Pattern Recognition Letters 19 (1998) 197–204

A new gray level based Hough transform for region extraction: An application to IRS images [1]

B. Uma Shankar, C.A. Murthy [2], S.K. Pal *

Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India

Abstract

A technique using the Hough transform is described for detection of homogeneous line segments directly from (i.e., without binarization of) gray level images. A definition of "region" in terms of these line segments, with constraints on its length and variance, is provided. The algorithm is able to extract gray level regions irrespective of their shape and size. The effectiveness of the method is demonstrated on Indian Remote-sensing Satellite (IRS) images.

Keywords: Hough transform; Region extraction; Satellite image

1. Introduction

The Hough transform (HT) (Hough, 1962; Rosenfeld and Kak, 1982) is used for finding straight lines or analytic curves in a binary image. There are several ways in which the HT for straight lines can be formulated and implemented. Risse (1989) has listed some of these forms and analyzed them along with their complexities. The most popular one is the (ρ,θ) form, which is given by Duda and Hart (1972). This parametric form specifies a straight line in terms of the angle θ (with the abscissa) of its normal and its algebraic distance ρ from the origin. The equation of such a line in the x-y plane is

$$x \cos\theta + y \sin\theta = \rho, \tag{1}$$

where θ is restricted to the interval $[0,\pi]$. Some of the advantages of this parametric form are its simplicity and ease of implementation.

Hough transforms are applied usually to binary images. Hence, one needs to convert, initially, the gray level image to a binary one (through thresholding, edge detection, thinning, etc.) to apply the HT. Note that, in the process of binarization, some information regarding line segments in the image may get lost. Thus, it becomes appropriate and necessary to find a way of making the HT applicable directly to gray level images.

This article is an attempt in that direction where we present a technique for extracting homogeneous regions of arbitrary shape and size in a gray level

* Corresponding author. Email: sankar@isical.ernet.in.

[1] Electronic Annexes available.
See http://www.elsevier.nl/locate/patrec.

[2] At present, Dr. Murthy is with the Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA.

image based on the Hough transform. The regions are defined in terms of homogeneous line segments. The technique includes some operations which are performed in a window to obtain homogeneous line segments. For every quantized (ρ,θ) cell in the Hough space (ρ represents radius, θ represents angle), the variance of the pixel intensities contributing to the (ρ,θ) cell is computed. The cells whose variances are less than a pre-specified threshold, are found. Each such cell would represent a homogeneous line segment in the image. The window is then moved over the entire image, so as to result in an output consisting of only the homogeneous line segments, thereby constituting different homogeneous regions. The performance of the method has been demonstrated on Indian Remote-sensing Satellite (IRS) images for different parameter values.

In this connection we mention the methods of gray scale Hough transform (GSHT) of Lo and Tsai (1995), and generalized Hough transform (GHT) of Ballard (1981). GSHT enables one to find thick lines (called bands) from gray scale images. Therefore, it can be used for detecting road like structures only in remote-sensing images. GHT is able to extract arbitrary shapes from the edge map of a gray level image using prototype information of the objects to be extracted. Note that our method does not need this information and is thus able to extract objects of irregular shapes and arbitrary sizes as found in remote sensing images.

There are many other methods (Duda and Hart, 1973; Pal and Pal, 1993; Richards, 1993; Rosenfeld and Kak, 1982) based on pixel classification and gray level thresholding which are used for segmenting remotely sensed images into meaningful homogeneous regions. The hard c-means (HCM) algorithm is one of them which is a widely used method based on the principle of pixel classification. Its performance on the IRS images for $c = 2$ (i.e., object background classification) is also shown as a comparison.

2. Definition and formulation

A region in a gray level image can be viewed as a union of several line segments, so that it consists of a connected set of pixels having low gray level variation. Therefore, to extract a region, we need to define a line segment in the gray level image. A line segment in a gray level image is defined using two threshold parameters, minimum length of the line (l) and maximum variation of the line (v). The mathematical formulation of the region in terms of line segments is stated below.

