SUMMARIES OF CERTAIN SPATIAL PATTERNS RETRIEVED FROM MULTIDATE REMOTE-SENSING DATA

HEMA NAIR

This paper presents an approach to describe patterns in remote-sensed images utilising fuzzy logic. The truth of a linguistic proposition such as "Y is *F*" can be determined for each pattern characterised by a tuple in the database, where Y is the pattern and *F* is a summary that applies to that pattern. This proposition is formulated in terms of primary quantitative measures, such as area, length, perimeter, and so forth, of the pattern. Fuzzy descriptions of linguistic summaries help to evaluate the degree to which a summary describes a pattern or object in the database. Techniques, such as clustering and genetic algorithms, are used to mine images. Image mining is a relatively new area of research. It is used to extract patterns from multidated satellite images of a geographic area.

1. Introduction

The objective of this paper is to propose an approach for developing linguistic summaries of certain spatial patterns or features in remote-sensed images. A linguistic proposition of the form "Y is *F*" is formulated using primary quantitative measures, such as length, area, perimeter, and so forth, of the pattern. From these measures, it is possible to compute secondary measures, such as circularity ratio and degree of self-affinity exponents of Hurst and Hack (see [7]). It becomes possible then to construct a spatiotemporal model if images over several time periods of the same geographic area are available.

In the past, much research has been focused on data mining or extracting implicit patterns in relational databases [6, 10, 13, 16], but data mining in multimedia environment has met limited success. This is mainly due to the fact that multimedia data is not as structured as relational data [17]. There is also the issue of diverse multimedia types, such as images, sound, video, and so forth. While one method of data mining may find success with one type of multimedia, such as images, the same method may not be well suited to many other types of multimedia due to varying structure and content. Some related work [17] has met success. In [17], the objective is to mine internet-based images and video. The results generated could be a set of characteristic features based on a topic (keyword), a set of association rules which associate data items, a set of comparison characteristics

that contrast different sets of data, or classification of data using keywords. Data mining techniques can be used in image mining [15] to classify, cluster, or associate images. Image mining is a relatively new research area with applications in many domains including space images and geological images. It can be used to extract unusual patterns, such as forest fires, from multidated satellite images of a geographic area.

This paper proposes an approach that utilises fuzzy logic to describe patterns in remote-sensed images. This method aims to extract some feature descriptors, such as area, length, and so forth, of objects in remote-sensed images and store them in a relational table. Data-mining techniques that employ clustering and genetic algorithms (GAs) are then used to develop the most suitable linguistic summary of each object stored in the table. The objective is to generate linguistic summaries of natural patterns, such as land, island, water body, river and so forth, in remote-sensed images. The approach is to use fuzzy logic to match actual image feature descriptors with feature definitions and to evolve the best-suited linguistic summary of the image object/pattern using GAs.

This paper is organised as follows. Section 2 describes the system architecture, Section 3 describes the approach, Section 4 discusses the implementation issues, and Section 5 discusses the conclusions and future work.

2. System architecture

The system architecture is shown in Figure 2.1. The data summarizer is the key component of the system. The input image is analysed and some feature descriptors are extracted by the image analysis component. Feature descriptors are extracted using MATLAB [5] and ENVI [4] which perform the functionality of the image analysis component. These descriptors are stored thereafter in a relational table in the database. The knowledge base uses geographic facts to define feature descriptors in a typical remote-sensed image. It interacts with a built-in library of linguistic labels. As new feature definitions are added into the knowledge base, corresponding linguistic labels are added in the built-in library. Likewise, as we see a need to expand the built-in library, we would add corresponding feature definitions based on geographic facts in the knowledge base. The built-in library also interacts with the summarizer as it supplies the necessary labels to it. The summarizer receives input from the database and the knowledge base. It performs a comparison between actual feature descriptors of the image stored in the database with the feature definitions stored in the knowledge base. After this comparison, the summarizer uses the labels supplied by the library to develop some possible linguistic summaries describing each object. From these summaries, the most suitable one is selected by interaction with the engine (GA). The GA evolves the most suitable solution which has the highest fitness value (Section 4) after several generations. This solution is passed back to the summarizer which translates it into its corresponding linguistic summary. Thus, this system is composed of two subsystems at this stage. The feature descriptor extraction using MATLAB and ENVI is a manual subsystem involving user interaction. After descriptors are extracted and stored in a relational table in the database, the automated subsystem consisting of summarizer, knowledge base, library, and engine evaluates the descriptors and compares them with feature definitions. An optimal linguistic summary of each object is then generated automatically.
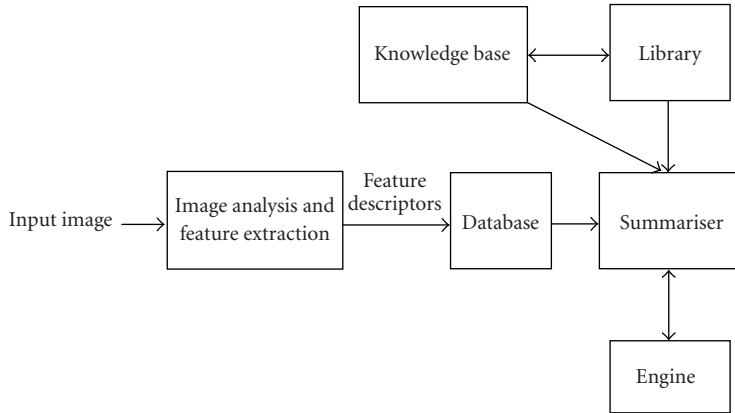
Figure 2.1. System architecture.

