Mapcodes: a new standard for representing locations

by Pierre Potgieter, EE Publishers

Mapcodes, a new location-referencing standard, are short codes, easily recognised, remembered and easy to communicate verbally. Mapcodes are also unambiguous, sufficiently precise, allow access to locations that have no address, and are international and universally supported.

Mapcode is a new standard for representing locations, and uses a short alphanumerical code consisting of two groups of letters and numbers, separated by a dot (see Fig. 1). TomTom founders Pieter Geelen and Harold Goddijn designed the system when the company expanded their operations to China and India, to address the problem of finding locations without addresses – which made postage, commerce and travel to these locations difficult. Today it is a free, brand-less, and international standard, with the Mapcode Foundation maintaining its algorithms and data tables.

The key idea of this standard is that every location can be represented by a nine-character mapcode, which can be shortened if used in a territory specific context. Furthermore, the shortest (four character) codes are reserved for big cities and densely populated areas. Easy to recognise, remember and communicate, and accurate up to 5 m, it is designed for everyday public use, and is therefore often described as "lending a human face to coordinates".

Since it is a patented universal standard, longitude/latitude coordinates can be converted to mapcodes, and the other way around. This conversion is not based on a formula, but instead uses an algorithm which makes use of additional data (as explained below). Being independent of local map data, and not restricted to a specific country (or politics), culture, or even alphabet, means that name changes, be it of towns or countries, have little effect on mapcodes.

Fig. 1: An example of a mapcode, in this case of the South African Parliament in Cape Town. The three-letter abbreviation ZAF denotes the country (usually left out in every-day use), and the code the precise location of a place within the country.

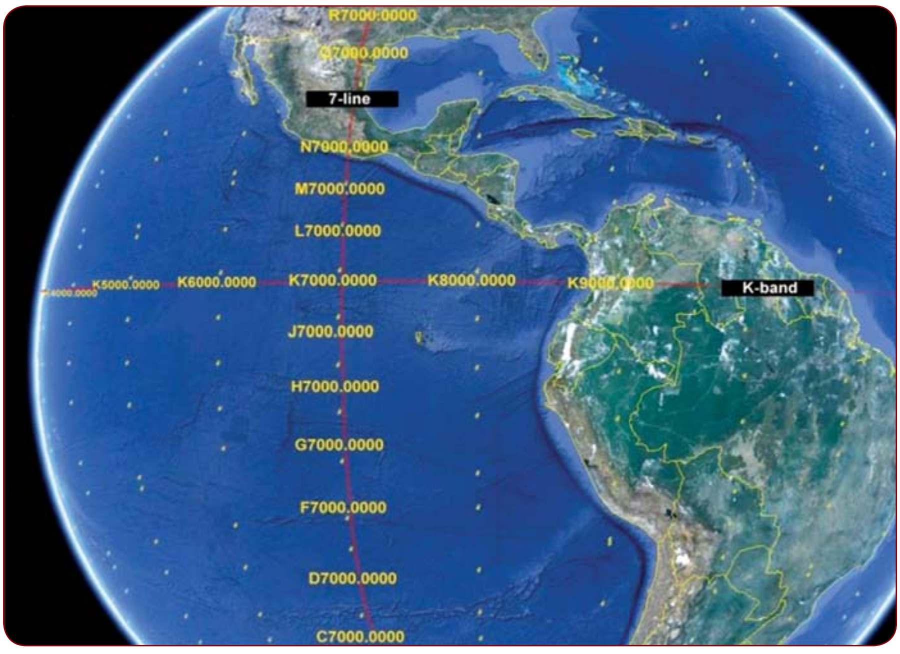

Fig. 2: Mapcodes work by dividing Earth's surface into (nine-character) numbered grids, which, when used alongside the data table of country names, can be shortened even further [1].

How it works

The mapcode system works by dividing Earth's surface into rectangles, or numbered grids (see Fig. 2). Basically, each grid (or area) is subdivided into sub-grids (or sub-areas), which are again divided into sub-sub grids and so forth until it produces a grid of sufficient precision (generally 5 m or less) to reference specific locations. These grids are called the encoding zone of the mapcode. The structure varies. The first two letters of a mapcode might divide the encoding zone into 31-by-31 subzones. The next letter may divide each subzone into even smaller 5 x 6 zones (see Fig. 3 a – e). In this manner, short codes, as opposed to long longitude/latitude coordinates, are used to represent any location.

