International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:12, 2011

A Strategy for Address Coding from HouseHold Registry Database

Yungchien Cheng, and Chienmin Chu

***Abstract*—**Address Matching is an important application of Geographic Information System (GIS). Prior to Address Matching working, obtaining X,Y coordinates is necessary, which process is calling Address Geocoding. This study will illustrate the effective address geocoding process of using household registry database, and the check system for geocoded address.

***Keywords*—**GIS, Address Geocoding, HouseHold Registry Database

I. Introduction

ADDRESS is common in all kind of files, using text to represent the location. In GIS, spatial information is crucial data base, text data can be representation spatially [1]. In early stage, the US used TIGER (Topologically Integrated Geographic Encoding and Referencing) system to represent address locations use census data transform into graphic appearance [2]. In Taiwan, there is no spatial database similar TIGER system. However, using GPS field measurements with basic household registry data from government to address geocoding is current practice in Taiwan.In Taiwan or other Asian countries, population household survey data is the most comprehensive address database. Using residence address information for address geocoding can build the most accurate database of address. However, when the residence address is missing, duplicate, and error, site investigation and follow-up in-house recover process is necessary. This study is to establish address database of twelve towns in Kaohsiung as an example, to bring up effective procedures of using census address coding and check system.

II. Data

A. Base map

The scale of base map is 1/1000, including building location, block boundary, cadaster line, road name, district boundary and district name, etc.

B. Administrative region house number list

Household information system- 20M house number list

C. Aerial photograph

This study us the latest aerial photograph to recognize the location of buildings.

Yungchein Cheng, Doctoral Candidate, is with the Institute of Remote Sensing and Geographic Information System, Peking University (phone: 86-10-62753220; fax: 86-10-62751150; e-mail: Mavin.cheng@gmail.com).

Chienming Chu is with the Tourism and Leisure Department, Hsing Wu Institute of Technology (phone: 886-02-26015310-2437; fax: 886-02-26010351; e-mail:100002@mail.hwc.edu.tw).

III. Study Aera

The study area includes twelve coastal towns in Kaohsiung city (Fig.1), and total area is 46,611 hectares. According to the house number statistics from Kaohsiung administrative region department, total number of household registry is 339,898.

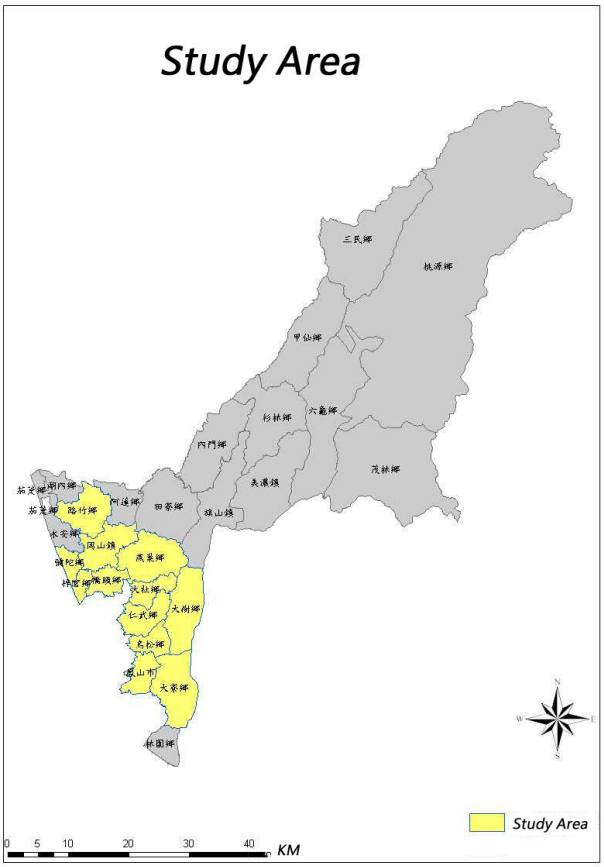

Fig. 1 Study Area

IV. Procedure Of Address Geocoding

There are four stages of address geocoding (Fig.2):

A. Survey Maps

Integrate the base map, address list and aerial photograph in advance. Draw the investigation map into A0 size (about 80cm*60cm) according to survey area with 1/500 or 1/1000 scale.

B. Field Survey

Main duties for field survey are house number confirmation, house number properties marks, existing house number information record and comparison with 20M list. Four results will be shown.

a. Existing house number
b. Not existing house number
c. New house number
d. Removed house number
e. Doubt house number (duplicate house number, wrong house number, etc.)

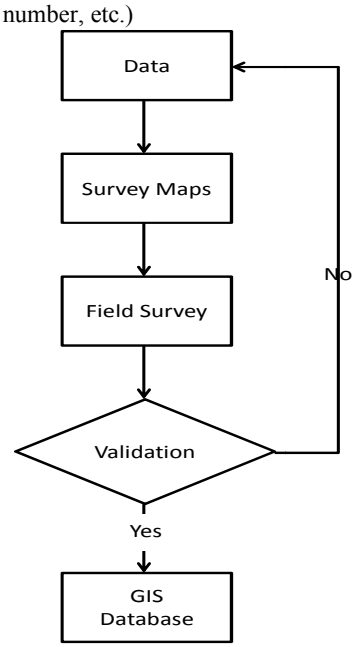

Fig. 2 Procedure of address geocoding

C. Check

There are two steps for checking house number information. If the correct ratio bellows 98% of each step, we need to re-investigation.

a. Systematic check

1. Quantity of house number and house number properties are correctness check: check the differences between existing house numbers and house registry numbers.
2. Buildings and the corresponding spatial location check: One is check the building number and the corresponding street segment, the other one is check the spatial sequence of building numbers. House address must be the same as the road name (Fig.3), and the house number must also be consistent with the order (Fig.4).

