Application of GIS Techniques in Hong Kong's Population Census

Annie Y W CHAN
Census and Statistics Department, Hong Kong Special Administrative Region

Abstract

It is an established practice in Hong Kong to conduct a population census every ten years, and a by-census in the middle of the intercensal period. Census/by-census is a large scale and complex operation, involving a few thousand to over twenty thousands of field workers making visits to households throughout the whole territory. The data collected in the census/by-census, in particular the small area statistics, are useful inputs to the Government in programme planning and also to the private sector for marketing and business research purpose.

In the 2001 Population Census (01C), Geographic Information System (GIS) technology was applied to support various aspects of work through a computer system named "Digital Mapping System (DMS)". This is the first time that GIS techniques were used to support a population census. On field operation, the tailor-made and up-to-date maps of good quality produced by the DMS allowed field operation to be conducted more efficiently. On data dissemination, data analysis on small geographic areas could be performed more conveniently through using thematic maps.

The 2006 Population By-census (06BC) will be conducted in July/August this year. It is estimated that around 300 000 quarters will be enumerated. GIS techniques will continue to be applied in various aspects of the 06BC. This paper discusses the applications of GIS techniques in the census/by-census in three broad areas: data collection, statistics compilation and data dissemination.

1. Introduction

GIS technology is used in a wide spectrum of official statistics activities nowadays, from data collection, to statistics compilation and data dissemination. The history of applying GIS technology in official statistics work can be traced back to the early 1990s, some 15 years ago, when we disseminated census/by-census results.

In Hong Kong, a population census is conducted every ten years, and a by-census is conducted in between two censuses. While a population census involves the enumeration of everyone in the population, the by-census is a large scale sample enquiry in which only a

fraction of the population is sampled for enumeration. The latest population census in Hong Kong was conducted in 2001 and the next by-census will be conducted this year.

With its large scale operation, population census/by-census is an important, and in fact the only source to provide socio-economic and demographic data of the population for small geographical areas, down to street blocks at the lowest level. It is particularly useful when the statistics for a particular geographical area are visualized, say on a map. To get full value of the census/by-census results, a CD-ROM product was produced in collaboration with private company to facilitate the analysis of census results in small geographical areas, for the first time in the early 90's. It bundled the data, digital maps and GIS software together. Using this product, the residents' profile of a particular area like a housing estate can be visualized on a map, and spatial analysis can be performed using the GIS tools embraced therein. For the government, the information is vital for considering the various kinds of facilities and services needed in the geographic areas concerned, like schools, recreational and transportation facilities and medical services. Businessmen can make use of the information to decide the kind of shops to be set up in the areas, the type of products and services to be sold there, and the prices to be charged. Similar CD-ROM products had been produced in the subsequent census and by-census, namely the 1996 Population By-census and the 01C.

2. Applying GIS in the 01C

Starting with data dissemination, GIS has been gradually engaged in other aspects of official statistics work. GIS techniques were employed to support data collection work for the first time in the 01C through the DMS. This system enables maps to be produced, maintained and updated much more efficiently, and the fieldwork of the census operation to be monitored more effectively. Basically, it contains map layers developed by various parties; for instance, building polygons, road centerlines, facility points, annotations, etc by the Lands Department, the boundary of Tertiary Planning Units and Street Blocks by the Planning Department and the boundary of segment, building polygons and landmarks, etc by the Census and Statistics Department (C&SD). It has customized GIS functions which permit the long-term maintenance and updating of map data, the production of maps of various types and scales and the spatial analysis of data to be performed efficiently.

This system had generated some 100 000 location maps to facilitate the enumerators to perform household visits in the 01C. Prior to that, maps were prepared manually in the previous censuses/by-censuses by doing the updating on the large paper maps through manual drawing, identifying the working areas of each fieldworker (i.e. the census enumerator) on the large paper maps, making photocopies of the relevant parts, and then cutting and pasting the areas required to arrive at a reasonably legible/decent map which

could be used by the census enumerator. All these are saved with the DMS in place; and most importantly, the efficiency of the project is enhanced. Nowadays we can retrieve a map for a particular enumerator as and when required from the system immediately without the need to search through the paper maps.

