

Recognition of 2-D Objects by Optimal Matching

Liu Lu, Fang Luo, N. J. Mulder

PB 162, P O Box 6, 7500AA, Enschede, The Netherlands

Abstract: This paper introduces an approach of recognizing 2-D objects by optimal matching. The method consists of two stages: object identification and object localization. Both of them are accomplished through optimal feature matching, in which the radiometric distribution of an object as a global feature extracted from an image is matched directly to the object model. A cost function is defined as a quantitative evaluation of the feature fitting and the recognition process is based on cost minimization. In this method, every subproblem in object recognition is formulated as an optimization problem and techniques of optimization are utilized to solve these problems.

Key Words: object recognition, optimization, cost function, object modelling.

1. Introduction

Object recognition is an important topic in computer vision. Today, at the time that techniques of 3-D object recognition have been widely developed, the recognition of 2-D objects is still of interest, because, on the one hand, 2-D object recognition has its own applications, such as the recognition of agriculture fields or rivers in remote sensing images; on the other hand, the 2-D object recognition could be the fundamental for some 3-D processing, for instance, in the matching of stereo images. In this paper, we concentrate on the recognition of agricultural fields in radar images. An object here means a land parcel of agricultural fields of a single crop type.

Object recognition concerns the recognition and localization of objects of interest in a scene from an image of the scene. According to their manners of dealing with these two issues, the approaches of object recognition can be categorized as: (1) recognition-followed-by-localization, or (2) recognition-via-localization [5]. Usually, in the recognition-followed-by-localization approaches, higher-order relational descriptions or global features of input are used to match to similar descriptions of objects. Therefore, this kind of approaches calls for extensive segmentation and their capabilities are limited by segmentation. On the contrary, the recognition-via-localization approach involves minimal effort in segmentation, since the primitive features extracted from the image are used to match directly to the object models.

In this paper, we present a method of recognizing 2-D objects. The method consists of two stages: object identification and object localization. In both stages, features extracted from an image are matched directly to the object

models that removes the restriction of segmentation. The radiometric distribution of an object as a global feature is used not only to identify the object but also to localize it. A cost function is defined as a quantitative evaluation of the feature fitting. In this way, the procedure of feature matching is transferred into a global minimum cost search. Problems in the optimization are (1) the existence of local minima in the cost, (2) the heavy computation burden. To solve these problems, specific techniques are developed in the proposed method. Local minima are avoided by using simulated annealing in the first stage, object identification, and its result serves as the initial state for the second stage, object localization. Multi-resolutions are applied to reduce the computation. As sizes of land parcels are not always known, the parameter estimation is also accomplished at the second stage. The final result includes identified objects and their parameters: location, rotation and scale as well as a confidence factor of the result.

The approach is object-based. All the processes, including the feature extraction and cost calculation, are executed at the object level.

Our approach is distinguished by the following respects:

1) The approach is composed of two stages. In both stages, the features extracted from an image are matched directly to the object models so that the capability of the approach is not limited by segmentation.

2) In addition to objects identified, the first stage offers a good initial state for the second stage to prevent the optimiser from failing into local minima.

3) Radiometric distributions are used as global features of objects in both object identification and object localization. The gamma distribution acquired with radar simulation is adopted as radiometric model to characterize the spectral properties of an object in a radar image.

4) Multi-resolutions are applied to speed up the optimization in the first stage.

5) Besides the recognition result, a confidence factor is offered which is related to the number of misclassified pixel.

Although the method is a problem-specific solution for agricultural field recognition, it is applicable for some other problems with trivial modifications.

This paper is organized into seven parts. Section 2 introduces our methodology and section 3 discusses the object model. The procedures of object identification and object localization are addressed in section 4 and 5 respectively. Section 6 presents experimental results, and Section 7 contains the summary.

2. The procedure of object recognition.

Consider the problem of 2-D object recognition that we are facing. Given a radar image of a scene, then the task of object recognition is to determine whether any of the objects of interest are present in the scene, and if they are, to determine their locations and scales.

Fig.1 shows our two stage procedure. In this procedure, object recognition is

divided into object identification and object localization. Both of them are accomplished through optimal feature matching. The optimal feature matching starts with a geometric prediction of an object. The radiometric feature of the object extracted from the predicted region in the image is compared with the object model. A cost function is calculated as a quantitative evaluation of the geometric prediction of the object in terms of its feature fitting. The single global minimum of the cost corresponds to the prediction with the best feature match. The feature matching is accomplished through optimization, the global minimum cost search. Fig.2 illustrates the procedure of optimal feature matching.

