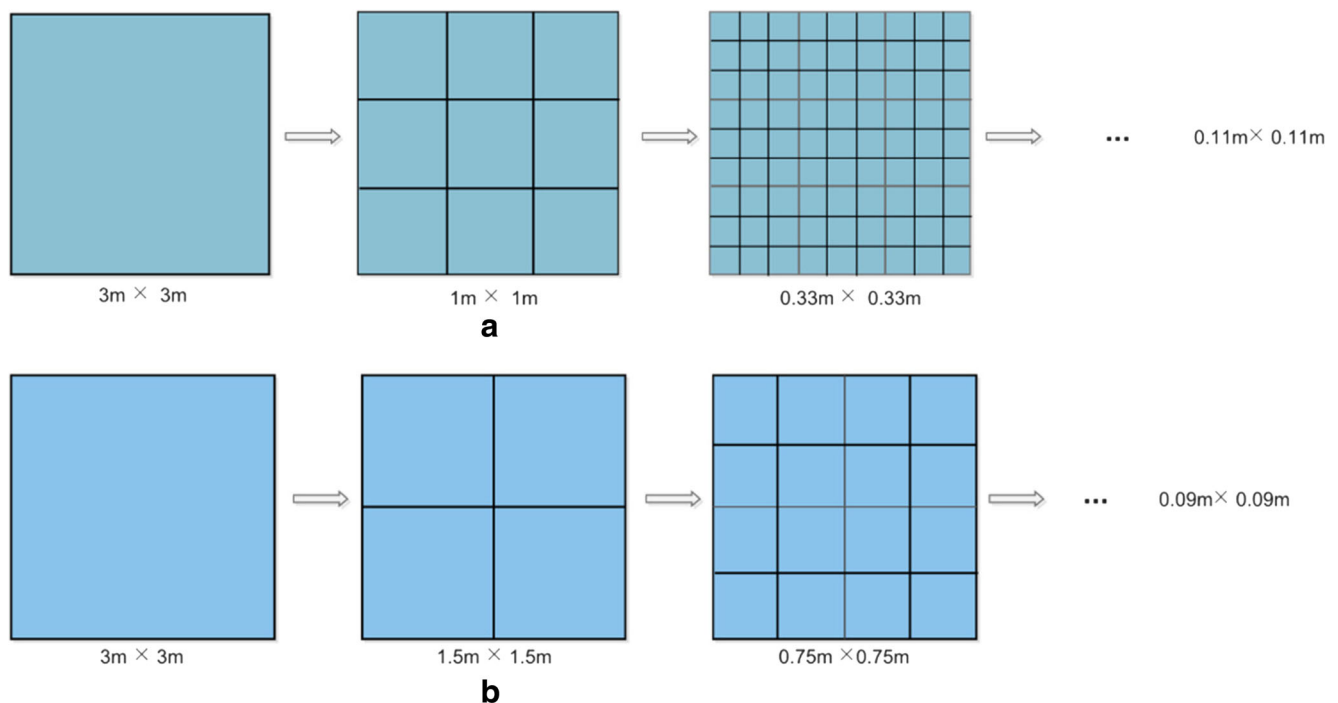
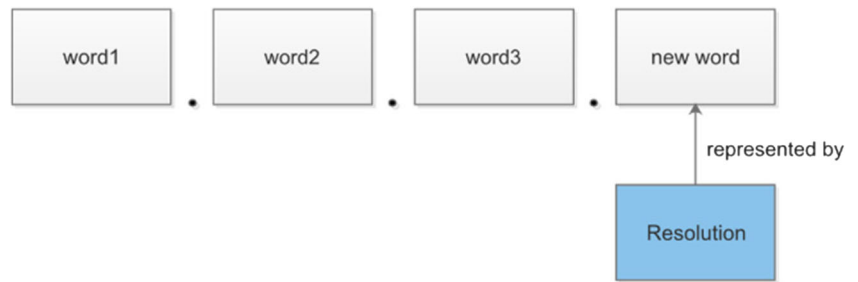


Fig. 4 Finer resolution extension models: **a** Temary-Tree Extension Model (TTEM) and **b** Quadtree Extension Model (QTEM)

Fig. 5 w3w resolution extension by adding an extra word



Extensions of what3words

This section describes how w3w geocoding system can be extended to support variable resolutions as well as the integration of elevation data into the alphanumeric code. In the original w3w system, resolution is fixed to 3 m × 3 m. With a variable resolution extension, w3w grid cells may have a smaller or larger size to satisfy the needs of a finer or coarser resolution, respectively. The finer resolution model could be used to represent any spatial points that fall into the same square beside the cell’s centroid (Fig. 1). The coarser resolution model could be used to represent areas with an extent larger than 3 m × 3 m, using a single code. The elevation could be described by absolute height values, relative height values (i.e., above or below the ground), or even floor labels and relative heights within the floor for indoor application purposes. The above models need to be combined as shown in Fig. 3 to represent specific areas or locations at variable resolutions in 3D space.

Variable Resolution Extension Models—Finer Resolution

Two methods to increase the resolution of the grid cells are considered (Fig. 4): Ternary-Tree Extension Model (TTEM) and Quadtree Extension Model (QTEM). The w3w square is divided into nine or four sub-squares, respectively. Each square is recursively divided into sub-squares, resulting in a cell size at the finest resolution equal to 11 cm by 11 cm and 9 cm by 9 cm, respectively.

The extended variable resolution model is represented by attaching a fourth element (a new word) to the w3w code, as shown in Fig. 5. At each resolution level, all new words start with the same letter so that the resolution level can implicitly be determined from the code (Figs. 6 and 7). This way, the fourth word conveys information about both the sub-cell size and location in the original w3w’s 3 m × 3 m square. The spatial relations between two cells can also be partially extracted from their codes.

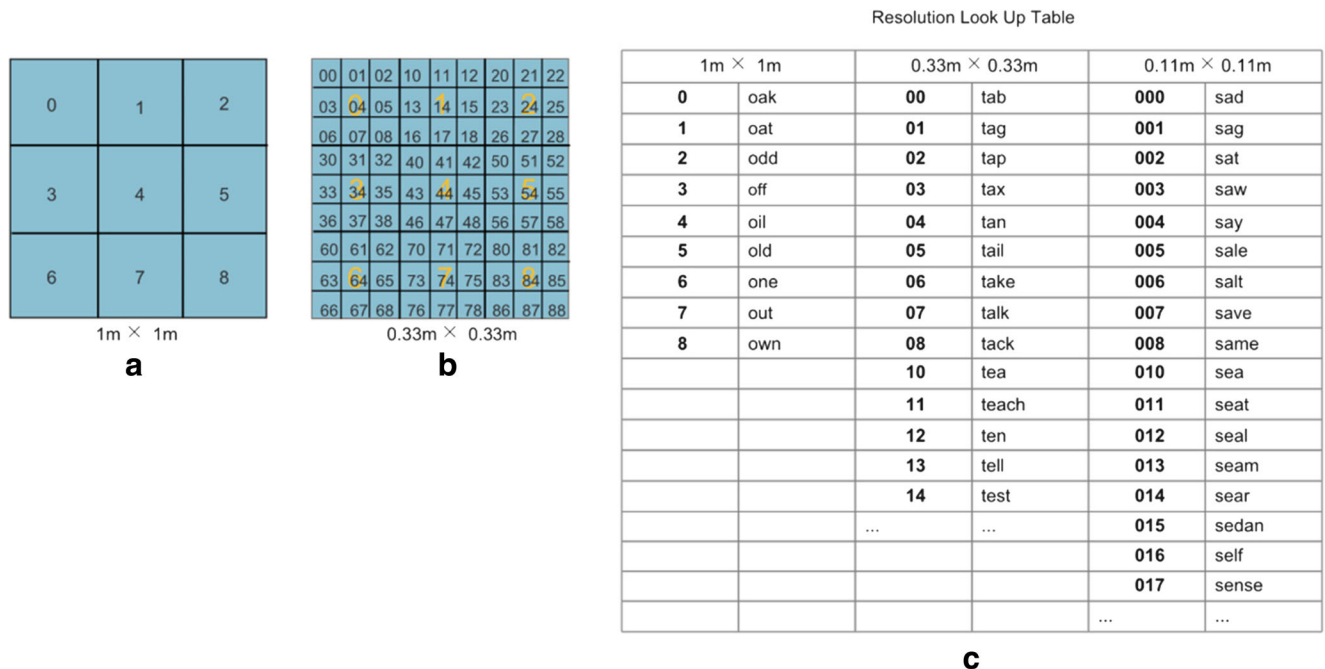


Fig. 6 Ternary-Tree Extension Model (TTEM): **a** resolution of 1 m × 1 m, **b** resolution of 33 cm × 33 cm, and **c** the fourth word for each sub-cell at three different resolution levels