3. Using GIS in 06BC

The DMS, though developed for the 01C, will be used more extensively to support the by-census to be conducted in July/August this year. A few main applications are highlighted below.

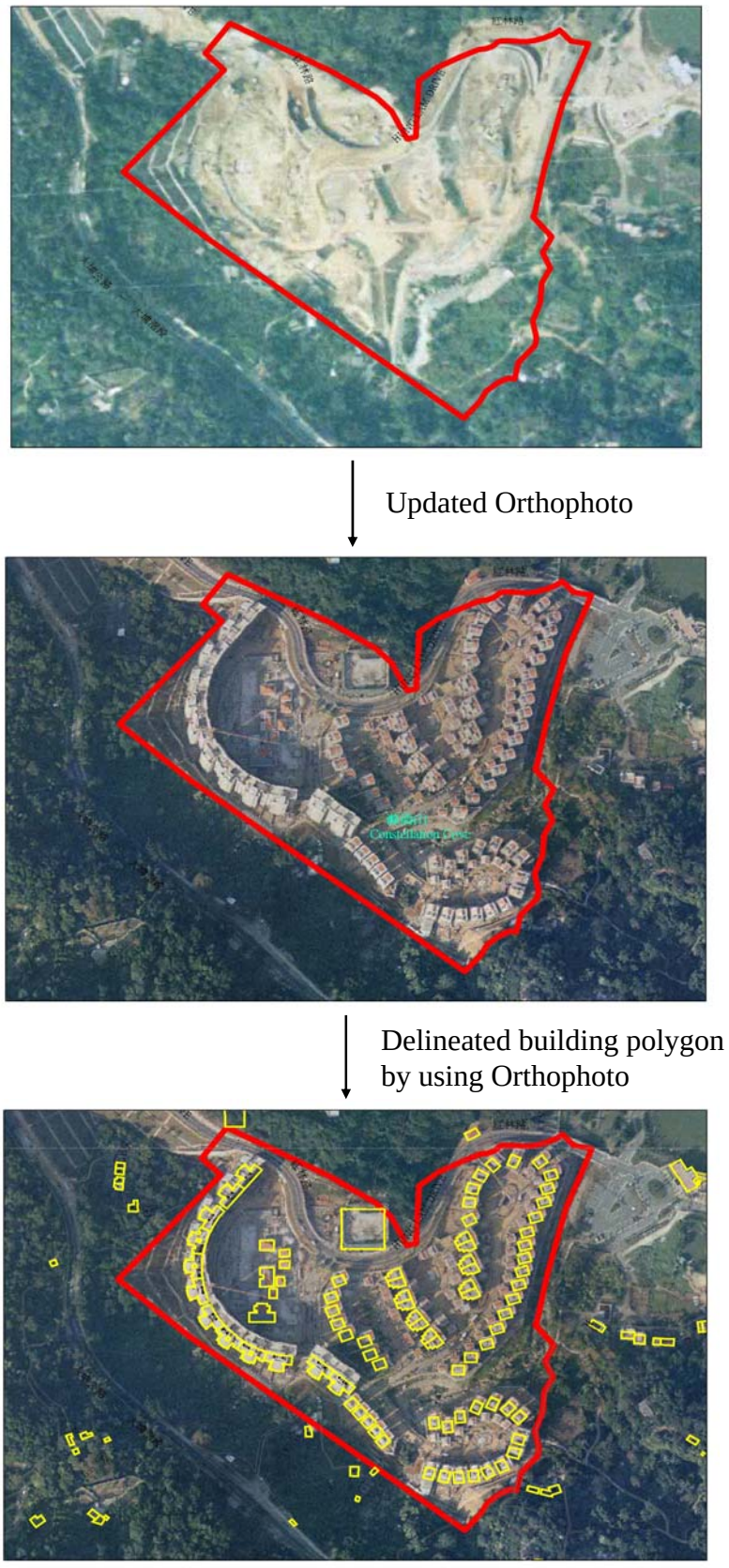
Updating of sampling frame

The availability of a complete sampling frame is a pre-requisite for conducting statistical sample surveys. For the purpose of population census/by-census and household surveys, a database containing all addresses in Hong Kong is maintained by the C&SD as the sampling frame. This database is made up of two parts or two registers, namely, the Register of Quarters (RQ) and the Register of Segments (RS). The RQ is a list of addresses of permanent quarters in built-up areas including urban areas, new towns and major developments in the New Territories. Each unit of quarters is identified by a unique address. The RS is a list of area segments in non-built-up areas (in other words, rural/remote areas). Segments are unstructured, and so are the arrangements or positions of the quarters contained in them. Each area segment contains some 8 to 15 quarters and is delineated by some physical or easily identifiable boundaries such as streams, footpaths, lanes and ditches. The use of area segments in non-built up areas is necessary because the quarters in these areas may not have clear addresses and cannot be easily identified.

In the 06BC, a one-tenth sample of the quarters in Hong Kong will be selected from the Frame of Quarters (FOQ) for enumeration. To eliminate coverage errors like omission of quarters, it is absolutely important to ensure that the FOQ is complete and up-to-date. With the implementation of the DMS, GIS techniques have been employed in updating the FOQ. Through overlaying the updated base maps or Digital Orthophotos[1] on a regular basis, information on new buildings, demolished buildings, and other changes can be revealed from the maps and photos, and fed into the FOQ for updating the quarters/buildings records. Figure 1 illustrates how the building records can be updated making use of the Digital Orthophotos.

