CHANGE DETECTION USING PRINCIPAL COMPONENT ANALYSIS AND FUZZY SET THEORY

by

P. GONG

RÉSUMÉ

Cet article présente deux nouvelles méthodes permettant de faire un meilleur usage de l'information multibande obtenue à partir de données de télédétection pour la détection des changements. Au lieu d'effectuer une analyse des composantes principales sur une combinaison d'images multibandes initiales, celle-ci a été effectuée sur des données résultant de différences d'images. Ainsi, la plupart des informations relatives aux changements ont été conservées sur les premières composantes principales. Des opérations fondées sur la théorie des ensembles flous ont été proposées en vue de combiner les informations sur les changements provenant des différents canaux en une seule image. Les zones ayant subi des changements peuvent alors être extraites de cette dernière image. Des images de Kitchener-Waterloo, en Ontario, acquises par le capteur thématique de Landsat pendant deux années successives sont utilisées pour illustrer ces méthodes. Un certain nombre de stratégies sur l'utilisation des opérations fondées sur la théorie des ensembles flous sont également présentées.

SUMMARY

In this paper two procedures that were developed to make better use of multispectral information from remotely sensed data for change detection are discussed. Instead of applying principal component analysis (PCA) to a combined data set of original multispectral images, PCA was applied to difference images. Thus, most change information was preserved in the first few principal component images. Operations based on fuzzy set theory were proposed to combine change information from different image channels into a single-image channel. Changed areas could then be extracted from this single image. Landsat Thematic Mapper (TM) images acquired in two successive years over Kitchener-Waterloo, Ontario, are used to illustrate these methods. Some strategies on the use of fuzzy set operations are discussed.

INTRODUCTION

In remote sensing, changes can be determined by comparing the spectral response differences at the same spatial location among a set of two or more multispectral images acquired at different times. These images are first spatially registered. A commonly used change detection procedure then follows in which changes are identified via thresholding a difference image that has been obtained by subtracting one band of image on one date from the same band of image on another date. However, it is usually not possible to detect changes occurring in a region using only one spectral band because different types of changes may be captured in different bands. Therefore, changes have to be enhanced and extracted from multispectral imagery.

Two types of procedures for using multispectral data in change component enhancement are commonly used. The first type involves simple image arithmetic between images of the same spectral band for two dates. For convenience, we define the two images obtained for two different dates of the same spectral band as band-pair images. Image arithmetic includes ratioing and differencing band-pair images. Images so generated are referred to as change component images. To locate and identify change automatically, change component images are thresholded or classified (Jensen, 1986; Fung and LeDrew, 1988; Pilon et al., 1988; Singh, 1989). To detect changes visually, ratioed or differenced band-pair images can be analyzed based on their colour displays on a video monitor or photographic products (Howarth and Wickware, 1981; Howarth and Boassan, 1983). Although logically straightforward due to increased data redundancy and display difficulties, this type of procedure becomes inefficient to use when the image dimension (that is, the number of spectral bands) exceeds three.

To overcome these difficulties, a second type of change component enhancement procedure employs image transformation
methods. Transformation methods include vegetation indexing (VI) (Tucker, 1979), tasseled cap analysis (KT-Transform) (Kauth and Thomas, 1976), change vector analysis (CVA) (Malila, 1980), and principal component analysis (PCA) (Lodwick, 1979; Byrne et al., 1980; Ingebritsen et al., 1985; Fung and LeDrew, 1987). By image transformation, the change information recorded in original multispectral data can be preserved in a relatively small number of components. Rather than serving as general enhancement tools, the VI and KT-Transform methods were developed specifically for such purposes as enhancing the vegetation or the soil component. While CVA has the potential of summarizing various types of change components and their magnitudes into separate image channels, it has rarely been applied since its introduction. Among these transformation techniques, the PCA method has been most commonly used. Researchers using PCA for change detection have reported that the minor component images are likely to contain most of the change information when multispectral images that have been obtained on two dates are applied as an integrated data set. The amount of change information contained in each principal component image, however, may vary from image to image. It may not be an easy task to determine which principal component image to work with. In addition, the use of PCA in such a manner is subject to the condition that the areas of changes have to be a small proportion of the entire study area (Richards, 1984; Fung and LeDrew, 1987).

