Improving the Graphical Cadastre Based on Genetic algorithm

Anna Shnaidman
Uri Shoshani
Yerach Doytsher

Mapping and Geo-Information Engineering
Technion – Israel Institute of Technology

Outline

- ***Introduction***
- ***The problem at hand***
- ***The proposed algorithm***
- ***Implementation of GA – cadastral analogy***
- ***Results***
- ***Statistical Analyses***
- ***Summary***
- ***Future work***

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- *Results*
- *Statistical Analyses*
- *Summary*
- *Future work*

Introduction

- One of the main objectives on the SOI agenda is transition from the existing physical cadastre to a coordinate based cadastre with legal validity
- A digital homogeneous cadastral system should be accurate, contain analytical consistent data, manage real estate in a smart, efficient and computerized way
- Digital cadastre is one of the concrete topics being discussed and researched in many countries
- Most customary solutions are based mainly on the Least Square (LS) method

- *Introduction*
- ***The problem at hand***
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- *Results*
- *Statistical Analyses*
- *Summary*
- *Future work*

The problem at hand

- The current cadastre in Israel is of a graphical nature and deals with surface properties only
- The products of the ground measurements are based on field surveys and are kept on paper
- The transition to an analytical cadastre is both crucial and inevitable

- *Introduction*
- *The problem at hand*
- ***The proposed algorithm***
- *Implementation of GA – cadastral analogy*
- *Results*
- *Statistical Analyses*
- *Summary*
- *Future work*

The proposed algorithm

- The conventional methods are mainly analytical and straightforward
- The proposed method is based on biological optimizations and is known as Genetic Algorithms (GAs)
- Characteristics:
  - ✓*stochastic method*
  - ✓*founded on evolutionary ideas and Darwin's principles of selection and survival of the fittest*
  - ✓*a natural selection which operates on a population of solutions –* ***chromosomes*** (individuals)

The proposed algorithm cont.

- The generic framework of GA :
  - An initial *population* of *n* vectors is randomly generated with a group of *individuals*
  - The individuals in the population are evaluated by a ***fitness*** function
  - A new population is created by applying variation-inducing operators: ***selection***, ***crossover*** (recombination) and ***mutation***

The proposed algorithm cont.

- Genetic operators - selection:
  - Two parent chromosomes are selected from a population according to their fitness to create new generation
  - Guiding principle – selection of the fittest
  - Superior individuals are of higher probability to be selected (survive)
  - Selection method – roulette wheel selection
  - Roulette slots' size is determined by the fitness value (the higher the grade, the wider the slot)

The proposed algorithm cont.

- Genetic operators - ***crossover***:
  - Two ***offspring*** are created using single point crossover

Parents chromosomes children chromosomes

- Genetic operators - ***mutation***:
  - The new offspring are changed randomly to ensure diversity

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- ***Implementation of GA – cadastral analogy***
- *Results*
- *Statistical Analyses*
- *Summary*
- *Future work*

Implementation of GA – *cadastral analogy*

- A generation represents turning points of parcels in a registration block
- Each individual in the population symbolizes a set of block coordinates stored in an array (vector) structure
- An objective function is defined to minimize the differences between the legal (registered) coordinates and those provided by the solution under the conditions specified (best suited to the predefined criteria)

Implementation of GA – *cadastral analogy* cont.

- With each generation the vectors are altered according to the best solution provided
- Every individual may assumed to be a set of coordinates, representing acceptable observations received from different sources
- The GA method was evaluated using synthetic data

Implementation of GA – *cadastral analogy*
cont.

Definitions:

- The preliminary population of *n* vectors is produced by randomly altering an "ideal" cadastral block
- A registered area criteria was chosen for analyzing the GAs' competence and effectiveness in the cadastral domain
- After the population is generated, a primary test is performed according to the SOI regulations

Implementation of GA – *cadastral analogy*
Cont.

Cadastral conditions:

- The fitness function ascribes a value to each solution vector using the desirable MSE of the parcels coordinates
- The vector's (individual's) grade is obtained by a weighted summation of parcel grades

$$f(u) = \Sigma u_i \cdot p_i \qquad p_i = \frac{S_i}{\Sigma S_i}$$

$$u_i = \frac{\Delta A_i - \Delta S_i}{\Delta A_i - T_i} \cdot 100$$

$$T = \sqrt{\Sigma[(\frac{\partial S_i}{\partial Y_i})^2 + (\frac{\partial S_i}{\partial X_i})^2] \cdot m_{XY}^2}$$

Implementation of GA – *cadastral analogy* cont.

