Int J Appl Earth Obs Geoinformation 63 (2017) 112–121

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Automatic mapping of event landslides at basin scale in Taiwan using a Montecarlo approach and synthetic land cover fingerprints

Alessandro C. Mondini[a,*], Kang-Tsung Chang[b], Shou-Hao Chiang[c], Romy Schlögel[d], Claudia Notarnicola[d], Hitoshi Saito[e]

[a] Istituto di Ricerca per la Protezione Idrogeologica, Consiglio Nazionale delle Ricerche, 06128 Perugia, Italy
[b] Department of Geography, National Taiwan University, Taipei 106, Taiwan
[c] Centre for Space and Remote Sensing Research, National Central University, 32001 Taoyuan, Taiwan
[d] EURAC Research, Institute for Applied Remote Sensing, 39100 Bolzano, Italy
[e] College of Economics, Kanto Gakuin University, 236-8501 Yokohama, Japan

ARTICLE INFO

Keywords:
Systematic event landslide mapping
Uncertainty estimation
Image analysis

ABSTRACT

We propose a framework to systematically generate event landslide inventory maps from satellite images in southern Taiwan, where landslides are frequent and abundant. The spectral information is used to assess the pixel land cover class membership probability through a Maximum Likelihood classifier trained with randomly generated synthetic land cover spectral fingerprints, which are obtained from an independent training images dataset. Pixels are classified as landslides when the calculated landslide class membership probability, weighted by a susceptibility model, is higher than membership probabilities of other classes. We generated synthetic fingerprints from two FORMOSAT-2 images acquired in 2009 and tested the procedure on two other images, one in 2005 and the other in 2009. We also obtained two landslide maps through manual interpretation. The agreement between the two sets of inventories is given by the Cohen's k coefficients of 0.62 and 0.64, respectively. This procedure can now classify a new FORMOSAT-2 image automatically facilitating the production of landslide inventory maps.

1. Introduction

Landslides occur when slopes are disturbed by earthquakes, storms, human activities, or a combination of these factors (Aleotti and Chowdhury, 1999). Landslides, which are frequent and widespread in the world, can cause fatalities and social and environmental damages. Landslides can involve movements of sliding, flowing, toppling, and/or falling (Cruden and Varnes, 1996; Cruden and Varnes, 1996; Sidle and Ochiai, 2006). A landslide inventory map records the location of mass movements (Guzzetti et al., 2012).

Mapping landslides is a difficult task. Conventional methods rely on the visual interpretation of stereoscopic aerial photographs or satellite images, aided by field surveys. They are time consuming and resource intensive (Guzzetti et al., 2012). Quantitative image analysis has facilitated the task in recent years. Landslides mapping is a type of image classification problem, as landslides represent one of many classes that constitute a land cover (Michie et al., 1994). Several semi-automatic and automatic methods have been used to map landslides. They can be grouped according to: (i) pixel based (Borghuis et al., 2007; Mondini and Chang, 2014; Parker et al., 2011) or object oriented (Cheng and Han, 2016; Lu et al., 2011; Martha et al., 2010, 2011, 2012; Stumpf and Kerle, 2011), (ii) change detection analysis (Yang and Chen, 2010) or single image approach (Borghuis et al., 2007; Mondini et al., 2013, 2014), and (iii) supervised or unsupervised classifications (Mondini et al., 2011a,b). In most of the cases, a combination of different approaches is used (Guzzetti et al., 2012). The training phase of these methods demands resources and time. Training samples can be obtained using site visits, maps or, more commonly, through photo interpretation of the satellite images (Richards and Jia, 2006; Gupta and Rajan, 2011).

This study investigates whether it is possible to use a set of training samples prepared from independent images capturing previous landslide events. Models of samples represent a *digital library* of an "a priori knowledge" to obtain a new landslide map once a new image is available on the area where the library has been prepared.

In this work, the library trains a supervised Maximum Likelihood (ML) classification. This choice dictates constraints in the definition of the elements of the library (Foody et al., 1992). ML assigns each pixel of

* Corresponding author.
E-mail address: alessandro.mondini@irpi.cnr.it (A.C. Mondini).

http://dx.doi.org/10.1016/j.jag.2017.07.016
Received 22 March 2017; Received in revised form 28 July 2017; Accepted 30 July 2017

an image to a land cover class (LCC) by measuring a distance between the pixel's spectral values and some knowledge on the land cover spectral properties (fingerprints), which are assumed to be normal multivariate distributed. Distribution parameters refer to the LCC population properties, which are unknown but can be estimated by modelling the statistical spectral behaviour of some training samples.

We obtained our fingerprints by selecting random samples in a set of multi-temporal images. Our sampling method and image datasets allowed having some LCC spectral behaviour variability, providing an uncertainty measure of the estimated parameters.

To measure the potential effects of training and classification data set shifts (Tuia et al., 2016), we adopted Montecarlo analysis to compute the error propagation in the final classification. If the propagated uncertainty is too high, the classification is to be refused, and the library needs to be upgraded.

The semantic landslide LCC is a bare soil class subsample. To distinguish landslides among other geomorphological features like riverbeds or run outs, an existing a priori geomorphological knowledge (Parker et al., 2011; Mondini and Chang, 2014) on where event landslides are expected to be more (or less) abundant in a region is introduced as filter.

We applied the proposed approach to mapping landslides triggered by typhoons in southern Taiwan. We prepared training samples from two satellite images post-Typhoon Morakot, and we classified landslides on two other images, one pre-Morakot and the other post-Morakot. We compared our final maps with inventories prepared through manual interpretation.