3. Approach

The following assumptions are made regarding the data model. R is a relational table defined as

$$R(A_1, A_2, \ldots, A_i, \ldots, A_n), \tag{3.1}$$

$A_1, A_2, \ldots, A_n$ are the attributes in the table R (i.e., the columns of the relational table) and $t_1, t_2, \ldots, t_k$ are the tuples or records or entries in the table R (i.e., the rows of the relational table).

A fuzzy set is the most natural representation of a linguistic variable. A linguistic variable is one whose value is not a number but a word or a sentence in a natural language [9]. As we are concerned with generating linguistic summaries of objects, we will define some fuzzy sets that represent our notion of what the object description or summary should look like.

The general form of a linguistically quantified proposition is "*QY's are F*", where Q is a fuzzy linguistic quantifier, Y is a class of objects, and *F* is a summary that applies to that class. *F* is defined as a fuzzy set in Y. Q represents a linguistic quantifier that groups objects in the class Y. An object/pattern in the image is characterised by a single tuple in the database, therefore, Q can be ignored in this analysis.

An example of such a linguistically quantified proposition in the domain of remote-sensed images would be as follows: *island is moderately large*.

In the above example, Y is *island* and *F* is *moderately large*. In terms of linguistics, this description is equivalent to: *moderately large island*.

The objects considered are river, expanse of water (other water body which is not river), land, and island. The attributes of the objects that are used to develop their linguistic summaries are

    (1) area,
    (2) length,

(3) location in image,

(4) additional information.

Area, length, and location (x- and y-coordinates in image) are extracted automatically by image analysis component in Figure 2.1. For river, the most significant feature descriptor that is extracted is its length. For land, island, and expanse of water, the most significant feature descriptor extracted is area.

If

$$y = y_1, y_2, \ldots, y_p, \tag{3.2}$$

then

$$\text{truth}(y_i \text{ is } F) = \mu_F(y_i) : i = 1, 2, \ldots, p, \tag{3.3}$$

where $\mu_F(y_i)$ is the degree of membership of y_i in the fuzzy set F and $0 \le \mu_F(y_i) \le 1$. The higher the degree of membership, the higher the truth value of the linguistic proposition. In this case, referring to (3.2) and (3.3), y_i could be *island*, *area of land*, *expanse of water*, or *river*. *Area of land* represents land other than island, *expanse of water* represents any water body that is not a river. For each object y_i, the degree of membership of its feature descriptor, such as area or length in corresponding fuzzy sets, is calculated. Fuzzy sets for area are *large, considerably large, moderately large, fairly large*, and *small* and fuzzy sets for length are *long, considerably long, relatively long, fairly long*, and *short*.

The linguistic description is calculated as follows:

$$T_j = m_{1j} \wedge m_{2j} \wedge m_{3j} \wedge \cdots \wedge m_{nj}, \tag{3.4}$$

where m_{ij} is the matching degree [6] of the ith attribute in the jth tuple. $m_{ij} \in [0,1]$ is a measure of degree of membership of the ith attribute value in a fuzzy set denoted by a fuzzy label. Referring to (3.4), T_j thus evaluates the truth value for each object y_i, as it matches the feature descriptors of that object with fuzzy set definitions by calculating the matching degrees and combining them together using logical AND operator. The logical AND ($\wedge$) of matching degrees is calculated as the minimum of the matching degrees [6].

$$T = \sum_{j=1}^{k} T_j, \quad \forall m_{ij} \ne 0. \tag{3.5}$$

Equation (3.5) means that we calculate the conjunction of only those matching degrees that are nonzero in order to evaluate T_j. This aids in computational efficiency. All such T_j's are added up to evaluate T. T is a numeric value that represents the truth of a possible set of summaries of all the objects in the database.

