Generating DEMs from Contours using Standard Sorting Routines

Poul Frederiksen and Ole Mærsk-Møller
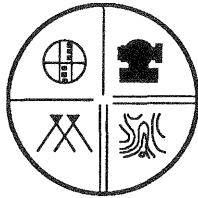
Institute of Surveying and Photogrammetry
Technical University of Denmark
DK-2800 Lyngby, DENMARK

Abstract:

A strategy for generating digital elevation models from digitized contours has been developed. The aim is to minimize the need for computer capacity and programming efforts utilizing standard sorting and merging routines. An interpolation procedure closely related to the data management is presented.

Introduction

The construction of a digital elevation data base is preferably based on new photogrammetrical measurements. This is due to the fact that most error-sources can be controlled and consequently, specific requirements for the data base can be defined in advance and the quality of the final product predicted.

The photogrammetrical measurements needed to create a data base for large areas - or even for a national data base - are time consuming and very expensive. Consequently, analogue contour maps considered as a source of data might prove interesting. Huge potential resources of elevation information is available in these maps. Automatic scanning and vectorizing facilitate the conversion into digital form and the time consuming process of labeling the contours is feasible by utilizing modern CAD-CAM-systems. Newly developed software has turned this process into a semi-automatic one, taking advantage of the systematic structure of the contours.

On the other hand, it must be realized that the quality of an elevation data base extracted from contours is difficult to predict since the genesis of the contours is often unknown or weakly documented. Quality assessments are restricted to the test of the final data base.

Compared to a photogrammetrically measured elevation data base the advantage of a contour based model is primarily to be found in a much shorter production time and a lower price. These factors, however, may be crucial, compensating for the disadvantages of a less accurate product.

In most cases, an elevation data base consists of spot heights organized in a grid. In order to achieve this pattern an interpolation procedure is introduced to transform the information given by the contours.

Due to the amount of data involved, a strategy for managing the data must be developed together with the interpolation procedure. In the following, such a strategy is described, and an example of an interpolation procedure closely related to the data management is presented.

From 2D to 2 * 1D

The digitized contours are available as sequentially stored polygons on magnetic tapes. Each contour is labelled by an index giving the elevation of the contour. Assuming the contour can be represented by a straight line between two adjacent points, the intersections between a polygon and gridlines with constant x- and y-coordinates are calculated.

The gridlines define the point spacing of the final data base and may be equal or differ in the x- and y-directions. Intersections between the contours and gridlines with constant x-values (x-gridlines) and y-gridlines are stored in two separate files (figure1). Each intersection is stored immediately after calculation which minimizes the need for computer capacity. So, in principle, only two adja-
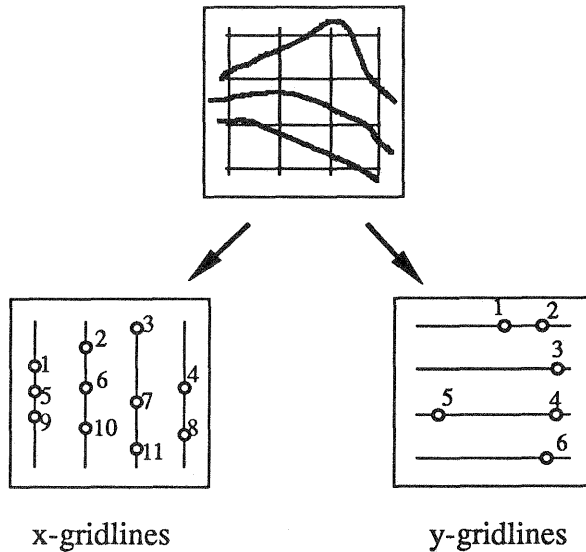
Figure 1. Intersections are stored in two separate files.

cent polygon points are necessary in the computer at the same time. Data are stored as x-, y- and z-coordinates.

From 2 * 1D to 2D

In order to avoid a time consuming search for corresponding points for interpolation, the data are reorganized by a sorting routine which is standard in most mainframe- or minicomputer environments.

The data on the x-gridlines are sorted primarily by the x- coordinates and secondarily by the y-coordinates, leading to a new file with column-structure.

Similarly, the data on the y-gridlines are reorganized by sorting primarily by y-coordinates and secondarily by x-coordinates. In this file a row- structure appears. The order of data is illustrated in figure 2.

After this reorganization of data any one-dimensional interpolation procedure is easily applied to the datasets to find the elevation in the gridpoints. The result is two files, each containing elevations in the gridpoints, but calculated from the data in the x- and y-directions separately.

The dataset with the interpolated points on the y-gridlines is sorted again, but this time primarily by x-coordinates and secondarily by y-coordinates, achieving the same order of data as in the x-gridline file.

