Hough Transform for Region Extraction in Color Images

Sarif Kumar Naik
Machine Intelligence Unit
Indian Statistical Institute
203 B. T. Road, Kolkata-700 108
sarif_r@isical.ac.in

C. A. Murthy
Machine Intelligence Unit
Indian Statistical Institute
203 B. T. Road, Kolkata-700 108
murthy@isical.ac.in

Abstract

This article aims to propose a method to use the idea of Hough Transform (HT) implemented in grey scale images to color images for region extraction. A region in an image is seen as a union of pixels on several line segments having the homogeneity property. A line segment in an image is seen as a collection of pixels having the property of straight line in Euclidean plane and possessing the same property. The property 'homogeneity' in a color image is based on the trace of the variance covariance matrix of the colors of the pixels on the straight line. As a possible application of the method, it is used to extract the homogeneous regions in the images taken from the Indian Remote Sensing Satellites.

1. Introduction

Hough Transform(HT) has been a widely used method for extracting arbitrary shapes from images. It was designed to extract straight line segments from binary images by Hough[3]. Usually HT is applied on binary images[4]. To extract line segments or any geometric structure using HT, from a gray scale image or color image one has to transform it to a binary image using thresholding or any other method. Usually, these binarized images are the edge maps of the original grey level or color image. There are very few works dealing with applying HT on grey scale images directly[8, 10, 5] i.e. without the binarization. To the best knowledge of the authors, no method has ever been proposed for implementing HT directly to color or multispectral images.

Hough transform is an extensively studied method of feature(line/curve) extraction for binary images[7, 6, 1] because of its robustness to image noise and the discrimination ability against unwanted shapes [4]. A survey of the HT for binary images appears in [2]. Thus only a brief introduction to HT is presented to introduce the notation and terminology used in the later part of the article and a short survey of it uses in grey level images is presented here.

Basically, HT is a voting process where each pixel of the image votes for all possible patterns passing through that point. The votes are accumulated in an accumulator array $\mathcal{A}(\rho, \theta)$, where, θ is the angle made by the normal to the straight line with the $x-$axis and ρ is the perpendicular distance of the straight line from the origin. The representation of the straight line in (ρ, θ) form is

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

This accumulator array $\mathcal{A}(\rho, \theta)$ is called the Hough Space. Number of votes for each cell in $\mathcal{A}(\rho, \theta)$ represents the number of pixels in the pattern passing through the point (x, y). In this process each pixel in the image space is mapped to a sinusoidal curve in the Hough space. So HT is a transformation from a point to a curve.

Lo *et al*,[8] proposed a technique to implement HT directly to grey scale image without thresholding, and without finding the edge map of the image. They used a four dimensional accumulator array, where G_i, the grey value of the pixel is the additional dimension along with ρ and θ as compared to conventional HT. Here each pixel (x_i, y_i) in the image space is associated with its grey value G_i and each accumulator cell $\mathcal{A}(\rho, \theta, G_i)$ in the grey-scale Hough parameter space. It finds line segments in grey-scale bands. This method is expensive in the sense of space complexity. Shapiro[11] proposed a method that replaces the original image by its digital half toning equivalent. A decision making procedure called Digital Half toning Hough Transform (DHHT), allowing the HT to work with arbitrary multilevel images is proposed in this paper. DHHT is an amalgam of the conventional HT and Digital Half toning. Kesidis *et al*[5] proposed a Grey-scale Inverse Hough Transform algorithm which is combined with a modified Grey-scale Hough Transform. The said method is implemented to detect line segments directly from a grey-scale image.

Uma Shankar *et al* [10] proposed a technique to extract line segments directly from a grey scale image without using thresholding or the edge map of the image. A region on a gray scale image has been defined as union of several connected line segments and a line segment is defined as a straight line with the property of uniformity among the pixels on the line. The property of uniformity is the value of variance of gray values of the pixels on the line. The formal definitions of these are discussed in the later part of the article. The said technique is used to find uniform regions from remotely sensed images.

The main idea behind this work is described in section 2. Definitions relating to this and some properties are described in section 3. Section 5 deals with implementation and results. Section 6 provides an analysis of the results and the scope for further research.