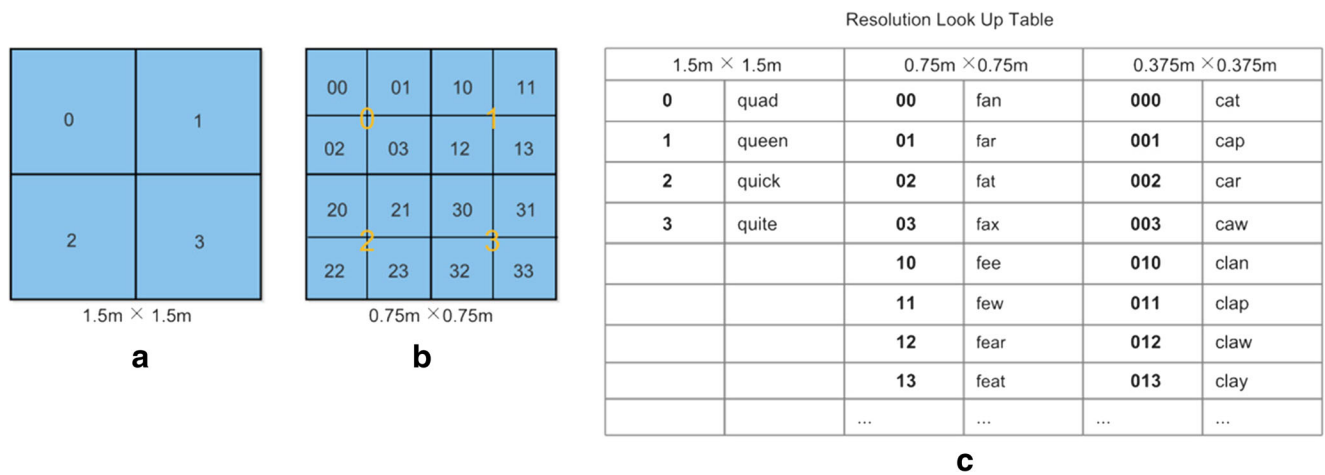


Fig. 7 Quadtree Extension Model (QTEM): **a** resolution of 1.5 m × 1.5 m, **b** resolution of 75 cm × 75 cm, and **c** the fourth word for each sub-cell at three different resolution levels

According to the nature of English vocabulary, the quantity of word starting with the letter “t,” “s,” “f,” or “c” is greater than words starting with “o” or “q” (Fig. 8). Therefore, these words were used for the sub-cells at the finer resolutions, where more words are demanded. Furthermore, when words starting with a certain letter are inadequate for a specific resolution (e.g., if words starting with “s” are not enough for TTEM at resolution 11 cm × 11 cm), words starting with another letter such as “p” could be utilized.

A fundamental principle of the w3w geocoding system is that the words assigned to a square give no clue as to the words of adjacent squares (Barr 2015). In other words, the w3w geocoding system is non-hierarchical and non-topological. In TTEM and QTEM models, though, the fourth word conveys some locational semantics as it corresponds to a given sub-square within each 3 m by 3 m square.

TTEM and QTEM models could be used to represent any point location. The accuracy of the point’s location increases with an increase in resolution. Given point A in Fig. 9, the four-word code for the sub-square including point A could be used to represent point A’s location through the cell’s centroid. The maximum offset from A’s actual location is 23 cm at

resolution 0.33 m × 0.33 m in the TTEM model. The higher the resolution level, the smaller the maximum offset is.

An alternative model to represent a point location is the binary code extension model (BCEM). This model was inspired by Geohash geocoding system (Geohash 2017), which recursively bisects the earth’s sphere along latitude and longitude. A binary code was then created for both coordinate values, while the location could be encoded by interleaving the binary representations of these coordinate values. BCEM recursively divides a w3w square in a quadtree fashion, with two binary digits (0 or 1) appended to each division to determine the quadrant. The divisions could continue until the error tolerance of the point is reached. Then, the string of appended binary digits is encoded into letters to form the fourth (new) word. Notice that the sequence of letters does not necessarily form a dictionary word. Figure 10 illustrates the process and results of the extension through the BCEM model. BCEM can be expanded or truncated in pairs to achieve higher or lower precision respectively (Table 1). The spatial relation of two points in the same w3w square could be partially retrieved by detecting the fourth word.

Fig. 8 The number of English dictionary words starting with each letter (from <http://which-english-letter-has-maximum-words.html>)

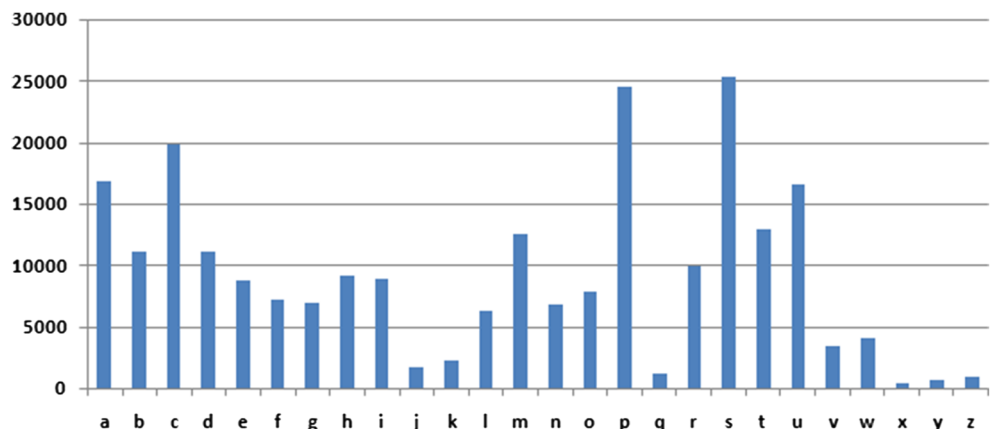
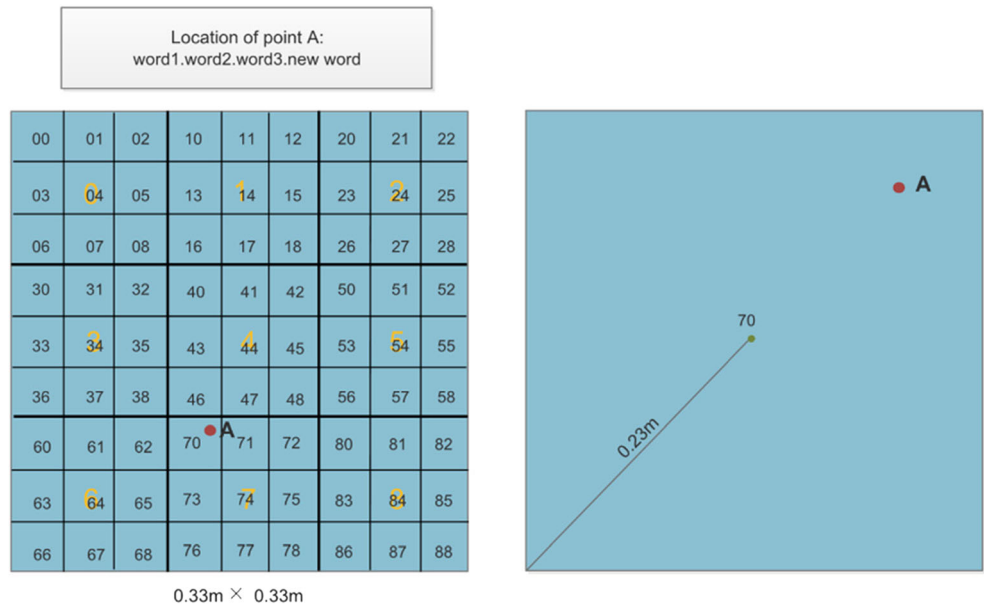


Fig. 9 Using Ternary-Tree Extension Model (TTEM) to represent an arbitrary point A



Variable Resolution Extension Models—Coarser Resolution

Three methods of coarser resolution extension are considered in this study. The first two methods use a $w3w$ square as a reference to decrease the resolution by expanding the size of the square. The first method, called radial expansion extension model (REEM), enlarges the square’s size by applying a radial expansion as shown in Fig. 11a. The second method, called diagonal expansion extension model (DEEM), uses a diagonal expansion as shown in Fig. 11b. Instead of adding another word to $w3w$ code to represent the lower resolution square, the fourth element for these models would be a label of a representative letter

and a number that defines the resolution itself (Table 2). Notice that the diagonal expansion could have four directions. Thus, the abbreviations for these directions need to be attached to the label.

