ISSN: 2664-9640

Expectation maximization algorithm segmentation of satellite images

Qingzhen Gong

School of Physics and Electronic Information Engineering, Jining Normal University, Ulanqab City, Inner Mongolia, 012000, China

Center for International Education, Philippine Christian University, Manila City, 1004 Metro Manila, Philippines

* Corresponding Email: xunuo1314_2012@163.com

Abstract

In this paper, we explain the definition, composition and different types of resolution of satellite images, and analyze and segment the images through the Expectation maximization (EM) algorithm, Gaussian hybrid model and K-Means method, so as to obtain the advantages and defects of various methods.

Keywords

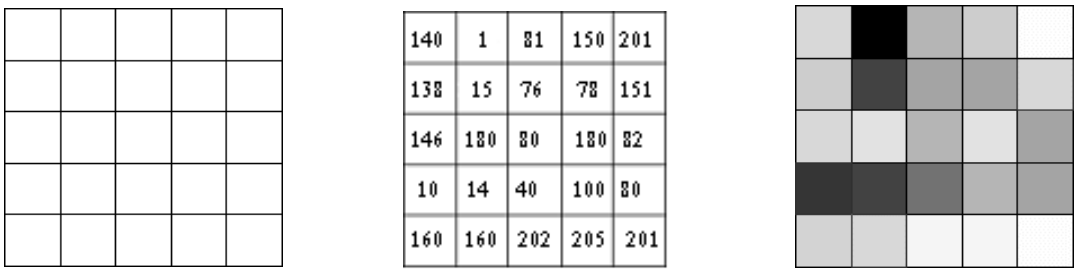
EM, Gaussiennes, K-Means, Segmentation.

1. Introduction

Image segmentation is to divide an image into multiple identical areas called categories. This spatial partition of the image is based on the statistical properties of pixels. The image is represented by a result (x1, x2, ...). , xn) the mixing of parametric laws. In this project, we suggest using EM algorithm to estimate these parameters. On the one hand, the required work is to implement the algorithm, and on the other hand, to segment satellite images with the algorithm.

2. Satellite image

A satellite image is a matrix of pixels. A pixel is the smallest uniformly composed area in a recorded image. Each value received by the ground station allows the drawing of a small square, called a "pixel," which is given a more or less intense shade of gray. Here is a chart showing an example of coding：



a) b) **c)**

Figure 1 Shows a schematic of an encoding example

Satellite images have different types of resolution: spatial resolution, spectral resolution, radiometric resolution and temporal resolution. In measurement systems, resolution is called

ISSN: 2664-9640

the ability of an instrument to distinguish between two objects or two adjacent values. In terms of images; It is also used to design the ability to distinguish between two adjacent points.

2.1. Spatial or geometric resolution

Spatial or geometric resolution is the minimum distance between two adjacent objects. This resolution measures the edges of a pixel in meters or kilometers. In fact, the detail that can be seen on the image depends on the spatial resolution of the sensor used. Spatial resolution is a function of the size of the smallest element that can be detected.

For example, a finer resolution (about 10 meters to 1 meter) can identify the finest communication networks. For example, for an image with a resolution of 20 meters, each pixel represents the corresponding area of 20 × 20 meters.

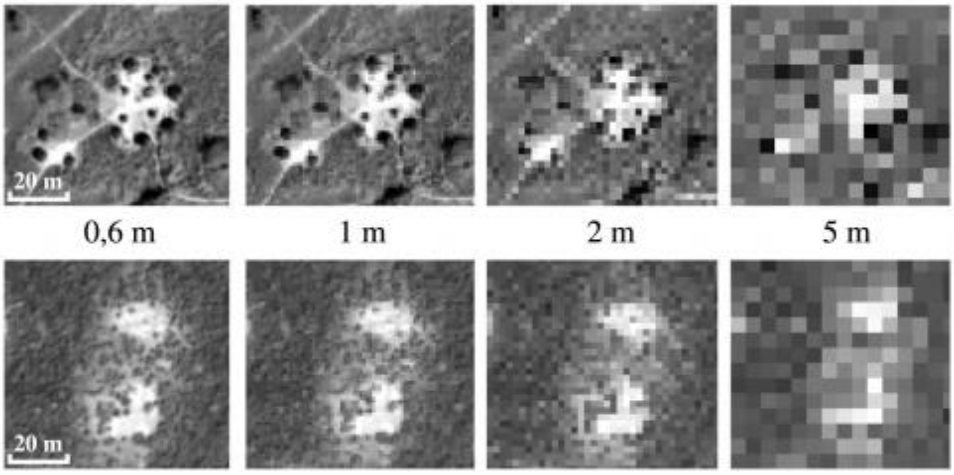


Figure 2 Images with different spatial resolutions

2.2. Resolution of spectrum

Spectral resolution describes the ability of a sensor to use a small wavelength window. If the spectral resolution is too coarse, it is impossible to distinguish between different minerals.

