Refinement of DEM boundaries using Point Distribution Criteria in Scattered Data Interpolation

Seung-Bum KIM
Satellite Technology Research Center, Korea Advanced Institute of Science and Technology, 373-1 Kusung, Yusung, Taejeon, S. Korea 305-701 (currently at Jet Prop. Lab./California Inst. Tech. USA) Tel: +82-42-869-8629. Fax: +82-42-861-0064. E-mail: sbkim@satrec.kaist.ac.kr

Abstract: Extrapolation off the boundaries of scattered data is an intrinsic feature of interpolation. However, extrapolation causes serious problems in stereo-vision and mapping, which has not been investigated carefully. In this paper, we present novel schemes to eliminate the extrapolation effects for the generation of a digital elevation model (DEM). As a first step, we devise point distribution criteria, namely COG (Center of Gravity) and ECI (Empty Center Index), and apply rigorous and robust elimination based on the criteria. Compared with other methods, the proposed schemes are computationally fast and applicable to a wide range of interpolation techniques.

Keywords: Interpolation, extrapolation, image segmentation, digital elevation model.

1. Introduction

A DEM is one of the most widely used data sets for analyzing the terrain and constructing a geographical information system. In the stereo-matching approach to produce a DEM, due to match failure, elevation values derived from stereo-match are scattered irregularly (compare Fig. 1a and Fig. 1b). A complete coverage is obtained mostly by interpolating the scattered data [1]. One of the problems in the interpolation is extrapolation, as illustrated in Fig. 1c. Comparison of the stereo-match results in Fig. 1b with the interpolated DEM in Fig. 1c reveals that incorrect elevation is derived outside the scattered data off the coast. These artifacts develop because interpolation is triggered by a minimum of a single scattered datum. Consequently elevation is extrapolated off the scene boundaries as wide as the interpolation radius. The presence of such extrapolation may introduce serious problems to civil property management and coastal engineering.

To remove the extrapolation, in this paper we present a series of novel schemes. They include the initial elimination of extrapolation by parameterizing the distribution of scattered input data, and the subsequent segmentation processes that compensate for excessive elimination and complete the initial elimination.

2. Data

We generate scattered elevation data using four pairs of stereo-images taken by SPOT (Satellite Probatoire d'Observation de la Terre) satellite. We choose two full-scale images in Seoul and Boryung regions and two sub-scenes in Donghae and Mokpo regions in S. Korea. Each full scene spans a 60 km × 60 km area and has the dimension of 6000 × 6000 pixels. In Seoul region the boundary of scattered data has complicated patterns whereas in Boryung region it is relatively simple. Donghae and Mokpo scenes cover 12 km × 14 km and 10 km × 10 km, respectively. These two areas have complex boundaries of scattered data due to harbors. Procedures for generating scattered data are as follows [2]. An area-based stereo-matching employs epipolarity and the normalized cross-correlation of stereo-patches. The match growth from seed points is accomplished by region growing. The cross-correlation is thresholded to produce a successful match.

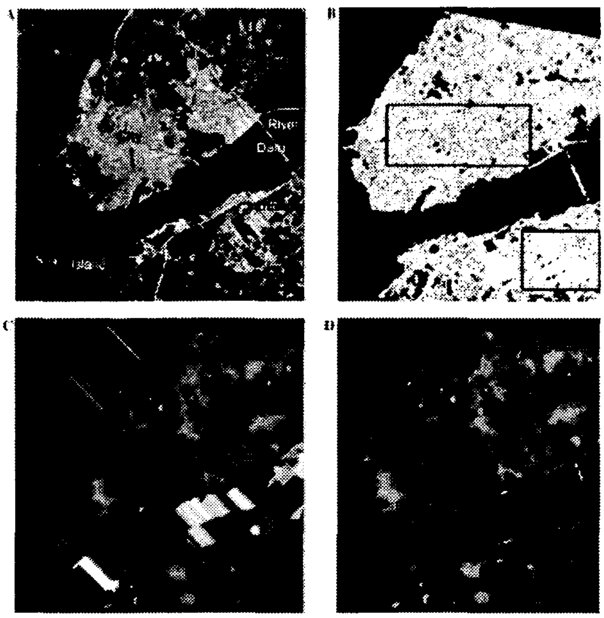

Fig. 1 Operation of the DEM boundary refinement over the 10km × 10km in Mokpo area, S. Korea. (a) one of the stereo-scenes from SPOT satellite. (b) stereo-matched points in white, used as the input to an interpolator. The rectangle defines the border of an area where the COG/ECI are computed in an ideal case. (c) DEM generated by Gaussian interpolator. (d) the final DEM after the elimination.

