Fuzzy Logic for Change Detection in Classified Images

Peter Deer and Peter Eklund
School of Information Technology, Griffith University
PMB 50, Gold Coast MC 9726, Queensland, Australia
p.deer@gu.edu.au, p.eklund@gu.edu.au

Abstract. The ability to detect changes in digital image sequences is an important application in remote sensing. A common approach to change detection is to independently classify each image, and then compare their corresponding class labels. This approach is known as Post Classification Comparison (PCC). Unfortunately, it exhibits poor accuracy. In this paper, we describe Fuzzy PCC, whereby we use fuzzy classification techniques, and then undertake the Post Classification Comparison using fuzzy logic operators. The choice of such operators is discussed, and results of Fuzzy PCC shown for an environmental monitoring application.

1 Introduction

Change detection in digital imagery is an important function in image processing. It has application in numerous areas, including environmental monitoring, surveillance, and detection of serious ailments in the medical domain. Although the term 'change detection' is commonly used in the literature, we are normally interested in more than the simple detection of changes: we also commonly seek where they occurred, their (spatial) extent, and what was the precise nature of change (often in terms of classes of interest).

A simple and intuitively attractive approach to the detection of changes in imagery is therefore to independently classify each image, and then compare their class labels. This approach is known as Post Classification Comparison (Singh, 1989). It will not only detect that changes have occurred, and where, but will also identify the precise nature of the changes. Unfortunately, traditional classifiers commonly suffer quite poor accuracy. This problem compounds in Post Classification Comparison (PCC) change detection (Pilon *et al.*, 1988; Quarmby and Cushnie, 1989). Pixels in which no change has occurred, but were classified incorrectly on the first date and not on the second, and vice versa, will show up as false change detections.

Traditional classification approaches seek to allocate each observation (e.g. pixel in an image) to one of a (hopefully exhaustive) set of mutually exclusive classes (or, more accurately, to associate a single class label with each observation). Unfortunately, classes commonly do not have distinct 'boundaries', and even human experts in the field will often disagree on class assignment e.g. when distinguishing between 'stressed' and 'not-stressed' vegetation, or between pasture and thinly wooded areas. It is argued that crisp class assignment is often inappropriate in the geographical and remote sensing sciences (see e.g. Burrough, 1989; Key *et al.*, 1989; Wang, 1990a,b; Fisher and Pathirana, 1990; Foody, 1992; van der Meer, 1995). Moreover, a single pixel in a remotely sensed image may contain a number of classes (i.e. it may be a 'mixed pixel', or 'mixel').

2 Fuzzy Sets and Image Processing

Fuzzy set theory (Zadeh, 1965) was introduced to address the issue of class or set 'vagueness'. Using fuzzy set theory, we can determine and reason with the grade of membership of a particular pixel in a number of classes. This provides an approach to (partial) class assignment in regions where there is a gradual transition from one class to another. It can also be used as an approach to the mixed pixel problem (Deer and Eklund, 2000).

The application of fuzzy set theory to pattern recognition and image processing finds considerable support in the literature. Zadeh states (in the Foreword to Bezdek and Pal, 1992), "the initial development of the theory of fuzzy sets was motivated in large measure by problems in pattern classification and cluster analysis". Bezdek and Pal (1992, p. 8) observe that "feature vectors (and the objects they represent) can and should be allowed to have degrees of membership in more than one class'.

The traditional approach to PCC (i.e. with 'hard' classification) is to simply compare class labels for the two (or more) separate dates. If they are different, it is concluded that a change has occurred. It makes no particular sense to undertake the comparison using arithmetic operations: even if the class labels are numbers, they are essentially arbitrary, and the same end result could derive from a number of situations. For example, an arithmetic difference of 3 could result from class 6 on t_1 (date 1) and class 3 on t_2 (date 2), or class 4 on t_1 and class 1 on t_2, and so on. An arithmetic operation on the class labels will therefore detect change, but will lose, unnecessarily, information about the nature of that change.

Traditional PCC is therefore more appropriately posed a problems in boolean logic. We seek those pixels in which the class label on t_2 is not the same as the class label on t_1. Alternatively, we might seek those pixels in which the class label on t_1 is, say, class 7, and the class label on t_2 is, say, class 4. If we use the latter approach, not only is change detected (and located), but the nature of the change is also known (i.e. the change is identified).