2. Motivation

The idea of the method proposed by Uma Shankar *et al.,* for region extraction using HT in grey-scale images. The work in [10], can be extended to color images. Their method has been revisited here parallelly with the development of our method with some changes to the definitions proposed by them. Uma Shankar *et al* [10], constructed a binary image consisting of all **homogeneous** line segments with the help of HT. A line segment is considered, its **homogeneity** is checked and all **homogeneous** line segments are found in this process. A generalization of the above is to

1. construct a property P,
2. check whether the pixels in the line segment satisfy P,
3. create a file of line segments possessing the property P.

The property P has to be defined according to the requirement. For region extraction the possible property can be the uniformity of grey values or color of the concerned set of pixels.

3. Definitions

A line segment in a grey level image should, in principle, represent a **homogeneous** set of pixels lying on the line segment, **homogeneity** is to be measured with respect to grey values of the concerned pixels. The word **homogeneity** may be viewed in different ways. One way is to make the variation in the grey values of the pixels on the line segment small. Another way is to define similarity between two pixels depending on the grey value difference, and making all the points on the line segment similar to each other. There are other ways of viewing **homogeneity**. In this article, we restrict ourselves to the above two ways only.

The above two concepts yield different expressions. For example, if $x_1, x_2, ..., x_n$ denote the grey values of the pixels lying on a line segment in a grey level image, in order to implement the first way we need to find the variance of $x_1, x_2, \ldots, x_n$ and check whether it is less than a threshold value. Second way is to consider a threshold value δ and define that the considered line segment is a valid one if $|x_i - x_j| < \delta \ \forall \ i, j$. These two ways of measuring homogeneity lead to different measures of evaluating the validity of line segments. Uma Shankar *et al*[10] incorporated the first way to find line segments and consequently regions in the grey level images. We shall generalize the above concepts for color images, as stated below.

Unlike, grey level images, a pixel in a color image is not represented by a scalar but by a vector (R, G, B), denoting for *red*, *green* and *blue* components of color respectively and each of them lies between 0 to 255.

Homogeneity of color pixels on a line segment may be viewed in two ways. One way is to define that a set of pixels possess homogeneity if for each of the three components, the pixels are homogeneous. Another way of defining homogeneity of the pixels is by combining the values of all the set of pixels without looking at them individually. Formal definitions of all these concepts are stated below.

Let I be a color image and the color of a pixel (i, j) is denoted by $(R(i, j), G(i, j), B(i, j))$.

Definition 1: Let $L = \{(i_1, j_1), (i_2, j_2), \ldots, (i_n, j_n)\}$ be an ordered set of pixels. L is said to be **connected** *if (i_{k-1}, j_{k-1}) is in the 8-neighborhood of (i_k, j_k) $\forall \ k = 2, 3, \ldots n$.*

Definition 2: A set of connected pixels $L = \{(i_1, j_1), (i_2, j_2), \ldots (i_n, j_n)\}$ is said to be a **possible set for a line segment (PSLS)** *if $\exists \ \rho$ and θ such that $i_k \cos\theta + j_k \sin\theta = \rho \ \forall \ 1 \leq k \leq n$.*

From the above discussion, **homogeneity** of a set of pixels for a color image can be defined in any of the three way stated below.

Definition 3: Let $L = \{(i_k, j_k) : k = 1, 2, \ldots n\}$ be a set of pixels and let the triplet $(R(i_k, j_k), G(i_k, j_k), B(i_k, j_k))$ be the color of the k^{th} pixel of the set. L is said to be **homogeneous** *if it satisfies any one of the following three criteria*:

1. $$\max_k(R(i_k, j_k)) - \min_k(R(i_k, j_k)) < \delta_R,$$
$$\max_k(G(i_k, j_k)) - \min_k(G(i_k, j_k)) < \delta_G,$$
$$\max_k(B(i_k, j_k)) - \min_k(B(i_k, j_k)) < \delta_B,$$
where, δ_R, δ_G and δ_B are the threshold values for $R, G,$ and B components respectively.