As an example, consider the area as the attribute in the single-attribute case. Possible fuzzy labels are *large, considerably large, moderately large, fairly large*, and *small*. If $k = 4$ (there are 4 objects in the table),

$$T_j = m_{1j}, \tag{3.6}$$

where m_{1j} represents the fuzzy membership value of area in fuzzy sets *large, considerably large, moderately large, fairly large*, or *small*,

$$T = \sum_{j=1}^{4} T_j, \quad \forall m_{1j} \neq 0. \tag{3.7}$$

Equation (3.7) means that each object has a nonzero fuzzy membership value for its area in any of the fuzzy sets mentioned above; that membership value is added cumulatively to the membership value calculated similarly for the next object in the table. Thus T is evaluated as the sum of membership values (nonzero) of the area of all 4 objects in the table. The next section discusses how the GA evolves the most suitable linguistic summary for all the objects by maximising T.

4. Implementation issues

This section explains the GA approach and then discusses the results from applying this approach to mining images.

4.1. GA approach.   A GA emulates biological evolutionary theories as it attempts to solve optimisation problems. The GA comprises a set of individual elements (the population) and a set of biologically inspired operators, such as selection, crossover, and mutation. According to evolutionary theories, only the most suited elements in a population are likely to survive and generate offspring, thus transmitting their biological heredity to new generations. In computing terms, a GA maps a problem onto a set of binary strings (the population), each representing a potential solution. Using selection, crossover, and mutation operators, the GA then manipulates the most promising strings (denoted by their high-fitness value from the evaluation function), as it searches for the best solution to the problem [2, 3, 14].

Given n attributes, each having m possible fuzzy labels, it is possible to generate $m^n + 1$ descriptions. The GA searches for an optimal solution among these descriptions. Each of these summaries is represented by a uniquely coded chromosome string (a string of 0's and 1's). The population of such strings is manipulated (using selection, crossover, and mutation operators [14]) and evaluated by the GA, and the most suitable linguistic summary that fits each object is generated. The evaluation function for the linguistic summaries or descriptions of all objects in the table is

$$f = \max(T), \tag{4.1}$$

where T in (4.1) is evaluated as shown in the previous section and f is the maximum fitness value of a particular set of linguistic summaries that have evolved over several generations of the GA.

4.2. Results.   In general, image objects are classified at the highest level into land and water. Land is further classified into island and other land. Water is further classified into river (characterised by its length) and other water body (characterised by area). Some of the fuzzy sets being considered are

(1) for land: *large, considerably large, moderately large, fairly large*, and *small*, based on degree of membership of area of the land in the respective fuzzy sets,

(2) for other water body or expanse of water (except river): *large, considerably large, moderately large, fairly large*, and *small*, based on degree of membership of area of the water body in the respective fuzzy sets,

(3) for river: *long, considerably long, relatively long, fairly long*, and *short*, based on degree of membership of length of the river in the respective fuzzy sets.

These fuzzy sets are defined based on geographic facts, such as

(i) largest continent is Asia with area of $44579000 \, \text{km}^2$,

(ii) largest freshwater lake is Lake Superior with area of $82103 \, \text{km}^2$,

(iii) smallest continent is Australia/Oceania with area of $7687000 \, \text{km}^2$,

(iv) longest river is the Nile with length $6669 \, \text{km}$,

(v) shortest river is the Roe with length $0.037 \, \text{km}$.

The fuzzy set for *large expanse of water* [11, 12] is defined in (4.2) referring to Figure 4.1(a) and the geographic fact describing the size of the largest freshwater lake, where $x_1 = 79900 \, \text{km}^2$, $x_2 = 82103 \, \text{km}^2$,

$$\mu_{large\ expanse\ of\ water}(x) = \begin{cases} 1, & 82103 \leq x, \\ \dfrac{x}{2203 - 36.27}, & 79900 \leq x < 82103, \\ 0, & x < 79900. \end{cases} \tag{4.2}$$

The fuzzy set for *considerably large expanse of water* is defined in (4.3) referring to Figure 4.2, where $x_1 = 28034.33 \, \text{km}^2$, $x_2 = 55068.66 \, \text{km}^2$, $x_3 = 82103 \, \text{km}^2$,

$$\mu_{considerably\ large\ expanse\ of\ water}(x)$$
$$= \begin{cases} 1 - \dfrac{(55068.66 - x)}{27034.33}, & 28034.33 \leq x \leq 55068.66, \\ 1 - \dfrac{(x - 55068.66)}{27034.33}, & 55068.66 < x \leq 82103, \\ 0, & x < 28034.33, \\ 0, & x > 82103. \end{cases} \tag{4.3}$$