Definition 1. A pixel $P = (i,j)$ is said to fall on a line segment joining the pixels $P_1 = (i_1,j_1)$ and $P_2 = (i_2,j_2)$ if $\exists\lambda_0$, $0 \leqslant \lambda_0 \leqslant 1$ such that

$$\lambda_0 P_1 + (1-\lambda_0) P_2 = P.$$

Definition 2. A collection of pixels $L(l,v)$ is said to be a line segment in a gray level image if: there exist $P_1 = (i_1,j_1)$ and $P_2 = (i_2,j_2)$, $P_1 \neq P_2$, such that

- $L(l,v) = \{P \mid P$ is a pixel falling on the line segment joining P_1 and $P_2\}$,
- the number of pixels in $L(l,v)$ is $\geqslant l$, and
- the variance of the gray values of pixels in $L(l,v)$ $\leqslant v$.

Let $\mathscr{A}_{l,v} = \{L(l,v) \mid L(l,v)$ is a line segment in the image$\}$. That is, $\mathscr{A}_{l,v}$ represents the collection of all line segments in the image.

Definition 3. A pixel P is said to be in the homogeneous region if $P \in L(l,v)$, at least for one $L(l,v) \in \mathscr{A}_{l,v}$.

Definition 4. A pixel P is said to be in the non-homogeneous region if $P \notin L(l,v)$, $\forall L(l,v) \in \mathscr{A}_{l,v}$.

Let $N_H = \{P \mid P$ is a pixel in the non-homogeneous region$\}$.

The region R is defined as follows.

Definition 5. Let $L_1(l,v), L_2(l,v) \in \mathscr{A}_{l,v}$. Then $L_1(l,v)$ and $L_2(l,v)$ are said to be directly connected if either $L_1(l,v) \cap L_2(l,v) \neq \emptyset$ or there exist pixels $P_1 \in L_1(l,v)$, $P_2 \in L_2(l,v)$ such that P_1 is one of the eight neighbours of P_2.

Note that a line segment $L(l,v)$ is directly connected to itself.

Definition 6. Two line segments $L_\alpha(l,v)$ and $L_\beta(l,v)$ belonging to $\mathscr{A}_{l,v}$ are said to be connected if they are directly connected or there exist $L_i(l,v) \in \mathscr{A}_{l,v}$, $i = 1,2,\dots,k$ where $k \geqslant 3$, such that $L_i(l,v)$ and

$L_{i+1}(l,v)$ are directly connected $\forall i = 1,2,\ldots,(k-1)$ where $L_1(l,v) = L_\alpha(l,v)$ and $L_k(l,v) = L_\beta(l,v)$.

Definition 7. Let $B_{L_\alpha}(l,v) = \{L(l,v) \mid L(l,v) \in \mathscr{A}_{l,v}$, and $L(l,v)$ and $L_\alpha(l,v)$ are connected$\}$, $L_\alpha(l,v) \in \mathscr{A}_{l,v}$.

Note that for $L_\alpha(l,v),\ L_\beta(l,v) \in \mathscr{A}_{l,v}$ either $B_{L_\alpha}(l,v) = B_{L_\beta}(l,v)$ or $B_{L_\alpha}(l,v) \cap B_{L_\beta}(l,v) = \emptyset$.

Note also that $\bigcup_{L_\alpha(l,v) \in \mathscr{A}_{l,v}} B_{L_\alpha}(l,v) = \mathscr{A}_{l,v}$. That is, $\mathscr{A}_{l,v}$ is partitioned into finitely many sets using $B_{L_\alpha}(l,v)$s.

Definition 8. Let $R_{L_\alpha}(l,v) = \bigcup_{L(l,v) \in B_{L_\alpha}(l,v)} L(l,v)$. Then $R_{L_\alpha}(l,v)$ is said to be a region generated by $L_\alpha(l,v)$. Note that $R_{L_\alpha}(l,v)$ is a set consisting of pixels in the given image. Observe also that the same region can be generated by different line segments (follows from Definition 7).