By applying a proper weight function or simply calculating the mean, the two data sets are merged to the final elevation data base.

If no sophisticated interpolation is introduced, this procedure requires a minimum of programming efforts or computer capacity, since most of the data manipulation is left to standard sorting routines already available in the computer environment.
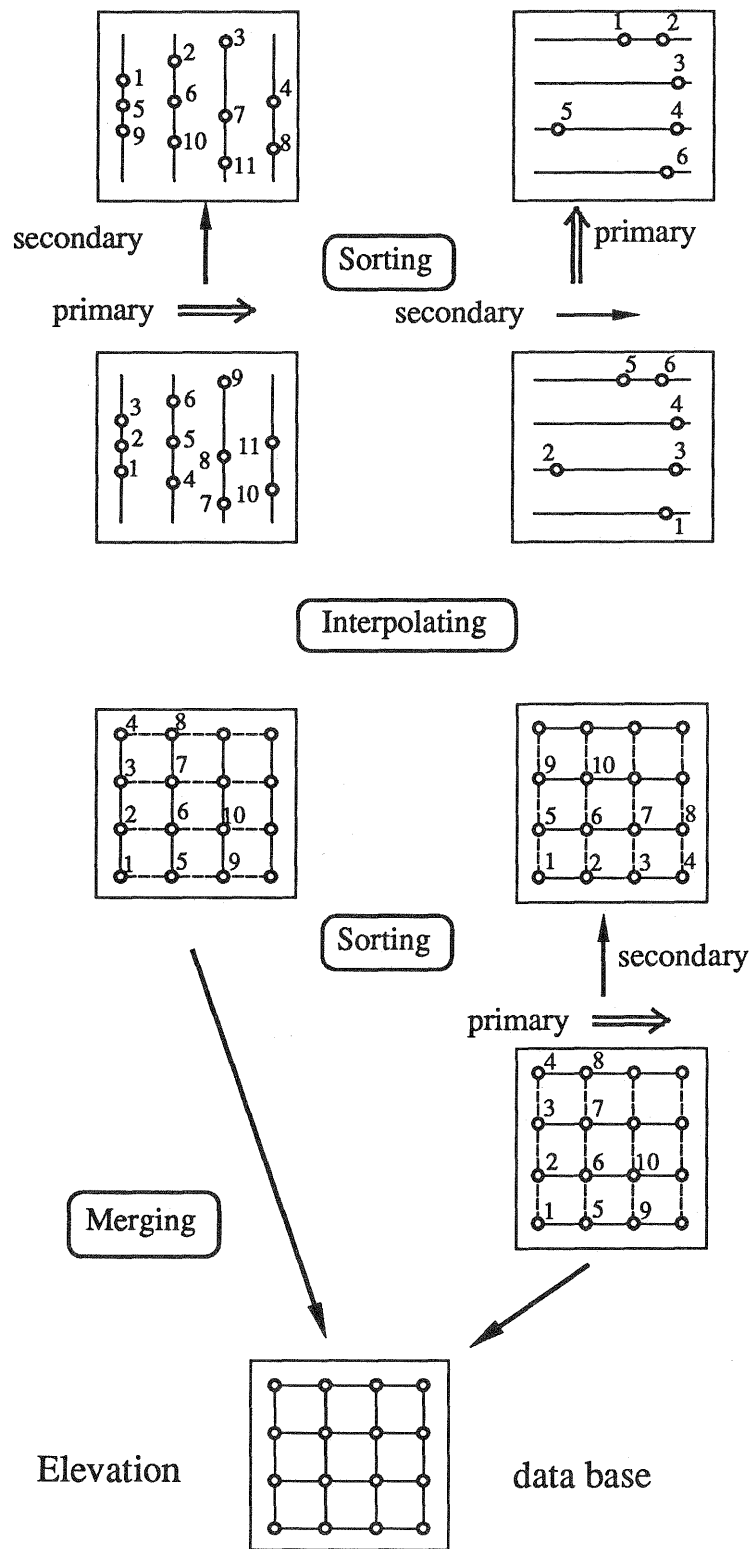
Figure 2. Strategy for sorting, interpolating and merging data to final elevation data base. Numbers indicate the order of data in the files

A national elevation data base

In the following, the construction of an elevation data base for Denmark is described. Based on the strategy mentioned above, the development of the data base was initiated by The Danish Posts and Telegraphs in order to facilitate the estimation of the field-strength level for a land mobile radio communication system.

The elevation data base serves as input for a radio simulation program (Storno, 1984) from which the positions of approximately 250 antenna-stations are determined.