3. DEM boundary refinement by eliminating extrapolation

To identify whether a mismatch area should be interpolated (e.g., inland) or a left uninterpolated (e.g., sea), we introduce COG and ECI parameter. COG is defined as:

$$COG = \frac{\sqrt{\left(\overline{\Delta x}\right)^2 + \left(\overline{\Delta y}\right)^2}}{\max\_dist} \qquad (1)$$

$$\overline{\Delta x} = \overline{(x_k - x_o)} \text{ and } \overline{\Delta y} = \overline{(y_k - y_o)}$$

where *max_dist* is the maximum distance between a scattered datum at (x_k, y_k) and the center of an interpolation disc (x_o, y_o), and the overbar denotes an average. The range of COG is [0,1] and a larger value of COG corresponds to the greater unevenness (e.g., along the coast) in the distribution of scattered data. If COG of a particular grid location is greater than 0.42, interpolation is not performed. The threshold 0.42 may be derived analytically, from real data (Fig. 2) or from a similated data [3].

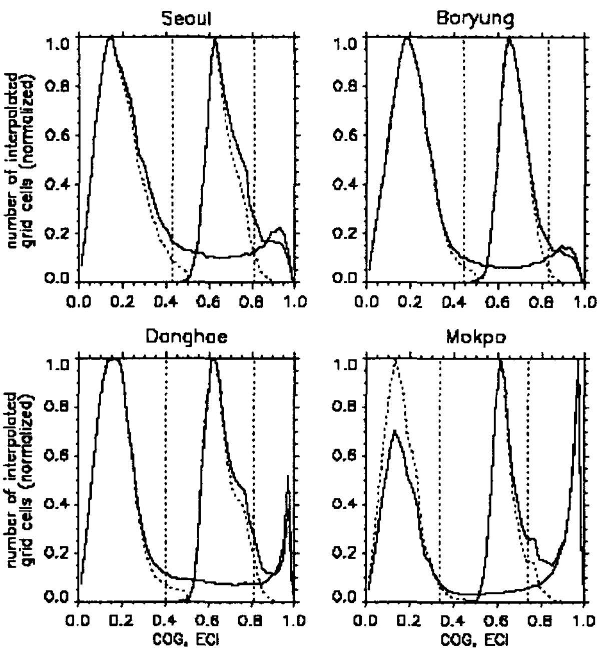

Fig. 2. Distribution of COG and ECI values. Curves whose peaks are located below (above) 0.4 correspond to COG (ECI). Dotted curves are derived from the ideal cases where scattered data are distributed evenly (marked by rectangles e.g. in Fig. 1b). Solid curves are computed over the entire regions. Dotted vertical lines mark three-sigma limit of COG and ECI distribution in the ideal cases.

COG operation leaves some undesired interpolation. To find a solution, we note that when an interpolation disc is located over a bay the center of the disc does not have any scattered data or becomes *empty*. To quantify the emptyness we devise an Empty-Center-Index (ECI), defined as

$$ECI = \frac{1}{\max\_dist} \frac{\sum_{k=1}^{N} r_k w_{eq}(r_k)}{\sum_{k=1}^{N} w_{eq}(r_k)}, \qquad (2)$$

where r_k is the distance between a scattered datum and the center of an interpolation disc and *max_dist* = max (r_k). w_{eq} is a weighting factor to equalize the distribution of scattered data with respect to *r*. Interpolation with ECI greater than 0.8 (i.e., center is too much empty) is removed. The threshold is determined from real (Fig. 2) or simulated data [3].