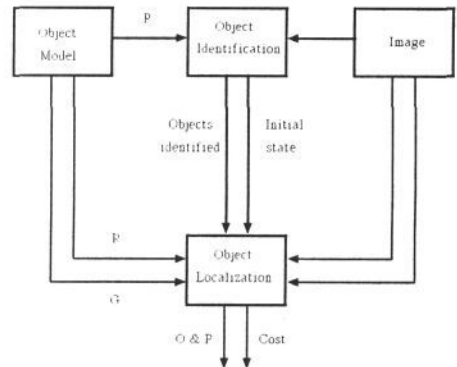


Fig. 1 The Two stage procedure for 2-D object recognition

P: Radiometric distributions of objects
 G: Geometric shape of objects
 O & P: Objects and their Parameters

is accomplished through optimization, the global minimum cost search. Fig.2 illustrates the procedure of optimal feature matching.

In optimization, the existence of local minima in the cost function is a big problem, because they make the optimiser fail into a false minimum and lead to a wrong result. There are three ways to solve this problem:

- 1) define a "smooth" cost function without or with less local minima.
- 2) choose a powerful optimization algorithm that guarantees to find the global minimum.
- 3) let optimization procedure start with a good initial state.

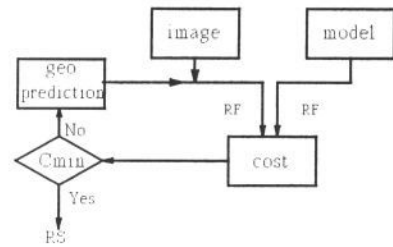


Fig. 2 The optimal feature matching

RF: Radiometric Feature
 RS: Result

The effort through the first way can be seen in the cost definition later. The function is smoothed by the integral. Because of their different tasks, the problem of local minima is dealt with through different ways in object identification and object localization. In object identification, the problem is solved by using simulated annealing, which guarantees to find the global minimum cost, while in object localization, local minima are avoided by starting with a good initial state. Besides, in object identification, the geometric prediction of an object is in the form of pixel set to suit simulated annealing. In object localization, parametric form of geometric prediction is adopted to get the localization and scales of objects.

Another advantage of using simulated annealing in the first stage is that it is insensitive to its initial state, which means the first stage can start with an arbitrary geometric prediction of the object.

3. The object model

An object model is a set of descriptions of objects or classes of objects that are expected to be required during the recognition procedure. These descriptions should be suitable for feature matching as well as sufficient to characterize the properties of objects to be recognized. Therefore, the structure of the model is dependent on the application domain.

The features of an object extracted from a radar image can be categorized as radiometric or geometric features. The radiometric features of an object corresponds to the land cover class, i.e. the crop type in a land parcel. The geometric features of an object contain the information about its shape, location and scales. Hence, the object model is composed of a radiometric part and a geometric part.

The radiometric model of an object described by the crop type and parameters of the distribution characterizes the spectral properties of an object in a radar image. It is acquired with the techniques of radar simulation and expressed as a gamma distribution:

$$P(I|A_{CT}) = \frac{1}{(\sigma^0/N)(N-1)!} \left(\frac{I}{\sigma^0/N}\right)^{N-1} \exp\left(-\frac{I}{\sigma^0/N}\right) \quad I \geq 0 \quad (3.1)$$

in which I is the intensity of an image pixel, A_{CT} is a region of the single crop type CT, N is the number of independent samples, σ^0 is referred to as the backscattering coefficient [2][4].

The geometric model is formed by geometric shape and parameters. As the land in Holland is rather flat and most fields have regular shapes, rectangle is employed and its parameters consist of one orientation, two translation and two scale parameters. These five parameters fix the localization and scale of a rectangle completely.

It should be noticed that the radiometric model, used to characterize the spectral properties of an object, is a statistical model. It implies this feature must be extracted from a region of enough size. As a matter of fact, in both stages of object identification and object localization, the radiometric feature of an object from the given image is extracted from the whole region of geometric prediction of the object.

