Global Journal of Computer Science and Technology
Graphics & Vision
Volume 12 Issue 10 Version 1.0 June 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Uncertainty Analysis for Spatial Image Extractions in the Context of Ontology and Fuzzy C-Means Algorithm

By Md.Sarwar Kamal, Sonia Farhana Nimmy & Linkon Chowdhury
BGC Trust University Bangladesh

Abstract - This paper emphasis on spatial feature extractions and selection techniques adopted in content based image retrieval that uses the visual content of a still image to search for similar images in large scale image databases, according to a user's interest. The content based image retrieval problem is motivated by the need to search the exponentially increasing space of image databases efficiently and effectively. It is also possible to classify the remotely sensed image to represent the specific feature of the target images. In this research we first imposed the Fuzzy C-means algorithm to our sample image and observed its value. After getting the experimental result from Fuzzy C-means we have had designed Ontological Matching algorithm which aftereffect better than the previous one. We have had espy that our Ontological Matching algorithm is twenty (20%) percent better than Fuzzy C-means algorithm.

Keywords : Ontological Matching Algorithm, Fuzzy C-means algorithm, Spatial Feature Extractions, feature selection, Fuzzy Logic.

GJCST-E Classification: I.4.8

UNCERTAINTY ANALYSIS FOR SPATIAL IMAGE EXTRACTIONS IN THE CONTEXT OF ONTOLOGY AND FUZZY C-MEANS ALGORITHM

Strictly as per the compliance and regulations of:

Uncertainty Analysis for Spatial Image Extractions in the Context of Ontology and Fuzzy C-Means Algorithm

Md.Sarwar Kamal [α], Sonia Farhana Nimmy [σ] & Linkon Chowdhury [ρ]

Abstract - This paper emphasis on spatial feature extractions and selection techniques adopted in content based image retrieval that uses the visual content of a still image to search for similar images in large scale image databases, according to a user's interest. The content based image retrieval problem is motivated by the need to search the exponentially increasing space of image databases efficiently and effectively. It is also possible to classify the remotely sensed image to represent the specific feature of the target images. In this research we first imposed the Fuzzy C-means algorithm to our sample image and observed its value. After getting the experimental result from Fuzzy C-means we have had designed Ontological Matching algorithm which aftereffect better than the previous one. We have had espy that our Ontological Matching algorithm is twenty (20%) percent better than Fuzzy C-means algorithm.

Keywords : Ontological Matching Algorithm, Fuzzy C-means algorithm, Spatial Feature Extractions, feature selection, Fuzzy Logic.

I. Introduction

The increase in computing power and electronic storage capacity has lead to an exponential increase of digital content available to users in the form of images which form the bases of many applications [1]. Consequently, the search for the relevant information in the large space of image databases has become more challenging. How to manage appropriate extracted outcome is still difficult problem and it is a proper field to make experiment. A typical image retrieval system includes feature extraction usually in conjunction with feature selection [2]. We can depict any image as a collection of color, texture and shape features. While several image retrieval systems rely on only one feature for the extraction of relevant images, but exact collection of relevant features can yield better retrieval performance [3]. The process of determining the combination of features that is most representative of a particular query image is called feature selection.

In case of analyzing real-world maps, the images shown there may not distinctly identify accurate and comprehensible information; rather lots of knowledge may be embedded in the domain in a hidden and unexplored form.

Author α σ ρ : Lecturer, Computer Science and Engineering, BGC Trust University Bangladesh.
E-mail α : sarwar.saubdcoxbazar@gmail.com
E-mail σ : nimmy_cu@yahoo.com
E-mail ρ : linkon_cse_cu@yahoo.com

a) *Fuzzy Logic*

The logic which works with approximation instead of exact and constant value is called fuzzy logic. The logic has been used from long back to solve various problem domains. The working value of fuzzy logic can be any value in between 0 and 1.Although the fuzzy logic is relatively young theory, the areas of applications are very wide: process control, management and decision making, operations research, economies and, for this paper the most important, pattern recognition and classification. An idea to solve the problem of image classification in fuzzy logic manner as well as comparison of the results of supervised and fuzzy classification was the main motivation of this work.