The construction of the elevation data base is a joint effort of The Danish Geodetic Institute, The Danish Posts and Telegraphs and The Technical University of Denmark (Frederiksen,1987).

Elevations are generated in grid points in geographical coordinates with a point spacing of 3 arc-seconds according to the USGS standard for digital elevation models. (USGS circular 895- B, 1983 and Olsen, 1987).

In addition, an elevation data base is generated in UTM coordinates with a grid spacing of a 100 metres. It has turned out, however, that the quality of the data base is better than originally expected, so that a 50 metres grid is now being generated.

Digitized contours

The elevation data base is constructed from digitized contour lines from 1:50.000 maps with a contour interval of 5 metres - Denmark being a relatively flat country. Approximately half the map series are digitized manually while the remaining part is automatically scanned.

Denmark is covered by 110 map sheets in scale 1:50.000 and typically, the contours of one map sheet are described by 150.000 digitized points

Contours are digitized in the UTM coordinate system, referring to zone 32 or 33 and organized with the contours of one map sheet in one file in accordance with a Dutch proposal for a standard format for the exchange of cartographic data (van Zuylen, 1980).

Since the primary use of the data base is for the radio simulation program, data are transformed to geographical coordinates.

Interpolation

The sorting and merging procedure is applied to each map sheet as a computing unit. In order to avoid boundary problems, data from the 8 surrounding map sheets are considered. A weighted linear interpolation is introduced along the gridlines.

In figure3, z_1, z_2, z_3 and z_4 are the contour values at the positions where the contours intersect the gridlines.
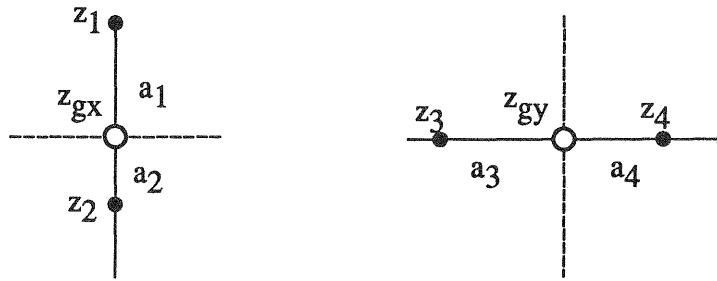
Figure 3. Interpolation on x- and y-gridlines separately

a_1, a_2, a_3 and a_4 are the distances from the intersections to the gridpoint.
The elevation of the gridpoint is estimated separately on the x- and y-gridlines.

$$z_{gx} = \frac{\frac{1}{a_1} \cdot z_1 + \frac{1}{a_2} \cdot z_2}{\frac{1}{a_1} + \frac{1}{a_2}}$$

and

$$z_{gy} = \frac{\frac{1}{a_3} \cdot z_3 + \frac{1}{a_4} \cdot z_4}{\frac{1}{a_3} + \frac{1}{a_4}}$$

When the data sets are merged, the final elevation is found by

$$z_g = \frac{(\frac{1}{a_1} + \frac{1}{a_2}) \cdot z_{gx} + (\frac{1}{a_3} + \frac{1}{a_4}) \cdot z_{gy}}{(\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \frac{1}{a_4})}$$

It has been discussed whether this least squares solution for the weighted mean leads to results which differ significantly from elevations estimated by

$$z_g^* = \frac{a_1 \cdot a_2}{a_1 \cdot a_2 + a_3 \cdot a_4} \cdot \frac{a_4 \cdot z_3 + a_3 \cdot z_4}{a_3 + a_4} + \frac{a_3 \cdot a_4}{a_1 \cdot a_2 + a_3 \cdot a_4} \cdot \frac{a_1 \cdot z_2 + a_2 \cdot z_1}{a_1 + a_2}$$

This estimator is achieved by a function which minimizes the energy of the terrain between the contours, having the contour values as boundary conditions. The function is harmonic and a linear expansion at the grid points yields the weighted mean z_g^* (Eeg, 1987).

Both estimators have been examined, revealing slightly different estimates.

It has been concluded that z_g^* in general gives better estimates, but unfortunately also bigger maximum errors, and it was decided to implement the simple weighted mean.

Both linear estimates have, however, the disadvantage of "smoothing" extreme points such as hilltops and valleys.

Hilltops in particular are crucial for the estimation of the field-strength level in the radio communication program. Since no spot heights were measured together with the contours, precautions have been taken to ensure a proper shape of profiles crossing hilltops and valleys.