The third method of coarser resolution extension, named rectangle extension model (RECTEM), can map the square into a rectangle of any size with dimensions denoted as m and n (Fig. 12). The new code is constructed by taking the centroid of a $w3w$ square, considering it as the center point of a rectangle, and using the length and width parameters delimited by a letter “v” to form the fourth element of the extended code (Table 2). Hence, the area covered by the rectangle could be computed by the fourth element as a product on $m \times n$.

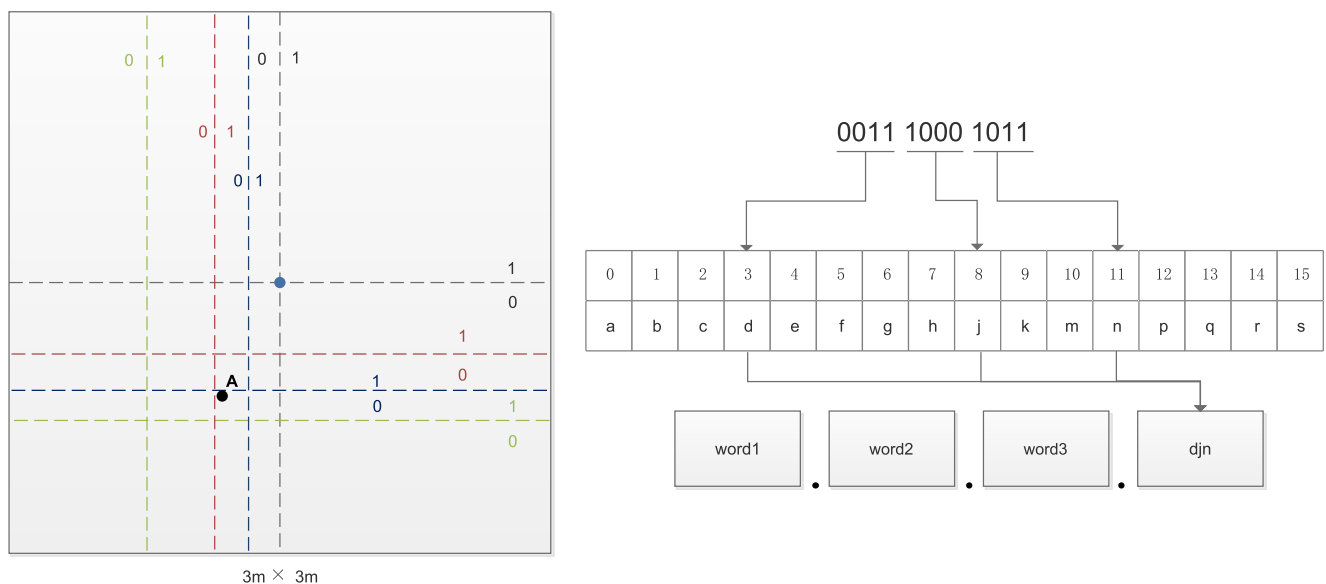


Fig. 10 Binary code extension model (BCEM). Modeling location A with three words plus a sequence of Latin characters

Table 1 The location precision for a point under BCEM

The 4th element length (number of letters)	Code bit	Max_vertical offset (m)	Max_horizontal offset (m)
1	4	0.75	0.75
2	8	0.1875	0.1875
3	12	0.0469	0.0469
4	16	0.0117	0.0117

The Third Dimension Extension Models—Elevation

Like many of the other geocoding systems, w3w does not consider the third dimension. Therefore, this study proposes an extended model for w3w that can distinguish between locations of different elevations. The elevation can be described in three ways: absolute height, relative height (ground reference; above or below the ground), and floor (floor label and relative height within the floor) for indoor applications. To represent the elevation, a label combining letters and numbers is attached to the w3w code (Table 3).

The letter(s) H, RH, and F represent the diverse types of elevation, i.e., absolute, relative, and floor, respectively. The number following the letters H and RH describes height values in units of meters. Letter F is followed by the floor label (e.g., 3, 0, -2, or E), a period, and

optionally a number describing the relative height from the surface of the floor in meters. Various types of floor description (i.e., Ground Floor, Floor C) need to be considered and standardized.

Combination of the Extension Models

To describe a precise location in space, the extension models addressing both variable resolution and elevation need to be integrated. Therefore, the full description includes five elements as shown in Fig. 13. A delicate balance exists between the descriptiveness and complexity of the code. While this new description loses some of the memory friendliness of the original w3w model, it retains many of the original aspects.

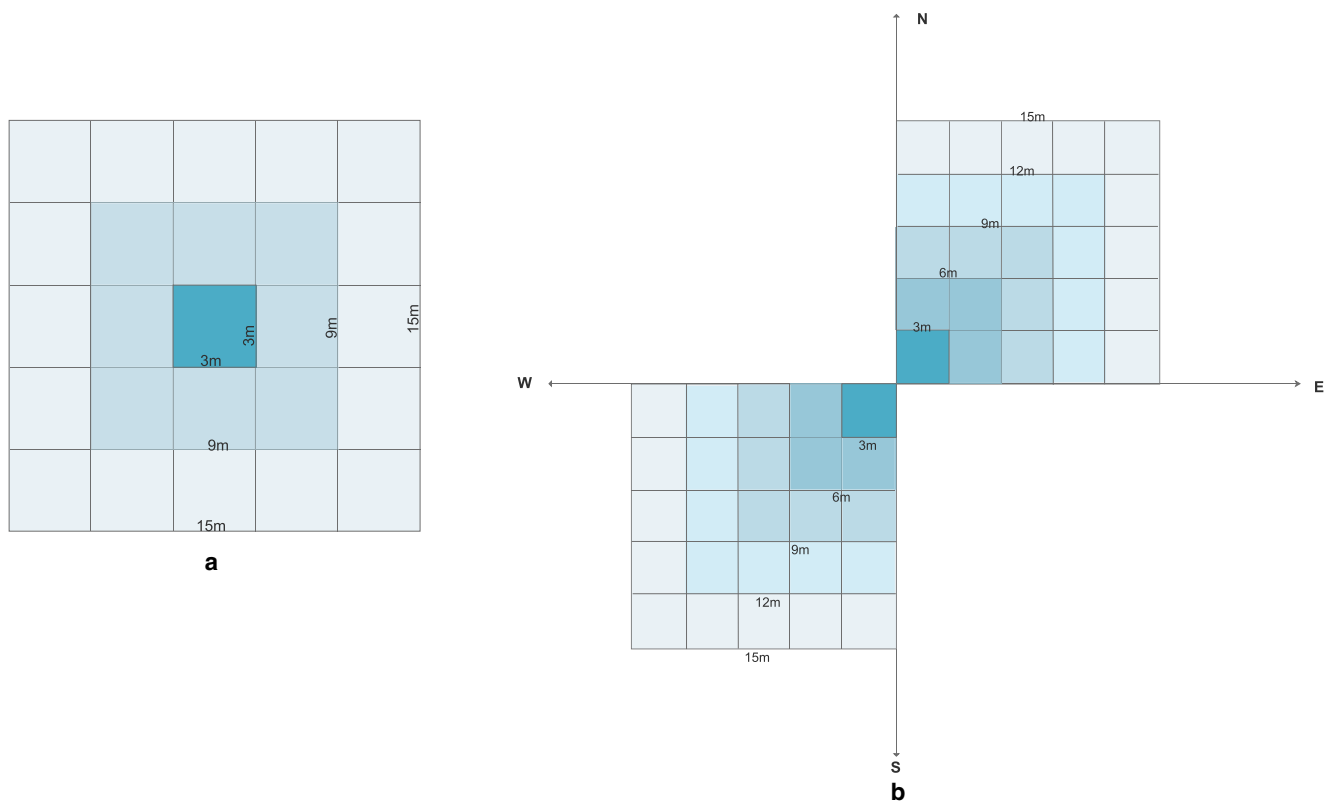


Fig. 11 Coarser resolution extension models: (a) radial expansion extension model (REEM) and (b) diagonal expansion extension model (DEEM)

Table 2 Alternative coarser resolution representations

<i>Method</i>	<i>Resolution Representation</i>			
Radial	R 9			
	R 15			
	...			
	D NW6	D NW6	D SE6	D SW6
	D NE9	D NW9	D SE9	D SW9
Diagonal	D NE12	D NW12	D SE12	D SW12
	D NE15	D NW15	D SE15	D SW15

	RECT7v5			
	RECTm _{vn}			
Rectangle	...			
	word1	word2	word3	R Num.
	word1	word2	word3	D (NW, SE,NE, SE) Num.
New code	word1	word2	word3	RECT Num.v Num.

Applications

The proposed extensions were able to accommodate a variable resolution and elevation, while still utilizing the w3w square as a basic unit. By adding a word and/or a label to a w3w code, these extensions retain a simple format and remain compliant with the advantages of the w3w geocoding system. Not only could they be used as input data for geocoding processing, but

also as a spatial index. Some potential applications that make use of the extension models for a university campus are listed in Table 4.