[1] Digital Orthophotos are digital images of ground surface with uniform scale and positional information. The set of Digital Orthophotos currently used by the C&SD is produced from the aerial photographs taken at a flying height of 8000 feet. Distortions of photograph images caused by tilting of aerial camera and terrain relief are rectified except for those of the building structures. It consists of 189 tiles covering all the land area of Hong Kong. The ground pixel size of this set of Digital Orthophotos is 0.5m x 0.5m.

Figure 1: Update Building Records using Digital Orthophotos



Allocation of assignments and itinerary planning

The census/by-census is a large scale and complex operation and involves a lot of enumerators, from a few thousand for a by-census to over twenty thousand for a full census. Each enumerator has to visit a number of households and interviews all members therein. Determining the number of households and also the particular households to be visited by each enumerator is a very difficult and time-consuming task. In the 06BC, the network analysis tool in GIS is used to tackle this task scientifically. The enumeration effort is estimated based on parameters like walking distance from one location to another to determine the optimal set of assignments (in terms of the number of assignments and also which assignments) for each enumerator; and thus, ensure that each enumerator's set of assignments will be comparable in terms of number of households as well as total walking distance. After the assignments are allocated, the DMS will also plan the itinerary for each enumerator. The assignments given for each enumerator, in the form of an assignment list, will be arranged in an order such that the shortest route (in terms of walking distance) to visit the assignments will be incurred. An enumerator block map, giving an overview of the assignment areas, will also be provided to each enumerator. Specimens of the assignment list and enumerator block map are in Figures 2 and 3 respectively.