This approach is illustrated in Fig. 1. Independent classifications were undertaken of a region on Kangaroo Island, South Australia, using Landsat Thematic Mapper (TM) images taken in 1989 and 1993, and a Mahalanobis Distance classifier (Deer, 1998). Fig. 1 shows, in grey, pixels that were assessed to be forest in 1989 AND pasture in 1993. Pixels that do not satisfy this particular condition are shown in white. Errors attributable to misregistration between the two images have been reduced using a 7×7 diamond shaped morphological opening filter (Deer, 1998).

Fig. 1. Traditional Post Classification Comparison, Forest 1989 AND Pasture 1993, After Morphological Filtering.

The approach to fuzzy classification adopted in this paper is based on the fuzzy c-means clustering algorithm (Bezdek *et al.*, 1984). This is an unsupervised fuzzy clustering, or classification, of the data. The fuzzy set memberships are determined by a distance measure (usually Euclidean, diagonal or Mahalanobis) from the class centroids, and the optimal clustering is determined by minimising an error function, subject to certain constraints. In this manner, the similarity (as measured by a distance metric) between an instance (i.e. the observation) and a prototypical class member (i.e. the class centroid) is used to determine the class memberships.

The unsupervised fuzzy c-means approach was used in Key *et al.* (1989) to classify clouds. The authors speculated that the number of classes could be specified and the class means supplied "in a manner analogous to using training sites to provide spectral statistics for a supervised classification", thereby greatly decreasing the program execution time. That is, it could be adapted to operate in a supervised manner. In this approach, the fuzzy set memberships would still be determined by a distance measure (in their case, diagonal) from the class centroids, but the class centroids were supplied, rather than being determined by the algorithm. This general approach has been used in Foody (1992) and Deer *et al.* (1996).

Training classes are selected, and used to determine the (squared) Mahalanobis Distance between each pixel and the mean of each class,

$$d_{ik}^2 = (\mathbf{x}_k - \mathbf{m}_i)^t \mathbf{\Sigma}_i^{-1} (\mathbf{x}_k - \mathbf{m}_i) \tag{1}$$

where d_{ik}^2 represents the (squared) Mahalanobis Distance between pixel $\mathbf{x}_k$ and the mean of the ith class, $\mathbf{m}_i$; superscript t denotes (vector) transpose; and $\mathbf{\Sigma}_i^{-1}$ is the inverse of the covariance matrix of the ith class.

The fuzzy class membership of each pixel in the image in each class is calculated according to

$$\mu_{ik} = \frac{(\frac{1}{d_{ik}^2})^p}{\sum_{j=1}^{c} (\frac{1}{d_{jk}^2})^p} \tag{2}$$

where

$$i = 1, \ldots, c \quad \text{and} \quad k = 1, \ldots, n$$

and where μ_{ik} is the fuzzy class membership of pixel x_k in class i, c is the number of classes, n is the number of pixels in the image, and p is a parameter used to control the 'degree of fuzziness' of the class membership allocation, related to m (sometimes denoted q) in earlier work on the fuzzy c-means clustering algorithm (e.g. Bezdek *et al.*, 1984) by

$$p = \frac{1}{m - 1} \tag{3}$$

Application of this approach to the two Landsat TM images (1989 and 1993) results in two fuzzy classified images. To undertake change detection with fuzzy classified images, we do not have single class labels to compare as we do in the traditional PCC approach. Instead, we have the degree of membership of each pixel, in each of the classes of interest. We wish to compare the corresponding class memberships to determine to what extent we believe that change has occurred.

Because the fuzzy set/class memberships are numbers, we can compare them, and determine the extent of change that has occurred, by simple arithmetic operations, on a "per pixel per class" basis. This is closely analogous to the traditional change detection methods known as Image Differencing or Image Ratioing (see e.g. Singh, 1989). It is, however, more appealing and useful to develop a logical approach to Fuzzy PCC.

Fisher and Pathirana (1993) report on the use of fuzzy classification in change detection, using an approach that might be described as the application of Boolean logic to ranked fuzzy class memberships. Change class labels were assigned for conditions:

1. the highest ranked cover class at time t_1 (i.e. the cover class with highest membership at time t_1) was the same as the highest ranked cover class at time t_2;
2. the highest ranked cover class at time t_1 was the second ranked cover class at time t_2;
3. the highest ranked cover class at time t_1 was the third ranked cover class at time t_2; and so on.

