Study on Algorithms for Automated Landforms Classification Extraction from DEM

Taewoon KIM, Byeongpyo JEONG, Osamu TAKIZAWA, Masafumi HOSOKAWA

Abstract
自然災害による被害を予測するため、日本では地形分類が使われたりするが、地形分類情報を構築するには膨大な時間や予算を要し、殆どの開発途上国には整備されていない。本稿では災害予測、特に地震被害予測に必要とされる基盤データの整備が遅れている地域への適用を目的に、数値標高モデル基づく地形分類の自動抽出のアルゴリズムを提案し、宮城県を対象として分析を行った。

Keywords
数値標高モデル (Digital Elevation Model), 地形分類(landforms classification), 宮城県(Miyagi prefecture)

1. Introduction
Landforms classification database is useful in natural disaster planning. In most developed countries such as in Japan, detailed landform classification maps are available to be used for earthquake damage assessment. For the areas or countries that this classification database is absent, building it by the traditional method including field survey that is expensive and time-consuming can be avoided if we have an automated method to classify landforms from readily available DEMs (Digital Elevation Model). The classification data made in this way tend to be considered inferior to the manual in terms of classification quality because it was constructed only from the elevation data, not including geological and social features of the area. On the other hand, it is much quicker and cheaper to run an automated algorithm than a manual procedures, and it is even preferred for the cases such that highly detailed classifications are unnecessary, resources like time and money are not available, and for the areas where conducting a field survey is difficult or impossible. Jeong et. al. (2008) asserted that the landform classification based on DEM would be adequate to be used in earthquake damage estimation which in turn plays an important role in the initial response to earthquakes.

Number of studies tried and showed how an automated landforms classification from DEM can be implemented. Long et. al. (2007) extracted basic landforms using a semi-automatic method with the aid of Landsat Thematic Mapper imageries and expert’s interpretation, resulting in the total precision of 76%. Mizukoshi and Aniya (2002) showed that the contour-based DEMs better preserve characteristics of slope compared to more common grid-based DEMs in the areas where contours are dense. The techniques of region-growing and use of surface spectral signature are presented by Millaresis and Argialas (2000) to identify alluvial fan toes that spans large area with little change in slope.

Their work successfully demonstrates many DEM-processing techniques as well as the usefulness of DEMs in topological analysis, but in some steps, they relies on semi-automatic procedure that requires manual intervention and often employ non-DEM data for classification, while our aim is to develop automatic algorithms that produce landform classification maps from DEM only. Matsuura, Aniya, and Yokohari (2004) developed an automated algorithm to find a class of valleys typical in Japan (“Yatsu-valley”) using only DEM as an input. The algorithm introduces the slope profiles of six types that appears to be also useful in.

Taewoon KIM is with the Disaster Management and Mitigation Group, Information Security Research Center, National Institute of Information and Communication Technology, 4-2-1 Nukui-Kitamachi, Koganei, Tokyo, 184-8795 Japan, phone: 042-327-5801, fax: 042-327-7941, email: tkim3@nict.go.jp.
identifying other landforms.

In this study, we developed an automated algorithm that identifies several selected basic landform classes, which are upland, lowland, and mountain, from a grid-based DEM to evaluate the performance and feasibility of the technique. The algorithm was demonstrated on the 50m-DEM of Miyagi prefecture (contains over 6.6 million cells in a grid format), Japan, to produce a landform classification map. Each cell stores an elevation value of the 50m by 50m area represented by the cell. Based on the fact that ground condition is closely related to the landform (Nakai, Bae, 2005), the result were then compared to the land condition map, 1:25,000 scale, by Geographical Survey Institute, Japan (called Landcondition hereafter), for the evaluation. Additionally, the result was used to increase details of a landform classification map in larger scale that consists of cells of size 1km such as JEGM (Wakamatsu et.al., 2005), and compared again to the land condition map. ESRI ASCII Grid is chosen for the container of DEM, computer programs were written in Java to process DEM, and ArcMap is used for visual inspection of the results.

2. Methodology

The position of a cell in a grid-base DEM can be referenced by (x, y) or (row, column) notations. The algorithm computes classification values on each cell from the elevation value and positional relations. JEGM and Landcondition used in this study has 17 and 10 classification criteria each, it is required to group those into upland, lowland, and mountain to enable comparisons with the classification results of the algorithm being presented. (Table 1) The grouping was performed manually according to the meaning implied by the classification names. Some criteria are excluded from the grouping and from the result comparisons for that they do not fit in any target categories.

![Cross profiles of upland, lowland, and mountain](image-url)

2.1. Classification Criteria

The algorithm classifies cells into three types, upland, lowland, and mountain (Fig. 1). The gradient of the cells in upland and lowland is less than or equal to 5 degree. The gradient threshold was used to filter out the cells with slope too steep to be considered as gentle or flat. 5 degree gradient threshold appears in other studies too (Matsuura, Aniya, and Yokohari, 2004; Millaresis and Argialas, 2000) in the purpose of determining flatness.

![Classification flowchart](image-url)

2.2. Algorithm

The computation steps are shown in Fig. 2. First, gradient is calculated for each cell from the 3 by 3 neighborhood window according to the estimator given by Burrough (1986). From these gradients, slope profiling is performed which assigns a cell one of six...
slope types introduced by Matsuura(2004). Slope types can be conceptually summarized into concave, convex, and straight. It is assumed that an upland is delineated by convex cells, and lowland by concave. Next, for the cells with gradient less than 5 degree and slope type of straight, determine if a cell belongs to upland, lowland, or neither. Finally, mark the cells which are not upland or lowland as mountain.