It would be desirable to deal with only one image channel and to extract most, if not all, change information from this channel of image. In this paper, a method is presented for transforming change information into one image channel from images of different spectral bands. The main objectives are:

- to introduce an alternative method on the use of PCA for change component enhancement with which change information is guaranteed to be preserved in the major component images regardless the proportion of changed area in a study area; and
- to demonstrate the effectiveness of fuzzy set theory in combining change information from different image channels into a single-image channel.

METHOD

The change detection procedure proposed in this study can be divided into six steps:

- spatially register images from two different dates;
- undertake a band-pair image differencing for each spectral band and reduce registration noise;
- apply PCA transformation to the multispectral difference image;
- determine change membership functions for a number of selected change component images;
- apply fuzzy operations to combine change information in different change component images into a single image;
- determine changed areas based on the image generated at step five.

At step one, two images from different dates can be registered using a geometric correction program. After the image-to-image registration, two images with the same coordinate system are obtained: $X = \{x_i | i = 1, 2, ..., n\}$ of date one and $Y = \{y_i | i = 1, 2, ..., n\}$ of date two, where $x_i' = [x_{1i}, x_{2i}, ..., x_{pi}]$ and $y_i' = [y_{1i}, y_{2i}, ..., y_{pi}]$. The parameter $n$ denotes the number of pixels in an image and $p$ the number of spectral bands.

At the second step, a difference image, $DIF_j$, for each spectral band $j$, can be created, where $DIF_j = (y_i' - x_i')$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., p$. This further reduces registration noise in each difference image (Gong et al., 1992).

In traditional change detection, one difference image of a specific spectral band is selected among these difference images. This difference image should contain more change information than the other difference images for a particular application. An image thresholding technique is then applied to detect changes using the selected difference image (Jensen, 1986; Fung and LeDrew, 1988). For example, DIF2, the red spectral band difference image of Landsat multispectral scanner (MSS) data is usually considered to contain more information on rural to urban landcover changes than the other three MSS bands. Thresholds $T_1$ and $T_2$ can be determined on the histogram of DIF2 using the mean (ave) and standard deviation (std) (Figure 1). Deciding whether a pixel has changed is simply a matter of testing whether $\delta y$ falls outside of range $[T_1, T_2]$.

Two problems are associated with the above-mentioned traditional method, and they will be overcome by subsequent steps three to five. As mentioned in the introduction, the first problem is that different types of change information are contained in different spectral bands; thus, the use of one spectral band usually does not allow all the types of changes to be detected. The second problem is that once thresholding is applied to a difference image, change information occurring at smaller magnitudes (that is, within range $[T_1, T_2]$) will be lost. Also, noise could be included as change if its magnitude falls outside range $[T_1, T_2]$.

At step three, the difference images are used to generate a variance-covariance matrix. This is used to find a new set of axes according to the eigenstructure of the feature space. With the variance-covariance matrix, PCA can be applied to the difference images. The resultant principal component images are called principal component difference images, denoted by $PCD_j = \{\delta x_i | i = 1, 2, ..., n\}$ where $j = 1, 2, ..., p$, $\delta x_i$ is a pixel value for pixel $i$.
the jth principal component which results from a linear transformation of the difference images with the transformation coefficients determined with PCA. PCD images are obtained from a p-dimensional data set \( \{ x_i, y_i \} \) instead of the traditional application of PCA in change detection where a combined two-date data set, \( \{ x_i', y_i' \} \), is used. Because the variance in a difference image represents primarily change information and the purpose of PCA is to preserve most variances into the first few principal components, the application of PCA to difference images will result in most change information preserved in the first few PCD images.

At step four, the first two or three PCD images containing change information are selected. The exact number of PCDs is determined according to the eigenvalues and the correlation matrix of difference images. Each selected PCD image and its histogram are analyzed, and a fuzzy membership function of change is empirically defined based on the analysis results. From Figure 1, it is reasonable to assume that the more distant a pixel value \( \delta \) is from the average, \( \text{ave} \), the more likely is that pixel to fall into the change class. Based on the shape of the histogram of a PCD image, parameters of a fuzzy membership function of change can be determined. While a fuzzy membership function may take a variety of forms (Zadeh, 1978), in the case of Figure 1 an inverse triangular-shaped function may be suitable. A more sophisticated change membership function may be determined based on the knowledge of the various types of change in the study area. For example, statistics on various change types can be estimated from selected training samples.