This is, as stated, essentially a Boolean logic approach: conjunctions of antecedent conditions were tested, in a nested IF THEN ELSE structure, to assign a change code. A particular code was assigned if the first test returned TRUE. Otherwise, some other test of conditions was applied. Again, some code was assigned if the test returned TRUE. This continued until all the antecedent conditions of interest (the sets of antecedent conditions were non-exhaustive) had been tested (except, as implied, that the testing stopped as soon as one test returned TRUE). If all tests returned FALSE, some other code was assigned. It is noted that exactly one change code was applied to each pixel. Although a number of tests may have returned TRUE, the tests were performed in a particular order, and as soon as one test returned TRUE, a code was assigned and the testing stopped.

In this approach, the actual fuzzy membership values were not used; only their rankings. To illustrate, consider a hypothetical pixel with class memberships of (0.95,0.05,0,0) at time t_1, and (0.05,0.95,0,0) at time t_2. Because the second ranked class on the second date is the same as the highest ranked class on the first date, this approach would determine that the likelihood that this pixel had changed between the two dates would be assessed as "not very likely". This result is not completely satisfactory.

This approach also loses information about what class a pixel was or is. For example, a pixel whose highest membership at time t_1 was forest, and whose second highest membership at time t_2 was also forest, would be assigned to the same change class as a pixel whose highest membership at time t_1 was sea, and whose second highest membership at time t_2 was also sea. This is so irrespective of what the highest membership class was on the second date, in either case.

Therefore, although this is a valid and useful approach for some applications, it provides no information about the change classes. The approximate measure of the 'likelihood' of change that is provided is also based only upon the rankings, not on the actual memberships involved, and we can hypothesise cases where this may produce unsatisfactory results. It is, as previously noted, essentially a classical Boolean approach to reasoning with fuzzy membership values. Despite the noted weaknesses (which, of course, may not be of concern in some applications), Fisher and Pathirana (1993) is important work in that it demonstrates an approach to identifying change through logical queries on fuzzy classifications. However, we seek a more general and powerful approach to logical querying with fuzzy classifications. Fortunately, this is available through the rich body of literature on fuzzy logic.

As noted in Dubois and Prade (1985), "the idea is to write down intuitively reasonable axioms which translate into functional equations, and then solve these equations in order to provide a mathematical representation of the class of connectives". This will give us the means of aggregating or combining information about membership in fuzzy sets, using 'fusion' or 'aggregation' operators such as conjunction (AND) and disjunction (OR). This idea can be followed in Dubois and Prade (1985), or Yager (1991), and leads to the *t-norm* and *t-conorm* operators. Parameterised families of operators satisfying these operators are given in Dubois and Prade (1985), Bonissone and Decker (1986), and Klir and Folger (1988).

To illustrate, the better known operators are:

t-norm

$$\begin{aligned} T_0(a,b) &= MIN(a,b) \text{ if } MAX(a,b) = 1 \\ &= 0 \text{ otherwise} \end{aligned} \tag{4}$$

$$T_1(a,b) = MAX(0, a+b-1) \tag{5}$$

$$T_2(a,b) = a \times b \tag{6}$$

$$T_3(a,b) = MIN(a,b) \tag{7}$$

t-conorm

$$\begin{aligned} S_0(a,b) &= MAX(a,b) \text{ if } MIN(a,b) = 0 \\ &= 1 \text{ otherwise} \end{aligned} \tag{8}$$

$$S_1(a,b) = MIN(1, a+b) \tag{9}$$

$$S_2(a,b) = a + b - (a \times b) \tag{10}$$

$$S_3(a,b) = MAX(a,b) \tag{11}$$

There are both theoretical and empirical studies that compare these operators (Tong and Shapiro, 1985; Whalen and Schott 1985; Bonissone and Decker, 1986; and Cross and Sudcamp, 1991). They are ordered:

$$T_0 \leq T_1 \leq T_2 \leq T_3 \qquad S_3 \leq S_2 \leq S_1 \leq S_0 \tag{12}$$

with T_0, T_3 forming the greatest lower and least upper bounds of all t-norms respectively, and S_0, S_3 forming the least upper and greatest lower bounds of all t-conorms respectively (Bonissone and Decker, 1986).