Figure 2: A specimen of the assignment list

2006 POPULATION BY-CENSUS **LIST L1EB**

ASSIGNMENT LIST – PERMANENT QUARTERS BY ENUMERATOR

CENTRE : 11 **DIVISION : 11-01** **ENUMERATION BLOCK (EB) : 01**

ENUMERATOR (ER) : 11-0101

ENUMERATOR (ER)	QUARTERS REFERENCE (QR)	HOUSE NO. (門牌)	FLOOR (層數)	BLOCK (座號)	FLAT/ ROOM (室／座)	QTRS TYPE	SEARCH	MERGER	SKETCH	REMARKS
	1F MAPLE STREET				楓樹街 1F 號					
11-0101	1013-01-1	1F	xx/F (xx 樓)		xx	411				
11-0102	1013-01-2	1F	xx/F (xx 樓)		xx	411				
	DRAGON PHOENIX MANSION, 64 SYCAMORE STREET				詩歌舞街 64 號 龍鳳大樓					
11-0101	1014-01-6	64	xx/F (xx 樓)		xx	411	HORIZONTAL			
11-0102	1014-01-7	64	xx/F (xx 樓)		xx	411				
	CLP POWER SYCAMORE STREET SUBSTATION, 30 SYCAMORE STREET				詩歌舞街 30 號 中華電力詩歌舞街變電站					
11-0101	1015-01-1	30	xx/F (xx 樓)		xx	611				
	42-44 SYCAMORE STREET				詩歌舞街 42-44 號					
11-0102	1015-01-2	43	xx/F (xx 樓)		xx	411				
11-0101	1016-01-6	43	xx/F (xx 樓)		xx	411				
11-0102	1016-01-7	44	xx/F (xx 樓)		xx	411				
	GOOD WORLD BUILDING, 197-199 TONG MI ROAD				塘尾道 197-199 號 好世洋樓					
11-0101	1017-01-1	197	xx/F (xx 樓)		xx	411				
11-0102	1017-01-2	197	xx/F (xx 樓)		xx	411				
	WINKI HOUSE, 123-125 LAI CHI KOK ROAD				荔枝角道 123-125 號 榮基樓					
11-0101	1018-01-6	125	xx/F (xx 樓)		xx	411				
11-0102	1018-01-7	125	xx/F (xx 樓)		xx	411				
	WIN TAK HOUSE, 159-161 LAI CHI KOK ROAD				荔枝角道 159-161 號 榮達樓					
11-0101	1019-01-1	159	xx/F (xx 樓)		xx	411	UPWARD			
	GODFREY CENTRE, 175-185 LAI CHI KOK ROAD				荔枝角道 175-185 號 嘉福中心					
11-0102	1019-01-2	179	xx/F (xx 樓)		xx	411	UPWARD			
	FOK KWAN BUILDING, 4-10 BOUNDARY STREET				界限街 4-10 號 福群大廈					
11-0101	1020-01-0	6	xx/F (xx 樓)		xx	411	BOTH	Main	511	
* 11-0101	1020-01-0	6	xx/F (xx 樓)		xx	411		Sub		