2.1. Observations on the above definitions

(a) A line segment may be termed as a *region* according to the above definitions.

(b) The variance of n points $x_1,x_2,\ldots,x_n$, is given by

$$\frac{1}{n}\sum\left(x_i-\overline{x}\right)^2 \quad \text{where } \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Now,

$$\text{variance} < v \Rightarrow \sum\left(x_i-\overline{x}\right)^2 < nv.$$

Observe that there may be a point (say, x') among $x_1,x_2,\ldots,x_n$, such that $(x'-\overline{x})^2$ may be high (say, $(x'-\overline{x})^2 > kv$, where $k>1$). Even then $\Sigma(x_i-\overline{x})^2$ can still be less than nv. This observation indicates the removal of noise up to some extent by the proposed variance based definition of line segment.

(c) The values for l and v are to be chosen "appropriately" to obtain the actual regions in an image. Some portions of the actual regions may not be obtained if the value of l is high. If the value of v is high, the number of pixels detected to constitute various regions increases, and therefore the possibility of some spurious collection of pixels being termed as region increases. Reducing the value of v, on the other hand, decreases the number of detected pixels, thereby increasing the possibility of losing some actual regions. Similarly if the value of l is low then some unwanted regions may arise.

(d) Observe that a pixel P does not fall into any region $\Rightarrow$ There exists no line segment passing through P with low gray level variation (i.e., variance of the gray level values of the pixels on any line segment passing through P is greater than v). Hence the above stated definitions intend to suppress the pixels in the non-homogeneous region of the image.

(e) Note that two adjacent collections of connected pixels, each collection having different average gray value, may fall into the same region, thereby losing their identity according to the definition. (However, in such cases, a further processing may be necessary to separate (partition) the said collection.)

The definitions regarding *regions* have been stated above. There may exist several other ways of obtaining the said regions from the image. Note that regions have been defined as a union of line segments, and the Hough transform is a standard method of obtaining line segments. But the Hough transform finds line segments in a binary image. In order to make the HT applicable directly to a gray level image, we formulate a method in the next section which is able to find line segments (and hence the regions) in a gray level image.

2.2. Strategies of region extraction in a gray level image using the Hough transform

Extraction of line segments.

(i) Consider the equation for a straight line to be $x\cos\theta + y\sin\theta = \rho$. Apply suitable sampling on θ and ρ, and construct the Hough accumulator. Transform each point of the image (pixel) using different values of θ (and its corresponding ρ values). Note that a point in the image space is mapped to more than one cell in the Hough space and each of these cells represents a line in the image space.

(ii) Compute, for each cell, the length of the corresponding line (ℓ) as the total number of image points (pixels) mapped into that cell, i.e., the cell count. Variance of the said pixel values may be termed as the variance (V) of the corresponding line.

(iii) For a cell in the Hough space, if the length of the line is less than l or variance of the line is larger than v, then suppress the cell in the Hough space.

(iv) Remap this Hough space (containing unsuppressed cells) to image space. This process of remapping preserves all those pixels which are not suppressed in at least one of the cells of the Hough space. Since this transformation preserves only the location of pixels, not their gray values, they may be restored from the original image.

Let us consider an image of size $M \times M$. If M is too large compared to the threshold parameter l, then ℓ values (cell counts) will be larger, and as a result, the variance of the gray levels on a line may exceed the threshold value v. Many genuine line segments, therefore, may not be detected in such a case. To avoid this, the search process for obtaining the line segments is to be conducted locally. That is, a window of size $\omega \times \omega$ needs to be moved over the entire image to search for line segments. Here ω may be taken as, $2l > \omega \geqslant l$, because $\omega \geqslant 2l$ may still lead to the suppression of actual lines in the image.

Extraction of regions. One can clearly see that the above mentioned process extracts line segments which are connected according to Definitions 5–8. Therefore the collection of these line segments will result in regions of different sizes and shapes.

Note.