Provided that a university campus extends in a relatively small and local area consisting of up to a few hundred thousands of 3 m × 3 m squares, a single word (e.g., word1 in w3w) may turn out to be sufficient for encoding these squares. This way the combined code in Fig. 13 may be abbreviated to

Fig. 12 Coarser resolution extension model using a rectangle (RECTEM)

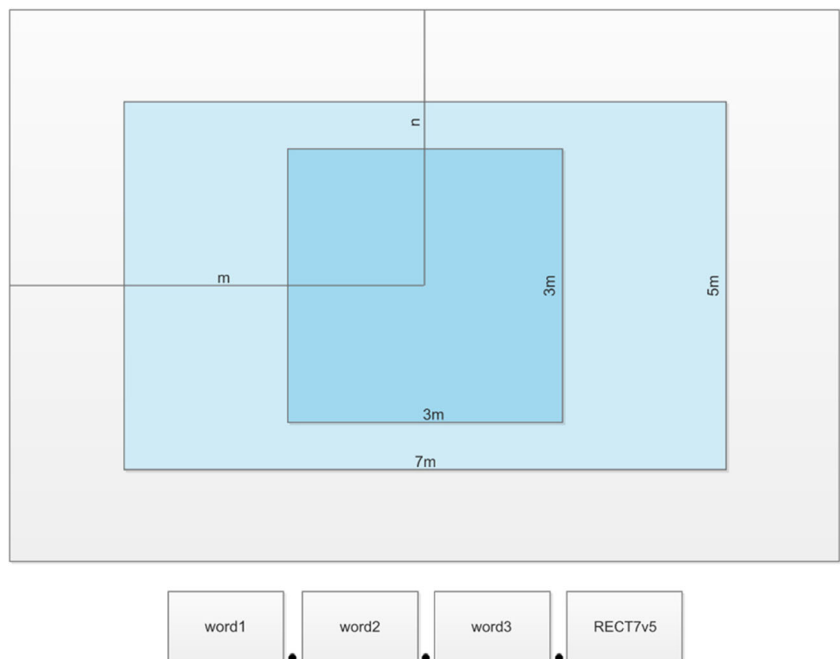


Table 3 Alternative extension models for the third dimension

<i>Elevation</i>	<i>Representation</i>	<i>Example</i>
Absolute Height	H	H 34.05
Relative Height	RH	RH 3.52
Floor Number	F	F 3.1.47
New Code	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; background-color: #4a7ebb; color: white; padding: 2px 5px; border-radius: 5px;">wordA</div> • <div style="border: 1px solid black; background-color: #4a7ebb; color: white; padding: 2px 5px; border-radius: 5px;">wordB</div> • <div style="border: 1px solid black; background-color: #4a7ebb; color: white; padding: 2px 5px; border-radius: 5px;">wordC</div> • <div style="border: 1px solid black; background-color: #4a7ebb; color: white; padding: 2px 5px; border-radius: 5px;">H(RH/F) Num.</div> </div>	

three words: “word1.resolution.elevation” for local use. Obviously, a search must be done to make sure that no two w3w squares in the campus share the same word1 of the original w3w code.

The following two example scenarios demonstrate how what3words and the extended models may enhance geocoding. The first scenario involves a mail delivery on a university campus. Specifically, a package must be dropped off at the desk of a person in room E5, which is located on the fifth floor (floor E) of Head Hall complex. The street address of Head Hall is “15 Dineen Drive, Fredericton, NB, E3B5A3”. As shown in Fig. 14, neither the street address nor the place name (i.e., “Head Hall, UNB Campus”) is very helpful to the postman, as they both fail to indicate an entrance to the building complex.

As explained in “Geocoding System and what3words” section, address-based geocoding systems may fail to represent large structures, such as a building complex, as they assign a single or a limited number of addresses to each structure. To overcome this limitation, the what3words geocoding system provides a unique address made up of three dictionary words to identify each 3 m by 3 m square on the earth’s surface. As shown in Fig. 15, the main entrance to Head Hall complex falls into the w3w square cell indexed by the three-word code: “tram.sullen.registration”. The centroid of this cell is located at -66.641686 W, 45.949803 N. Obviously, the w3w address code may help the postman locate the entrance to the building and reach out the delivery spot more efficiently. The

navigation in the outdoor space will be supported by satellite positioning systems.

The package must be dropped off at the desk of a person in room E5, which is located on the fifth floor (floor E) of Head Hall complex. Hence, the postman needs additional information to navigate within the building complex. As shown in Fig. 16, the horizontal location of room E5 is indexed by the three-word code: “psychic.rolling.recital”. However, the room is located on the fifth floor (or floor E; according to the floor numbering scheme of this complex). Hence, the extended code for room E5 is either “psychic.rolling.recital..FE” when using the floor number or “psychic.rolling.recital..RH12” when using the relative height (see Table 3; for a floor height equal to 3 m). The latitude and longitude of the square cell can be retrieved from the first three elements (w3w code), while the height value can be retrieved from the fifth element. The extended code may better assist the postman to reach out the delivery spot. The navigation within the building will be supported by indoor positioning systems.

Notice that there is no fourth element included in the latter two codes. The postman can only locate the right desk in E5 to drop off the package at, if the fourth element became available. As shown in Fig. 17, a finer resolution can be achieved by subdividing the 3 m × 3 m square into 81 33 cm × 33 cm sub-squares. Those sub-squares will be identified through the fourth element (see Fig. 6). Two alternative extended (five-word) codes for the desk surface are “psychic.rolling.recital.tan.RH13” and “psychic.rolling.recital.tack.RH13”. Notice that the relative

Fig. 13 The combined five-word code

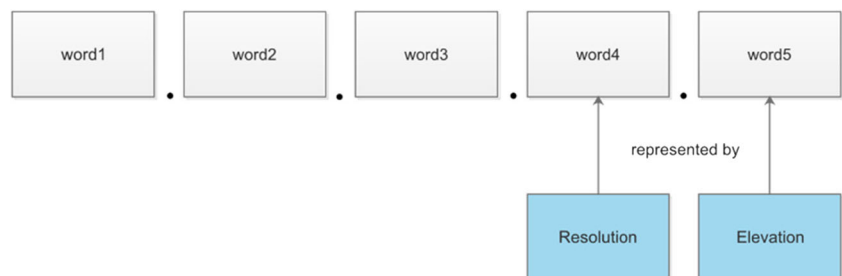


Table 4 Applications of the proposed geocoding extension models to a university campus

Location	Application
Indoor	(1) Describe a path between rooms and support navigation
	(2) Label rooms and important sites, such as exit, entrance, and extinguisher
	(3) Track equipment, books, and personnel
	(4) Locate facility management points of interest (POI) for maintenance or repair
Outdoor	(1) Define safety area for emergency evacuation
	(2) Describe paths between various locations (buildings) on campus
	(3) Manage student survey data

height is increased by 1 m compared to the floor height of 12 m in Fig. 16, to include for the height of the desk itself. Obviously, the extended five-word code may assist the postman to reach out the delivery spot accurately.

The second scenario handles the definition of an evacuation zone within a university campus. The evacuation zone is an open space that has accessibility to roads and is intended to accommodate students, staff, faculty, and visitors in case of emergency. Figure 18 shows an example evacuation zone. The size of the zone is such that encloses hundreds or thousands

of 3 m by 3 m square cells. Hence, this zone may be encoded with too many different w3w codes, while none of them convey the extent of the zone. By introducing a coarser resolution, the evacuation zone can be encoded more efficiently. The rectangular zone in Fig. 18 can be explicitly described through the extended code “nurse.marginally.animation.RECT76.5v34.5” and according to Fig. 12. Notice that the area covered by the rectangle is embedded in the code and can be extracted as the product of the two numbers reported in the fourth element.