II. Fuzzy C-Means Algorithm

Fuzzy C-means Algorithm capitalizes image segmentation under consideration of pixels values. It bring the pixels into multiple classes under the value of membership function. Fuzzy C-means Algorithm is formulated as the minimization of the following objective function:

$$J_m(U,\mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} D_{ik}^{2} \qquad (1)$$

Where, $U \in M_{fcn}$, $V = (v_1, v_2, \ldots, v_c)$, $v_i \in R^p$ is the i^{th} prototype $m>1$ is the fuzzifier and

$$D_{ik}^{2} = \left\| \mathbf{x}_i - \mathbf{v}_k \right\|^2$$

The objective is to find that U and V which minimize J_m

The Steps fro FCM Algorithm:

1. Choose: $1 < c < n$, $1 < m < \infty$, ϵ = tolerance, max iteration = N
2. Calculation of membership values as according to the equation (2)

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{D_{ij}}{D_{ik}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad \forall i, j \qquad (2)$$

3. Computer the centroids values according to the equation (3)

$$\mathbf{v}_i = \left(\frac{\sum_{k=1}^{n} u_{ik}^{m} \mathbf{x}_k}{\sum_{k=1}^{n} u_{ik}^{m}} \right) \forall i \qquad (3)$$

4. Selection of new multiplier fields.
5. Repeat the step 2 until the algorithm has converged.

III. Fuzzy Matching

Let us consider the fuzzy matching for the mixing images on the input images [10]. .The degree to which the input target images satisfy the conditions of fuzzy rules and conditions .Suppose IMAGE X is defined by rules R1 and IMAGES Y is defined by rules R2.In this case the matching degree will be represented by as follows:

Matching Degree (IMAGE X,R1)=μ(IMAGE X)

Matching Degree (IMAGE Y,R2) =μ (IMAGE Y)

Where μ is the fuzzy membership function.

The fuzzy matching determines the actual outcome for fuzzy optimization which is accomplished here by fuzzy matrix. Here is a graphical view of fuzzy matching degree for IMAGE Y as follows:

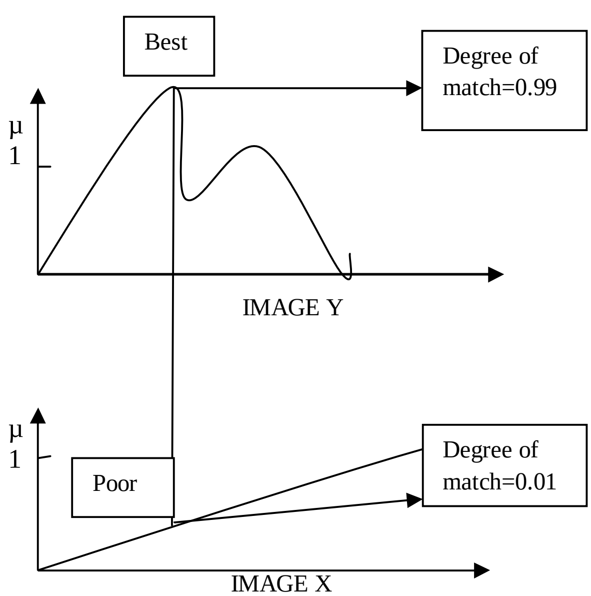


Fig. 1: The matching degree of fuzzy images

IV. Classification Procedure

In our previous work we have done the classification by projecting the maximum classifier without NULL classifier is used. We implied a normal distribution and evaluate the variance and correlation of spectral response during the classification of the unknown pixel.

Here we have had fixed the partitioning as follows:

Let we have a data set $X = \{x_1, x_2 \ldots, x_n\} \subset R^p$ and A classification of X is a $c \times n$ matrix $U = [U_1 U_2 \ldots U_n] = [u_{ik}]$, where U_n denotes the k-th column of U. We have found three classifications efficient and suitable for our research activity. The labeled vectors for these classifications are:

1. $N_{pc} = \{ y \in R^c : y_i \in [0\ 1]\ \forall i,\ y_i > 0\ \exists i\}$ Possibility Label
2. $N_{fc} = \{y \in N_{pc} : \Sigma y_i = 1\}$ Fuzzy Label
3. $N_{hc} = \{y \in N_{fc} : y_i \in \{0,1\}\ \forall i\}$ Hard Label

The Fuzzy classification =

$$M_{fcn} = \{U \in M_{pcn} : \mathbf{U}_k \in N_{fc} \forall k\}$$

V. Ontology and Knowledge Base

According to Ehrig (2007), ontology contains core ontology, logical mappings, a knowledge base, and a lexicon [3]. Core ontology, S, is defined as a tuple of five sets: concepts, concept hierarchy or taxonomy, properties, property hierarchy, and concept to property function.