A fuzzy membership function of change, \( \mu G(\delta_0) \), can be defined as:

\[
\mu_3(\delta_0) = \begin{cases} 
1 & 0 \leq \delta_0 < L_0 \\
(\delta_0 - \text{ave})/(L_0 - \text{ave}) & L_0 \leq \delta_0 < \text{ave} \\
(\delta_0 - \text{ave})/(H_0 - \text{ave}) & \text{ave} \leq \delta_0 < H_0 \\
1 & H_0 \leq \delta_0 \leq 255 
\end{cases}
\]

where \( \mu_0(\delta_0) \) represents the degree of pixel value \( \delta_0 \) in image PCD, belonging to a fuzzy set of change, \( \text{C} \). \( L_0, \text{ave}, \) and \( H_0 \) are the three parameters defining the inverse triangular-shaped function. \( L_0 \) and \( H_0 \) can be determined empirically by examining the histogram distribution of image PCD, and \( \text{ave} \) is the average pixel value in image PCD. A graphical form of \( \mu_0(\delta) \) is shown in Figure 2. After applying a fuzzy membership function, \( \mu_0(\delta_0) \), to an image, PCD, a change membership (CM) image, \( CM = [ \mu_0(\delta_0) ] \), is obtained.

At the fifth step, various change information from different CM images can be combined into one image, CCM, by applying the fuzzy set theory (Zadeh, 1965). Most operations based on the fuzzy set theory can be realized by using three basic types of fuzzy set operators: fuzzy union (\( \lor \)), fuzzy intersection (\( \land \)), and fuzzy complement (\( \bar{\mu} \)). While there are a number of definitions for fuzzy union and fuzzy intersection, the maximum and minimum rule are used in this study. Therefore, fuzzy union of \( \mu_1 \) and \( \mu_2 \) is equivalent to \( \mu_1 \lor \mu_2 \), and their fuzzy intersection is \( \mu_1 \land \mu_2 \). A fuzzy complement of \( \mu_1 \) is \( 1 - \mu_1 \). For example, if one wishes to combine change information in three CM images in such a manner that both the subtle changes in CM1 and changes in either CM1 or CM2 are included in the final resultant image, CCM = \( \mu_1(\delta_0) \lor (\mu_2(\delta_0) \lor \mu_3(\delta_0)) \)

\[
\mu(\delta) = (1 - \mu_3(\delta_0)) \land (\mu_1(\delta_0) \lor \mu_2(\delta_0))
\]

where \( 1 - \mu_3(\delta_0) \) represents the complement of \( \mu_3(\delta_0) \) because subtle changes have lower degrees of change membership compared to change membership in image CM3. Fuzzy set operations should be defined according to the characteristics of changes. In most cases, successful change information extraction requires that knowledge about the study area and expert knowledge on various change types be properly represented with fuzzy membership functions and fuzzy set operations.

Once the desirable change information from different image channels has been combined into one image CCM, at the final step, the CCM itself can be stored in a database to represent change information. One can also apply the thresholding or classification technique to determine and identify areas of change and to make a change map.

TEST AND RESULTS

In this section, steps one to five on the use of PCA and fuzzy set theory in change detection in the previous section are illustrated using an example. However, no attempt was made for step six to undertake a thorough change detection of the study area, instead, the purpose is to combine change information from different image channels into a newly created image channel. The principal component analysis module, PCA, image arithmetic, ARI, and image display programs in the EASI/PACE image analysis software package (PCI Inc., 1991) were used in this study. Programs implementing the grey-scale mapping, the fuzzy membership functions, and the fuzzy set operations have been developed by the author as additional EASI/PACE modules.

The Study Area and Data Preparation

The study area consists of a large sector of the twin cities of Kitchener-Waterloo, Ontario, and a small part of their surrounding rural area. A number of change detection studies have been carried out for this area with both Landsat Multispectral Scanner and Thematic Mapper data (Fung and LeDrew, 1987, 1988; Fung, 1990).