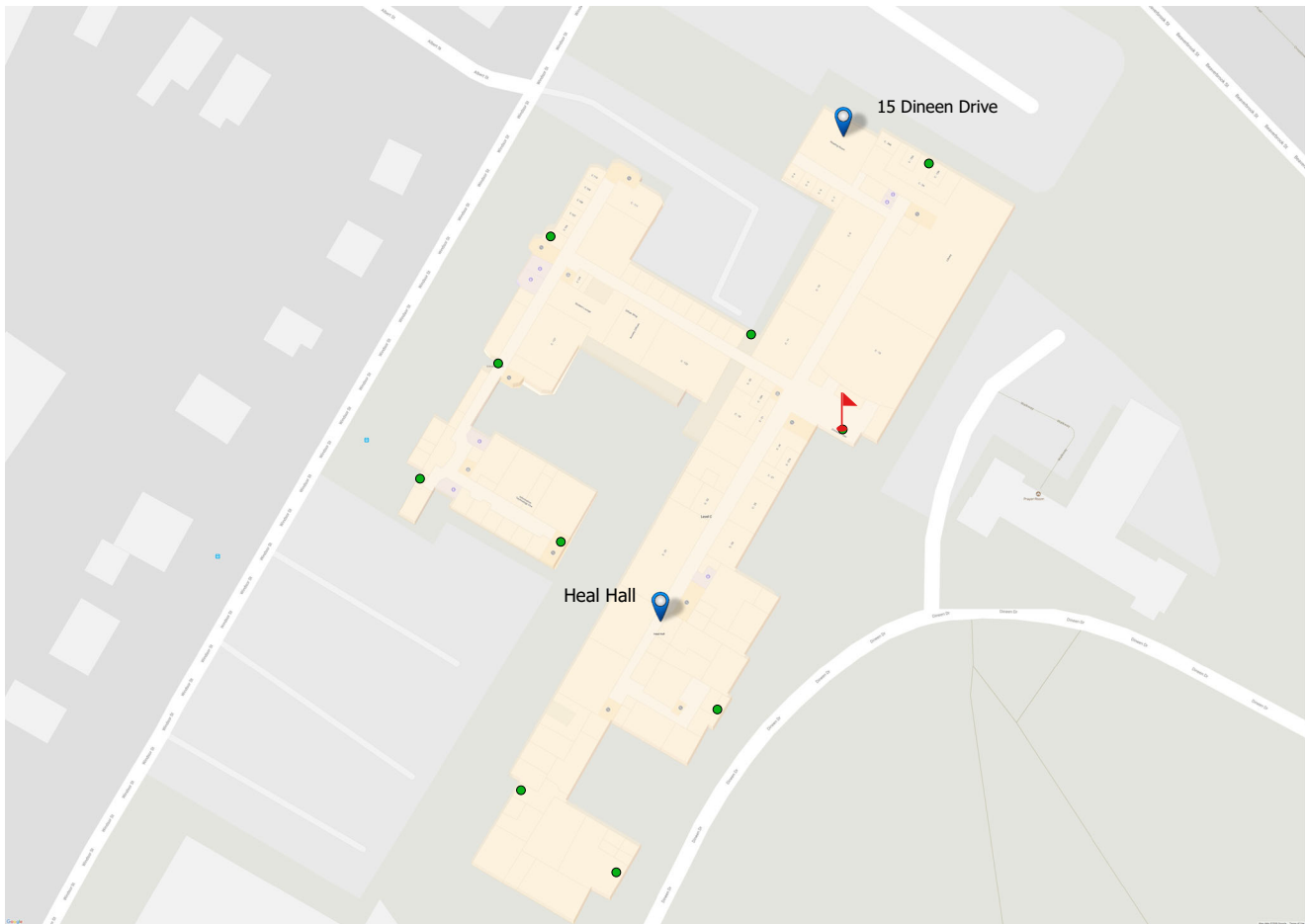


Fig. 14 The Head Hall complex. The red markers indicate the geocoding results based on street address and place name. The blue markers represent the entrances to the building complex, while the black flag indicates the main entrance

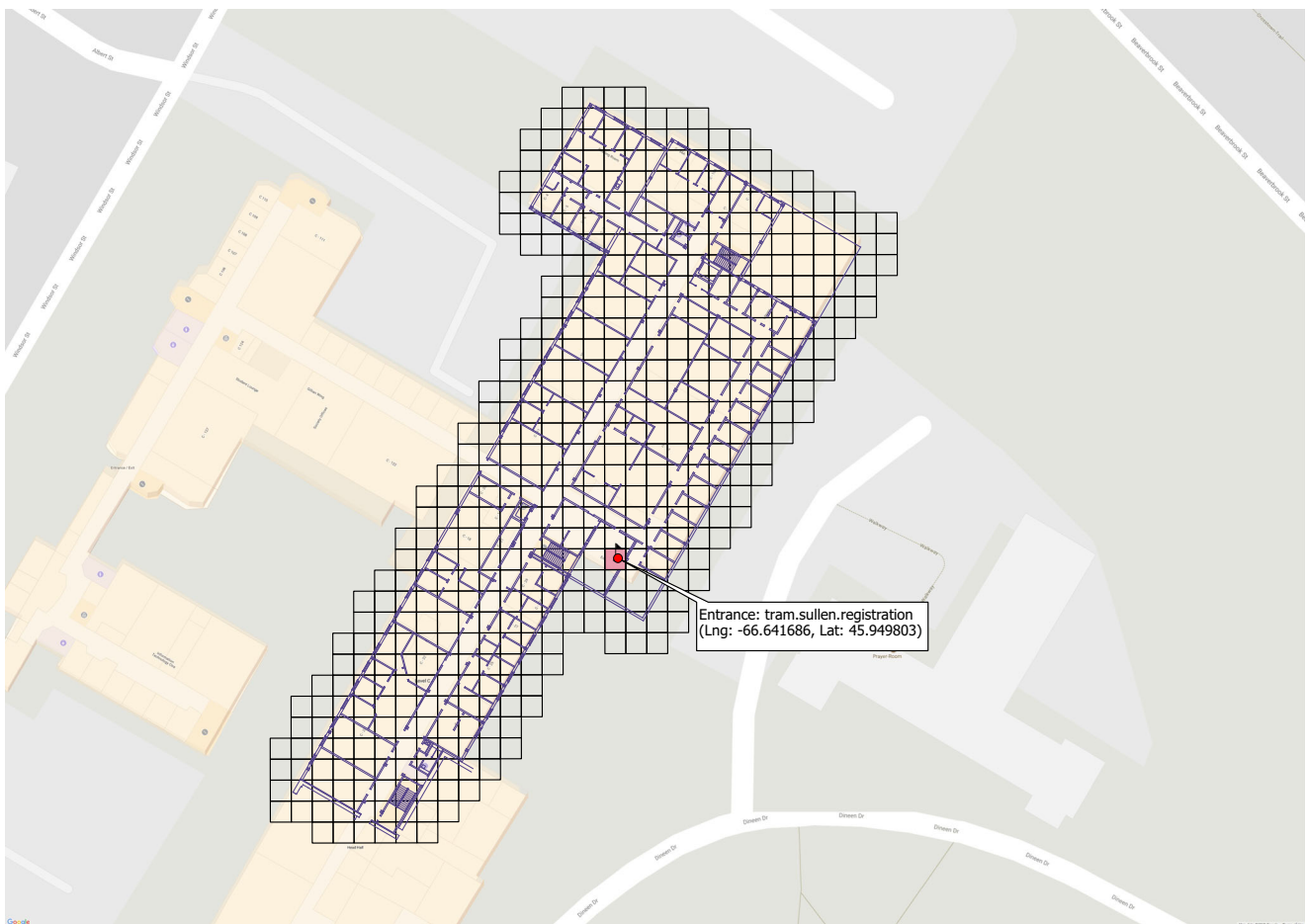


Fig. 15 The original what3words grid over Head Hall complex. The annotated $3\text{ m} \times 3\text{ m}$ square is used to encode the location of the main entrance

Clearly, both the w3w encoding and the proposed extensions assume a positioning precision that goes beyond the capabilities of common outdoor or indoor positioning systems mounted to mobile devices, such as cell phones. There are, however, systems that can reach the required precisions, such as the relative kinematic positioning (RKP; Teunissen and Montenbruck 2017) using satellite systems, or the ultra-wideband technology for 3D positioning (Pozyx 2017). It is anticipated that systems like those will be in wide use in the future.

A geocoder software tool is currently being developed. The tool will implement the two-way transformation between five-element codes (using the various w3w extension models introduced in “[Extensions of what3words](#)” section) and local or universal coordinates in 3D space for individual locations. Also, it will support the two-way mapping of basic geometric elements, such as lines, surfaces, and solids (used to model the entities of interest in various university campus applications) to extended w3w codes (Table 4).

Forward and Reverse Geocoding Transformations

The following paragraphs describe the two-way transformation between an extended code and the corresponding geographic coordinates. The forward geocoding converts an extended code into geographic coordinates including elevation. The reverse geocoding converts geographic coordinates and elevation into an extended code. Both transformations make use of the what3words API (Application Programming Interface; w3w API 2017) as an integral component.