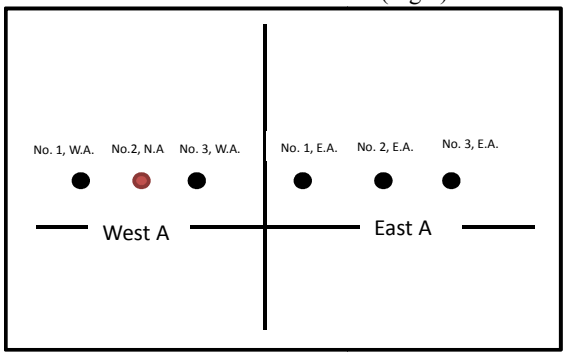

Fig. 3 Check house number and road name (No.2 N.A. is wrong)

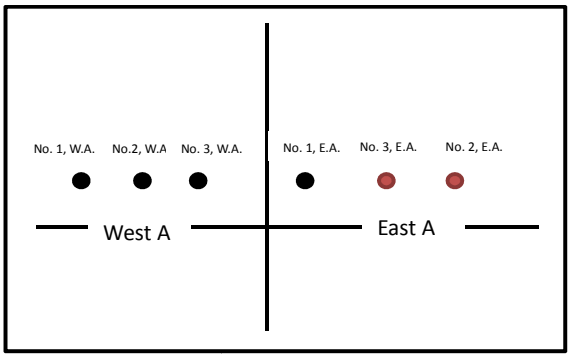

Fig. 4 Check house number sequence (No.3 E.A. is wrong)

a. Spatial location check
Use a town as a unit, extracting 5% house numbers for on-site check.
1. Check the house number's XY coordinates and its GPS coordinates
2. Check the house number point within the building or not (Fig.5)

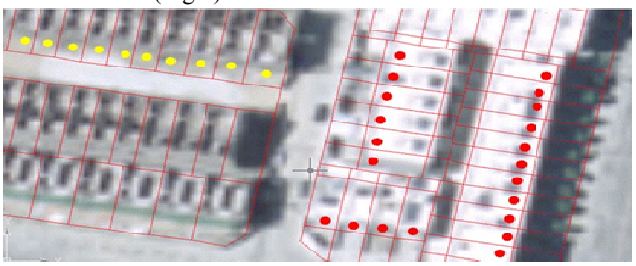

Fig. 5 Check the house number point

D. Build the database

Import data into GIS database under GIS data standard.

V. Result

According to the household registry list, total house number is 339,893. However, after field investigation, the real existing house number is 325,202 and the distribution is shown as Table I. There are four towns below 98% at the first check, and all reach the standard after the second check. All house numbers are built as database with GIS standard, and establish WebGIS for public services (Fig. 6).

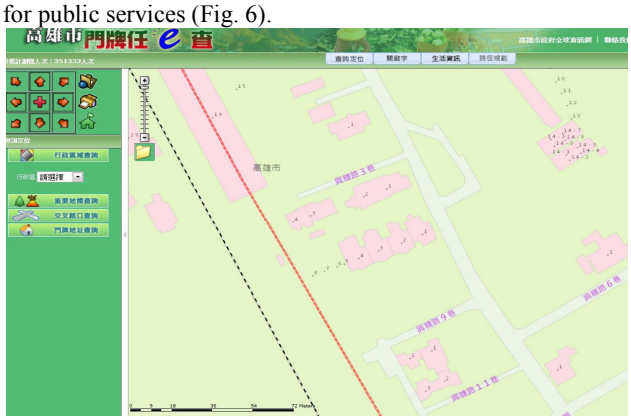

Fig. 6 Interface of Kaoshiung City Address Webgis

VI. Conclusion

Nowadays, information of house number is very important for Location-Based Services (LBS) [4]. Like other Asian countries, street segments in Taiwan are chaotic, not regular. Hence, the field survey is indeed necessary. Based on the basic maps, people can check the correctness of household registry quickly. and also can through this study to re-check the quality and accuracy for house number information. To other Asian countries without address location spatial information system, the study can provide an effective way to build the address geocoding and establish a reliable GIS address database.

References

[1] McElroy, J. A., Remington, P. L., Trentham-Dietz, A., Robert, S. A., & Newcomb, P. A. (2003). Geocoding addresses from a large population-based study: lessons learned. *Epidemiology*, 399-407.

[2] Gore, A., Paranjpe, S., Rajgopalan, G., Kharshikar, A., Joshi, N., Watve, M., & Gogate, M. (1993). Tiger census: Role of quantification. *Current science. Bangalore, 64*(10), 711-714.

[3] Marx, R. W. (1986). The TIGER system: automating the geographic structure of the United States census. *Government Publications Review, 13*(2), 181-201.

[4] Steiniger, S., Neun, M., & Edwardes, A. (2006). Foundations of location based services. *Lecture Notes on LBS*.