2.3. Resolution of radiation

The pixel arrangement describes the spatial structure of the image, while the radiation features describe the information contained in the image. Whenever an image is captured by a film or sensor, its sensitivity to the intensity of electromagnetic energy determines the radiative resolution. The radiative resolution of a remote sensing system describes its ability to identify small differences in electromagnetic energy. The better the radiation resolution of the sensor, the more sensitive the sensor will be to small differences in the intensity of the energy received. The range of wavelengths the sensor is sensitive to is called the dynamic range.

The image data is represented by a numerical value, which varies between 0 and 2, with a certain power reduced by 1. This range corresponds to the number of bits used to encode the value into binary format. Each bit represents an exponent of base 2 (for example, 1 bit = 21 = 2). The maximum number of intensity levels available depends on the number of bits used to represent the recorded intensity. For example, a sensor that uses 8 bits to record data will have 28 = 256 strength levels available because it will have 256 values available, ranging from 0 to 255. If you use only four bits, then only values between 24 = 16 0 and 15 are available. As a result, the radiation resolution will be reduced. The recorded data is usually shown in gray, with black representing the value "0" and white representing the maximum value. Comparing a 2-bit image of the same scene to an 8-bit image shows a huge difference in the amount of detail that can be distinguished according to radiometric resolution.

ISSN: 2664-9640

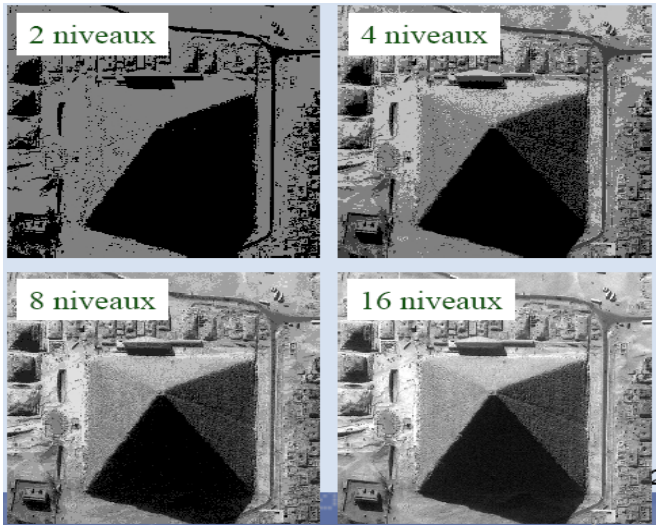


Figure 3 Satellite images with different radiological resolutions

2.4. Resolution of time

Temporal resolution is the time it takes a satellite to observe the same scene from the same point in space. Therefore, the absolute time resolution of the remote sensing system is equal to this time. However, certain areas of the surface can be observed more frequently because there is overlap between adjacent corridor overlays, and these overlap areas become increasingly large as they approach the poles. Some satellites can also point sensors at the same point for different satellite transits. One of the main advantages of satellite exploration is its ability to collect information on the same part of the Earth on a regular basis. The spectral characteristics of the observed region may change with time. These changes can be detected by comparing multi-temporal images.

3. EM algorithm and Gaussian mixture model

3.1. Gaussian mixture model

Gaussian mixing (GMM, GM or MoG, depending on the reference) is a widely used tool in the literature and computer engineering. It can be used to model maximum likelihood numerical data for probabilistic model parameters when the model depends on unobservable latent variables.

In mixed models, which are often used for automatic classification, data samples are not considered to follow the usual probability laws, but rather to follow a law where the density function is the mixed density.

While any law can be used, the most common is that the density function is Gaussian's normal law. This is known as a Gaussian mixture.

3.2. The EM algorithm

The EM algorithm is an iterative procedure that allows the values to be close to the maximum likelihood.

EM algorithms are a class of algorithms that allow finding the maximum likelihood of the parameters of a probabilistic model when the model depends on unobservable latent variables.

Expectation maximization is commonly used in data classification, machine learning, or artificial vision.

Expectation maximization alternate steps:

Step E: The expectation evaluation phase, calculating the probability expectation considering the last observed variable.

ISSN: 2664-9640

Step M: The maximization phase, which estimates the maximum likelihood of the parameters by maximizing the likelihood found in step E, then uses the parameters found in M as the starting point for a new expectation evaluation phase, and repeats this step.

3.3. EM algorithm for mixed models

One of the main applications of EM is to automatically classify the estimation of mixture density parameters in Gaussian mixture models.

In this problem, the sample $(x_1, \ldots, x_n)$ of $\mathbb{R}^p$ characterized by p continuous variables actually comes from G different groups. Considering that each group G_k follows the f law with parameters θ_k, its proportion is given by the vector $(\pi_1, \ldots, \pi_g)$.