Figure 19 presents the flowchart of the forward transformation. The input value is a code consisting of three, four, or five words (or elements). After parsing the code and depending on its content, one or more of following three procedures will be performed:

- (1) If a three dictionary words code were provided, the geographic coordinates of the centroid point (Lng, Lat) for the corresponding 3 m by 3 m square cell could be

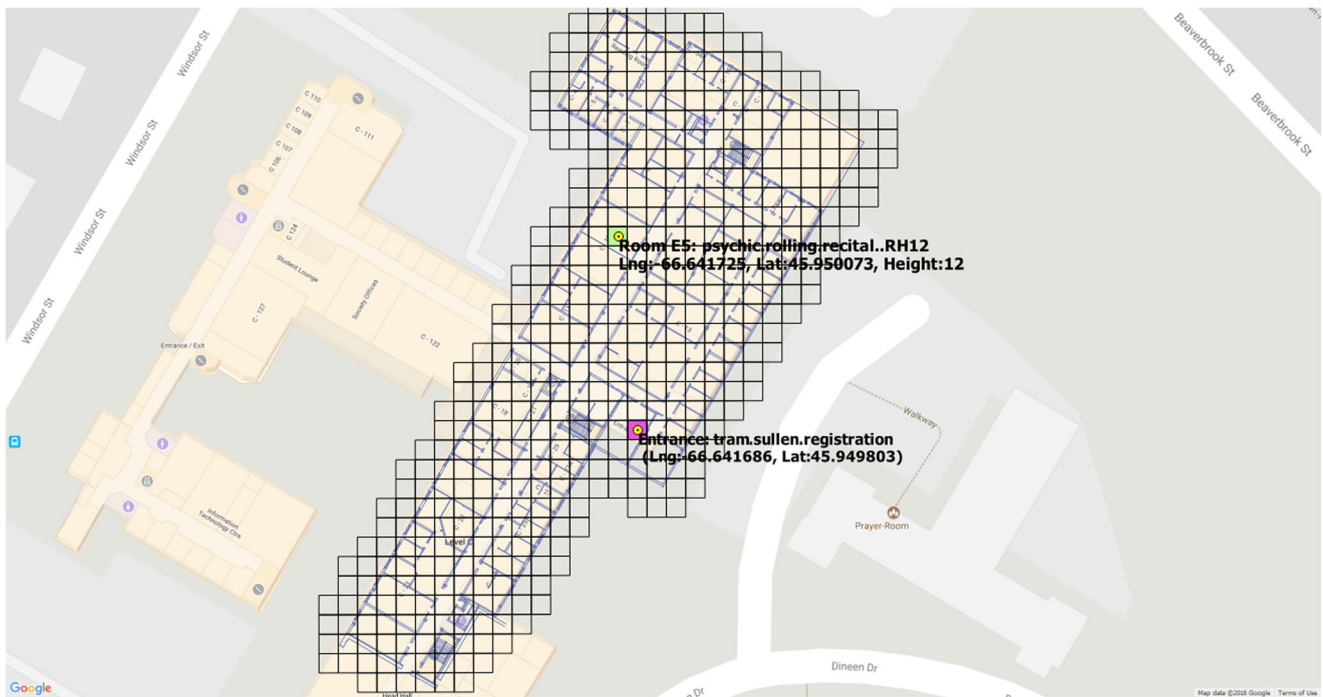


Fig. 16 The plan of floor E in Head Hall complex and the location of room E5. The red $3\text{ m} \times 3\text{ m}$ square corresponds to one w3w square cell full contained in the horizontal extent of room E5

retrieved using the what3words API. These coordinates will be reported as a result of the transformation.

- (2) If an extended code with the fourth element (word4) were provided, the first three words would feed the what3words API to retrieve the geographic coordinates of the centroid point (Lng, Lat) for the corresponding 3 m by 3 m square cell. Then, the fourth word or element will be used to calculate the offset from the centroid point (ΔX , ΔY). Depending on the extension model in use, the corresponding calculations will be carried out, and the geographic coordinates of the location will be retrieved and reported.
- (3) If an extended code with the fifth element (word5) were provided, that element would be processed according to the adopted model for the third dimension and the height (Z) of the location will be retrieved to complement the horizontal coordinates.

Figure 20 presents the flowchart of the reverse transformation. The input values of the reverse transformation include mandatory and optional items. The mandatory items are the geographic coordinates (Lng, Lat) of the location. The optional items are the resolution level (r) and the height (Z). By feeding the what3words API with the mandatory items, a three-dictionary-word code

will be retrieved to identify the 3 m by 3 m square cell enclosing the geographic location.

The optional item r takes a value that determines the resolution level expressed in centimeters. If the value of $r < 300\text{ cm}$, a finer resolution is needed. A series of calculations will be applied to compute the bottom left corner of the 3 m by 3 m square enclosing the location (denoted as LngMin and LatMin, respectively), the offset of the input location from that corner (denoted as ΔX and ΔY , respectively), and ultimately the sub-square that encloses the location at the requested resolution (identified by: row and column). A lookup table (see Figs. 7 and 9) will be used to retrieve the fourth element (word4) that corresponds to that cell. If the value of $r > 300\text{ cm}$, depending on the actual value of r , the fourth element (word4) will be retrieved (see Table 2). If a code representing an arbitrary point is requested (e.g., for $r = -1$), the quadtree sub-division algorithm will be used to retrieve the string of binary code which will next be transformed into a letter code (Fig. 10). That letter code composes the fourth element of the code (word 4).

The optional item Z takes a value that determines the height of the location. According to the adopted model for the third dimension, the height (Z) of the location will be transformed into the fifth element (word5) of the extended code.

The computational complexity of the above transformations is anticipated to be in the magnitude of the algorithm complexity

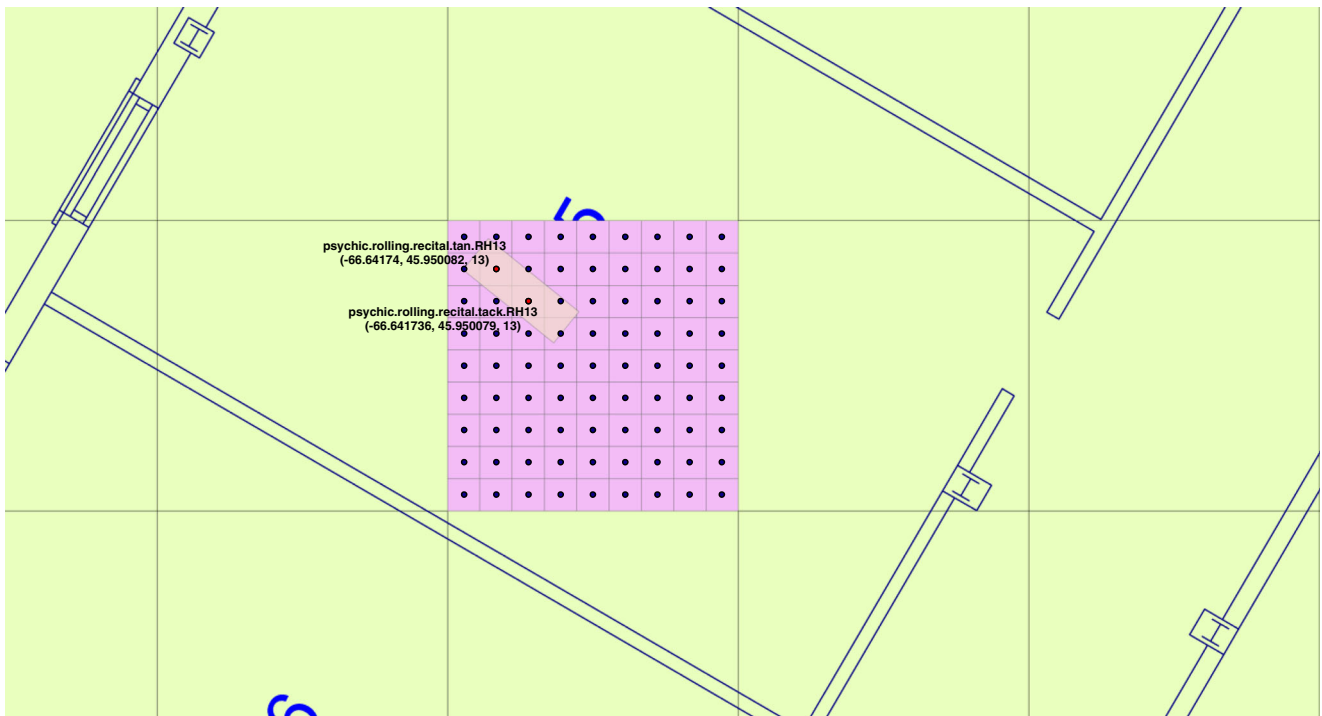


Fig. 17 A 3 m × 3 m square in room E5 further sub-divided into 33 cm × 33 cm sub-squares. A desk in the room encoded with an extended five-word code

of w3w API. The processing of the fourth and fifth elements involves very simple mathematical operations and/or a search in a lookup table of a few hundred of words at most (the number depends on the resolution level; see lookup tables in Figs. 6 and 7), which can easily be ordered and indexed so that a logarithmic performance is achieved. After the implementation and

validation of the geocoder software tool, a series of experiments will be carried out to prove the above argument and extract solid formulas describing the computational complexity of both the forward and reverse transformations.