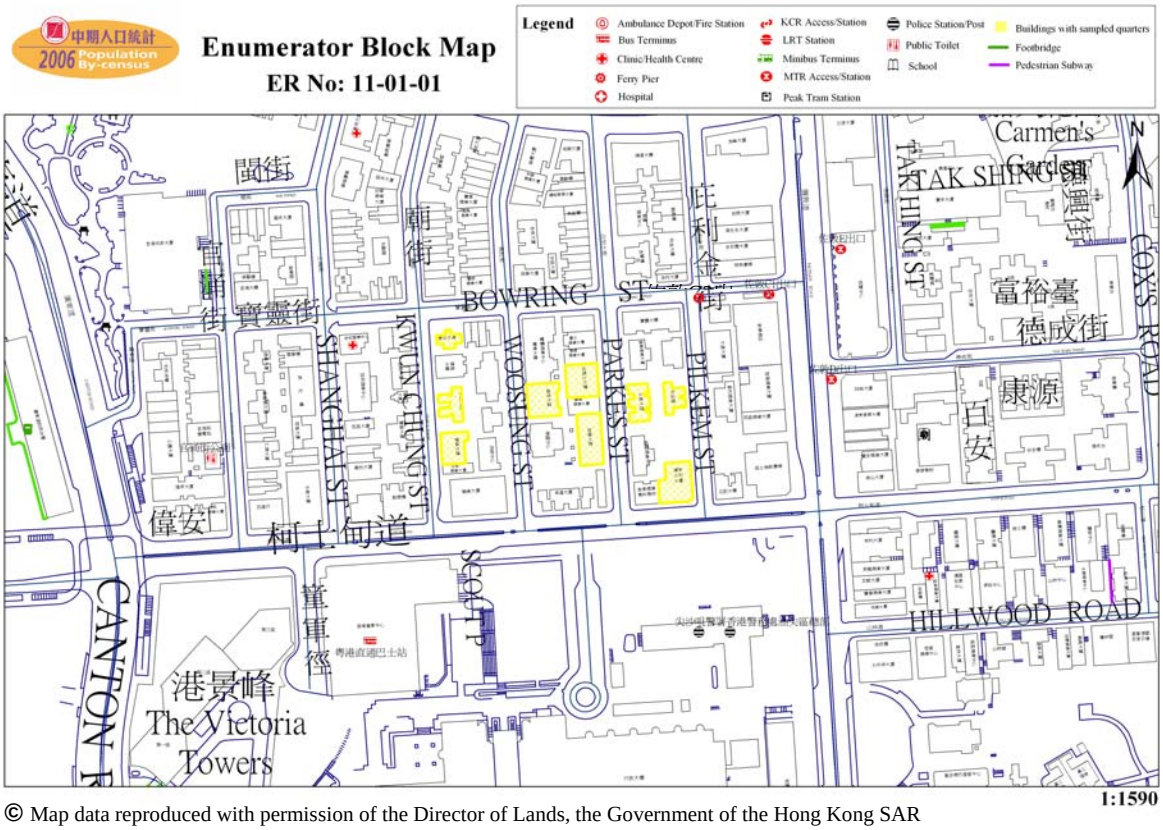
---------- End for the Enumeration Block ----------

Notes: * Denotes sub-merger 副打通單位 @ Householder letter has not been sent 沒有發出住戶通知信 Date: 2006..06.15 Page: 1

Liaise with Building Management Office before enumeration to confirm whether there are persons usually living in any quarters of the building. 須於訪問前先與大廈管理處聯絡，確定該樓宇內的單位是否有人經常居住。

Shaded addresses refer to the assignments of the pair-wise enumerator in the same enumeration block (EB). 反暗底份爲伙伴統計員在同一統計工作範圍內負責訪問的編宇單位。