The fuzzy set for *moderately large expanse of water* is defined in (4.4) referring to Figure 4.2, where $x_1 = 1000 \, \text{km}^2$, $x_2 = 28034.33 \, \text{km}^2$, $x_3 = 55068.66 \, \text{km}^2$,

$$\mu_{moderately\ large\ expanse\ of\ water}(x)$$
$$= \begin{cases} 1 - \dfrac{(28034.33 - x)}{27034.33}, & 1000 \leq x \leq 28034.33, \\ 1 - \dfrac{(x - 28034.33)}{27034.33}, & 28034.33 < x \leq 55068.66, \\ 0, & x < 1000, \\ 0, & x > 55068.66. \end{cases} \tag{4.4}$$
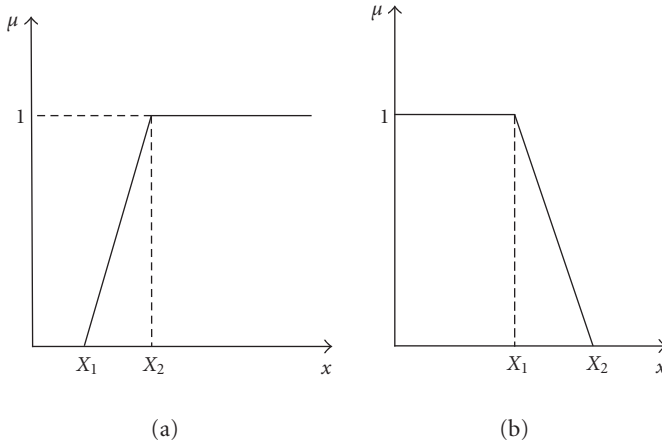
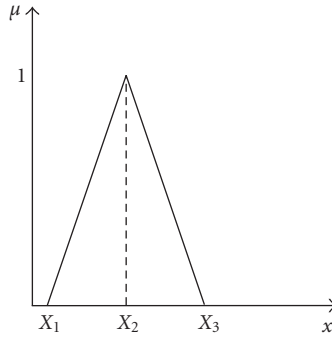Figure 4.1. (a) Fuzzy set for large. (b) Fuzzy set for small or short.

Figure 4.2. Fuzzy set for considerably large, moderately large, or fairly large.

The fuzzy set for *fairly large expanse of water* is defined in (4.5) referring to Figure 4.2, where $x_1 = 100 \, \text{km}^2$, $x_2 = 1000 \, \text{km}^2$, $x_3 = 28034.33 \, \text{km}^2$,

$$\mu_{\text{fairly large expanse of water}}(x) = \begin{cases} 1 - \dfrac{(1000 - x)}{900}, & 100 \leq x \leq 1000, \\ 1 - \dfrac{(x - 1000)}{27034.33}, & 1000 < x \leq 28034.33, \\ 0, & x < 100, \\ 0, & x > 28034.33. \end{cases} \qquad (4.5)$$

The fuzzy set for *small expanse of water* is defined in (4.6) referring to Figure 4.1(b):

$$\mu_{small\ expanse\ of\ water}(x) = \begin{cases} 1, & 0 < x \leq 600, \\ \dfrac{-x}{400} + 2.5, & 600 < x \leq 1000, \\ 0, & \text{otherwise.} \end{cases} \quad (4.6)$$

The set for *small area of land* is defined in (4.7) referring to Figure 4.1(b):

$$\mu_{small\ area\ of\ land}(x) = \begin{cases} 1, & 0 < x \leq 7687000, \\ \dfrac{-x}{313000} + 25.56, & 7687000 < x \leq 8000000, \\ 0, & \text{otherwise.} \end{cases} \quad (4.7)$$

4.2.1. An example. Consider the data in Table 4.1. For the second tuple which is an expanse of water with approximate area $6.683\,\mathrm{km^2}$, the possible fuzzy labels are *large expanse of water, considerably large expanse of water, moderately large expanse of water, fairly large expanse of water*, or *small expanse of water*. The truth value (fuzzy membership value) for each of these cases is evaluated as shown below by substituting $x = 6.683$ in (4.2), (4.3), (4.4), (4.5), and (4.6), respectively:

(i) $\mu_{large\ expanse\ of\ water}(6.683) = 0,$

(ii) $\mu_{considerably\ large\ expanse\ of\ water}(6.683) = 0,$

(iii) $\mu_{moderately\ large\ expanse\ of\ water}(6.683) = 0,$

(iv) $\mu_{fairly\ large\ expanse\ of\ water}(6.683) = 0,$

(v) $\mu_{small\ expanse\ of\ water}(6.683) = 1.$

Thus this pattern can be described as a *small expanse of water*.