The storage complexity is driven by the needs of the w3w API. The w3w system uses a proprietary algorithm in

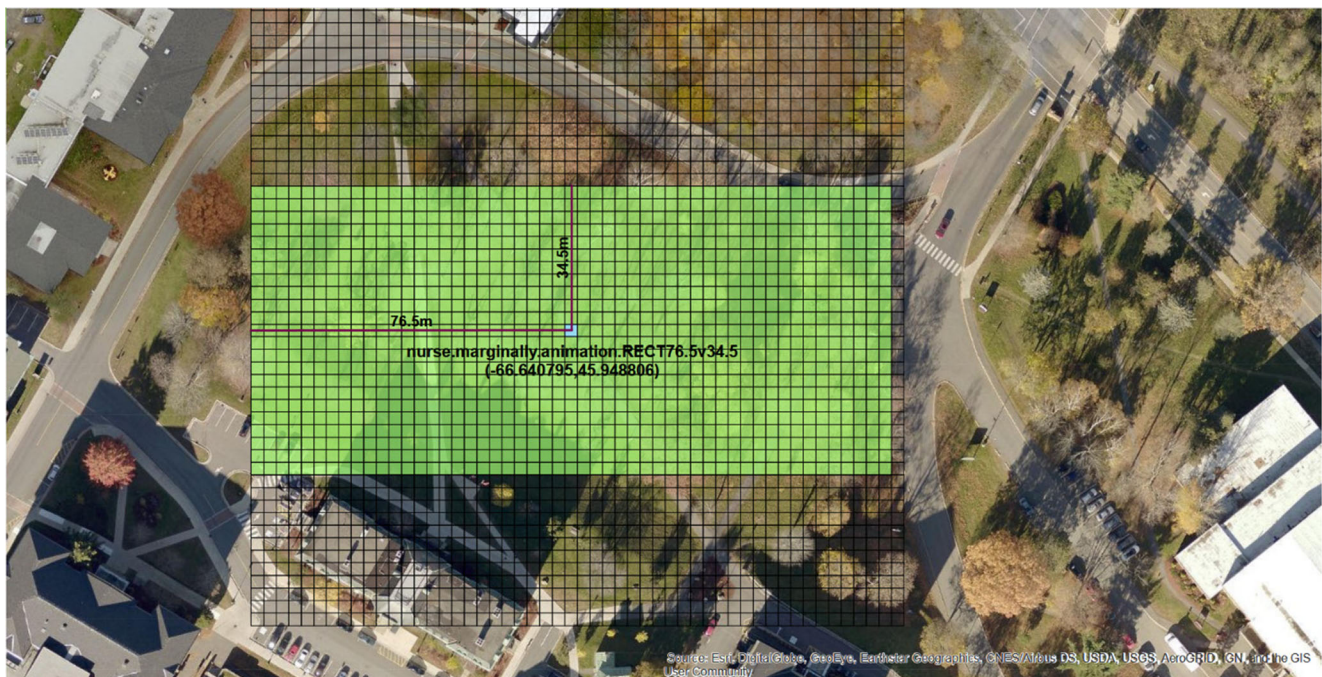


Fig. 18 An extended rectangular area aligned to the w3w grid encoded using the coarser resolution extension model

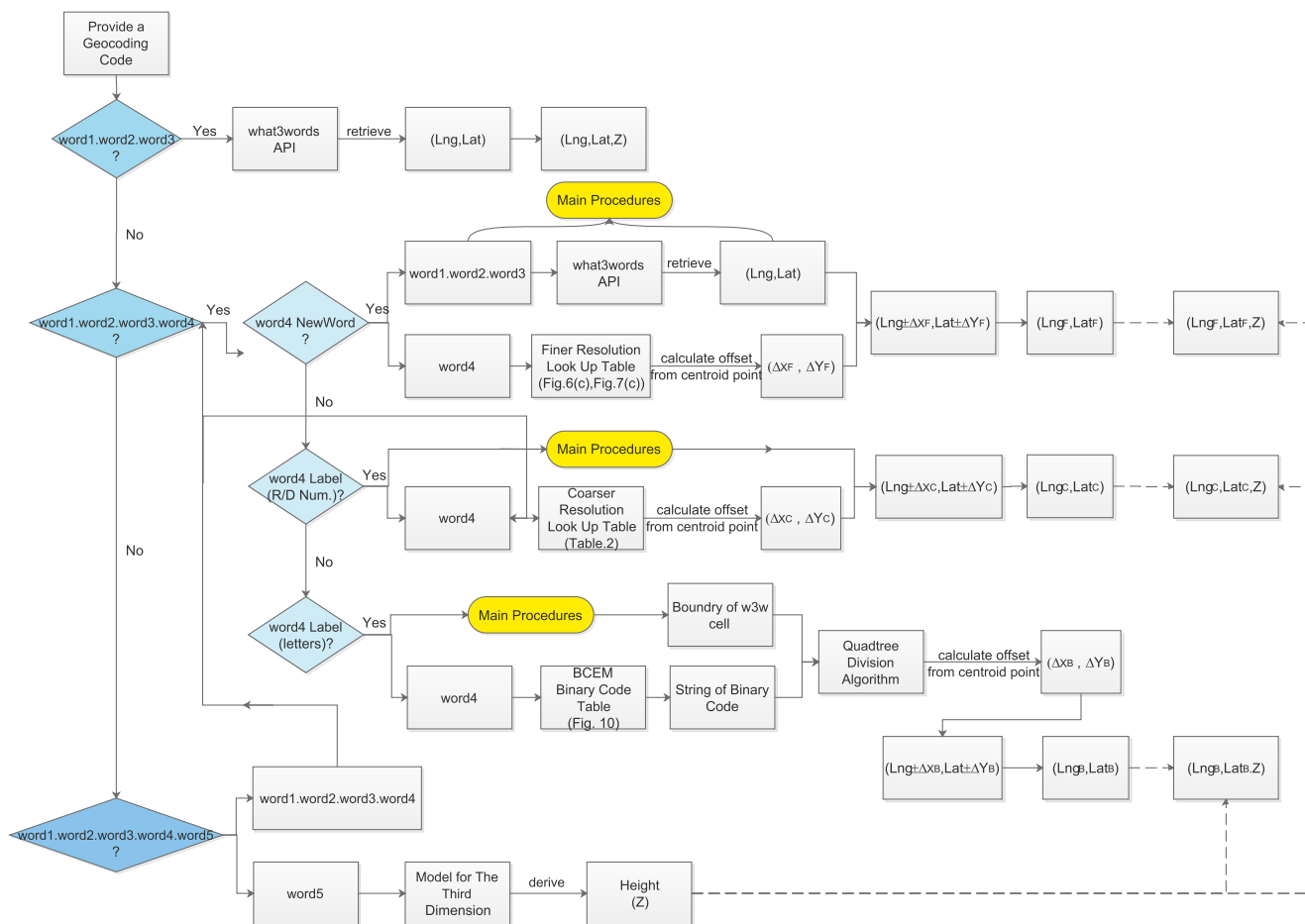


Fig. 19 Flowchart of the forward geocoding transformation. The what3words API is adopted for the transformation between the three dictionary words and geographic coordinates (denoted as “Main Procedures”). The chart is not extensive as regards to the calculation of the third dimension

combination with a database that stores the triplets of dictionary words for all $3\text{ m} \times 3\text{ m}$ square cells. The w3w core technology is contained with a file around 10 MB in size that can be run even when an internet connection is not available (Barr 2015). To support some of the extended models introduced in this study, some additional space is also required to store the corresponding look up tables (e.g., Figs. 6 and 7). Obviously, this space is minimal compared to the one needed for the w3w core database.