Figure 3: A specimen of the enumerator block map



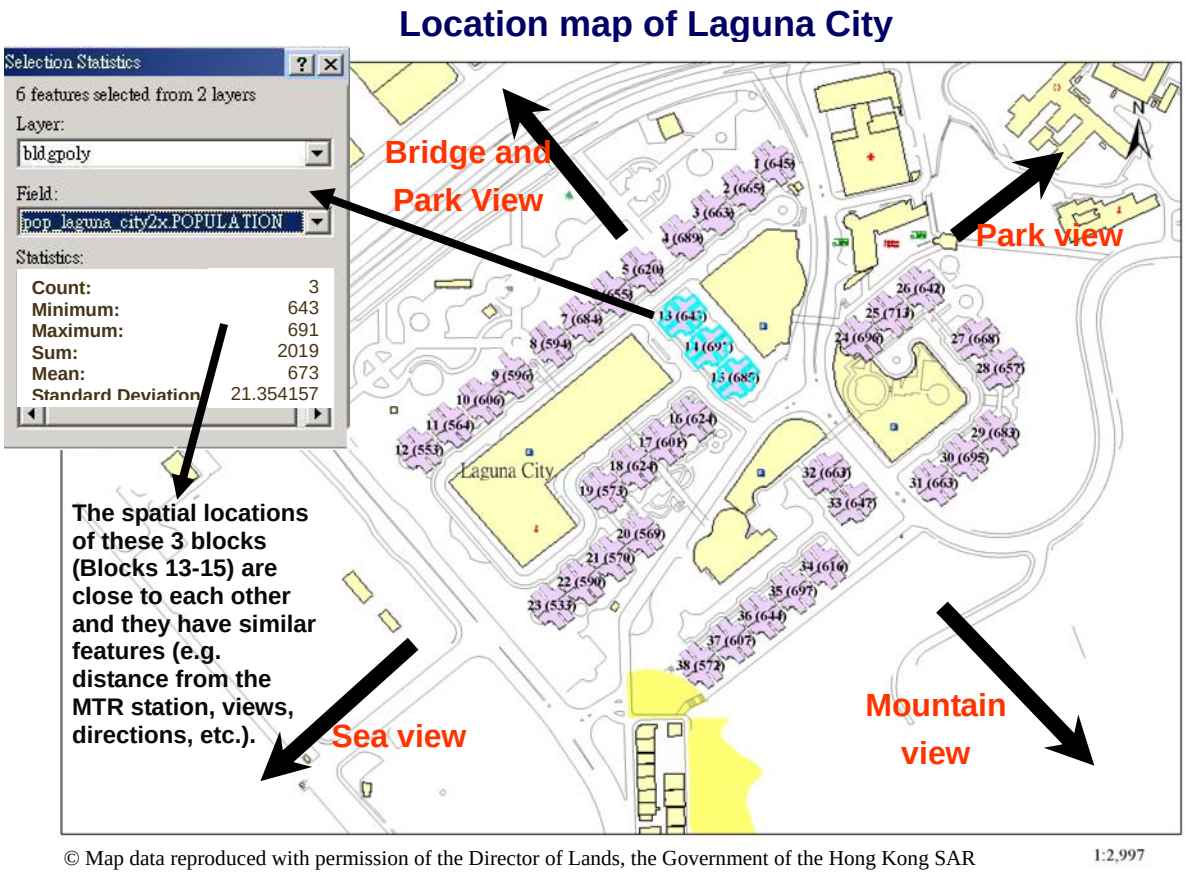
Statistics compilation

In Hong Kong, there are two basic geographical demarcation systems, namely, the District Council/Constituency Area system for district administration and election purposes, and the Tertiary Planning Unit/Street Block system for town planning purposes. However, it may not be possible to release the statistical data using these two systems directly. First, statistics of an area with very small population cannot be released in order to safeguard the confidentiality of data pertaining to individual persons and households. Second, it is necessary to ensure that the statistics released are precise. Last but not least, the requirements of individual users have to be considered with a view to meeting their needs as far as possible. For those cases where statistics cannot be released using the standard systems, C&SD will have to delineate a set of boundaries for data dissemination purpose.

GIS techniques can help us to perform these jobs efficiently based on the number of residents therein and the spatial locations. For some geographical units (e.g. for those with very few residents), they may have to be grouped with adjacent geographical units into broader groups for data dissemination. The formation of these groups is based on some threshold values of number of residents, types of housing and the spatial locations of these geographical units. In doing the formation, we need to know the orientation of the buildings

and group those nearby buildings together so that the small area statistics are meaningful and useful. An example is the formation of building groups in a large estate development "Laguna City" in the eastern part of Kowloon (see Figure 4). Blocks 13-15 are grouped together instead of with the other Blocks for data dissemination, taking into account their distinct geographical locations. Such groupings can be done more easily and efficiently with the availability of GIS tools.

Figure 4: Illustration of how buildings can be grouped for data dissemination using GIS techniques