Further discussion in determination of upland/lowland follows. The basic idea is to count number of convex and concave cells surrounding a flat area to see which slope type dominates in number. As in Fig. 3, for a given cell of straight slope type and gradient less or equal to 5 degree, the number of convex and concave cells delineating the area are counted. Then the cell can be determined to be in lowland if concave cell count is bigger than convex count, and upland otherwise.

With this strategy, the area A in Fig. 2 can be processed. The cells closer to the edge of many convex cells will be classified as upland while the cells closer to concave cells as lowland. By changing the way of comparing concave and convex counts, we can control the upland/lowland proportion in the area A. For example, if we require a lowland cell to have twice more concave cell count than convex cell count, more cells in the area A will be identified as in upland. However, it is inevitable to have some lowland cells in the area A near the neighboring uphills, and this is expected to contribute to error of the result.

The classification result from the DEM (let us call it DEM classification) is also merged with JEGM in the way that increases the border details of the classification areas in JEGM. Since JEGM has lower resolution due to the larger grid size and DEM classification has more inconsistent areas such as small group of cells with divergent classifications, the classification area borders of JEGM are adjusted by referring DEM classification iteratively until no further change occurs, attempting to increase the border details of JEGM while excluding the inconsistent areas in DEM classification.

The visual representation of the landform classification results are shown in Fig. 4.

<Fig. 3> Determining upland / lowland

With this strategy, the area A in Fig. 2 can be processed. The cells closer to the edge of many convex cells will be classified as upland while the cells closer to concave cells as lowland. By changing the way of comparing concave and convex counts, we can control the upland/lowland proportion in the area A. For example, if we require a lowland cell to have twice more concave cell count than convex cell count, more cells in the area A will be identified as in upland. However, it is inevitable to have some lowland cells in the area A near the neighboring uphills, and this is expected to contribute to error of the result.

The classification result from the DEM (let us call it DEM classification) is also merged with JEGM in the way that increases the border details of the classification areas in JEGM. Since JEGM has lower resolution due to the larger grid size and DEM classification has more inconsistent areas such as small group of cells with divergent classifications, the classification area borders of JEGM are adjusted by referring DEM classification iteratively until no further change occurs, attempting to increase the border details of JEGM while excluding the inconsistent areas in DEM classification.

The visual representation of the landform classification results are shown in Fig. 4.

<Fig. 4> (a), (b), (c), and (d): Visual representation of classification
3. Results and Conclusion

DEM classification, JEGM, and DEM classification merged with JEGM were compared with Landcondition. For each comparison, every cell in Landcondition was compared to the corresponding cell of the target classification grid and the ratio of classification match to the total number of cell was recorded and regarded as the precision of the comparison. This precision was computed for the whole classification map and each classification criteria for analysis (Table 2).

<table>
<thead>
<tr>
<th>Classification</th>
<th>DEM classification</th>
<th>JEGM</th>
<th>DEM classification + JEGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upland (10)</td>
<td>34.8</td>
<td>49.6</td>
<td>33.9</td>
</tr>
<tr>
<td>Lowland (48)</td>
<td>91.2</td>
<td>81.8</td>
<td>90.8</td>
</tr>
<tr>
<td>Mountain (42)</td>
<td>87.4</td>
<td>86.8</td>
<td>91.4</td>
</tr>
<tr>
<td>Overall (100)</td>
<td>83.7</td>
<td>80.5</td>
<td>85.1</td>
</tr>
</tbody>
</table>

The total precision of DEM classification is 84%, 3% higher than JEGM’s 81%, and the individual precisions of lowland and mountain were also higher or at least equal to those of JEGM. But in upland category, the precision is 35% and it is obviously lower than JEGM’s 50%. DEM+JEGM(DM classification merged with JEGM) has about 5% higher precision than JEGM alone, due mainly to the precision improvement in lowland and mountain categories. Upland precision, however, dropped to the value close to that of DEM classification. In summary, higher error rate is observed in upland while the other two categories and the overall precision were satisfactory.

The upland inconsistencies are generally located in the wide areas with very gentle change in elevation as well as the areas with jagged surface including series of narrow and small terraces. For the errors in the areas of gentle change, the errors are thought to be caused by the limited size of 3 by 3 processing window used for slope profiling. Because the size of the window is small, it cannot recognize very gentle concavity or convexity over large area. Lowland classification is not affected by this as much as upland for that lowlands are delineated mostly by steeper slopes such as mountains. To correct this, a larger processing window must be used in the areas with small changes in elevation to be flexible and accurate in computing slope types. The upland classification algorithm also need to be modified to incorporate the model for gently sloped uplands.

The areas with jagged surface also raise error rate causing the algorithm end up identifying many small uplands and mountains. This is due to the lack of perspective and comprehensive ability of the algorithm which does not consider the surroundings of the area on which it operates. Developing and running another algorithm as a second pass which unifies or removes these inconsistent small classification areas would improve the overall precision.

In this study, we demonstrated the identification of some basic landforms from a DEM to show the feasibility of the automated landform classification. The overall precision was 84%. The classification was combined with JEGM to produce a better result. Further research on improving the algorithm is requested for higher precisions and wider range of classification criteria.

References