Conclusion

The address-based geocoding systems have been used for decades. However, these are not universally applicable in large-scale GIS applications. Alternative geocoding systems have been developed to overcome some of the constraints of address-based systems. Alphanumeric code geocoding systems divide earth surface into cells and assign each cell unique alphanumeric codes to represent the location. Each implementation of alphanumeric code geocoding utilizes its geocoding

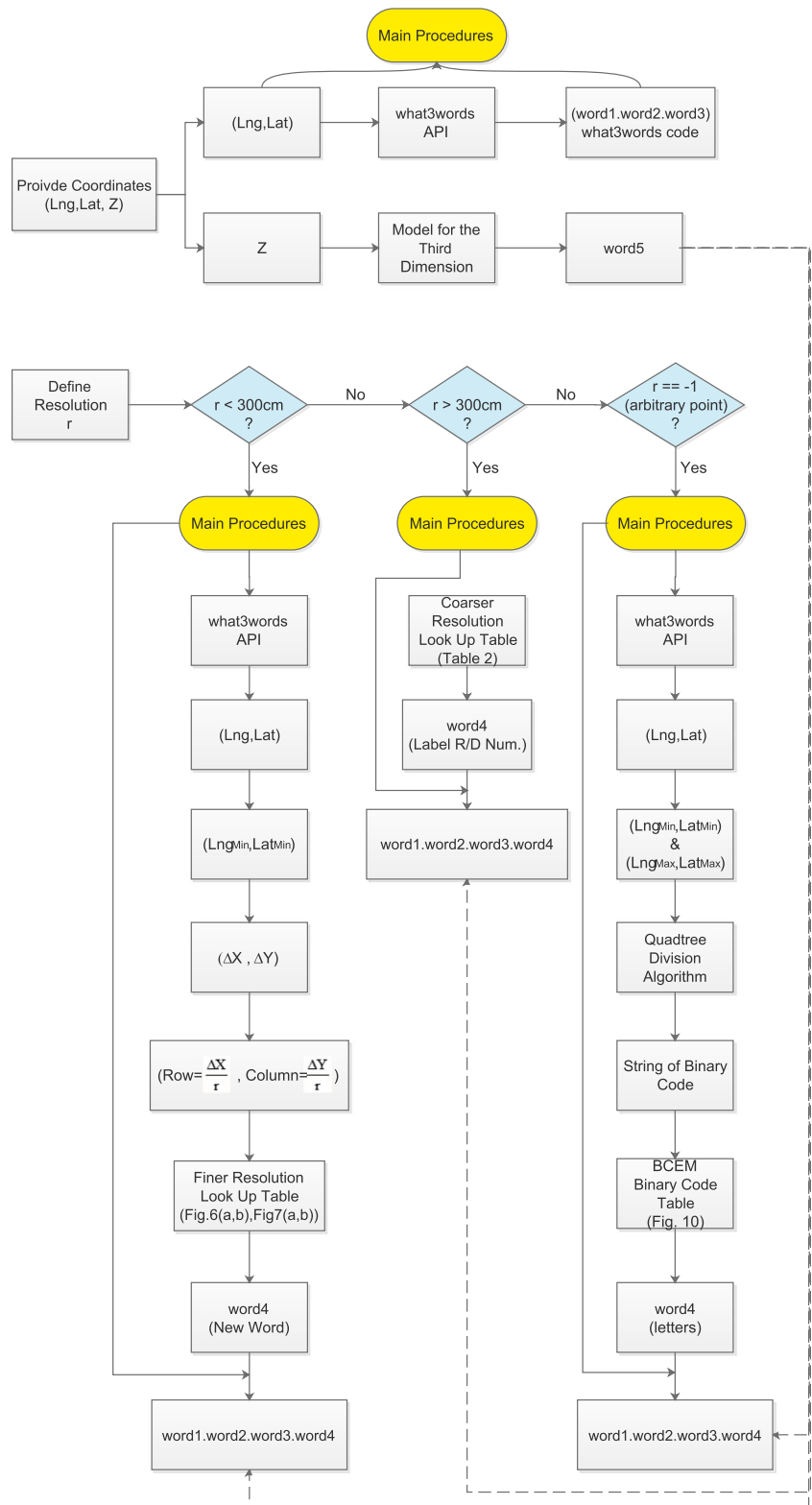
algorithms to transform between alphanumeric codes and the corresponding coordinates.

The w3w geocoding system has several advantages over other geocoding systems. However, due to its limitations, it needs to be extended, so that it can better support geographic applications. To support indoor applications, a finer resolution of squares is required, and elevations must be supported. For outdoor applications, a coarser resolution may also be needed. Therefore, this study proposes a series of extension models to the w3w geocoding system focusing on two main aspects, variable resolution, and elevation.

The variable resolution and elevation extension models are represented by the addition of elements to the standard three elements (words) used by the w3w model. The first three elements of the extended code are still the three words provided by the w3w geocoder. The fourth element is a word representing a variant resolution or an offset from the centroid of the standard w3w square. The fifth and final element is a label to representing elevation. These extensions comply with and retain the advantages of the w3w geocoding system.

The extension models proposed in this study may further enhance the w3w geocoding system to better support

Fig. 20 Flowchart of the reverse geocoding transformation. The what3words API is adopted for the transformation between the three dictionary words and geographic coordinates (denoted as “Main Procedures”). The chart is not extensive as regards to the calculation of the third dimension



geographic applications ranging from business and marketing to social and economic development of countries. The level of extension required in each application is variable and subject to the application needs as well as the decisions taken by both developers and end users.

Notice that, the first four elements (words) in the combined code refer to locations on the earth’s surface and their values are considered static. However, the fifth element refers to elevation, and its value (except the absolute height) depends on changes in infrastructure (e.g., building or road construction).

For this reason, a sixth element could be added to provide a temporal reference. The extension of the combined five-word code with a sixth word representing time may also serve in modeling spatiotemporal applications. This is another direction for future research.

Acknowledgments The authors wish to thank the editors and anonymous reviewers for their valuable comments on the initial manuscript.

Funding This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC-DG).

Compliance with Ethical Standards

The manuscript is compliant with ethical standards.

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval Not needed.

Informed Consent Not needed.

References

- Barr R (2015) About what3words. Retrieved from [what3words.com: https://what3words.com/about/](https://what3words.com/about/) [visited on October 31, 2017]
- Chen D, Cheng C, Tong X, Yuan J (2016) Research on the multi-scale spatial location coding model for address. *Geogr Inf Sci* 18(6):727–733. <https://doi.org/10.3724/SP.J.1047.2016.00727>
- Geelen P (2015) Mapcode downloads and developer resources. Retrieved from [mapcode.com: http://www.mapcode.com/downloads.html?iso3=112&mapcode=49.4V](http://www.mapcode.com/downloads.html?iso3=112&mapcode=49.4V) [Visited on October 31, 2017]
- Geohash (2017) Tips & trick—geohash.org. Retrieved from [geohash.org: geohash.org/site/tips.html](http://geohash.org/site/tips.html) [Visited on October 31, 2017]
- Goldberg DW (2011) Advances in geocoding research and practice. *Trans GIS* 15(6):727–733. <https://doi.org/10.1111/j.1467-9671.2011.01298.x>
- Goldberg D, Wilson J, Knoblock C (2007) From text to geographic coordinates: The current state of geocoding. *URISA Journal* 19(1):33–46
- Karimi HA, Sharker MH, Roongpiboonsopit D (2011) Geocoding recommender: an algorithm to recommend optimal online geocoding services for applications. *Trans GIS* 15(6):869–886. <https://doi.org/10.1111/j.1467-9671.2011.01293.x>
- Kellison MT (2012) Address points and a master address file: improving efficiency in the City of Chino. Master dissertation. Los Angeles, California, USA: University of Southern California
- Lee J (2009) GIS-based geocoding methods for area-based addresses and 3D addresses in urban areas. *Environ Plan B Plan Des* 36(1):86–106. <https://doi.org/10.1068/b31169>
- Pozyx (2017) Pozyx accurate positioning. Retrieved from <https://www.pozyx.io/> [Visited on October 31, 2017]
- Rhind G (2015) How Google could improve open location code. Retrieved from [pcapredict.com: http://resources.pcapredict.com/index.php/google-open-location-codes/#comment-3294363901](http://resources.pcapredict.com/index.php/google-open-location-codes/#comment-3294363901) [Visited on October 31, 2017]
- Stefanakis E (2016) Location encoding systems—could geographic coordinates be replaced and at what cost? *GoGeomatics Magazine*. Retrieved from <https://gogeomatics.ca/location-encoding-systems-could-geographic-coordinates-be-replaced-and-at-what-cost/> [Visited on October 31, 2017]
- Teunissen P, Montenbruck O (Eds) (2017) Springer handbook of global navigation satellite systems (Springer Handbooks) 1st edn. Springer, Cham, p 1327
- The Ethicalist (2017) How three words are making the world a safer place. *Social Business*. Retrieved from <https://theethicalist.com/three-words-making-world-safer-place/> [Visited on October 31, 2017]
- Verbree E, Zlatanova S (2007) Positioning LBS to the third dimension. In: Cartwright W, Gartner G, Meng L (eds) *Location based services and telecartography*. Springer, Berlin, pp 107–118. https://doi.org/10.1007/978-3-540-36728-4_8
- w3w API (2017). *what3words restful API*. Retrieved from <https://docs.what3words.com/api/v2/> [Visited on October 31, 2017]
- Zandbergen PA (2008) A comparison of address point, parcel and street geocoding techniques. *Comput Environ Urban Syst* 32(2008):214–232. <https://doi.org/10.1016/j.compenvurbysys.2007.11.006>