$$S = (C, \leq c\, R, \sigma, \leq R)$$

where **C** and **R** are two disjoint sets called concepts" and relations" respectively. A relation is also known as a property of a concept. A function represented by σ(r) =< dom(r); ran(r) > where r € R, domain is dom(r) and range is ran(r). A partial order ≤R represents on R, called relation hierarchy, where $r_1 \leq R\ r_2$ iff dom $(r_1) \leq_C$ dom (r_2) and ran $(r_1) \leq_C$ ran (r_2). The notation $\leq_C$ represents a partial order on C, called concept hierarchy or taxonomy". In a taxonomy, if $c_1 <_C c_2$ for c_1; $c_2 €_C$, then c_1 is a sub concept of c_2, and c_2 is a super concept of c_1. If c1 $<_C c_2$ and there is no $c_3 €_C$ with c1 $<_C c_3 <_C c_2$, then c1 is a direct sub concept of c_2, and c_2 is a direct super concept of c_1 denoted by $c_1 \prec c_2$. The core ontology formalizes the intentional aspects of a domain. The extensional aspects are provided by knowledge bases, which contain asserts about instances of the concepts and relations. A knowledge base is a structure KB = (C,R, I, ιC, ,ιR) consisting of

_ two disjoint sets C and R as defined before,

_ a set I whose elements are called instance identifiers (or instance for short),

_ a function C : C→ Ə(I) called concept instantiation,

_ A function { R: R → Ə(I2) with (r) ⊆ ι_C (dom(r)) × ι_C (ran(r)), for all r ε R. The function R is called relation instantiation.

VI. Ontological Instance Matching Algorithom

The operational block of the instance matching integrates ontology alignment, retrieves semantic link clouds of an instance in ontology and measures the terminological and structural similarities to produce matched instance pairs. Pseudo code of the Instance Matching algorithm

Algo. InstanceMatch (ABox ab1, ABox ab2, Alignment A)
for each insi element of ab1
cloudi=makeCloud(insi,ab1)
for each insj element of ab2
cloudj=makeCloud(insj,ab2)
if $\forall$a(c1; c2) elements of A|c1 elements of Block(ins1:type) ^
c2 elements of Block(ins2:type)
if Simstruct(cloudi; cloudj) $\geq \delta$
imatch=imatch$\bigcup$makeAlign(insi; insj)

VII. Experiments With Real-World Data

For the procedures of image classification was used to gather images from "Google Earth" on the Bangladesh region (Chittagong zone).). It uses this as a case study for implementing feature extraction. The collected images contain some common features such as roads, water, field, agriculture, buildings. The features will be separated based on the pixel intensity value selected for the individual features. It has been chosen as an application area because number of spatial features can be extracted from the map images of Forestry complex.

This image contains three channels recorded in three bands: the first band for green, the second for red and the third for blue. In the figure below, we present a fragment of this image and some statistics for the whole image.

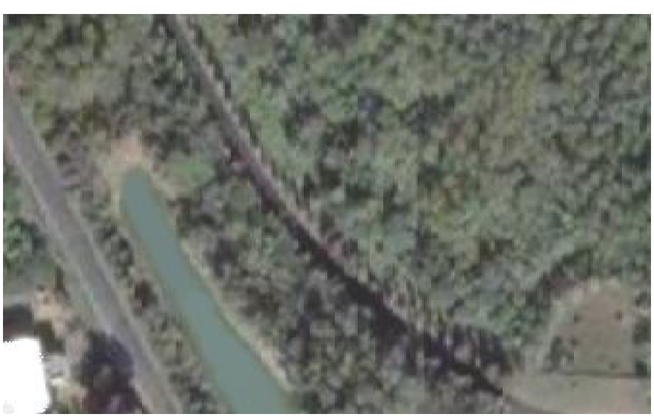
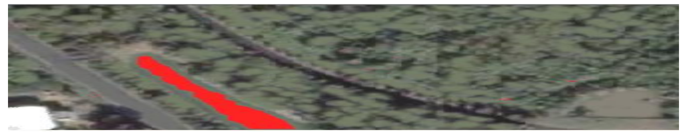


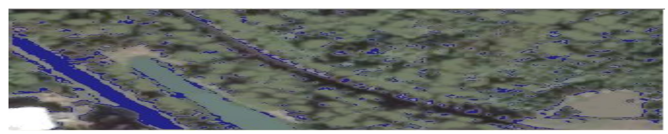
Figure 2 : Forestry Complex Area

After performing thresholding [5] based on color intensities defined for each and every feature, the features are higlighted with individual colors. Therefore, the highlighted feature area is clearly distinguished from the background. The thresholding process finally extract number of spatial features from the particular region such as road, water, field, building and forest.