It is interesting to note that, with the inclusion of just one additional desirable characteristic, idempotence, all choice about fusion operators is removed. Idempotence ensures that if we have certain evidence or support for something, and we receive further evidence of the same value, then we neither increase nor decrease our belief in it. That is,

$$T(a,a) = a \text{ and } S(a,a) = a \tag{13}$$

If we impose this requirement, the only conjunction and disjunction operators satisfying the desirable characteristics are MIN and MAX respectively. The same effect is achieved if we alternatively introduce the requirement for mutual distributivity between T and S (Dubois and Prade, 1985). Mutual distributivity requires that

$$T(a, S(b,c)) = S(T(a,b), T(a,c)) \text{ (and } S(a, T(b,c) = T(S(a,b), S(a,c))) \tag{14}$$

To illustrate the use of the MIN and MAX operators (for conjunction and disjunction) for combining fussy image class memberships,

> from "*feature1* AND *feature2* AND .... AND *featureN*" (where *featureN* represents the fuzzy membership in the Nth class), we would compute MIN(*feature1*, *feature2*, ... , *featureN)*,
>
> and from "*feature1* OR *feature2* OR .... OR *featureN*", we would compute MAX(*feature1*, *feature2*, ... , *featureN*).

If we wish to pose a query of the nature "*feature1* AND *feature2* BUT NOT *feature3*", we need to introduce a negation operator.

$N : [0,1] \to [0,1]$ is called the negation operator:

$$\begin{array}{lll} \text{a.} & N(1) = 0; & N(0) = 1; \\ \text{b.} & N(N(a)) = a & \text{— involution;} \\ \text{c.} & \text{if } a \geq b \text{ then } N(a) \leq N(b). & \end{array} \tag{15}$$

These axioms do not uniquely determine a negation operator (Bellman and Giertz, 1973) and, again, there are parameterised families of operators that satisfy them e.g,

$$N_\lambda(a) = (1-a)/(1+\lambda a),\ \lambda > -1 \text{ (Sugeno, 1977), and} \tag{16}$$

$$N_w(a) = (1-a^w)^{1/w},\ w > 0 \tag{17}$$

(Klir and Folger, 1988, term this latter family the Yager class of fuzzy complements).

It is common to select $N(a) = 1 - a$ as the negation operator. This is a special case of both the Sugeno and Yager complements. If either of two additional (sometimes questioned) axioms (see Gaines, 1976, and Bellman and Giertz, 1973, for details) are added to the basic set of negation axioms, all choice is again removed: $N(a) = 1 - a$ becomes the only possible negation operator.

For negation, we would therefore compute NOT *feature3* = 1 - *feature3*. The example "*feature1* AND *feature2*, but NOT *feature3*" would be computed as MIN(*feature1*, *feature2*, 1 - *feature3*).

Discussion A number of authors offer a treatment of the relationship between t-norms, t-conorms and negation operators under a generalised De Morgan's Law (which requires that $N(T(a,b)) = S(N(a), N(b))$ and $N(S(a,b)) = T(N(a), N(b))$), and note that the three cannot be independently defined (see Bonissone and Decker, 1986; Klir and Folger, 1988; or Yager, 1991).

Zadeh's original paper on fuzzy sets (Zadeh, 1965) proposed MIN, MAX and $1 - a$ for fuzzy set intersection, union and complementation respectively. Bonissone and Decker (1986) observed that these were a common selection for most expert systems. Klir and Folger (1988) identify a highly desirable feature of these operators: their inherent prevention of error compounding. This feature is highly desirable for Fuzzy PCC (recall that error compounding was noted to be a problem in classical PCC), and is lacking in most other candidate fuzzy set operators. The literature shows that MIN, MAX and $1-a$ remain valid choices, from an axiomatic viewpoint, for aggregation operators, but it is noted that they may not be a unique choice.

The arguments in support of their use are solid (in particular, the issue of error compounding), but by no means overwhelming. We could, under certain conditions, argue for the use of $S_1(a,b) = MIN(1, a+b)$ for disjunction. If we consider the fuzzy classification of a particular pixel on a particular date, the unity sum condition for memberships means that the membership in class a OR class b would reduce to the sum of the memberships in the two classes ($S_1(a,b) = MIN(1, a+b) = a+b$, because $a + b \leq 1$). Given our approach to determining fuzzy memberships, this is intuitively quite reasonable, even desirable, as computing disjunction over all possible classes of a particular pixel would return a membership of unity. S_1 would thus appear to be superior to computing the membership in two or more classes as the maximum of the memberships in the separate classes.

