Computers & Geosciences ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Computers & Geosciences

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

Case study

Reconstruction of binary geological images using analytical edge and object models

Mohammad J. Abdollahifard*, Sadegh Ahmadi

Electrical Engineering Department, Tafresh University, Tafresh, Markazi Province, Iran

ARTICLE INFO

Article history:
Received 14 September 2015
Received in revised form
29 November 2015
Accepted 29 December 2015

Keywords:
Prior model
Training image
Gradient descent
Inverse problems

ABSTRACT

Reconstruction of fields using partial measurements is of vital importance in different applications in geosciences. Solving such an ill-posed problem requires a well-chosen model. In recent years, training images (TI) are widely employed as strong prior models for solving these problems. However, in the absence of enough evidence it is difficult to find an adequate TI which is capable of describing the field behavior properly. In this paper a very simple and general model is introduced which is applicable to a fairly wide range of binary images without any modifications. The model is motivated by the fact that nearly all binary images are composed of simple linear edges in micro-scale. The analytic essence of this model allows us to formulate the template matching problem as a convex optimization problem having efficient and fast solutions. The model has the potential to incorporate the qualitative and quantitative information provided by geologists. The image reconstruction problem is also formulated as an optimization problem and solved using an iterative greedy approach. The proposed method is capable of recovering the image unknown values with accuracies about 90% given samples representing as few as 2% of the original image.

1. Introduction

Characterization of environmental variables is accomplished using samples acquired by either in situ data-acquisition or remote sensing. Although with the development of complex acquisition methods the amount of data available for characterization of geological variables is becoming abundant, the complete sampling of a field is performed very rarely in practice due to financial and practical limitations. In other words there are always unsampled regions between acquired data which should be estimated using interpolation techniques.

Reconstruction of missing image data can be considered as an ill-posed inverse problem with many possible solutions. Therefore it is necessary to confine the solution space to geologically realistic patches using either regularization techniques (Lee and Seinfeld, 1987; Calderon et al., 2015) or probabilistic prior models. A multi-Gaussian distribution can be easily parameterized by mean and spatial covariance. Although mathematically convenient properties of multi-Gaussian distributions make them popular for modeling spatial variables (Chu et al., 1995; Li et al., 2003; Emery, 2007; Mariethoz et al., 2009; Abdollahifard and Faez, 2013b), their limited variability results in overly smoothed maps not consistent with realistic heterogeneities. Such methods are unable to reproduce the connectivity patterns appropriately for modeling flow and transport processes (Western et al., 2001; Knudby and Carrera, 2005; Bastante et al., 2008; Klise et al., 2009; Green et al., 2010). To alleviate this problem, some nonlinear and non-Gaussian high-order statistics models were developed based on spatial connectivity measures including spatial cumulants (Dimitrakopoulos et al., 2010) and copulas (Bárdossy and Li, 2008).

Object-based methods are able to produce realistic patterns with good spatial connectivity through defining basic shapes representing geobodies and placing them in the model domain based on a probability model (Deutsch and Tran, 2002; Allard et al., 2005; Keogh et al., 2007; Pyrcz et al., 2009; Michael et al., 2010). As an important advantage, object-based methods are able to control the parameters of geobodies (e.g. the channel width and orientation) to some extent. However, the conditioning to observed samples is usually achieved using a trial and error process resulting in a significant increase in computational complexity (Lantuéjoul, 2002; Allard et al., 2005).

The use of training images (TIs) has recently gained significant popularity for modeling environmental variability (Strebelle, 2002; Feyen and Caers, 2006; Zhang et al., 2006; Honarkhah and Caers, 2010; Mariethoz et al., 2010; Mariethoz and Renard, 2010; Abdollahifard and Faez, 2013a; Abdollahifard, 2015). The TI is an interesting tool for geologists allowing them to represent the desired geological concept in a direct manner (Feyen and Caers,

* Corresponding author.
E-mail address: mj.abdollahi@tafreshu.ac.ir (M.J. Abdollahifard).

http://dx.doi.org/10.1016/j.cageo.2015.12.018
0098-3004/

2006). To reconstruct an incomplete patch, multiple-point simulation (MPS) methods seek the TI to find a patch consistent with local samples. MPS methods are capable of producing realistic images conditioned to observed samples. The problem of finding a suitable patch among thousands of training patches (known as template matching problem) should be solved several times. As a result, MPS methods are CPU-intensive. Extensive effort was devoted to overcome this problem by using search trees or lists (Strebelle, 2002; Straubhaar et al., 2011), approximate gradient descent (Abdollahifard and Faez, 2013a), search space reduction (Abdollahifard, 2015), and training pattern clustering (Zhang et al., 2006; Honarkhah and Caers, 2010). Although the efforts have brought significant improvements in CPU-time, the approaches are still remarkably slower than their two-point predecessors because of their search-based nature.

The problem of selecting a proper training image is also a challenge in MPS approaches, especially when enough information is not available for such a decision. Selection of an inadequate TI may lead to realizations incompatible either with observed data or real field variations (Pyrcz et al., 2008; de Almeida, 2010). Even when the geological context is clear, constructing a complex 3D training image that adequately represents the complexity of geological structures requires lengthy computations (Mariethoz and Kelly, 2011).