4. Object identification

The task of object identification is to determine what objects are present in an image, in the other word, is to identify the crop type of each object.

The first stage starts with an arbitrary geometric prediction of an object. A geometric prediction of an object is an estimates about the geometric appearance of the object in an image. The radiometric feature is extracted from the predicted region in the image and compared with the object models. A cost function is calculated as a quantitative evaluation of the geometric prediction of the object in terms of its feature fitting. The single global minimum cost corresponds to the prediction with the best feature match. The feature matching

is accomplished through optimization, the global minimum cost search. After the global minimum cost is found, if it is under a threshold, the object is identified. In order to solve the problem of local minima, in this stage, simulated annealing is used to guarantee the global minimum of cost is found. Another advantage of using simulated annealing is that it is insensitive to its initial state, which allows the procedure start with an arbitrary geometric prediction of the object. The weakness of simulated annealing is its long computation time. A multi-resolution technique is performed to speed up the convergence.

Suppose there are L objects in an image. Every object in the image is identified individually. While an object is identified, all the others in the image are regarded as the background of it. In this stage, a geometric prediction of a object is a set of pixel supposed to belong to the object. If we denote a geometric prediction of the l th object O_l as P_{O_l} , $e(x,y)$ is an element of an image with the coordinate (x,y) , in original resolution it means a pixel, then a geometric prediction of the l th object can be expressed as: $P_{O_l} = \{e(x,y) \in O_l\}$. A cost function is defined to evaluate the geometric prediction in terms of its feature fitting.

4.1 A cost function for object identification

Since the cost serves as a quantitative evaluation of a geometric prediction in terms of its feature fitting and will guide the feature matching, in order to avoid an endless search or an unrepeatable outcome of the matching, the cost function must meet the following requirements:

- The cost function assign a lower cost for a prediction with better feature fitting.
- The cost function has a global minimum and its global minimum corresponds to the prediction with the best feature fitting.

Keeping these requirements in mind, we, now, define a cost function for object identification. If CT_l is the supposed crop type for the l th object, the cost function of a geometric prediction of the l th object is defined as:

$$C[P_{O_l}] \triangleq 2 + \int_I |p(I|P_{O_l}) - p(I|O_{CT_l})| dI - \int_I |p(I|P_{B_l}) - p(I|O_{CT_l})| dI \quad (4.1.1)$$

where $C[P_{O_l}]$ is the cost of a geometric prediction of the l th object P_{O_l} , $p(I|P_{O_l})$ is the radiometric distribution obtained from a geometric prediction of the l th object, $p(I|P_{B_l})$ is the radiometric distribution obtained from the geometric prediction of the background of the l th object, CT_l is the supposed crop type for the l th object, $p(I|O_{CT_l})$ is the radiometric distribution of supposed crop type for the l th object in the model.

It is easy to prove that $C[P_{O_l}]$ meets the requirements above and its single global minimum corresponds to the geometric prediction with the radiometric features matching perfectly to that of the supposed crop type.

4.2 The optimization process for object identification

After a proper cost function is defined, an optimization is performed to find

the prediction of the global minimum cost, which means the geometric prediction with the best feature match. The goal of the optimiser is to find the P_{O_i} that satisfies: $\forall P_{O_i}, C[P_{O_i}'] \leq C[P_{O_i}]$. The problem in optimization is the existence of local minima in the cost function that make the optimization fail into a false minimum and lead to a wrong identification.

In optimization theory, there are numerous algorithms, such as gradient, simplex and conjugate direction set etc, based on iterative improvement by comparing the current point with its neighbourhoods. Unfortunately, these algorithms do not guarantee to find the global minimum. Unlike these algorithms, simulated annealing accepts a function increase with a probability and does not stuck in a local minimum with any initial state [3],[6]. Therefore, simulated annealing is adopted to search the prediction of the global minimum cost.