This way, the mapcode system overlays the Earth's surface. Nine characters can thus represent any location on the Earth's surface, but providing context (such as a specific country) makes for even shorter codes. For example, the "full" (international) mapcode for Soccer City is 9KXGV.ZDMD. By stating that Soccer City is in South Africa, it can be shortened to ZAF H7B.N95. When used in South Africa, with the country context assumed, it can be communicated merely as H7B.N95, as shown in Fig. 4 (a) and (b).

Two key ideas in the design of the mapcode system are (i) that all mapcodes needn't be the same length, and, (ii) that short codes (fewest characters) are reserved for the most densely populated areas such as Manhattan in New York for example (Fig. 5). As such, mapcodes typically contain between four and seven digits.

The mapcode system uses 31 characters, and excludes vowels and letters such as I and O which can be confused for 1 and 0. To further avoid possible confusion, letters are also not case sensitive, and are written in uppercase. A data table defines mapcodes for every "context" on Earth (Fig. 6), with areas claimed by two countries included in both references.

The dot separator too fulfils an important function in indicating location, and moving it even in a seemingly similar mapcode such as from PQ.RST to PQR.ST changes the location referred to. Also, longer mapcodes, say a four-character as opposed to a five-character mapcode, can refer to several times as many locations, as there are more positions for the dot to be placed.

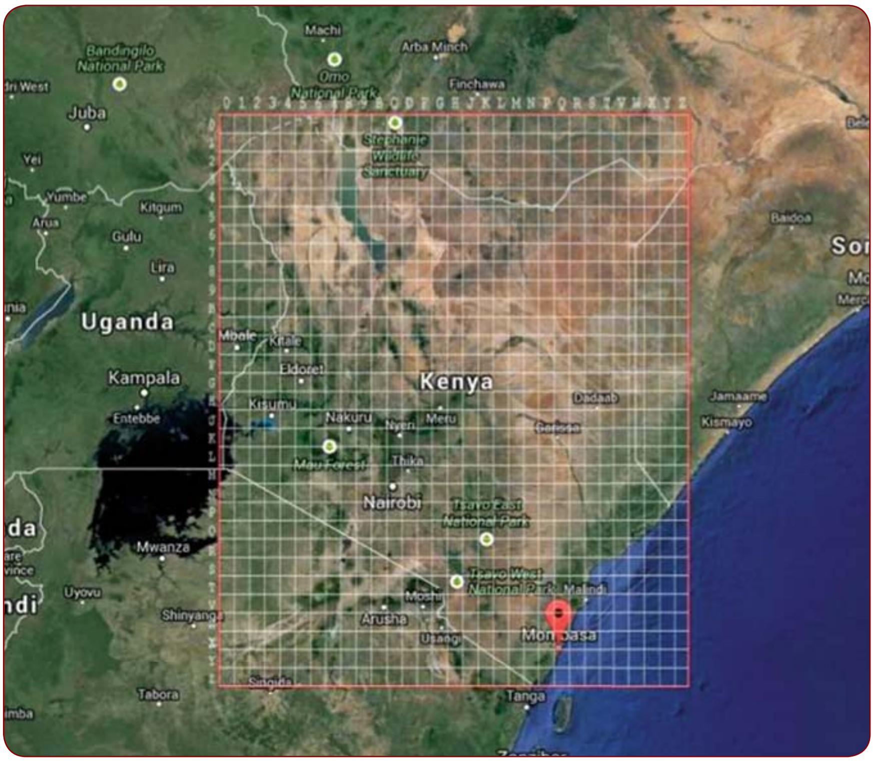

Fig. 3 (a): A virtual grid demonstrating the mapcode grid system [2].