Furthermore, unlike multi-Gaussian models or object-based models, it is not straightforward to parameterize the training images. In other words, the problem of TI selection is a discrete decision (either image A or B) and there is no parameter in a specific TI to be controlled continuously. Suzuki and Caers (2008) proposed a parameterization allowing several discrete choices of geological architectures within the prior. Another interesting TI parameterization is proposed by Mariethoz and Kelly (2011) by using small elementary training images as basic structural elements of the field and applying some parameterized transformations (e.g. rotation and affinity) on the lag vectors. Although the TIs selected are smaller than traditional ones, the additional continuous freedom degree (e.g. rotation) causes a remarkable increase in the search space.

The aim of this paper is, first, to introduce a single model capable of modeling a wide range of binary images and then, to exploit this model for reconstruction of missing image values. Consider the binary images depicted in Fig. 1(a)–(c). Although the images seem very different in macro-scale, their micro-scale building blocks are very similar as depicted in Fig. 1(d)–(f). In the selected scale, the patches have very simple structures and can be modeled effectively using one or a combination of two linear edges. This holds true for every binary geological image (either simpler or more complicated), if the proper patch size is considered.

In this paper an analytical model for linear edges is suggested with parameters controlling the edge orientation and sharpness. Thanks to the convenient mathematical properties of the suggested model, for any given incomplete patch (or a complete noisy patch) its corresponding match in the model space can be found in a few iterations using classical optimization techniques. This is in contrast to TI-based approaches whereby CPU-intensive exhaustive search is usually required to find a match in the TI. To achieve further flexibility, we have enriched the model space by allowing a combination of two linear edges capable of approximating narrow channels and nonlinear edges. Geologists' knowledge can also be incorporated by controlling the edge orientation range and setting relational constraints on two edges in

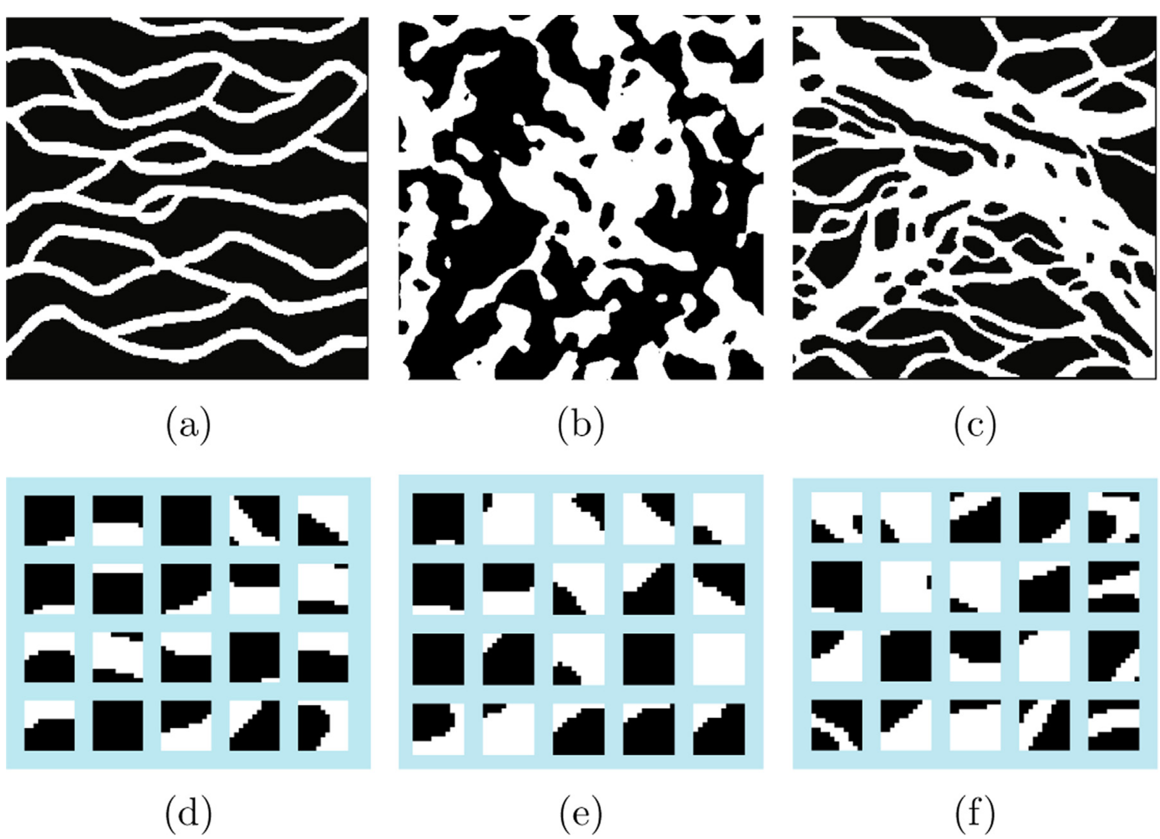

Fig. 1. Top row: three binary geological images (all images are obtained from the website of the book, Mariethoz and Caers, 2014). (a) A simplification of a channelized depositional system with white as sand and black as shale (Strebelle, 2002) of size 251×251, (b) a 245×245 image obtained from truncated Gaussian simulation, (c) a 243×243 image constructed based on a satellite image of the Ganges delta (Bangladesh), with soil properties classified as channel (white) and alluvial bars (black). Bottom row: 10×10 patches randomly extracted from binary images of top row.