- ***Iterations – creation of successive generation:***
  - For each parcel in the synthetic block two parents are selected at a time according to the roulette wheel method
  - A single point crossover is performed to generate two new individuals
  - The process continues until the original population size is reached
  - Mean coordinates are calculated
  - Mutation

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- ***Results***
- *Statistical Analyses*
- *Summary*
- *Future work*

Simulation results

- The proposed method's quality and accuracy were examined by performing simulations on the synthetic data
- The main purpose of these simulations is to test the ability of the GA to return to the initial theoretical state - an ideal, errorless solution
- For comparison an Least Square iterative adjustment was applied as well

Simulation results

- A characteristic example of the solution accuracy (meters)

Param - eters	Min dX	Min dY	Max dX	Max dY	Mean X	Mean Y	σ X	σ Y
Initial	-0.656	-0.703	0.590	0.691	-0.047	-0.009	0.244	0.199
GA	**-0.231**	**-0.176**	**0.235**	**0.249**	**-0.008**	**0.002**	**0.080**	**0.084**
LS	-0.574	-0.629	0.478	0.550	-0.047	-0.009	0.208	0.184

- The following parameters have been used:
  - Standard deviation error - 0.25 meter
  - An expected MSE of the coordinates - 0.03 meter
  - Maximum generations (iterations) - 25

Simulation results Cont.

- ***Geometric quality***

  - An adjustment process may conclude with distorted areas of the original shape or with an overall shift of the entire area
  - Even though the differences in the coordinates may be small, they do not have any indication on the matter
  - To ensure area shape preservation and oversee shifting of coordinates, additional analyses were applied
  - For each parcel were calculated: mean coordinate differences, shifted coordinates, union and intersection areas (of the original errorless and final GA computed parcels)

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- *Results*
- ***Statistical Analyses***
- *Summary*
- *Future work*

Statistical Analyses

- A series of statistical tests were carried out to examine solution's quality and facilitate statistical decisions regarding the population
- The population parameters are: null H_0 and the alternative H_1 hypotheses
- The hypotheses complement one another - if one is true the other one is false

Statistical Analyses cont.

- The following equations were used to perform the tests:
  - Expected value test

$H_0 : \mu_i \leq |\mu_0|\ m$

$H_1 : \mu_i > |\mu_0|\ m$

$reject\ H_0 \rightarrow t_o > t_\alpha^\nu$ $\qquad t_0 = (\frac{\bar{X}_i - \mu_0}{s_i})\sqrt{n}$

  - Standard deviation test

$H_0 : \sigma_i \leq \sigma_0\ m$

$H_1 : \sigma_i > \sigma_0\ m$

$reject\ H_0 \rightarrow \chi_0^2 > \chi_{\alpha,\nu}^2$ $\qquad \chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$

Statistical Analyses cont.

- A statistical test (computed sample value) and significance level must be chosen
- Based upon those values a decision is made to accept or reject the null hypothesis
- The test were carried out for the mean differences and STD of the coordinates as well as for the area differences
- The test results conformed: GAs‘ results are better than LS

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- *Results*
- *Statistical Analyses*
- ***Summary***
- *Future work*

Summary

- The proposed method examines a new approach for achieving homogeneous coordinates by using an evolutionary algorithms - GAs
- GAs imitate the natural process of evolving solutions
- Applying the GA to synthetic data yields satisfactory results
- Repeated simulation executions showed similar results (GA vs. LS)
- The GAs method is more accurate and provides better results than those of the traditional LS approach

- *Introduction*
- *The problem at hand*
- *The proposed algorithm*
- *Implementation of GA – cadastral analogy*
- *Results*
- *Statistical Analyses*
- *Summary*
- ***Future work***

Future work

- The simplicity of the algorithm enables considering additional cadastral and geometric conditions without altering its fundamental mechanism:
  - *road width*
  - *perpendicularity*
  - *parallelism*
  - *straight line requirement*

  etc…
- Future research will include implementation of the algorithm on "real" data (using different constrains and several cadastral blocks)

Thank you