TM data for this area were acquired on August 3, 1985, and July 21, 1986 (Path-Row =18-30). During this period, major changes occurred at the rural-urban fringe where agricultural land was either cleared for construction or has undergone crop rotation (Fung, 1990). Illumination difference is ignored because these two images were obtained around the same time of the year. The atmospheric difference is assumed to be horizontally homono-

Figure 2.
An example fuzzy membership function of change. The degree of a pixel value, \( \delta_0 \), belonging to change denoted by \( \mu(\delta_0) \) ranges from 0 to 1. The further is \( \delta_0 \) away from \( \text{ave} \), the higher is \( \mu(\delta_0) \). \( \mu(\delta_0) \) saturates as \( \delta_0 \) falls outside of range \( (L_0, H_0) \).
Figure 3. The 1985 (top) and 1986 (bottom) Landsat TM images of Kitchener/Waterloo, Ontario. Both images are black-and-white versions of colour displays by assigning TM bands 2, 3, and 4 with blue, green, and red colour pans, respectively.

Figure 4. The first four PC image channels obtained by applying the PCA to the two-date data set. The PCA is based on the variance-covariance matrix obtained from the entire image. PC3 mainly contains the change information.

Traditional Use of Principal Component Analysis

As described in the introduction section, in change detection, PCA has been traditionally applied to a multi-date multispectral data set. With the two-date TM images of this study, the problem associated with this procedure can be illustrated. Normally, the variance-covariance matrix obtained from the entire image (or a systematic sampling procedure is made if the image data volume is too large) is used to determine the eigenvectors (that is, the principal component axes) of the feature space. After applying PCA to the two-date TM image set, a set of 12 principal component images were obtained, and the first four principal component images are shown in Figure 4. It can be seen from Figure 4 that, while PC1 and PC2 primarily contain information about the non-

consistent at each image acquisition time because the size of the study area is relatively small. This assumption implies that the actual change in the difference image is not modified significantly by the atmosphere. Because TM images often manifest striping effects due to the sensor radiometric calibration problem, a destriping program has been applied to both images.

Part of the 1986 TM image (approximately 460 by 600 pixels) covering the study area was geometrically registered to part of the 1985 one using 12 ground control points. A first-order polynomial transformation with a nearest neighbour sampling scheme was used. The standard error of geometric registration estimated from the 12 ground control points is 0.000 and 0.198 pixels across-track and along-track, respectively. Finally, subscenes of 250 by 512 pixels of the registered images were used. A black-and-white version of the colour composite of Bands 2, 3, and 4 of this subscene is shown in Figure 3, with the 1985 image at the top and the 1986 one at the bottom. During the experiment, TM bands 3 to 5 and 7 from both dates have been used. They constitute a 12-channel data set.

Traditional Use of Principal Component Analysis

As described in the introduction section, in change detection, PCA has been traditionally applied to a multi-date multispectral data set. With the two-date TM images of this study, the problem associated with this procedure can be illustrated. Normally, the variance-covariance matrix obtained from the entire image (or a systematic sampling procedure is made if the image data volume is too large) is used to determine the eigenvectors (that is, the principal component axes) of the feature space. After applying PCA to the two-date TM image set, a set of 12 principal component images were obtained, and the first four principal component images are shown in Figure 4. It can be seen from Figure 4 that, while PC1 and PC2 primarily contain information about the non-
change information in different PC images in the traditional use of PCA is a major drawback.

Principal Component Analysis of Band-Pair Difference Images, an Alternative

Difference images were derived by subtracting the 1985 image from the 1986 one. This resulted in six difference images corresponding to Landsat TM Bands 1 to 5 and 7. They were processed using the grey scale mapping program to reduce registration noise. Each band-pair difference image usually contains specific change information. A correlation matrix constructed for all six difference images will reveal the extent of change information redundancy among these difference images. An absolute correlation coefficient of 1 indicates that change information contained in two band-pair difference images is completely redundant. On the other hand, a correlation of 0 means there is no change information duplicating between two difference images. Between the two extreme cases of 0 and 1, an absolute correlation close to 1 means that there is a high level of change information redundancy between two difference images; a low absolute correlation indicates that there is a low level of change information redundancy. It can be seen from the correlation matrix (Table 1) of the six difference images that change information varies from one band to another. While most of the difference images are positively correlated to a high degree, indicating they have a high level of change information redundancy, DIF 5 has very low negative correlations with the rest of the difference images. This suggests that two transformed images would result from the PCA procedure containing most of the change information.