2. $\sigma_R^2 < \delta_R$ *and* $\sigma_G^2 < \delta_G$ *and* $\sigma_B^2 < \delta_B$, *where* σ_R^2, σ_G^2 *and* σ_B^2 *are the variances and* δ_R, δ_G *and* δ_B

are the threshold values for $R, G, and B$ components respectively.

3. *$\lambda_1 + \lambda_2 + \lambda_3 < \delta$, where, λ_1, λ_2 and λ_3 are the eigen values of the variance covariance matrix Σ of the set of pixels. δ is a threshold value of the sum of the eigen values. Alternatively, Trace(Σ) $< \delta$ is the equivalent condition because the Trace(Σ) is sum of the eigen values.*

Note 1: Though homogeneity could be defined for any color space, the definition is given for RGB color space here because the values for each of the three components vary in the same interval of 0 to 255, which is necessary to compute the variance covariance matrix.
Note 2: The choice of taking the sum of the eigen values is proper because of the property of the eigen value that the i^{th} largest eigen value of the variance covariance matrix gives the variance of the i^{th} principal component[9].

Definition 4: A possible set of line segment **(PSLS)** *is said to be a* **valid line segment (VLS)** *if it is a set of pixels with* **homogeneous** *color and the cardinality of the set is $\geq l$. l is the minimum number of pixels on a straight line segment.*

Mathematically, a line segment can be defined in terms of two threshold parameters l and δ, where l is the minimum number of pixels in the **PSLS** and δ is the maximum allowed value for the sum of the eigen values of the variance covariance matrix. With these parameters a mathematical way of defining VLS using definition 3(iii) is given below.
Definition 5: A **PSLS** $L(l, \delta) = \{(i_k, j_k), k = 1, 2, \ldots n\}$ *is said to be a* **VLS** *if*

1. *number of pixels on the line segment L(l, δ) $\geq l$,*

2. *Trace(Σ) $< \delta$ where, Σ variance covariance matrix.*

Definition 6: Two line segments $L_1(l, \delta)$ and $L_2(l, \delta)$ are said to be **connected lines** *if $L_1(l, \delta) \cap L_2(l, \delta) \neq \phi$ or $\exists$ any two pixels $P \in L_1(l, \delta)$ and $Q \in L_2(l, \delta)$ such that P is an 8-neighbor of Q.*
Definition 7: Let $\Re = \{L_k(l, \delta), k = 1, 2, \ldots .K\}$ is a set of K line segments. $\Re$ is said to be a **connected set of line segments** *if L_{k-1} is connected to L_k $\forall k = 2, 3, \ldots, K$.*
Definition 8: Let $\mathcal{L} = \{L_k(l, \delta), k = 1, 2, \ldots, K\}$ is a connected set of line segment and

$$\Re = \bigcup_{k=1}^{K} L_k, \quad L_k \in \mathcal{L}.$$

Then $\Re$ is said to be a **region** *if $L_k(l, \delta) \in \mathcal{L}$ is a* **valid line segment** *$\forall k$.*
Note 3: In definition 3, three ways of defining homogeneity are given. Among them, the definition which provides a combined affect of the three components R, G, B is the definition 3(iii). Secondly, the number of threshold values to be chosen is less for definition 3(iii) compared to definitions 3(i) and (ii). Thus, we shall confine ourselves to definition 3(iii) in the remaining portion of this article.

Figure 1. Original Kolkata Image taken from IRS-1A

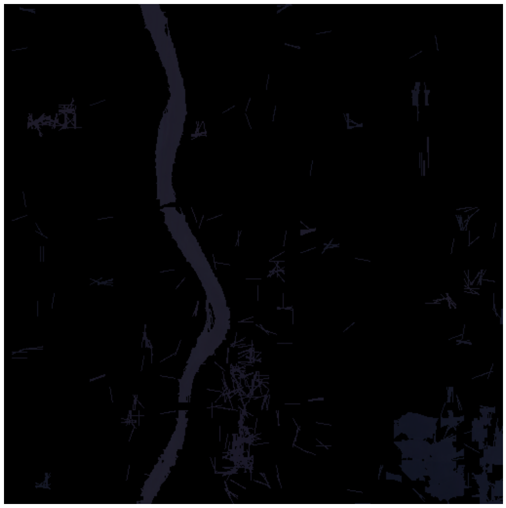

Figure 2. Extracted regions in Kolkata Image(IRS-1A) with w= 16, l= 14, δ = 1.2

4. Formulation and Implementation

HT can be implemented on the entire image directly but it is not wise to apply on the whole image at a time, because, it is not going to be computationally efficient. Additionally, it requires more memory space. If the size of an image is $N \times M$ and N and M are very large then threshold value l for the length of the line has to be chosen larger, thereby