- There is no restriction on the shape of the "region" thus obtained. The only restriction, we used on the size of the region (i.e., length of the line $\geqslant l$), is a weak one.
- The method does not need any prior representation of the shape of the region to be detected. Therefore, it can extract regions of arbitrary shape and size.

3. Algorithm and implementation

It has been mentioned in the earlier section that values for l (length of the line), v (variance threshold) and ω (window size) are to be selected and the obtained lines are to be remapped to the image domain to procure regions. This process of remapping the cells in the Hough space is to be carried out on every window. The steps of the entire algorithm are stated below.

Step 1. For a window ($\omega \times \omega$) of the image obtain the Hough accumulator values for different ρ and θ. ρ values and θ values are sampled suitably in their respective domains. For each cell in Hough space, mean and variance of corresponding pixel values in the image domain are computed using two more accumulators (one for the sum of gray values Σx and one for the sum of squares of gray values Σx^2). These sum and sum of squares along with the count of cells (ℓ) will be used for computing the variance $V = (\frac{1}{\ell}\Sigma x^2) - (\frac{1}{\ell}\Sigma x)^2$ of the cell. If the cell count is $< l$, then replace the cell count by zero (i.e., the cell is suppressed). If $V > v$ then also replace the cell count by zero. The cells with count non-zero are remapped to the image domain preserving the position of window.

Step 2. Repeat Step 1 for all possible windows of size $\omega \times \omega$ in the image.

Step 3. Restore the gray values of remapped pixels from the original image.

The number of computations in the aforesaid algorithm can be reduced drastically in the following way.

(i) Note that for a given window size ω, Hough transformation of the pixel does not change with its location, because the reference frame for computing θ (and hence ρ) remains the same. Thus, the Hough accumulator values can be calculated only once and be used for every position of the window.

(ii) Again, all possible windows of size $\omega \times \omega$ need not be considered. The window can be moved by half of its size, both horizontally and vertically. This process, though marginally reducing the accuracy of the regions obtained, decreases computations drastically.

(iii) Keeping point (iv) of Section 2.2 in mind, the process of restoring gray values in the aforesaid Step 3 can be combined with Step 1 by preserving only those pixels in the image domain which are not getting suppressed in at least one cell of the Hough space.

4. Results

We have applied the proposed method on to (Indian Remote-sensing Satellite) images to demonstrate its usefulness. The IRS images considered here have a spatial resolution of 36.25 m × 36.25 m, wavelength range 0.77 μm – 0.86 μm and gray level values in the range 0–127 (Richards, 1993). The size of the images is 512 × 512. An enhanced (linearly stretched) image is provided in Fig. 1(a) (city of Calcutta) and Fig. 2(a) (city of Bombay) for the convenience of readers, since the original images are poorly illuminated. However the method has been applied to the original images.