Similarly, for the other object in Table 4.1, the truth values are evaluated. The fuzzy label with highest truth value is selected to form the most suitable linguistic summary for the corresponding object/pattern.

An example pair of Spot multispectral images to be analysed is shown in Figures 4.3 and 4.4. These are subimages of the original images, used here due to file-size limitations. The geographic coordinates of the original images are approximately $3°17'$U-$3°48'$U latitude and $100°58'$T-$101°38'$T longitude referring to the topographic map. The scale of the images is approximately $1 : 0.0003764$. This means that 1 pixel square represents $0.0003764\,\mathrm{km^2}$. Table 4.1 shows a small sample data set of feature descriptors from some of the objects in the image (Figure 4.3). Figure 4.4 shows an image of the same geographic area taken on a later date. Table 4.2 shows a small sample data set of feature descriptors from some of the objects in the image (Figure 4.4). Area is in $\mathrm{km^2}$ and length is in km. Additional information attribute denotes numbers as follows: $0 =$ river, $1 =$ other water body, $2 =$ island, $3 =$ land, and $4 =$ fire. Location indicates x- and y-coordinates of centroid of object. $X, Y = 0$ indicates the remaining part of image as location. The grey level values are from the R band as this band shows all the patterns clearly. For objects where area is considered as the most significant parameter in calculations, their length is set to 0. Fire is considered as a separate pattern.

The objective of this paper is to describe patterns/objects, such as river, land, island, expanse of water, and so forth, quantitatively in terms of measures, such as area or length.

Table 4.1. Feature descriptors from Figure 4.3. Location is in pixel coordinates.

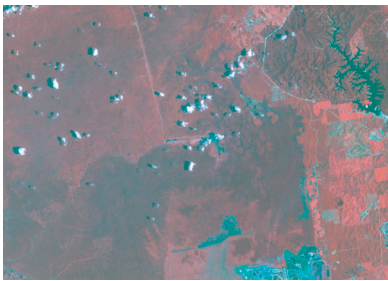| Grey-level value (R band)	Approximate area km^2	Location in image		Additional information
		X	Y	
150	3300.84	1606	1457	3
0	6.683	1546	1132	1

Figure 4.3. Image of area in peninsular Malaysia acquired on March 6, 1998. Scale is approximately 1 : 0.0003764.
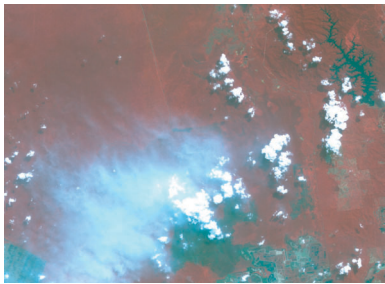

Figure 4.4. Image of the same area in peninsular Malaysia acquired on July 10, 2001. Scale is approximately 1 : 0.0003764.

The additional information attribute of each object in the table is calculated using a clustering technique that implements the k-means algorithm [8]. Tables 4.3 and 4.4 list the data recorded from the images in Figures 4.3 and 4.4 for this purpose. The feature vector in this case consists of X, Y, R, G, B values of several points on the patterns/objects and their corresponding envelopes in the images. Additionally, a pattern/object is classified using the following rules.

(1) If a pattern is to be classified as an island, it should have a water envelope surrounding it such that it has a uniform band ratio of at least eight points on this envelope (corresponding to directions E, W, N, S, NE, NW, SE, SW). Also, grey-level values on the envelope could be lower than the grey-level values on the pattern.

(2) If a pattern does not have an envelope in all directions as described in the first rule above, then it is classified as land.

Table 4.2. Feature descriptors from Figure 4.4. Location is in pixel coordinates.

Grey-level value (R band)	Approximate area km²	Location in image		Additional information
		X	Y	
150	2874.38	1899	1150	3
27	6.683	2506	976	1
166	—	1550	1587	4

Figure 4.5. Indications of the smoke plume prominently in purple and the associated burnt scar in dark green in the green slice range (32 to 63) or blue slice range (64 to 95).

(3) If a pattern is to be classified as water body (or other water body or expanse of water), it is necessary that it should have a uniform band ratio.

The above rules hold for multiband images.