However, simulate annealing is a very time-consuming algorithm. A technique of multi-resolution is developed to reduce the convergence time. Assume the full resolution of an image is $2^m \times 2^m$, $m > 1$. If we denote an image element in this resolution as $e_m(x,y)$, for an identified object O_i , every element has two states: $e_m(x,y) \in O_i$ or not. A geometric prediction of O_i is expressed as $P_{O_i}^m = \{e_m(x,y) \in O_i\}$. All the possible geometric predictions of an object is denoted as $\Omega_m = \{P_{O_i}^m\}$. Let us consider the image of a lower resolution of $2^{m-1} \times 2^{m-1}$, its element $e_{m-1}(x,y) = \{e_m(2x,2y), e_m(2x+1,2y), e_m(2x,2y+1), e_m(2x+1,2y+1)\}$ so that its whole set of prediction $\Omega_{m-1} \subset \Omega_m$. Furthermore, we get $\Omega_0 \subset \Omega_1 \subset \dots \subset \Omega_{m-1} \subset \Omega_m$. It shows that the reduction of image resolution yields the shrink of the prediction set which means a smaller search space and a shorter convergence time for optimization.

For each object to be identified, the optimization starts with a low resolution $2^n \times 2^n$, $1 < n < m$. Simulated annealing is executed to find the global minimum prediction $P_{O_i}^n$. Then the resolution is increased to $2^{n+1} \times 2^{n+1}$, another search of simulated annealing starts with $P_{O_i}^n$, which is quite close to $P_{O_i}^{n+1}$). This procedure is repeated until a predetermined image resolution $2^N \times 2^N$ is reached. Since the localization of an object is not determined by the result at this stage, the predetermined image resolution need not to be the original image resolution, i.e. $n \leq N \leq m$.

After the final minimum cost $C[P_{O_i}^{N_i}]$ is found, it is compared with a threshold. If it is under the threshold, the object is identified, i.e. the l th object is of the supposed crop type CT_l . Otherwise, it is aborted. In turn, all the objects in an image are identified.

5. Object localization

After all the objects in the scene have been identified, objects in the scene and their radiometric features are known. The task of the second stage, then, is to estimate the geometric parameters of each object, including the parameters of position, orientation and scale.

In fact, in object localization, there are two subtasks to be accomplished: 1) From the labelled image supplied by the object identification, get the initial parameters of each object for the object localization. 2) From the given image, get the optimal parameters of objects. Each of them is expressed as an optimization problem. Similar to the object identification, the localization is accomplished through the optimal feature matching. However, it is different from the first stage in the following aspects: 1) The initial parameters of each object is extracted from the minimum cost prediction of the object obtained in the first stage and they forms the initial prediction of the scene in the second stage. 2) The geometric prediction of an object is in a parametric form, i.e. a geometric prediction of the l th object in this stage is represented as $P_{O_l} = R(\theta, x_t, y_t, x_s, y_s)$, among them θ is the parameter of orientation, x_t, y_t are parameters of translation and x_s, y_s are parameters of scale. 3) Two cost function are defined individually for initial parameter extraction and the final parameter estimation. 4) A direction set algorithm is applied as the optimization algorithm.

As a consequence of the linear relation between a conditional probability and its geometric region, the value of the cost is related to the number of misclassified pixel. On the other word, the global minimum cost corresponds to the prediction with not only the best feature match but also the least misclassified pixel.

5.1 A cost function for the initial parameter extraction

The initial parameter extraction for each object is executed from the labelled image formed by the minimum cost prediction of the object. If we use O'_1 to express the minimum cost prediction of the l th object obtained from object identification, i.e. $O'_1 = P_{O_l}^{N_s}$, a parametric prediction of the l th object is $P_{O_l} = R(\theta, x_t, y_t, x_s, y_s)$, the cost for the prediction is defined as:

$$C[P_{O_l}] \triangleq N(P_{O_l} \cup O'_1) - N(P_{O_l} \cap O'_1) \quad (5.1.1)$$

where $N(A)$ is the number of pixel belonging to the set A . In fact, the cost is the sum of the number of pixel which belong to O'_1 but is not in P_{O_l} and that does not belong to O'_1 but be predicted as P_{O_l} . In short, The cost is the total number of misclassified pixel. The global minimum cost corresponds to the parameters with the geometric prediction matching to the labelled image completely.

5.2 A cost function for object localization

In object localization, a geometric prediction of the scene is formed by the predictions of the objects in the scene and denoted as $P_s = \{P_{O_1}, \dots, P_{O_L}\}$. A cost function is defined as an evaluation of the prediction in terms of feature fitting:

$$C[P_s] \triangleq \sum_{l=1}^L \int_I |p(I|P_{O_l}) - p(I|O_{CT})| dI \quad (5.2.1)$$

among them, $C[P_s]$ is the cost of a geometric prediction of the scene, $p(I|P_{O_l})$ is the radiometric distribution obtained from the geometric prediction of the l th

object P_{O_l} , L is the number of objects in the scene, $p(I|O_{CTl})$ is the radiometric distribution of the crop type of the l th object in the model.