Unfortunately, we wish to be able to construct and simplify logical queries involving conjunction, disjunction and negation, in accordance with the conventions of logic. For change detection, these queries will span memberships on two dates. As noted, we cannot define conjunction, disjunction and negation separately: conjunction and disjunction must be "De Morgan duals".

The dual of $S_1(a,b) = MIN(1, a+b)$ is $T_1(a,b) = MAX(0, a+b-1)$. Because $a + b \leq 1$ (for any two classes for a particular pixel on a particular date), $T_1(a,b)$ will be identically equal to 0. That is, any conjunction of two classes (same pixel, same date) would have zero membership. While this is satisfactory in traditional classification, it does not accord well with the philosophy of fuzzy classification, in which a particular pixel can have a (non-zero) degree of membership in two or more classes simultaneously, and any conjunction of two (non-zero) class memberships should be non-zero.

Arguments can also be formulated for the use of other possible operators. These arguments are largely based on an intuitive appeal to have disjunction return a value higher than the maximum of the input values e.g. to have $S(0.8, 0.9)$ return a higher value than $S(0.1, 0.9)$. Palubinskas *et al.* (1995) used this line of reasoning in choosing to investigate the use of Yager functions for combining the outputs of fuzzy classifiers, with an aim of increased classifier accuracy.

The Yager function for disjunction and conjunction are, respectively

$$u_w(a,b) = MIN\left[1, (a^w + b^w)^{1/w}\right] \tag{18}$$

$$i_w(a,b) = 1 - MIN\left[1, (\{1-a\}^w + \{1-b\}^w)^{1/w}\right] \tag{19}$$

(the Yager function for negation was given earlier as Eqn. 17)

Klir and Folger (1988) show that, in the limit as $w \to \infty$, $u_w(a,b) \to MAX(a,b)$ and $i_w(a,b) \to MIN(a,b)$. It can also be shown that, for $w = 1$, the Yager functions reduce to $u_1(a,b) = S_1(a,b) = MIN(1, a+b)$, $i_1(a,b) = T_1(a,b) = MAX(0, a+b-1)$.

Not withstanding the intuitive appeal of these functions, Palubinskas *et al.* (1995) found that MIN and MAX performed comparatively with the Yager functions with $w = 2$. No other values of w were tested, but we note that there would seem little point in investigating values of $w > 2$, as they would return values increasing close to MIN and MAX.

Further, as stated earlier, the penalty for choosing some intuitively appealing disjunction operator (other than MAX) is that we should accept an obligation to choose the De Morgan dual as the conjunction operator, and this will return a value less than the minimum of the two values.

The above discussion on choosing aggregation operators has been based upon values being numeric. In some cases, it may be more appropriate to select (non-continuous) values from an ordered linguistic set. In this case, Yager (1979) has shown that MIN and MAX are the only acceptable operators for conjunction and disjunction respectively. The selection of a negation operator may, however, be problematic.

Choice of Aggregation Operators for Change Detection It has been shown that general class of the aggregation operator is determined by the axioms that it is required to satisfy. We have followed the literature in choosing a reasonable set of axioms for such operators. We have noted that the operator may not be uniquely specified by the chosen axioms. We have, however, chosen to use MIN, MAX and $1 - a$ for this current work: they are sound, easily computed, and do not compound errors. As far as is known, there has been no previous use of fuzzy logic operators for change detection to influence this decision.

2.1 Fuzzy Post Classification Comparison

We now have the machinery to compute any set of compound conditions, in a multi-valued fuzzy logic, that we can pose in a logical form. The statement "pixel x is in class A at time t_1" is, in such a fuzzy logic, TRUE, to the extent $\mu_A(x, t_1)$. We can therefore compute the extent to which the statement "pixel x is in class A at time t_1 AND in class B at time t_2" is TRUE: $MIN(\mu_A(x, t_1), (\mu_B(x, t_2))$. This is the membership of the pixel x in the specific change class (A, B). Because, as has been noted, MIN forms the least upper bound on the t-norm conjunction operators, it will provide the highest reasonable (i.e. in accordance with our stated desirable axioms) estimate of membership in such a change class.