Comparing the images in Figures 4.3 and 4.4, it is noted that there are a few major changes in the spatial patterns. A fire is indicated prominently in Figure 4.4 on the left side of the image. Fire is considered as a separate pattern characterised by its bluish white smoke plume and burnt scar area close to it. These patterns become clearly visible if density slicing is performed on the image in the R-band. The colour image of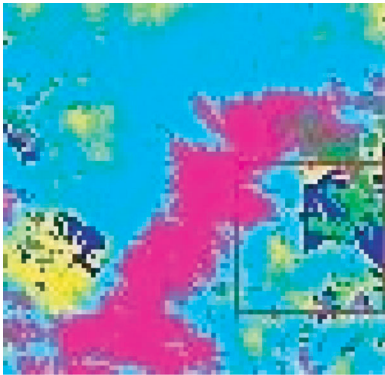 Figure 4.5 indicates the smoke plume prominently in purple and the associated burnt scar in dark green in the green slice range (32 to 63) or blue slice range (64 to 95). Histograms of the area near the fire (Figures 4.6 and 4.7) corresponding to Figures 4.3 and 4.4 also indicate that most of the pixels are of lower intensity for the burnt scar area from the fire image (Figure 4.4) when compared with the same area without fire (Figure 4.3).

Fire is identified in Table 4.2 with additional information attribute equal to 4. Future work (Section 5) will focus on developing rules for classifying river.

The spatial location attribute in Tables 4.1 and 4.2 is given a linguistic value, such as centre, left, top, left, and so forth, using the following calculation. *Centre span* is a variable defined to denote a circular distance around the x- and y-coordinates of the centre of an image. The value of centre span may vary from image to image as it is subjective. It is a number that is obtained by measuring the distance around the centre of the image, which can be used to denote an area that still represents the centre of the overall image.
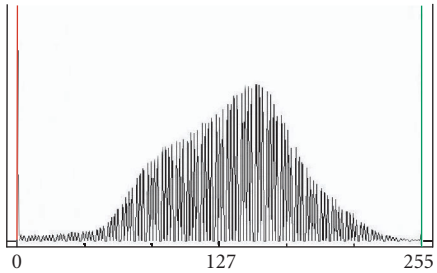
0   127   255

Figure 4.6. Histogram of area from Figure 4.3 where fire is later detected.
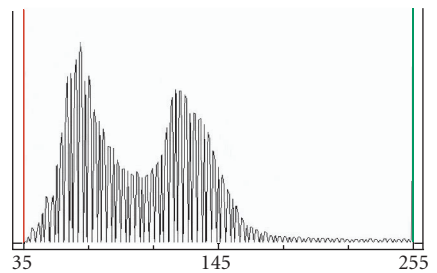

35   145   255

Figure 4.7. Histogram of same area from Figure 4.4 near fire and burnt scar.

This value is evaluated by user interaction with the image. All objects, whose centroids [1] lie within the range of centre span from the centre of the image, are still located at the centre of the image. If the difference between x- and y-coordinates of the centroid of the object and the centre of the image is greater than centre span, then the object is located at lower right (diagonally from image centre). If the reverse is true, then the object is located at top left (diagonally from image centre). If the difference between the x-coordinate of the object and the x-coordinate of image centre is greater than centre span and the difference between the y-coordinate of image centre and the y-coordinate of centroid of the object is greater than centre span, then the object is located at the top right of the image. Similar calculations are used to evaluate the locations lower left, right, left, top, and bottom of image. An x- and y-coordinate of 0, 0 evaluates the location as *remainder of image*, because the actual coordinates of the image have an origin greater than 0, 0.

It is to be noted that patterns, such as urban area settlements, are ignored as trivial in this analysis. The main concerns are natural patterns, such as water bodies, land, island, and so forth. Additionally, patterns that signal calamities, such as fire, are also extracted and described. The linguistic summaries are generated with reference to the scale of land and water defined in the geographic facts from which the fuzzy sets are developed, even though the area of land in the images may be large compared to the expanse of water.

The GA is run with the following input parameter set. These parameter values are set after several trial runs. With other values, the GA produces the summary of only one object/pattern in the table:

Table 4.3. Data recorded from Figure 4.3 for clustering.

Object		Envelope		Object			Envelope		
X	Y	X	Y	R	G	B	R	G	B
1721	1549	114	2917	146	103	94	0	127	175
1723	1521	165	2933	135	115	113	0	127	175
1677	1482	109	2889	137	120	122	0	123	162
1636	1511	237	2973	226	103	105	0	127	175
1633	1540	300	2908	255	99	107	0	123	175
1645	1574	105	2884	155	103	105	0	118	167
1680	1595	31	2905	153	106	102	0	123	158
1727	1572	61	2969	155	120	116	0	125	169
1504	1090	1497	1104	0	94	85	130	108	109
1533	1086	1528	1061	0	96	85	221	125	125
1450	1039	1413	1017	0	94	91	151	106	109
1606	1210	1567	1221	0	89	78	169	115	118
1553	1187	1567	1211	0	94	82	112	111	109
1592	1136	1593	1120	0	89	85	92	115	109
1611	1198	1658	1226	0	91	78	96	99	94
1574	1119	1581	1103	0	89	78	103	96	96

Table 4.4. Data recorded from Figure 4.4 for clustering.