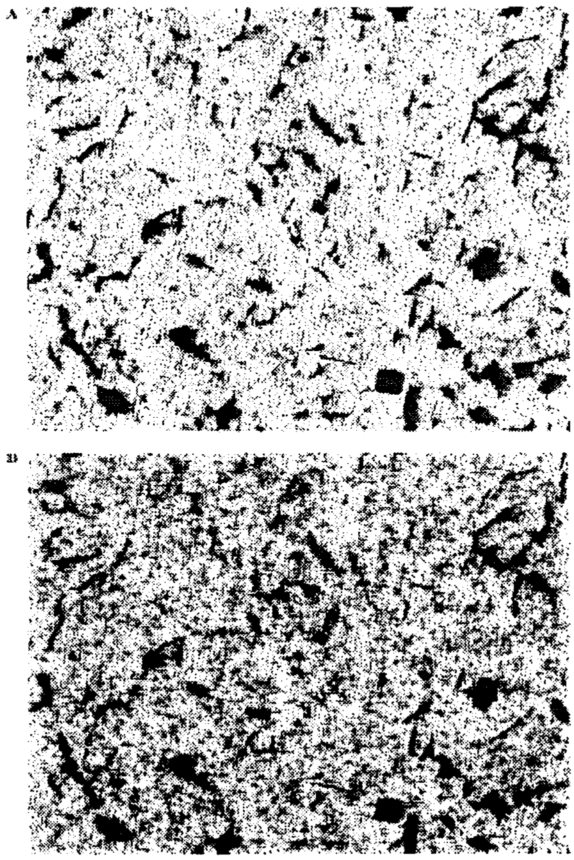

Fig. 3 Location of interpolated elevation (in red plus) when their ECI is (a) in [0.85, 1.0] and (b) in [0.745, 0.755]. The background image is the stereo-matched scattered data.

An interesting feature in the distribution (dotted lines) is a local maximum of ECI at 0.75 (Fig. 3). ECI higher than 0.75 is found mainly in the areas where the center is empty (Fig. 3a). Therefore ECI higher than 0.75 represents extrapolation. In contrast, ECI of 0.75 is distributed almost randomly over *non-empty* areas (Fig. 3b), indicating that ECI of 0.75 is representative of interpolation. The local maximum is attributed to is a hole in the scattered data with the size of 3-4 pixels. Such hole develops commonly by stereo-match failure. In comparison, the objects where extrapolation occurs, namely water surfaces in a bay or a river, tend to have much bigger sizes than 3-4 pixel or 20 m × 20 m.

Final DEMs after eliminating extrapolation are contrasted to DEMs before the elimination in Fig. 1c-d. The elimination schemes successfully restore the detailed boundaries of scattered data.

Table 1. The size of the area of extrapolation expressed in terms of the width along the coast, before and after eliminating the extrapolation. The width is estimated by comparing with a 1:50,000 topographic map. Units in km.

Region	Seoul	Boryung	Donghae	Mokpo
Before	3.7	1.7	1.8	1.3
After	1.7	0.52	0.12	0.12

4. Performance analysis

One way to assess quantitatively the performance of the elimination schemes is to compare with a topographic map. The map is digitized from a 1:50,000 original issued by the National Geography Institute of Korea. After the elimination, the surplus of land areas with respect to the topographic map is 144 km^2 and 44 km^2 for Seoul and Boryung respectively. These correspond to 1.7 km and 0.52 km on average when converted into the width of the land surplus. These are significant improvements, noting that before the elimination the width are 3.7 km and 1.7 km respectively for Seoul and Boryung (Table 1) Similar improvements are found in Donghae and Mokpo regions as well. The remaining surplus is due mostly to the topographic map's being outdated. A considerable amount of reclamation had taken place since the map was generated (between 1987 and 1995) until the acquisition of SPOT images in 1997-1999. A map of the land surplus reveals that except for the reclaimed land, most of the surplus has been removed by the elimination schemes (not shown).

5. Conclusions

In this paper, we developed the novel schemes to eliminate extrapolation off the boundaries of scattered data. The novel schemes consist of a series of the COG and ECI elimination and further segmentation. The values of COG and ECI threshold are determined rigorously through the analytical and empirical methods. By applying the schemes to DEM generation from real and simulated data, we have demonstrated that the schemes eliminate the extrapolation correctly. We expect that the developed schemes will apply most effectively to the accurate representation of the topography near the water boundary, where the presence of extrapolation stands out as a serious problem. Furthermore they can be integrated into existing photogrammetric tools for DEM generation at a minimal amount of computational cost.

Acknowledgement

Discussions with W. Park, T. Kim and H.Y. Lee were very helpful. The Ministry of Science and Technology, S. Korea, is acknowledged for funding the research.

References

[1] Schenk, A.F., Automatic generation of DEMs, in *Manual of Photogrammetry addendum*, C. Greve, ed., Amer. Soc. Photogram. Remote Sens., Bethesda, MD, 1996, pp. 145-150.

[2] Lee, H.Y., W. Park, S.B. Kim, T.J. Kim, H.K. Lee and T.-g. Kim, The development of an accurate DEM extraction strategy for satellite image pairs using epipolarity of linear pushbroom sensors & intelligent interpolation schemes, *Int. Arch. Photogramm. Remote Sensing*, 33, 2000, 705-712.

[3] Kim, S.B., 2004: Eliminating extrapolation using point distribution criteria in scattered data interpolation, *Comput. Vision Image Understand.*, **95,** 30-53.