Similarly, we can compute any other change class of interest. For example, "pixel x is in class A OR in class B at time t_1, AND in class C at time t_2" and so on (including the ability to compute negations, such as "AND NOT in class D"). We are thereby able to aggregate classes into superclasses (using disjunction), and to thus allow (to some extent) for the presence of unwanted classes such as cloud and shadow.

This approach lets us concentrate on particular change classes or superclasses, whereas:

1. an arithmetic approach such as differencing or ratioing only lets us consider whether a pixel has become more or less like a particular class.
2. the approach of Fisher and Pathirana (1993) aggregates change classes and deals only with rankings, and thereby loses valuable information on change class, and 'likelihood' of change.

We now use these results to compute changes between our two (fuzzy classified) Kangaroo Island images.

We can reason with the extent to which pixels have changed from one particular class to another. To illustrate, "to what extent has change occurred from forest to pasture?" This is shown in Fig. 2. A grayscale morphological filter has been applied to this image to remove error effects of isolated 'misclassified' pixels and misregistration between the images.

The detection of a conversion in land use from forest to pasture was a major objective in this study. Regions that were not clearly defined, or not shown at all in the traditional PCC image (Fig. 1) are clearly apparent in the fuzzy PCC image (Fig. 2). The large regions above the centre of the figures are

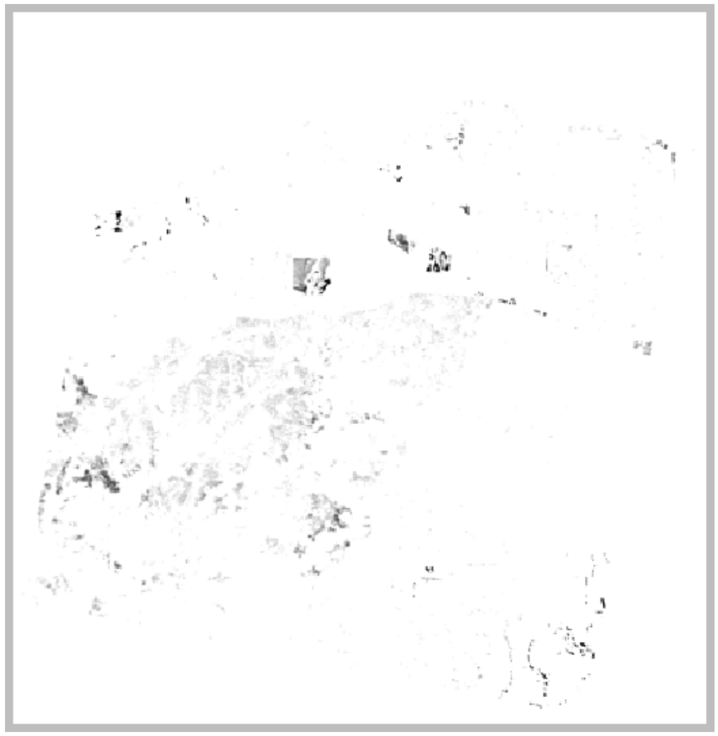

Fig. 2. Fuzzy Post Classification Comparison, Forest 1989 AND Pasture 1993, $p = 1.25$, after Morphological Filtering.

of particular interest. The straight edges of the boundaries of these regions draws attention to them, strongly suggesting human-induced change. The regions at the far left centre, although displaying similar memberships in the change class, do not exhibit straight boundaries, and are more suggestive of natural cause agents for change. In fact, records reveal that a major fire occurred in this area between the two study dates (but no detailed mapping of the fire damaged region was undertaken).

We can also pose queries of the form "was some particular class, but is now NOT that class" i.e. the extent to which membership in a particular class has diminished. This is illustrated in Fig. 3, which shows areas that have become 'less like' forest i.e. the extent to which pixels were forest in 1989, but not forest in 1993.