![Figure 5](image)

The first three PC image channels obtained by applying the PCA to the two-date data set with PC1, PC2, and PC3 located at the top, middle, and bottom of the figure, respectively. In this case, the PCA is based on the variance-covariance matrix obtained mainly from parts of the changed area in the image. PC1 contains much of the change information.

<table>
<thead>
<tr>
<th>DIF 1</th>
<th>DIF 2</th>
<th>DIF 3</th>
<th>DIF 4</th>
<th>DIF 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIF 2</td>
<td>0.940</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIF 3</td>
<td>0.951</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIF 4</td>
<td>-0.115</td>
<td>0.126</td>
<td>-0.280</td>
<td></td>
</tr>
<tr>
<td>DIF 5</td>
<td>0.720</td>
<td>0.762</td>
<td>0.784</td>
<td>-0.079</td>
</tr>
<tr>
<td>DIF 6</td>
<td>0.741</td>
<td>0.797</td>
<td>0.945</td>
<td>-0.260</td>
</tr>
</tbody>
</table>

* DIF 1–5, 7 stand for difference images generated from each TM band-pair.

The PCA was applied to the processed difference images. As a result, six PCD images were obtained. The eigenvalues for all PCD image channels are listed in Table 2, where it can be seen that the first two PCD images contain over 90 percent of the variation of the six difference images. Because variances in different images represent different degrees of change, this 90 per-

<table>
<thead>
<tr>
<th>PC Channel</th>
<th>Eigenvalue</th>
<th>% Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>228.62</td>
<td>61.9</td>
</tr>
<tr>
<td>2</td>
<td>95.62</td>
<td>27.1</td>
</tr>
<tr>
<td>3</td>
<td>15.5</td>
<td>4.0</td>
</tr>
<tr>
<td>4</td>
<td>7.5</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>2.55</td>
<td>0.7</td>
</tr>
<tr>
<td>6</td>
<td>0.80</td>
<td>0.5</td>
</tr>
</tbody>
</table>
cent in variation indicates that most of the change information has been preserved in the first two PCD images. The first three PCD images are shown in Figure 6. From image PCD1 it can be seen that the changed area contains a large portion of “pepper and salt” type of noise. For subsequent tests, only PCD1 and PCD2 are used.

Change Membership Images and Their Fuzzy Set Operations

The histograms for images PCD1 and PCD2 are shown in Figure 7. Because both histograms have distributions close to a normal one, fuzzy membership functions of change in a symmetrical inverse triangular form, as shown in Figure 2, were used. It should be noted that histogram distributions in Figure 7 only helped to determine whether the inverse triangular function should be of symmetrical form. A histogram distribution does not offer any help for determining which type of membership function (for example, linear or non-linear) should be used. To construct fuzzy membership functions of change, two sets of parameter values (LB, LRC, and RB for PCD1 and PCD2) were determined from their histograms. They were 97, 127, and 157, and 5, 120, and 190 for PCD1 and PCD2, respectively. Fuzzy membership functions were then applied to PCD1 and PCD2 to create CM1 and CM2 (Figure 8). Image CM1 mainly shows land use changes from rural to urban conversion. On the other hand, it seems that crop rotation in agricultural land is primarily contained in image CM2. Since no ground truth knowledge about changes in the study area was available, detailed change detection was not attempted. A fuzzy union was applied to image CM1 and image CM2 to generate image CCM (Figure 9). It can be seen from Figure 9 that the change components in image CM1 and image CM2 have been successfully integrated into image CCM. Further change information analysis and extraction may then be done using only one image: image CCM.

DISCUSSION

Traditionally, PCA is directly applied to multi-date multispectral images. In traditional use of PCA for change component enhancement, variances and covariances are calculated not only from the changed part in multi-date original images but also from
the new use of the PCA procedure and to demonstrate the use of fuzzy set operations in combining change information from a number of image channels, which contain change information.