Object		Envelope		Object			Envelope		
X	Y	X	Y	R	G	B	R	G	B
1913	1159	131	1662	123	92	89	5	100	118
1906	1146	218	1781	126	92	87	5	103	121
1837	800	315	1944	145	86	77	6	105	123
1624	900	538	2177	135	87	79	6	98	118
1550	1037	723	2476	127	86	87	9	113	134
1924	1962	1059	2775	88	145	128	22	130	167
2312	1600	1363	2910	135	92	87	22	124	155
1917	1165	1510	2954	127	95	92	20	121	144
2572	963	2596	946	36	82	71	102	81	66
2518	898	2539	887	41	81	69	116	76	68
2496	887	2492	870	44	81	71	185	82	76
2402	792	2390	897	38	76	61	110	79	66
2417	984	2383	787	38	86	76	126	76	63
2501	1054	2505	1078	43	90	90	134	94	89
2551	1132	2546	1154	40	95	89	126	92	84
2597	1046	2614	1043	36	84	77	113	84	77

(1) number of bits in a chromosome string of the population = 10,
(2) generations per cycle = 26,
(3) population size = 200 strings,
(4) probability of cross-over = 0.53,
(5) probability of mutation = 0.001.

After 208 generations, the linguistic summaries generated from the image in Figure 4.3 (no fire) are

 (i) a small area of land at the centre,
(ii) a small expanse of water at the top right.

The GA input parameters are varied to obtain the linguistic summaries of patterns in Table 4.2. The parameters used are

(1) number of bits in a chromosome string of the population = 10,
(2) generations per cycle = 10,
(3) population size = 200 strings,
(4) probability of cross-over = 0.53,
(5) probability of mutation = 0.001.

After 80 generations, the linguistic summaries generated from the image in Figure 4.4 are

  (i) bluish white smoke indicating fire at the left,
 (ii) a small expanse of water at the top right,
(iii) a small area of land at the centre.

Thus, comparing the results of the GA after mining the images of the same geographic area without fire and with fire taken on two dates separated by a period of more than three years, we can see that the GA can correctly describe an unusual pattern, such as the fire indicated in the image in Figure 4.4. Referring to the corresponding topographic map, it is possible to conclude that this fire could be the result of burning in a paddy field or a nearby primary forest.

Thus, with two attributes such as length and area, each having five possible fuzzy labels, it is possible to generate $5^2 + 1$ descriptions. The GA has searched for an optimal solution among these descriptions within a very short time.

5. Conclusions and future work

This paper has presented a new approach to describing images using linguistic summaries that use fuzzy labels. A GA technique has been employed to evolve the most suitable linguistic summary that describes each object/pattern in the database. This method can be extended to an array of images of the same geographic area, taken over a period of several years, to describe many other interesting and unusual patterns that emerge over time. Some directions for future work include the following.

(1) Development of a user friendly tool with graphical interface to ease the task of extracting and calculating feature descriptors, such as area, length, grey-level intensity, colour, and so forth, stored in the tables. Currently, both MATLAB and ENVI are required in order to populate the tables. Each has its own limitations.

(2) Development of rules to identify and classify objects/patterns, such as river, correctly.

References

[1]   B. P. Buckles, F. E. Petry, D. Prabhu, and M. Lybanon, *Mesoscale feature labeling from satellite images*, Genetic Algorithms for Pattern Recognition (S. K. Pal and P. P. Wang, eds.), CRC Press, Florida, 1996, pp. 167–175.

[2]   J. L. R. Filho, P. C. Treleaven, and C. Alippi, *Genetic-algorithm programming environments*, IEEE Computer **27** (1994), no. 6, 28–43.

[3]   E. D. Goodman, *An introduction to GALOPPS—the genetic algorithm optimized for portability and parallelism system (Release 3.2)*, Tech. Report 96-07-01, Genetic Algorithms Research and Applications Group, Michigan State University, Michigan, 1996.

[4]   Research Systems Inc, *ENVI Version 3.0 User's Guide*, Better Solutions Consulting, 1997.