An important feature of this logical approach to fuzzy change detection is that we can pose queries in a manner to focus on particular classes and changes, and reduce or remove error due to unwanted classes such as cloud or shadow. Let us suppose that a pixel was forest at time t_1, and obscured by cloud at time t_2. The simple difference of the class memberships at these two times will show a significantly diminished membership of forest, thereby

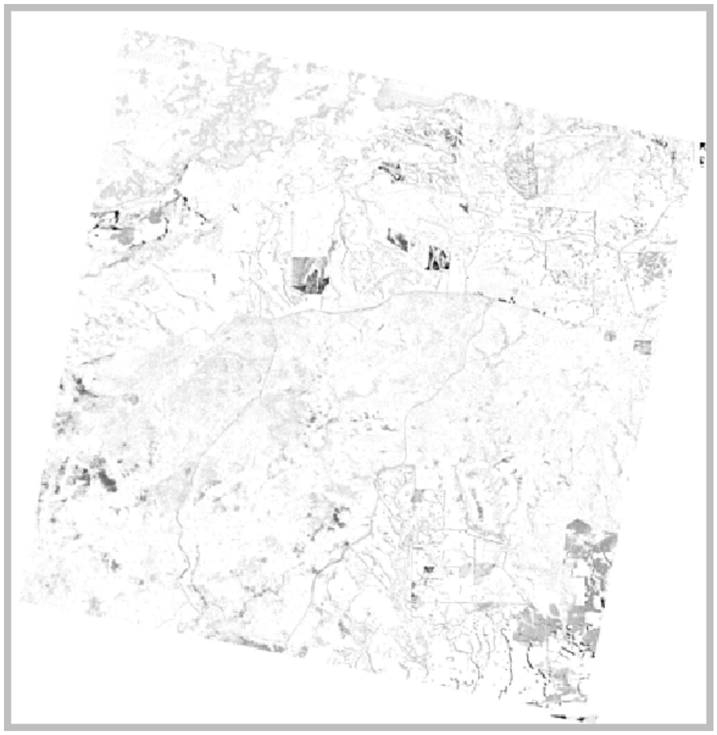

Fig. 3. Fuzzy Post Classification Comparison, Forest 1989 AND NOT Forest 1993, $p = 1.25$.

leading to some measure of (potentially erroneous) belief that change has occurred. The logical query "forest at time t_1 AND (forest OR cloud) at time t_2" would not cause this 'error'. Furthermore, the results can now be interpreted as memberships of a meaningful fuzzy set (i.e. the change class "less like (some particular class)").

3 Conclusion

This paper posed traditional PCC change detection as a problem in Boolean logic, and then showed that a similar approach could be taken to PCC with fuzzy classifications. The choice of aggregation operators for fuzzy PCC was discussed. Some results were shown, for particular choices of fuzzy classifier (and parameters) and aggregation operators. It is suggested that this approach is less prone to the compounding error problem experienced by traditional approaches, and reveals more subtle changes. Change detection problems posed in this manner provide a general, powerful and intuitive approach to formulating and computing any query relating to fuzzy class memberships on two (or more) dates. It is noted that some or all of the classified images do not even strictly need to be fuzzy, and that some or all of the inputs do not even strictly need to be images: they could be any other form of geospatial data, fuzzy or not.

References

Bellman R. and M. Giertz (1973). On the Analytic Formalism of the Theory of Fuzzy Sets. *Information Sciences*, 5, pp. 149-156.

Bezdek J.C., R. Ehrlich and W. Full (1984). FCM: The Fuzzy *c*-Means Clustering Algorithm. *Computers and Geoscience*, 10, pp. 191 - 203.

Bezdek J.C. and S.K. Pal (eds.) (1992). *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data.* IEEE Press.

Bonissone P.P. and K.S. Decker (1986). Selecting Uncertainty Calculi and Granularity: An Experiment in Trading-Off Precision and Complexity. *Uncertainty in Artificial Intelligence*, Kanal L.N. and J.F. Lemmer (eds.), North-Holland, pp. 217-247.

Burrough P.A. (1989). Fuzzy Mathematical Methods for Soil Survey and Land Evaluation. *Journal of Soil Science*, 40, pp. 477-492.

Cross V. and T. Sudcamp (1991). An Empirical Study of Fuzzy Compatibility Measures and Aggregation Operators. *SPIE Vol. 1607*, pp. 415-428.

Deer P.J., P.W. Eklund and B. Norman (1996). A Mahalanobis Distance Fuzzy Classifier. 4th Australian and New Zealand Intelligent Information Systems Conference (ANZIIS 96), Adelaide, Australia, pp. 220-223.