Let $\Phi = (\pi_1, \ldots, \pi_g, \theta_1, \ldots, \theta_g)$ denote the mixing parameter, and the density function of the sample is

$$g(x, \Phi) = \sum_{k=1}^{g} \pi_k f(x, \theta_k),$$

Therefore, the log-likelihood of the parameter Φ is given by

$$L(x, \Phi) = \sum_{i=1}^{n} \log \left(\sum_{k=1}^{g} \pi_k f(x_i, \theta_k) \right).$$

This function maximizing data Φ is very complicated. For example, if we want to determine the parameters corresponding to 2 groups in a 3-dimensional space according to the normal law, which is small, we have to optimize the nonlinear function of $\mathbb{R}^{26}$.

At the same time, if we knew which group each individual belonged to, then the problem would be a very simple and classical estimation problem.

The EM algorithm has the advantage of using this data for estimation. If x_i belongs to group G_k, then the z_{ik} value is 1, otherwise it is 0, then the log probability of completing the data is:

$$L(x, z, \Phi) = \sum_{i=1}^{n} \sum_{k=1}^{g} z_{ik} \log \left(\pi_k f(x_i, \theta_k) \right).$$

And then we get it very quickly

$$Q\left(\Phi, \Phi^{(c)}\right) = \sum_{i=1}^{n} \sum_{k=1}^{g} E\left(z_{ik} | x, \Phi^{(c)}\right) \log \left(\pi_k f(x_i, \theta_k) \right)$$

By recording the quantity t_{ik} given by $t_{ik} = E\left(z_{ik} | x, \Phi^{(c)}\right)$, it is possible to divide the EM algorithm into two phases, classically referred to as the estimation phase and the maximization phase in the case of mixed models. These two steps are repeated until convergence.

Step E: Calculate t_{ik} using Bayesian inversion rule:

$$t_{ik} = \frac{\pi_k f(x_i, \theta_k)}{\sum_{\ell=1}^{g} \pi_\ell f(x_i, \theta_\ell)}$$

Step M: Determine that Φ is maximized

$$Q\left(\Phi, \Phi^{(c)}\right) = \sum_{i=1}^{n} \sum_{k=1}^{g} t_{ik} \log \left(\pi_k f(x_i, \theta_k) \right)$$

ISSN: 2664-9640

The advantage of this approach is that the problem can be broken down into *g* basic problems, which are usually relatively simple. In all cases, the optimal ratio is given by

$$\pi_k = \frac{1}{n}\sum_{i=1}^{n} t_{ik}$$

The estimate of θ also depends on the chosen probability function *f*. In the normal case, these are the μ_k mean and the variance-covariance matrix Σ_k.

$$\mu_k = \frac{\sum_{i=1}^{n} t_{ik} x_i}{\sum_{i=1}^{n} t_{ik}}$$

$$\Sigma_k = \frac{\sum_{i=1}^{n} t_{ik}(x_i - \mu_k)(x_i - \mu_k)'}{\sum_{i=1}^{n} t_{ik}}$$

M' is M transpose, and let's say μ_k is a column vector.

3.4. The characteristics and disadvantages of EM algorithm

The characteristics of EM algorithm

Determine convergence to a possible local optimum.

The final result depends on the initialization.

The convergence of the algorithm can be slow.

The defects of EM algorithm

One of the main defects of EM algorithm is that it needs to select the value of *K* a priori. It depends on the experimenter knowing the number. This is problematic because the larger *K*, the higher the probability, and we want it to be "just big enough". The simplest solution is to use criteria that penalize too much complexity, i.e. *K* is too high. However, this requires tuning the algorithm for multiple rotations because we want different values of K to compare.

In particular, as mentioned before, EM converges to a local maximum; But we can't guarantee that this maximum is "good," i.e. Close to the global maximum in value. To achieve this goal, a further increase in computation is necessary.

Finally, standard EM runs the risk of finding "degenerate" solutions where at least one component is suitable for a single individual; However, this Gaussian possibility may be infinite (i.e. close to the Dirac peak), thus breaking our algorithm.

4. K-means

One of the most commonly used unsupervised classification (clustering) techniques. Given an integer *K*, *K*-means divides the data into *K* non-overlapping groups or "clusters" or "classes". This is done by placing *K* "prototypes" or "centroids" in the most populated area of space. Each observation is then assigned to the nearest prototype (the so-called minimum distance rule). Thus, each class contains observations that are closer to a stereotype than any other stereotype (Figure below).