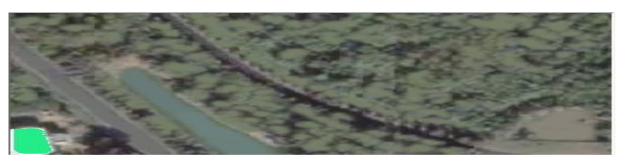
(a)Water

(b)Field



(c) Road

(d) Forest

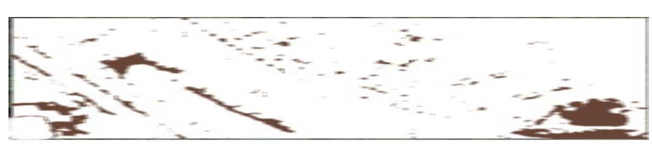


(e) Building

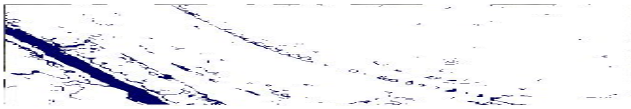
Figure 4: Extracted Features from Forestry Area Image

The extracted features are further threshold for separating them from the background. This has been done by setting the background to all white form, thus displaying the particular feature is.

(a)Water



(b) Field



(c) Road

(d) Forest

(e) Building

Figure 5: Separated Features from Forestry Area Image

VIII. Fuzzy Matrix Optimization

Comparing between two given matrix and finding out the optimum values between them.

Steps To Solve the Problem:

1. Taking values of first matrix as input into the first array from a file for iterative comparisons.
2. Taking values of second matrix as input into the second Array from a file for iterative comparisons.
3. The values of both arrays will be compared than.

IX. Pseudo Code for Optimization Process

```
Fuzzy optimization (x.finput1 [], y.finput2 [])
for the value i,j where i ≠ j
float matrixone[][] = new float[][];
float matrix two[][] = new float[][];
iteration up to the i=n and j=n
{matrix one[k][i]=finput.nextInt();
matrix two[k][j]=finput.nextInt();
(matrixone[m][n]>=matrixtwo[m][n])
System.out.print(" "+matrixone[m][n]);}
```

X. Result Evaluations for Fuzzy Cmeans Classification

One way of the result evaluation was through the accuracy assessment. The classification results are compared to the raw image data and the report is created. This process is done during the random sample selection. The idea of the accuracy assessment is: point is highlighted in the sample list and observation [9] was done where it is located on the image.

The following table shows the mean and standard deviation for the classified classes:

Channel	Mean	Standard
water (from 50 samples)		
Green	73.53	12.32
Red	52.47	9.53
Blue	67.64	14.71
Forest (from 75 samples)		
Green	143.12	22.12
Red	58.77	18.12
Blue	44.12	17.11
Agriculture (from 50 samples)		
Green	122.77	15.50
Red	62.47	13.53
Blue	65.45	17.31
Buildings (from 50 samples)		
Green	52.23	13.21
Red	39.12	8.56
Blue	44.12	10.11
Road (from 75 samples)		
Green	83.35	16.00
Red	29.37	9.12
Blue	41.12	12.19

Creation of the membership functions for the output variables is done in the similar manner. Since this is Sugeno-type inference, constant type of output variable fits the best to the given set of outputs (land classes). When the variables have been named and the membership functions have appropriate shapes and names, everything is ready for writing down the rules.

Class	Output variable
water	1
Forest	2
Agriculture	3
Buildings	4
Roads	5

Based on the descriptions of the input (green, red and blue channels) and output variables (water, agriculture, forest, buildings, and roads), the rule statements can be constructed:

Rules for image classification procedure in verbose format are as follows:

IF (GREEN is a1) AND (RED is a1) AND (NIR is a1) THEN (class is water)

IF (GREEN is a2) AND (RED is a2) AND (NIR is a2) THEN (class is agriculture)

IF (GREEN is a3) AND (RED is a3) AND (NIR is a3)THEN (class is forest)

IF (GREEN is a4) AND (RED is a4) AND (NIR is a4)THEN (class is buildings)

IF (GREEN is a5) AND (RED is a5) AND (NIR is a5)THEN (class is roads)

XI. Result Evaluations for Ontological Classification

Ontological classification is different that fuzzy classification. The idea of the accuracy assessment is: point is highlighted in the sample list and observation was done where it is located on the image.