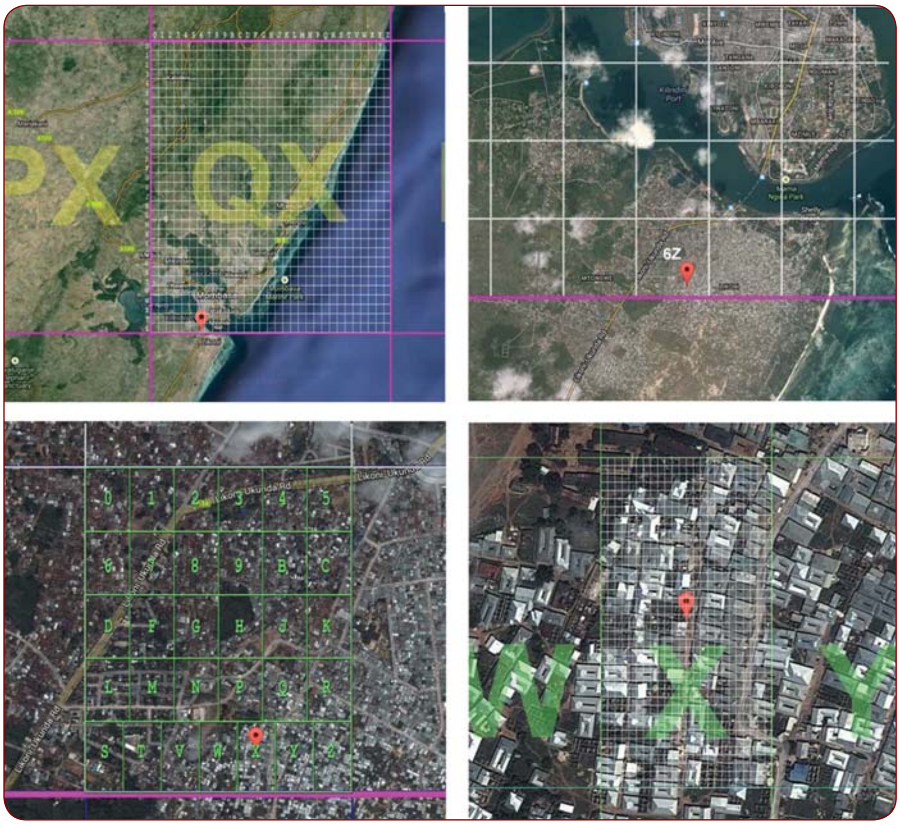

Fig. 3 (b – e): A series of sub-divided grids on different zoom levels demonstrate how the mapcode is made up and how it refers to a specific area. Inside sub-area QX is 6Z, denoted as QX6Z. Each level of drilling down further for more accuracy simply adds characters to the mapcode, until a complete mapcode such as QX6Z.XJ3 is formed [2].

Most locations are referenced using between four and seven characters. A few very large territories, such as Australia or Russia, would require eight-character mapcodes. In most cases these consist of sub-contexts such as (states or republics), that people living there will know, so that even there mapcodes are seven characters or less.

Although the source code is open and freely distributed, the Mapcode Foundation manages, maintains and regulates the algorithms and data tables, which remain crucial to mapcodes' ability to function as a standard system. Any change to the data table or algorithm can either result in different coordinates (to be avoided at all costs), or certain locations getting mapcodes that cannot be recognised on older systems (something that might be necessitated by changes to the world order, i.e. new countries coming into being).

Conversions and accuracy

As a standard, mapcode allows conversion from and to longitude/latitude coordinates. Unlike many other standards, it doesn't use a formula. Instead, algorithms using a 32-bit integer arithmetic are used for the conversion between latitude/longitude coordinates and mapcodes, as shortness of the code and its alphanumerical combination depends on this, and remains essential. It also makes the software efficient and easy to port to other computer languages.

A four-character mapcode covers a territory of 100 km^2, while a five-character mapcode can cover an area of 6000 km^2, and a six-character mapcode covers 250 000 km^2. Mapcodes are accurate up to 5 m to keep it practical for everyday use, making it accurate enough to distinguish between two entrances on the same property for example.