ISSN: 2664-9640

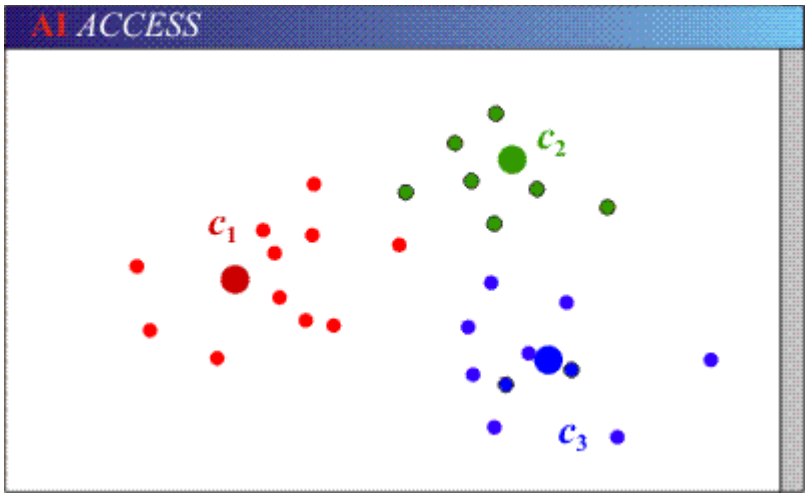


Figure 4 Example

K-means is so popular because:

Simplicity of concept.

High speed, low memory size requirements.

But it also has some disadvantages:

The user must pre-select the value of *K*, the number of classes. For two-dimensional data, this can be achieved by simple visual inspection, but not for high-dimensional data. In general, there is no clear indication of the appropriate number of classes, and the "wrong choice" of *K* values will lead to typologies that are disconnected from reality.

For a given value of *K*, the resulting class depends heavily on the initial configuration of the prototype, making it difficult to interpret the class.

K-means is an objective technique, which means that it minimizes the value of a numerical criterion. So, this is an optimization technique. As is often the case in optimization, the *K*-means algorithm stops when it cannot lower the standard value. However, it is likely that different configurations of the prototype will result in lower standard values. In the optimization vocabulary, *K*-means is said to reach a local minimum, but does not guarantee to reach the global minimum (the lowest possible value) of the criterion.

5. Segmentation

Clustering is a basic step in image processing. The purpose of this operation is to separate different homogeneous regions in the image in order to organize objects into groups with different properties (intensity, color, texture, and so on). Segmentation methods can be divided into two categories: unsupervised segmentation and supervised segmentation:

Supervised segmentation:

This is based on the knowledge of each class defined by probabilistic methods.

Unsupervised segmentation:

It aims to automatically segment images into natural clusters, that is, without any prior knowledge about classes.

Different approaches to unsupervised classification

Faced with such an imperfectly defined problem, it is natural that a plethora of technologies should emerge.

Partition "hard"

ISSN: 2664-9640

The general idea is to divide the observation space into a number of separated regions defined by boundaries and to stipulate that all observations located in the same spatial region belong to the same class.

There are many partitioning and classification techniques, the best known being *K*-means and its variants.

Partition "soft"

Mixed model: This approach uses the concept that classes have multiple normal distributions. Thus, for a given number of hypotheses of *K* classes, the dataset is actually *K* samples of multiple normal distributions, each with a prior probability. We then look for the covariance centers and matrices of these distributions, as well as prior probabilities, to maximize the likelihood of the data, which is done by the classical algorithm "EM" .

6. Conclusion

By comparing with several algorithms, we learn about the most common segmentation algorithms in the families of algorithms based on EM algorithms and Gaussian mixing. The proposed study involves the integration of K mean algorithm to ensure better identification of mixmixture.

References

[1] Romualdo Douglas Colauto, Paulo Savaget Nascimento, Bruna Camargos Avelino et al. Evidenciação de Ativos Intangíveis Não Adquiridos nos Relatórios da Administração das Companhias Listadas nos Níveis de Governança Corporativa da Bovespa[J] Contabilidade Vista & Revista, 2009, 20(1).

[2] C.C.R. Cerqueira, L.S. Bicalho Composed “piggy back” splints for maxilla segmentation in orthognathic surgery: a novel technique[J] International Journal of Oral & Maxillofacial Surgery, 2019, 48(Supl.1).

[3] ML Montandon, F Assal, O Ratib et al. MRI-guided voxel-based partial volume effect correction in brain PET: Assessment of the impact of 3 MR image segmentation algorithms[J] Neuroimage, 2009, 47(S1).

[4] D. Pasquier, L. Peyrodie, F. Denis et al. Segmentation automatique des images pour la planification dosimétrique en radiothérapie[J] Cancer / Radiothérapie, 2010, 14(Suppl.1).

[5] Fei Zheng;Stéphane Derrode;Wojciech Pieczynski Semi-supervised optimal recursive filtering and smoothing in non-Gaussian Markov switching models [J] Signal Processing,2020.