Deer P.J. (1998). A Quantitative Comparison of Traditional Classification Approaches. 9th Australasian Remote Sensing Conference, Sydney, Australia.

Deer P.J. and P.W. Eklund (2000). On Parameter Values For a Mahalanobis Distance Fuzzy Classifier. 8th Int. Conf. on Information Processing and the Management of Uncertainty (IPMU). Madrid, Spain.

Dubois D. and H. Prade (1985). A Review of Fuzzy Set Aggregation Connectives. *Information Sciences*, 36, pp. 85-121.

Fisher P.F. and S. Pathirana (1990). The Evaluation of Fuzzy Membership of Land Cover Classes in the Suburban Zone. *Remote Sensing of Environment*, 34, pp. 121-132.

Fisher P.F. and S. Pathirana (1993). The Ordering of Fuzzy Land-Cover Information Derived from Landsat MSS Data. *Geocarto International*, 3, pp. 5-14.

Foody G.M. (1992). A Fuzzy Sets Approach to the Representation of Vegetation Continua from Remotely Sensed Data: An Example from Lowland Heath. *Photogrammetric Engineering and Remote Sensing*, 58, pp. 221 - 225.

Gaines B.R. (1976). Foundations of Fuzzy Reasoning. *International Journal of Man-Machine Studies*, 8, pp. 623-668.

Key J.R., J.A. Maslanik and R.G. Barry (1989). Cloud Classification From Satellite Data Using a Fuzzy Sets Algorithm: A Polar Example. *International Journal of Remote Sensing*, 10, pp. 1823-1842.

Klir G.J. and T.A Folger (1988). *Fuzzy Sets, Uncertainty, and Information.* Prentice Hall.

Palubinskas G., R.M. Lucas, G.M. Foody and P.J. Curran (1995). An Evaluation of Fuzzy and Texture-Based Classification Approaches For Mapping Regenerating Tropical Forest Classes From Landsat-TM Data. *International Journal of Remote Sensing*, 16, pp. 747-759.

Pilon P.G., P.J. Howarth, R.A. Bullock and P.O. Adeniyi (1988). An Enhanced Classification Approach to Change Detection in Semi-Arid Environments. *Photogrammetric Engineering and Remote Sensing*, 54, pp. 1709-1716.

Quarmby N.A. and J.L. Cushnie (1989). Monitoring Urban Land Cover Changes at the Urban Fringe from SPOT HRV Imagery in South-East England. *International Journal of Remote Sensing*, 10, pp. 953-963.

Singh A. (1989). Digital Change Detection Techniques Using Remotely-Sensed Data. *International Journal of Remote Sensing*, 10, pp. 989-1003.

Sugeno M. (1977). Fuzzy Measures and Fuzzy Integrals: A Survey. *Fuzzy Automata and Decision Processes*, Gupta M.M., G.N. Sardis and B.R.Gaines (eds.), North-Holland, Amsterdam, pp. 89-102.

Tong R.M. and D.G. Shapiro (1985). Experimental Investigations of Uncertainty in a Rule-based System for Information Retrieval. *International Journal of Man-Machine Studies*, 22, pp. 265-282.

Van der Meer F. (1995). Spectral Unmixing of Landsat Thematic Mapper Data. *International Journal of Remote Sensing*, 16, pp. 3189-3194.

Wang F. (1990a). Fuzzy Supervised Classification of Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing*, 28, pp. 194-201.

Wang F. (1990b). Improving Remote Sensing Image Analysis Through Fuzzy Information Representation. *Photogrammetric Engineering and Remote Sensing*, 56, pp. 1163-1169.

Whalen T. and B. Schott (1985). Alternative Logics for Approximate Reasoning in Expert Systems: A Comparative Study. *International Journal of Man-Machine Studies*, 22, pp. 327-346.

Yager R.R. (1979). A Measurement Informational Discussion on Fuzzy Union and Intersection. *International Journal of Man-Machine Studies*, 11, pp. 189-200.

Yager R.R. (1991). Connectives and Quantifiers in Fuzzy Sets. *Fuzzy Sets and Systems*, 40, pp. 39-76.

Zadeh L.A. (1965). Fuzzy Sets. *Information and Control*, 8, pp. 338-353.