

国内向けに調整した GeoHash による位置特定手法

概要

GeoHash (特定の位置を、初期ボックスから 4 分割とボックス選択を規定回繰り返し、位置情報コードを生成する)を基に、初期ボックスを、日本全国をカバーする (122° E, 20° N)–(149° E, 46° N) とする。base32 で符号化すると 10 文字で 1 辺 10cm 未満のボックスとなるので、建物の識別子に十分活用できる。

課題

位置情報に基づく地物の識別子であって、文字列長が冗長でないものを生成する手法。

課題解決のための手法

GeoHash の概要を図 1 に示す。位置 (経度、緯度) を引数に取り、ボックスを 4 分割して、引数位置が属する分割ボックスを選択して、対応する 2 ビットの数を追加することを繰り返す。分割・割当回数を Z とすると、ボックスを特定するためのビット列の長さは $2z$ となる。ビット列を base32 (5 ビットごとに分けて、対応する文字を割当て) で符号化すると、符号化文字数は $\text{Ceil}(\frac{2z}{5})$ 文字となり、個々の文字は、英字大文字と数字 (2 から 7 に限る) に限られ、見間違えることが少ない。

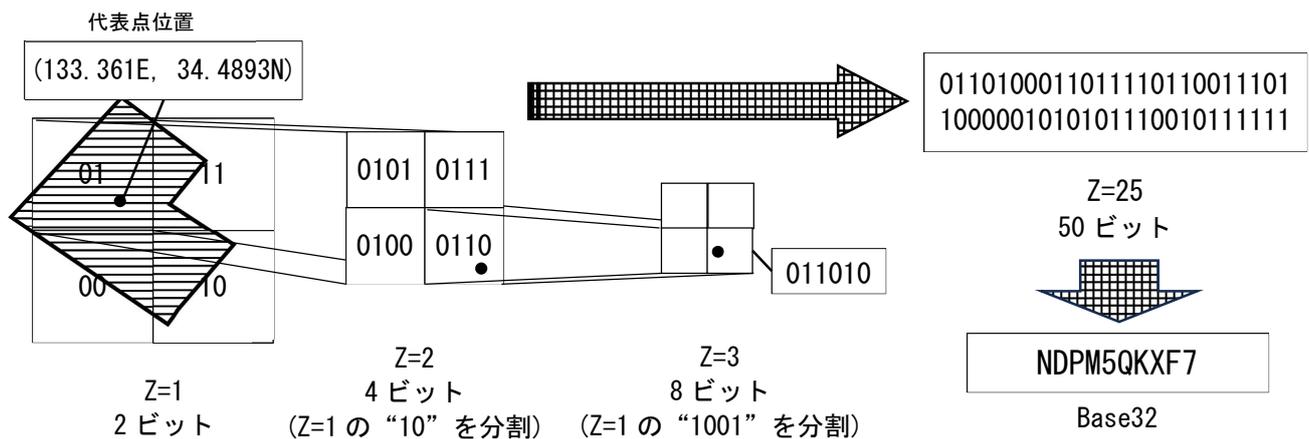


図 1 GeoHash 概要

本来の GeoHash の初期ボックスは地球全体の範囲とするが、提案手法では日本国内に限定するため、初期ボックスを (122° E, 20° N)–(149° E, 46° N) とする。符号化文字数を 10 文字とすると、 $Z = 25$ だから、経度、緯度単位のボックスサイズは、 $(27/2^{25}, 26/2^{25}) \approx (8.0466 \times 10^{-7}, 7.7486 \times 10^{-7})$ となる。

ボックスサイズをメートル単位で簡易的に計算する。地球を半径 6378137[m] の球面とし、初期ボックス内で最悪のボックスサイズになる 22° N で計算すると、メートル単位のボックスサイズは (0.0831, 0.0863) となる。

1 辺が 10cm 未満であり、建物を対象とする場合には、十分に一意に識別することができる。

利用法

建物のポリゴンの代表点を位置から提案手法を用いて 10 文字の英数字に変換したものを ID とする。代表点は重心を利用する。PostGIS の `ST_PointOnSurface()` を利用すると代表点がポリゴン内に存在することが保証される。