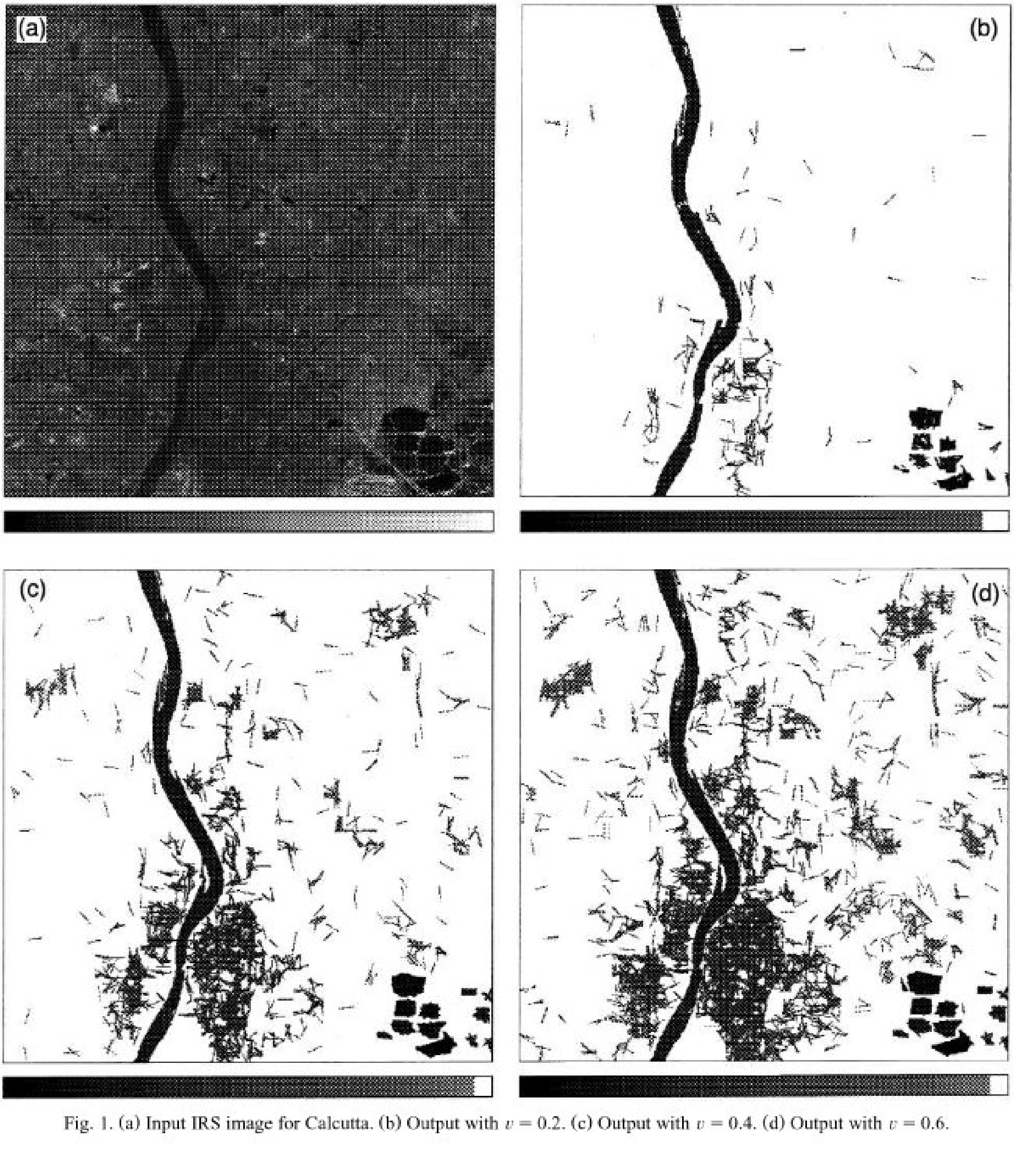

Fig. 1. (a) Input IRS image for Calcutta. (b) Output with $v = 0.2$. (c) Output with $v = 0.4$. (d) Output with $v = 0.6$.