Using the linear relation between radiometric distributions and their geometric regions, the formula (5.2.1) is easily to be rewritten as:

$$C[P_s] = \sum_{l=1}^L \sum_{j=1, j \neq l}^L N(P_{O_l} \cap O_j) \times \frac{1}{N(P_{O_l})} \int_I |p(I|O_{CTl}) - p(I|O_{CTj})| dI \quad (5.2.2)$$

in which O_j is the set of pixel belonging to the j th object, $N(P_{O_l})$ is the number of pixel being predicted as the l th object. $N(P_{O_l} \cap O_j)$ is the number of pixel being predicted as the l th object but, in fact, belonging to the j th object, i.e. the number of misclassified pixel of P_{O_l} .

It is readily to prove that the cost of a scene prediction also meets the requirements stated in section 4.1. Moreover, its only global minimum, theoretically zero, corresponds to the prediction without any misclassified pixel, which supplies the best geometric parameters of objects.

5.3 The optimization process in object localization

As the geometric predictions in object localization, in both the initial parameter extraction and the final parameters estimation, are parametric forms, the cost of a prediction is a function of the parameters of the object(s).

Since the initial parameters are estimated in the labelled image and object localization has good initial parameters, the optimization in this stage is quite easy. A direction set algorithm [1] is applied in both the initial parameter estimation and object localization to got the parameters of objects with minimum cost.

6. Experimental results

In this section, the proposed method will be tested with a radar image of agriculture fields. The radar image in the experiment were acquired from the 1991 NASA/JPL SAR campaign.

The test site was located in southern Flevoland [4]. The land surface is flat. The soils are homogenous over large areas and are classified as fine-textured Calcaric Fluvisol (World Soil Map FAO). Southern Flevoland is an agriculture area with rectangle-shaped fields of ± 80 ha. Farmer have subdivided these fields into smaller parcels and planted a single crop in each parcel. The crops there are mainly sugar-beet, potato and winter wheat. Fig.3 presents an example of it. It is a radar image of agriculture field in Flevopolder, the Netherlands. There are seven land parcels in the scene and they are numbered in the Fig.8. Fig.4 shows the minimum cost predictions of objects acquired in the first stage by using simulated annealing. Fig.5 illustrates the convergence of simulated annealing in the first stage. From Fig.5, it is obviously that there are quite a few local minima in the cost function and the convergence of simulated annealing is not a process of iterative decrease of the cost. Fig.6 is the initial prediction in the second stage formed by the initial parameters of objects extracted from

the minimum cost predictions in the first stage. Fig.7 is the final prediction of the scene in the second stage. The crop type of each land parcel and its geometric parameters are listed below. Comparing Fig.6 with Fig.7, the difference between them is not big. It indicates the first stage does supply a good initial state for the second stage, which is quite close to the final result.

Here list the coefficients for the computation of radiometric distributions in formula (3.1). The number of independent samples is 16. The σ^0 for sugar beet, potatoes, winter wheat and grass are -12.68 dB, -11.04 dB, -15.05 dB and -14.07 dB, respectively. Following are the detailed results in the procedure:

Field No.	$C[P_{oi}']$	Initial parameters	$C[P_{oi}^0]$	Crop type	Geometric parameters
#1	0.4071	(1, 48, 84, 75, 151)	690	grass	(1, 51, 83, 75, 151)
#2	0.3562	(1, 59, 204, 100, 87)	453	potato	(2, 59, 205, 101, 89)
#3	0.4595	(0, 123, 39, 75, 59)	425	potato	(-1, 128, 40, 75, 63)
#4	0.3850	(0, 128, 118, 75, 85)	477	wheat	(2, 129, 117, 80, 87)
#5	0.4288	(0, 140, 206, 53, 83)	674	wheat	(3, 139, 207, 54, 83)
#6	0.5273	(-1, 209, 82, 65, 145)	1174	sugar beet	(-1, 209, 84, 71, 146)
#7	0.4420	(1, 207, 201, 75, 77)	364	potato	(1, 207, 204, 75, 87)

Among them, $C[P_{oi}']$ is the global minimum cost obtained in the first stage. Each of them corresponds to the geometric prediction of the minimum cost shown in Fig.4. The prediction of object with the initial parameters in the second stage is displayed in Fig.6. $C[P_{oi}^0]$ is the cost to extract the initial parameters from P_{oi}' . The final results includes the crop type and geometric parameters of each parcel. The cost of the final prediction is 3.965.