[5]   The Mathworks Inc, *MATLAB Image Processing Toolbox User's Guide*, 1997.

[6]   J. Kacprzyk and A. Ziolkowski, *Database queries with fuzzy linguistic quantifiers*, IEEE Trans. Systems Man Cybernet. **3** (1986), 474–479.

[7]   B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, California, 1982.

[8]   P. M. Mather, *Computer Processing of Remotely-Sensed Images: An Introduction*, John Wiley & Sons, New York, 1999.

[9]   J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice Hall, New Jersey, 2001.

[10] A. Motro, *Intensional answers to database queries*, IEEE Trans. Knowl. Data Eng. **6** (1994), no. 3, 444–454.

[11] H. Nair, *Developing linguistic summaries of patterns from mined images*, Proceedings of 5th International Conference on Advances in Pattern Recognition (ICAPR'03), Allied Publisher, Calcutta, 2003, pp. 261–267.

[12] H. Nair and I. Chai, *Linguistic description of patterns from mined images*, 6th International Conference on Enterprise Information Systems (ICEIS '04), vol. 2, Universidade Portucalense, Portugal, 2004, pp. 77–83.

[13] H. Nair, R. George, R. Srikanth, and N. A. Warsi, *Intensional answering in databases—a fuzzy logic approach*, Proceedings of the Joint Conference on Information Systems, North Carolina, 1994.

[14] R. E. Smith, D. E. Goldberg, and J. A. Earickson, *SGA-C: A C-language implementation of a simple genetic algorithm*, TCGA Report 91002, The University of Alabama, Alabama, 1991.

[15] B. Thuraisingham, *Managing and Mining Multimedia Databases*, CRC Press, Florida, 2001.

[16] R. R. Yager, *On linguistic summaries of data*, Knowledge Discovery in Databases (G. Piatetsky-Shapiro and W. J. Frawley, eds.), AAAI/MIT Press, California, 1991, pp. 347–363.

[17] O. R. Zaïane, J. Han, Z. Li, and J. Hou, *Mining multimedia data*, Proc. CASCON '98, Meeting of Minds, Toronto, 1998.

Hema Nair: Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia

E-mail address: hema.nair@mmu.edu.my

Special Issue on
Intelligent Computational Methods for Financial Engineering

Call for Papers

As a multidisciplinary field, financial engineering is becoming increasingly important in today's economic and financial world, especially in areas such as portfolio management, asset valuation and prediction, fraud detection, and credit risk management. For example, in a credit risk context, the recently approved Basel II guidelines advise financial institutions to build comprehensible credit risk models in order to optimize their capital allocation policy. Computational methods are being intensively studied and applied to improve the quality of the financial decisions that need to be made. Until now, computational methods and models are central to the analysis of economic and financial decisions.

However, more and more researchers have found that the financial environment is not ruled by mathematical distributions or statistical models. In such situations, some attempts have also been made to develop financial engineering models using intelligent computing approaches. For example, an artificial neural network (ANN) is a nonparametric estimation technique which does not make any distributional assumptions regarding the underlying asset. Instead, ANN approach develops a model using sets of unknown parameters and lets the optimization routine seek the best fitting parameters to obtain the desired results. The main aim of this special issue is not to merely illustrate the superior performance of a new intelligent computational method, but also to demonstrate how it can be used effectively in a financial engineering environment to improve and facilitate financial decision making. In this sense, the submissions should especially address how the results of estimated computational models (e.g., ANN, support vector machines, evolutionary algorithm, and fuzzy models) can be used to develop intelligent, easy-to-use, and/or comprehensible computational systems (e.g., decision support systems, agent-based system, and web-based systems)

This special issue will include (but not be limited to) the following topics:

- **Computational methods**: artificial intelligence, neural networks, evolutionary algorithms, fuzzy inference, hybrid learning, ensemble learning, cooperative learning, multiagent learning

- **Application fields**: asset valuation and prediction, asset allocation and portfolio selection, bankruptcy prediction, fraud detection, credit risk management

- **Implementation aspects**: decision support systems, expert systems, information systems, intelligent agents, web service, monitoring, deployment, implementation

Authors should follow the Journal of Applied Mathematics and Decision Sciences manuscript format described at the journal site http://www.hindawi.com/journals/jamds/. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at http://mts.hindawi.com/, according to the following timetable:

Manuscript Due	December 1, 2008
First Round of Reviews	March 1, 2009
Publication Date	June 1, 2009

Guest Editors

Lean Yu, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; yulean@amss.ac.cn

Shouyang Wang, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; sywang@amss.ac.cn

K. K. Lai, Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; mskklai@cityu.edu.hk