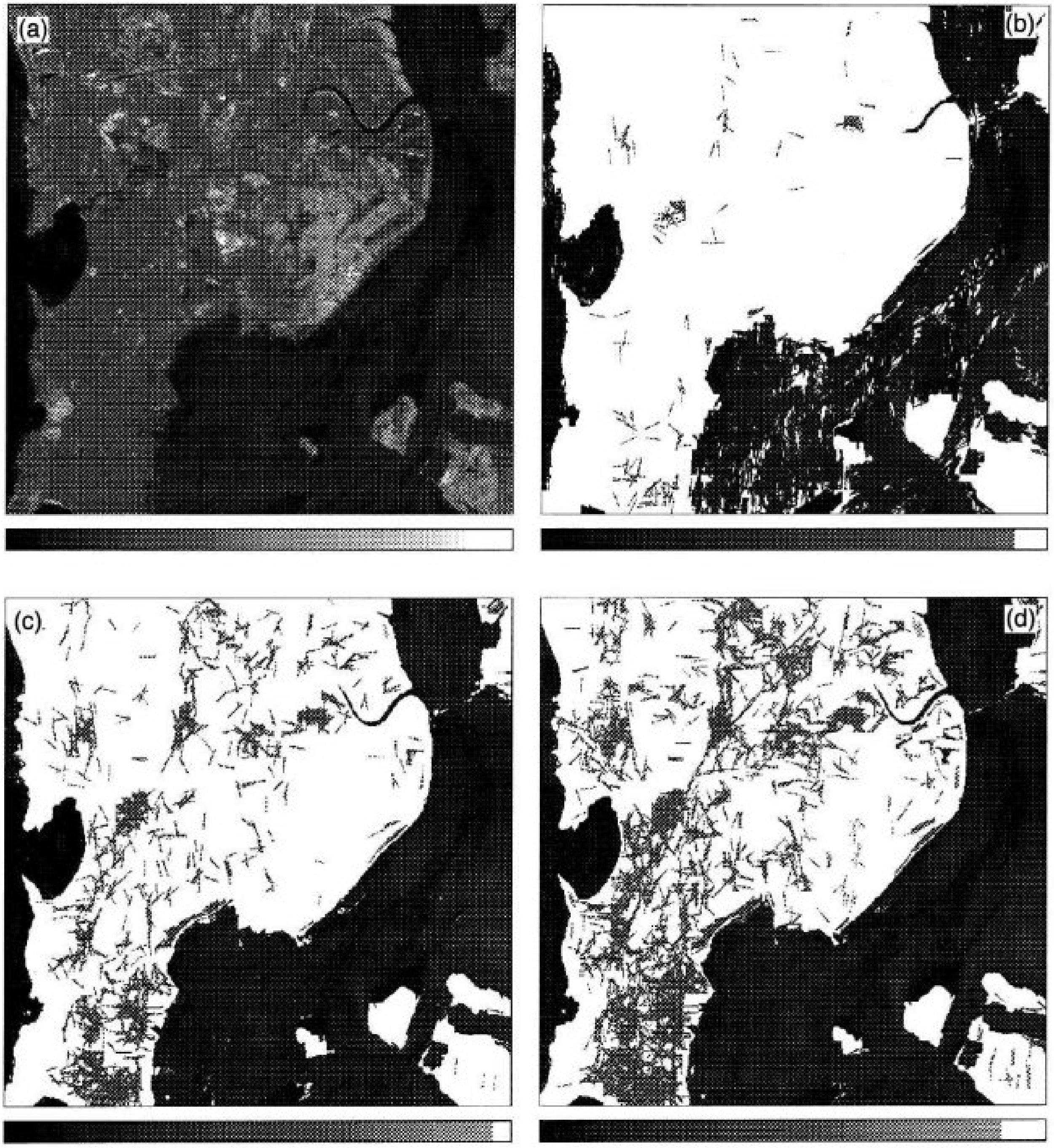

Fig. 2. (a) Input IRS image for Bombay. (b) Output with $v = 0.2$. (c) Output with $v = 0.4$. (d) Output with $v = 0.6$.

In the present investigation, we have used $w = 16$ and $l = 14$ for various values of v. The output corresponding to Fig. 1(a) and Fig. 2(a) for $v = 0.2$, 0.4 and 0.6 are shown in Fig. 1(b), (c) and (d), and Fig. 2(b), (c) and (d), respectively. As stated in Step 3 of the algorithm in Section 3, the extracted regions have their original gray values restored. Justification of the result (output image) is evident from the observations laid down in Section 2.1 and under the note in Section 2.2. For example, as v increases, the number and size of the detected regions are seen to increase. (Note that the set of pixels in Fig. 1(b) (Fig. 2(b)) is a subset of that of Fig. 1(c) (Fig. 2(c)), and that Fig. 1(c) (Fig. 2(c)) is a subset of Fig. 1(d)