The experiment with a SAR image of agricultural fields indicates the proposed method works well in the practical case. More experimental results about its insensitivities to the initial state and the signal to noise ratio of the image are reported in other publications.

7. Summary

We have described an approach for the recognition of 2-D objects. In this method, every subproblem in object recognition is formulated as an optimization problems, including object identification, the initial parameter estimation and object localization, and techniques of optimization are utilized to solve these problems. Cost functions are calculated as quantitative evaluations of object prediction in terms of their feature fitting and the recognition process is based on cost minimization. As simulated annealing is applied in object identification and its results are used to supply the initial parameters for object localization, our approach does not suffer from initialization sensitivity. Moreover, the approach is parallelizable and allows a degree of flexibility in cost function selection.

Reference:

- [1] Fang L. and Mulder N.J. "Model Based Object Recognition through Hypothesis and Parameter Matching", Proceeding of IGARSS'93, Vol.1 pp.162-164, 1993
- [2] Frost V.S., Shanmugan K.S., Holtzman J.C. and Stiles J.A. "A Statistical Model for Radar Images of Agricultural Scenes", Proceeding of IGARSS'82, TA-5.4, 1982
- [3] Gemen S. and Gemen D., "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images", IEEE Trans. Patt. Anal. Machine Intell., Vol. PAMI-6, pp.721-741, 1984
- [4] Hoekman D. H. "Radar Remote Sensing Data for Applications in Forestry" PhD. Dissertation, Wageningen Agricultural University, 1990
- [5] Suk M. and Bhandarkar S.M. "Three-Dimensional Object Recognition from Range Images" Springer-Verlag, Tokyo, 1992
- [6] Van Laarhoven P.J.M. and Aarts E.H.L. "Simulated Annealing, Theory and Applications", D. Reidel Publishing Company, 1986,

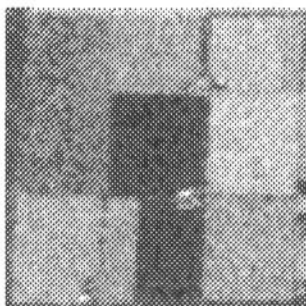


Fig.3 A SAR image

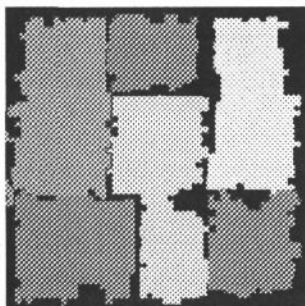


Fig.4 The minimum cost predictions in object identification.

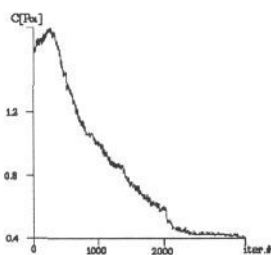


Fig.5 Convergence of simulated annealing

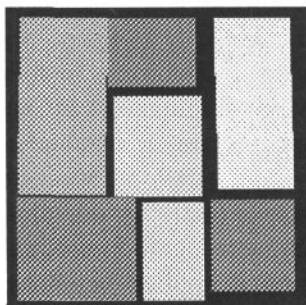


Fig.6 The initial prediction in object localization

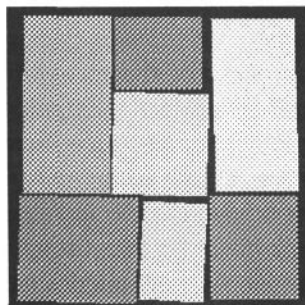


Fig.7 The final prediction in object localization

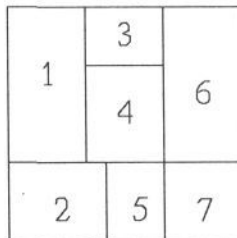


Fig.8 The illustration of field number