2. Method

The framework includes two main blocks: the library preparation and the classification of new images (Fig. 1). The ML choice constrains the preparation of the library, establishing a dependence of the first block to the second.

2.1. Maximum likelihood classifier

In a Bayesian framework (Richards and Jia, 2006), the probability for the identification of a LCC w_i in a pixel, given its reflectance, is:

$$p(w_i|\boldsymbol{x}) = p(\boldsymbol{x}|w_i)p(w_i)/p(\boldsymbol{x}) \tag{1}$$

where $p(w_i)$ is the probability that class w_i occurs in the image, assumed to be equal for all classes; $p(\boldsymbol{x}|w_i)$ is the probability that a pixel with its reflectance belongs to a class w_i given the class w_i characteristics; and $p(\boldsymbol{x})$ is the probability to find a pixel from any class of M LCCs at its location, obtained with the formula:

$$p(\boldsymbol{x}) = \sum_{i=1}^{M} p(\boldsymbol{x}|w_i)p(w_i) \tag{2}$$

Using a ML classifier, the probability distributions describing the statistical spectral properties of each LCC must be normal multivariate. The class membership probability (CMP) of pixel x, given its spectral values, to class w_i, is then:

$$p(\boldsymbol{x}|w_i) = (2\pi)^{-N/2}|\Sigma_i|^{-1/2} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^t\Sigma_i^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i)\right\} \tag{3}$$

where $\boldsymbol{\mu}_i$ and Σ_i are the mean (central tendency) and the covariance (dispersion) of the multivariate normal probability distribution representing the class w_i, N is the number of training classes, and t stands for "transpose." $\boldsymbol{\mu}$ and Σ are unknown, but they can be estimated (($\boldsymbol{\mu}$, Σ)- > ($\boldsymbol{m}$, S)) from a representative set of training samples pixels selected in the image to classify.

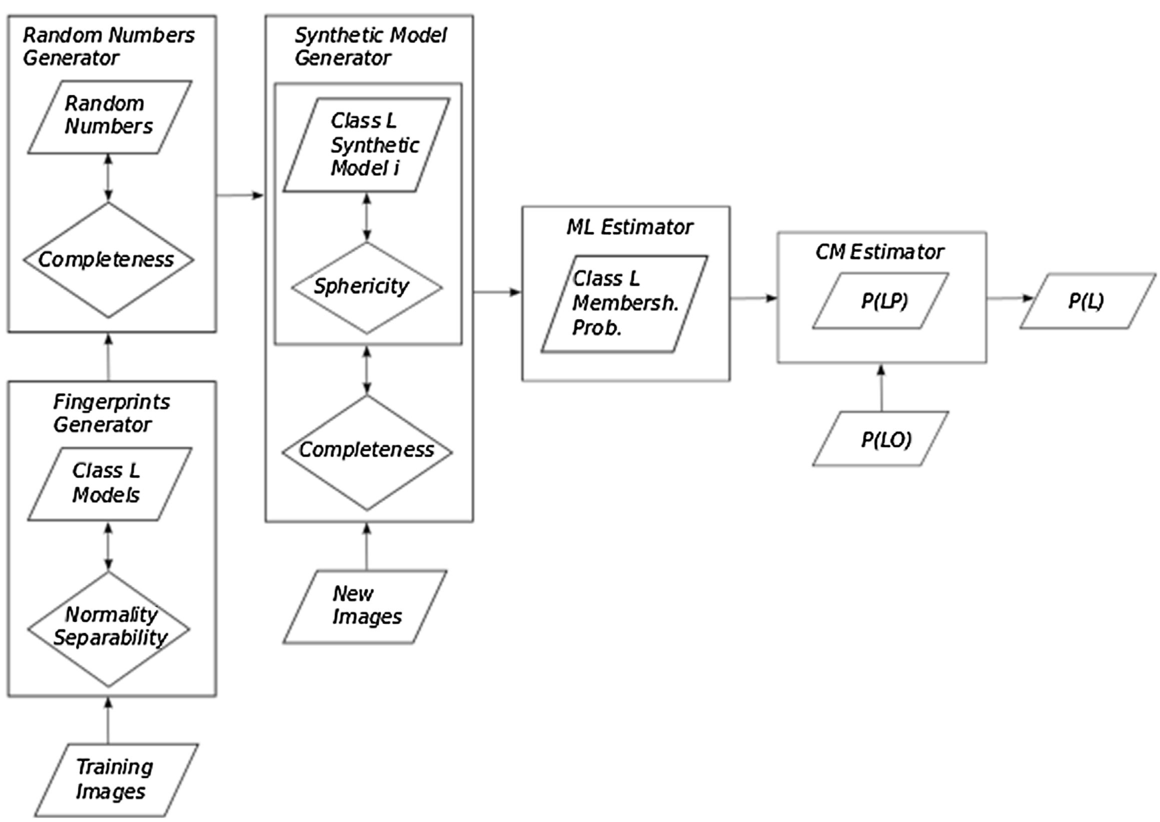

Fig. 1. procedure workflow. It includes five blocks (rectangles): a Fingerprints generator, a Random numbers generator, a Synthetic models generator, the ML estimator, and the CM estimator. Parallelograms represent input/output data: training images, new images, a susceptibility map, landslide map. Rhombuses represent completeness, normality, separability and spericity tests.