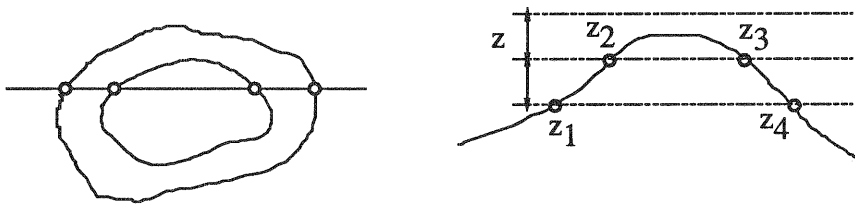Data on the gridlines are tested for convex or concave behaviour as illustrated in figure 4.

Figure 4. Prediction interpolation on hilltops.

When a sequence of elevations of 4 points on a gridline meets the requirement

$$z_1 < z_2 \land z_2 = z_3 \land z_3 > z_4 \land z_4 = z_1$$

a local maximum is located.

It is known that the terrain elevations between the positions of z_2 and z_3 cannot exceed $z_2 + \Delta z$, where Δz is the contour interval. If they did, another contour with the value of $z_2 + \Delta z$ would exist and furthermore, the condition stated before would not be met.

A prediction interpolation along the gridlines and on 4 intersection values is introduced. The covariance function has the form

$$cov(a) = e^{-1.8^{\alpha} \cdot a^2}$$

where a is the distance-dependent parameter. α is adjusted with the condition that z_g may not exceed the highest contour value by more than $k*\Delta z$, $0 < k < 1$. On the basis of several experiments, it has proved a pragmatic solution to fix k at 2/3.

Starting with $\alpha = 1$, the profile between z_2 and z_3 is described by a smooth curve that turns more flat as α increases. This procedure is repeated with the stop-criterion

$$z_g = z_2 + \frac{2}{3} \cdot \Delta z$$

A local minimum is treated in the same way.

Comments

When it was decided to create the elevation data base for Denmark, the production time was of the utmost importance, the reason being that the construction of the radio communication system was already in progress and the simulation program was just waiting for the elevation data to indicate the optimal positions of the antenna-stations.

The data management strategy described ensured a minimum of programming and developing expenses and a short production time. The first 10 files each containing 1.4 million elevations in geographical coordinates in accordance with the USGS-standard were delivered to The Posts and Telegraphs less than 6 months after the start of the project.

In order to control the product and in particular to detect blunders, contours were generated from the data base to compare with the original contours. A few label-errors of the digitized contours were discovered and corrected. This control procedure proved to be extremely effective since any anomalous course of the generated contours was immediately revealed by the human eye.

Some knowledge of the quality of the digitized contours was available (Bojsen and Jensen, 1983). Considering the relatively simple interpolation procedure, it was expected that a standard deviation in the final data base of approximately 2 metres should be realistic. This was acceptable for use in the radio simulation program, since no information about land-usage such as woods, urban areas or other obstacles was introduced at that stage. Later on, land-usage figures will be added to the elevation data base (Juul-Nyholm and Andresen,1986). Some 2000 points distributed in geologically different areas of the country were checked against maps of scale 1:25.000 with 2.5 metres contours. An average standard deviation of 1.3 metres was found, and it must be concluded that the accuracy of the data base is well on the right side of 2 metres.

So, beyond the primary application as input for the radio simulation program, the elevation data base is suitable for general planning purposes and for ortho-photo production of scale 1:10.000 or less.

References

Bojsen, T. and A.Jensen, 1983: En undersøgelse af Geodætisk Instituts 1:25.000 kurveplaner som grundlag for oprettelsen af en digital højdedatabase i Danmark, Geografisk Institut, København.

Eeg, J., 1987: A Simple Method for Interpolating Heights from Digitized Contour Lines. ISPRS-Proceeedings of the International Colloquium 'Progress in Terrain Modelling' WG III/3, Lyngby, Denmark.

Frederiksen, P., 1987: A Digital Elevation Model for Radio Communication. ISPRS-Proceedings of the International Colloquium 'Progress in Terrain Modelling' WG III/3, Lyngby, Denmark.

Juul-Nyholm, G. and J.E.Andresen, 1986: Land Mobile Radio Communication. Land-usage Figures for use in Field-strength Prediction as Measured in Denmark. COST 207, Denmark.

Olsen, R.W., 1987: Standards and Specifications for USGS Digital Terrain Data. ISPRS-Proceedings of the International Colloquium 'Progress in Terrain Modelling' WG III/3, Lyngby, Denmark.

Storno, 1984: Racops - Radio Coverage Prediction System, Denmark

USGS Digital Cartographic Data Standards, 1983. Digital Elevation Models. National Mapping Program. Geological Survey Circular 895-B, Reston, Virginia.

van Zuylen, L.,1980: A Standard Format for the Exchange of Cartographic Data. ITC-Journal 1980-1.