The example on the two ways of using PCA clearly indicates that in traditional use of PCA, the proportion of changed area in an image determines which principal component images will contain change information. Rigorously speaking, it is the relative amount of variance between the changed area and the unchanged part in an image that determines which particular PCs contain change information. If the amount of variance from changed areas in an image dominates the total variance of the image, the first PC image will contain the change information. As the amount of variance from changed areas in an image decreases, change information will be preserved in minor PC image channels.

The fuzzy set procedure proposed in this paper can be applied to convert other types of images containing change information into change membership (CM) images. For example, instead of using PCD images, band pair difference images and/or varied images can be used if one wishes to integrate different types of change information into one image channel. To do so, fuzzy membership functions have to be determined for each difference or ratio image channel. For a ratio image channel, the change information may not present normal distribution. Therefore, different parameters or even forms of fuzzy membership functions may be used than a symmetrical inverse triangular function, as used in this research. As for difference images or the PCD images, fuzzy membership functions with an inverse triangular form are applicable, according to the results from this study.

The advantage of using CM images is that changes of various types and magnitudes are preserved. Therefore, almost no change information is lost. Subsequent fuzzy set operations can be defined according to the characteristics of specific types of changes, and these operations can be applied to CM images to extract these changes.

In the process of CM image generation, however, care has to be taken when parameter values $\alpha, \omega, \text{ and } 16\%$ are used to be determined from the difference images or PCD images. The actual data variance in each band of difference image or each channel of PCD image may be modified by scaling, which is often taken as a procedure to force the data range to fit in between 0 and 255 to produce a byte image. The scaling effect can be noted by comparing the two histograms in Figure 7. Contrary to the fact that PCD$_1$ contains more data variation than PCD$_2$, the effective standard deviation of PCD$_2$ (35) is greater than that of PCD$_1$ (21).

In order to achieve a balanced contribution of change information from each difference image or each PCD image during the change information integration stage, the scaling effect can be compensated by adjusting $L_0$ and $L_8$ for each image according to the relative magnitudes of the effective standard deviations and actual standard deviations of all the image channels under consideration. For example, as indicated in Table 2, the actual standard deviations of PCD$_1$ and PCD$_2$, which are square roots of their corresponding eigenvalues, are 15.1 and 9.75, respectively. Due to the scaling effect, however, the effective standard deviations are 21 and 35, respectively. Therefore, a narrower range between $L_0$ and $L_8$ for PCD$_2$ than that for PCD$_1$ should be used. To balance the change information from PCD$_1$ and PCD$_2$ in CM, the range ratio of $L_0, L_8$ between PCD$_2$ and PCD$_1$ should be 0.987. The range ratio in the previous section for PCD$_1$ and PCD$_2$ is selected according to 0.987.
CONCLUSIONS

Generating difference images before undertaking the principal component analysis (PCA) for enhancing and extracting change information from multispectral images is a valuable approach. In this approach, change information is always accumulated in the first few principal component difference images, and thus makes it a trivial task for an image analyst to determine which PCA image contains change information. However, difference images tend to be noisy and noise reduction needs to be undertaken (for example, Gong et al., 1992).

Operations based on the fuzzy set theory can be used to further integrate change information from different image channels into a single-image channel. With the procedure illustrated in this paper, change information of various land cover types can be eventually preserved and stored in one image channel. With this image, subsequent image information extraction methods, such as thresholding, classification, and various quantitative analysis algorithms, can be applied when specific change information is to be derived from multispectral images acquired from different dates.

Further research is needed to apply the methods proposed here to practical change identification and quantitative change information extraction tasks. A training sample of a changed area should be collected as a basis for change membership function determination, while a test sample is needed for quantitative accuracy assessment.

ACKNOWLEDGEMENTS

The author is grateful to PCI Inc., Richmond Hill, Ontario, for providing access to its EASI/PACE image analysis software. The Landsat TM data of Kitchener-Waterloo were made available by Dr. E.F. LeDrew and Dr. T. Fung. This research was funded by an NSERC research grant and a starter grant from the University of Calgary.

REFERENCES