(Fig. 2(d)).) Similarly, decreasing v enables one to detect tiny homogeneous regions as separate classes, even when they are embedded in a different wide homogeneous region. That is why the two bridges over the river Ganges in the Calcutta image (Fig. 1(b)) and a bridge on the Arabian sea (Thane creek) in the Bombay image (Fig. 2(b)) became prominent as separate regions for $v = 0.2$. For $v = 0.4$ and 0.6, they disappeared in Fig. 1 and became faint in Fig. 2. (This makes ''the evidence'', for the existence of a bridge, available for a further stage of the vision process to recognize or deal with.) In other words, as v increases the Ganges in the Calcutta image comes out as a single region and the sea in the Bombay image becomes smoother. Similarly, note from the lower parts of Fig. 1 and Fig. 2 that the two dense city areas (namely, Howrah and Calcutta) on two sides of the river (Fig. 1), and the dense city area of Bombay (Fig. 2) become prominent because of the increase in the number of constituting pixels with higher values of v. Note further that these extracted city areas did not get merged with the river (Fig. 1) and the sea (Fig. 2). Experiments were also conducted for other values of l such as $l = 15$ and 16, but the results were not much different.

Fig. 3. Segmentation of the Calcutta image using HCM with $c = 2$.

Fig. 4. Segmentation of the Bombay image using HCM with $c = 2$.

As a comparison of the performance, we consider the hard c-means (HCM) algorithm (with $c = 2$, i.e., object and background classification) which is a widely used segmentation algorithm based on pixel classification (Duda and Hart, 1973; Pal and Pal, 1993; Richards, 1993). Here the input features are considered to be the gray level value of the pixel and the average value over its 3×3 neighbourhood. (This method with $c = 2$ is chosen for comparison because this also provides object background classification like ours.) From the segmented output (Fig. 3 and Fig. 4) one may note that the algorithm failed to isolate the respective city areas from the river and sea. It also could not, unlike Fig. 1(b) and Fig. 2(b), enhance the bridge regions.

5. Conclusions and discussion

A method of extracting regions in a gray level image using the principle of Hough Transform has been described. A definition of ''region'' in terms of line segments is provided. Since the methodology does not involve any representation (such as parametric, template, etc.) of the shape of regions, it has the ability to detect regions of arbitrary shape and size.

To restrict the size of the article, we have presented the results corresponding to $v = 0.2$, 0.4 and 0.6 for $\omega = 16$ and $l = 14$ only, although the experiment was also conducted for other values of l (e.g., 15 and 16) and ω (e.g., 8, 24 and 32). The variance of pixel values is used here as a measure of homogeneity of line segments. One may use any other homogeneity measure for this purpose. Like many other pattern recognition and image processing algorithms, the selection of v is problem dependent. An investigation for detecting automatically v constitutes a part of further research.

Although the effectiveness of the method is demonstrated on IRS images, one can apply it to any kind of images for region extraction. The results will, however, be governed by the characteristics as laid down under Section 2.1 and under the Note of Section 2.2.

References

Ballard, D.H., 1981. Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition 13, 111–122.

Duda, R.O., Hart, P.E., 1972. Use of the Hough transform to detect lines and curves in pictures. Comm. ACM 15, 11–15.

Duda, R.O., Hart, P.E., 1973. Pattern Classification and Scene Analysis. Wiley, New York.

Hough, P.V.C., 1962. A method and means for recognizing complex patterns. U.S. Patent 3, 069, 654.

Lo, R., Tsai, W., 1995. Gray-scale Hough transform for thick line detection in gray-scale images. Pattern Recognition 28, 647–661.

Pal, N.R., Pal, S.K., 1993. A review on image segmentation. Pattern Recognition 24, 1277–1294.

Richards, J.A., 1993. Remote Sensing Digital Image Analysis: An Introduction, 2nd ed. Springer, New York.

Risse, T., 1989. Hough transform for line recognition: Complexity of evidence accumulation and cluster detection. Comput. Vision Graphics Image Process. 46, 327–345.

Rosenfeld, A., Kak, A.C., 1982. Digital Picture Processing, vol. II. Academic Press, New York.