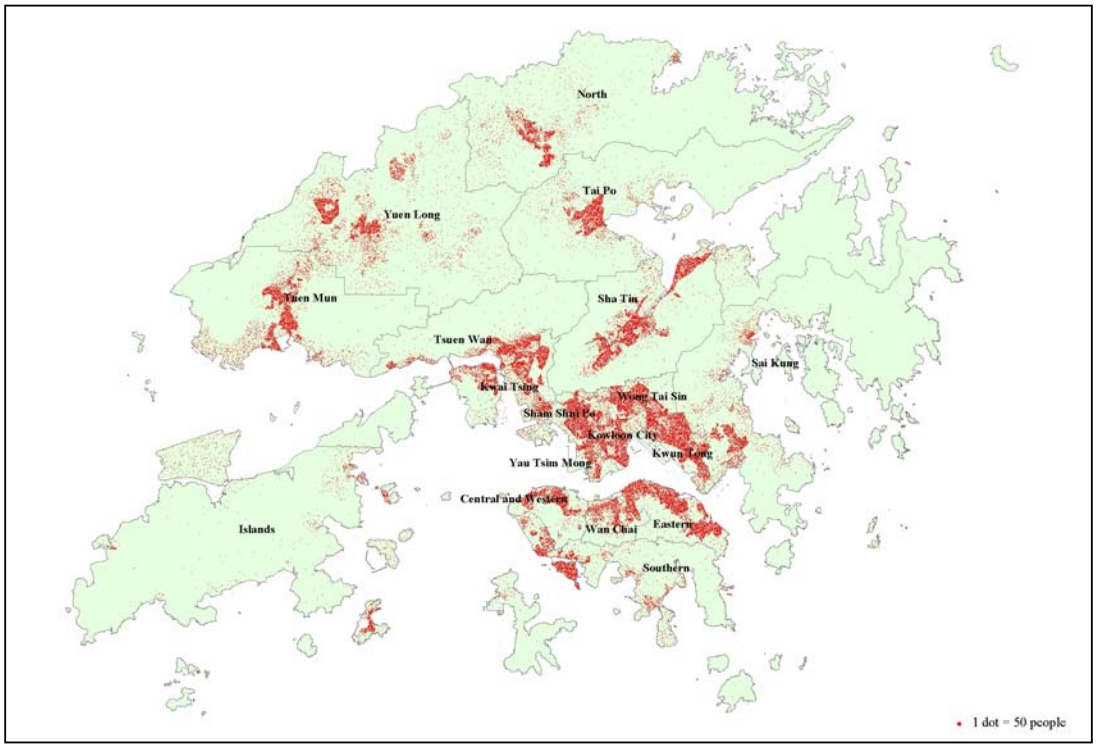
Data dissemination

A population census/by-census, with its large scale, provides the most detailed statistical database of the population with detailed spatial reference to the locations where people reside. The extent of detail of both the data and the spatial reference renders statistics from population census/by-census most conducive to the utilization of GIS applications. Statistics from population census/by-census have all along been an important fundamental data layer in many GIS applications, both within the private and public sectors, supporting various modeling functions and better informed decision making with a spatial dimension.

Hong Kong has been using GIS technology to support the data dissemination work

of population census/by-census since the early 90's. CD-ROM products containing detailed census/by-census data and digital maps, both produced by C&SD or in collaboration with private companies, allow users to perform spatial analysis using with GIS techniques. Thematic and area maps are produced using GIS technology to disseminate census/by-census results in publications and presentations. Such maps show the spatial distribution of one or more specific themes for standard geographic areas and bring out the content clearly. Figure 5 is a thematic map showing the distribution of population by small street block group. It can be easily visualized from this map that the population density is higher in those areas with more red dots.

Figure 5: Dot density map showing the distribution of population by Small Street Block Group within the boundary of District Council district based on 01C results



4. GIS - a tool for wider application in statistics work

Apart from being useful in population census/by-census, GIS is a useful tool for helping researchers and managers better understand problems, interpret data and make decisions involving a spatial dimension. For the general public, it enables them to better understand the community that they live in. Applying GIS techniques to perform spatial analysis on statistical data is certainly a potential area for further development in various kinds of statistical work.

References

David Martin (2002). Geography for the 2001 Census in England and Wales. Department of Geography, University of Southampton, United Kingdom.

Naoki Kurihara (2005). Improvements of the Census Operation of Japan by Using Information Technology. Statistics Bureau, Japan.