The following table shows the mean and standard deviation for the classified classes:

Channel	Mean	Standard
water (from 50 samples)		
Green	63.11	22.12
Red	32.01	18.31
Blue	47.58	24.14
Forest (from 75 samples)		
Green	120.54	32.31
Red	43.35	27.02
Blue	33.19	25.33
Agriculture (from 50 samples)		
Green	92.12	19.31
Red	98.58	35.64
Blue	69.11	21.65
Buildings (from 50 samples)		
Green	71.55	25.35
Red	96.25	21.98
Blue	48.56	20.28
Road (from 75 samples)		
Green	89.32	19.22
Red	64.10	5.63
Blue	49.54	18.16

XII. Accuracy Assessments by Fuzzy Cmeans Classification

Idea for accuracy assessment of fuzzy C-Means classification results comes from the manner the maximum likelihood accuracy assessment was performed: select random sample areas with known classes and then let fuzzy logic 'say' what these samples are. With 100 random selected samples, results were as following:

Correctly classified samples: 72
Misclassified: 28
Accuracy: 72%

XIII. Accuracy Assessments by Ontological Classification

Idea for accuracy assessment of ontological classification results comes from the manner the maximum likelihood accuracy assessment was performed: select random sample areas with known classes and then let fuzzy logic 'say' what these samples are. With 100 random selected samples, results were as following:

Correctly classified samples: 92
Misclassified: 08
Accuracy: 92%

The both experiments and observations clearly showed that Fuzzy Logic classification is better than ontological knowledge base classification for Histo for Spatial Feature Extractions.

XIV. Discussion and Conclusion

This paper aimed for extracting the spatial features for providing a fundamental abstraction for modeling the structure of maps representing various raster images. The central part of this paper is an established procedure that is carried out for spatial Historical Heritages classification. As the work continues, it tries to implement every part of the procedure so as to establish its effectiveness and efficiency. It involved the use of supervised learning, assigning membership functions and discovery of pattern feature phases for successfully classifying an image. In the knowledge base, it must be well known whether selected sample forest area or water area.

References Références Referencias

1. Benjamins, V., Contreras, J., Corcho, O. & G_omez-P_erez, A. (2004).Six Challenges for the Semantic Web. AIS SIGSEMIS Bulletin, 1, 24-25.
2. Berners-Lee, T., Fischetti, M. & Dertouzos, M. (1999). Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web.Harper San Francisco.
3. Caracciolo, C., Euzenat, J., Hollink, L., Ichise, R., Isaac, A., Malais_e, V., Meilicke, C., Pane, J., Shvaiko, P., Stuckenschmidt. _Sv_ab-Zamazal, O. & Sv_atek, V. (2008). Results of the Ontology Alignment Evaluation Initiative 2008. Proceedings of Ontology Matching Work-shop of the 7th International Semantic Web Conference, Karlsruhe, Germany, 73{119.
4. A.K Shackelford, and C.H. Davis. A hierarchical fuzzy classification approach for highresolution multispectral data Over urban areas. *IEEE Trans. Geo. Rem. Sen.* 41(9): 1920-1932. 2003.
5. Rob J. Decker. Texture analysis and classification of ERS SAR images for map updating of urban areas

in the Netherlands. *IEEE. Trans. Geo. Rem. Sen.,* 41(9):1950-1958, 2003.
6. B. Solaiman, R. Fiset and F. Cavayas.Automatic road extraction using fuzzy concepts. IGARSS'98, Seattle, USA. 1998.Sunnyvale, CA, 2004.
7. Stefanakis, E.; Sellis, T.- Enhancing a Database Management System for GIS with Fuzzy Set Methodologies, Proceedings of the 19th International Cartographic Conference, Ottawa, Canada, August 1999.
8. M. Törmä. Introduction to neural networks and their use in remote sensing. Photogrammetric Journal of Finland. 13(1): 87-102. 1992.
9. Liu, H., Li. J., Chapman, M. A.;- Automated Road Extraction from Satellite Imagery Using Hybrid Genetic Algorithms and Cluster Analysis, Journal of Environmental Informatics 1 (2) 40-47 (2003)
10. Guesgen, H.;Hertzberg, J.; Lobb, R.;Mantler, A. { Bu_ering Fuzzy Maps in GIS, Spatial Cognition and Computation, 2003
11. W.K. Pratt. *Digital Image Processing*. John Wiley & Sons. NY. 1991
12. Eastman, R { Idrisi 32 Release 2, Guide to GIS and Image Processing, Volume I- II, Clark University, May 2001
13. Benz, U.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. - Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, Journal of Photogrammetry & Remote Sensing, 2004