There is however an extended version of mapcodes available for higher accuracy requirements, producing 16 cm accuracy, though it arguably defeats the general purpose and practicality of mapcodes. The high-precision extension uses a hyphen, and adds additional characters to the mapcode for higher accuracy. The addition of one character provides accuracy of less than a metre, with the addition of a second character providing 16 cm accuracy. Although this idea seems easily extendable by adding more characters for higher accuracies, the current mapcode algorithms do not support this, as it was designed around 32-bit integer arithmetic for efficiency's sake, and to ensure that calculations do not suffer from floating-point errors.

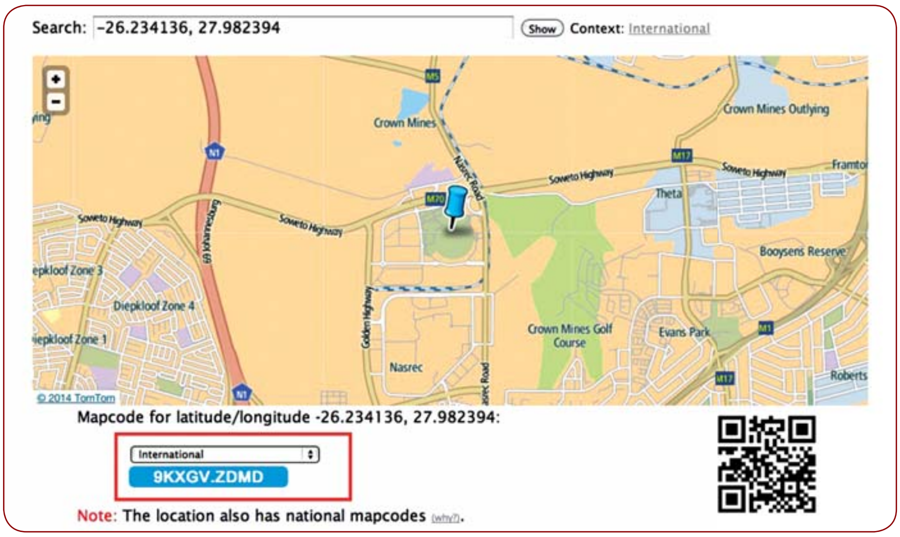

Fig. 4 (a): Besides contextual mapcodes, any location also has a context-less nine-digit mapcode, also referred to as the international mapcode.

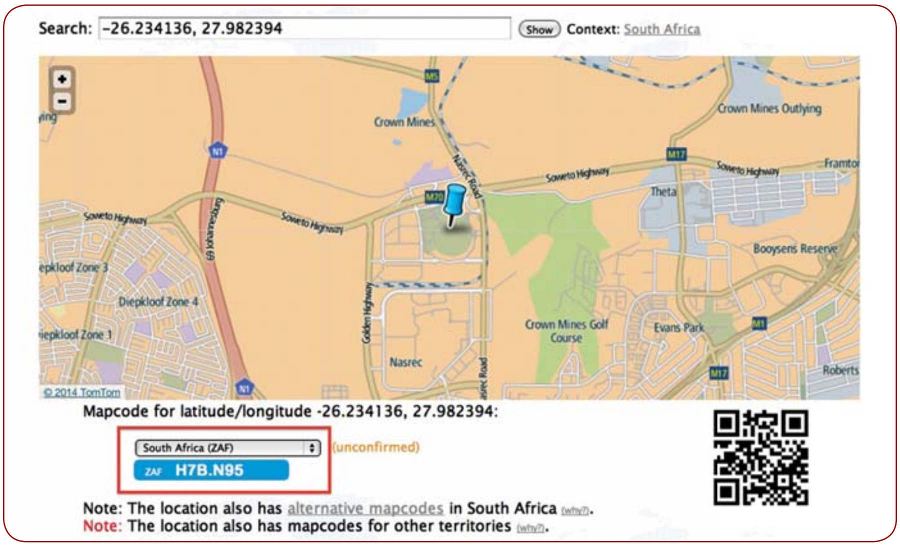

Fig. 4 (b): The nine-character mapcode for Soccer City in Johannesburg, South Africa, as shown in Fig. 4 (a), can be shortened to six characters by providing a country context.

Fig. 5: Short codes (four characters as opposed to five or more) are reserved for high-density population urban areas such as Manhattan in New York [1].

As a standard

The mapcode system is being filed as an ISO Standard at the International Organisation for Standardisation (ISO). It uses ISO 3166, the international standard for country codes and codes for their subdivisions, issued by ISO. To become a standard, mapcodes should be unambiguous and easily recognisable. The dot separator in the middle of the code is used, along with the lack of vowels, to avoid a mapcode looking like words. Likewise, consonants are incorporated to prevent a mapcode from looking like numbers. This is also important to avoid confusion with other content in online searches.

Furthermore, the default mapcodes system, based on the Roman alphabet, can easily be translated into several different alphabets by specifically chosen character substitutions. Great care was also taken with regard to inter-alphabet ambiguities, and the alphabets were encoded so that similar-looking symbols result in the same mapcode locations.

Substitution characters with few strokes and no accents were favoured, along with characters that are short and easy to pronounce. So too were characters that sound distinctly different, or don't require the use of special keys on computer keyboards. As with the default system, vowels were excluded whenever possible. The design staff also set some practical limits to the mapcode design, using only rectangular areas in their definitions, and limiting the size of the data table to 16 000 records.

The Mapcode Foundation, formally known as the Stichting Mapcode Foundation was established in April 2013. A non-profit organisation, it is based in the Netherlands. The foundation manages and maintains the algorithms and data tables (which include patents, rights, brands, designs, properties, collateral, algorithms, data tables etc.).

Territory	Parent	Short code	Full code	Code standard
Aaland Islands		ALA	ALA	ISO 3166-1 alpha-3
Afghanistan		AFG	AFG	ISO 3166-1 alpha-3
Albania		ALB	ALB	ISO 3166-1 alpha-3
Bonaire, St Eustasuis and Saba		BES	BES	ISO 3166-1 alpha-3
Bosnia and Herzegovina		BIH	BIH	ISO 3166-1 alpha-3
Botswana		BWA	BWA	ISO 3166-1 alpha-3
South Africa		ZAF	ZAF	ISO 3166-1 alpha-3
South Georgia and the South Sandwich Islands		SGS	SGS	ISO 3166-1 alpha-3
South Korea		KOR	KOR	ISO 3166-1 alpha-3
South Sudan		SSD	SSD	ISO 3166-1 alpha-3
Spain		ESP	ESP	ISO 3166-1 alpha-3
Texas, USA	USA	TX	US-TX	ISO 3166-2:US
United States Minor Outlying Islands, USA	USA	UM	US-UM	ISO 3166-2:US
US Virgin Islands, USA	USA	VI	US-VI	ISO 3166-2:US
Utah, USA	USA	UT	US-UT	ISO 3166-2:US
Vermont, USA	USA	VT	US-VT	ISO 3166-2:US
Virginia, USA	USA	VA	US-VA	ISO 3166-2:US
Washington, USA	USA	WA	US-WA	ISO 3166-2:US
West Virginia, USA	USA	WV	US-WV	ISO 3166-2:US
Wisconsin, USA	USA	WI	US-WI	ISO 3166-2:US

Fig. 6: Excerpts from the mapcode data table, or ISO Table of territories and their standard abbreviations. Most codes, with the exception of a few "legacy" codes are based on the latest ISO 3166 standards.

Fig. 7: Examples of substitution characters in the Hebrew and Cyrillic alphabets, which the mapcode system uses to make translation between alphabets easy without introducing problems.

References

[1] Stichting Mapcode Foundation.

[2] Etienne Louw: MapIT presentation: "A Short Address for Any Location on Earth", at GISSA Gauteng meeting & GISSA National AGM, 14 March 2014.

[3] Pieter Geelen. "The Mapcode System – reference document Version 1.04, 2013," on The Mapcode Website. Accessed online at www.mapcode.com/downloads.html?iso3=112&mapcode=49.4V (April, 2014).

[4] Interview with Mapcode Foundation managing director Kewal Shienmar at the mapIT-TomTom Africa User Conference, 21 February 2014.

[5] Etienne Louw, speaking at the MapIT-TomTom User Conference Africa, 21 February 2014.

Send your comments to: positionit@ee.co.za