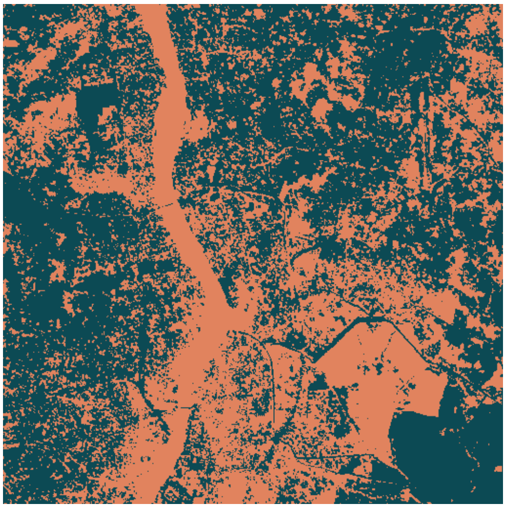

Figure 3. Extracted regions in Kolkata image(IRS-1A) by *K-means* **algorithm with** $K = 2$

eliminating finding smaller uniform regions. To avoid this problem a small window of size $w \times w$ needs to be moved on the entire image and valid line segments are to be found in each of the $w \times w$ windows. It may also be noted that suitable discretization of θ and ρ values needs to be done for computer implementation. Thus, the steps to be followed to extract valid line segments in a window are described below.

1. Consider a $w \times w$ window.

2. Construct the Hough accumulator space by transforming the each pixel of the window using different values of θ and their corresponding ρ values. Value ρ for each θ can be found using (1). The values of θ varies from 0° to $180^\circ - \theta_{step}$ with a step length θ_{step}. Note that θ_{step} is to be chosen in such a way that 180° is a multiple of θ_{step}. The obtained value of ρ is approximated to the nearest integer.

3. After constructing the Hough Space for a window, scan the Hough Space and consider those cells where count is greater than or equal to l and re-map this cell of Hough space to image space. This is a possible line segment. The cells with count $< l$ are simply ignored. In this way spurious line segments whose lengths are less than l are eliminated.

4. Compute Trace(Σ) of the color triplets of the pixels corresponding to each of the cells with count $\geq l$. If it is less than δ then consider it as a valid line segment else the line is ignored. This step eliminates those lines which do not possess homogeneous color pixels.

5. Restore the color triplet from the original image in a new image for the valid line segments detected in the above step.

Without considering the original color triplet in the image and for a fixed window size w, the possible set of line segments are same for all the possible windows in the image. Thus computing the steps (ii) and (iii) once is sufficient. This reduces the computation time significantly. In order to find regions in the image, line segments are found in all possible $w \times w$ windows. A file of these lines are made. This file of lines form a region if there is a homogeneous set of connected pixels present in the image. Thus in the process of finding the line segments regions are located in the image.

5. Results

The parameters to be tuned in this algorithm are w, l, and δ. w is the height (number of rows) and width (number of columns) of the window. The windows taken here are all square windows. l is the minimum number of pixels on a possible line segment and δ is the maximum allowed value of Trace(Σ). Values of these parameters are to be chosen properly.

In this section, results provided by the proposed algorithm on Indian Remote Sensing Satellites (IRS) are given. Two image frames, one from IRS-1A and one from IRS-1C, are considered for this purpose. A pixel in an IRS-1A image occupies, approximately, $36.25m \times 36.25m$ area on earth whereas, in IRS-1C, it occupies, approximately, $23.5m \times 23.5m$. Both the images have red, green and blue bands. In IRS-1A, the grey values of each band vary from 0 to 127 whereas in IRS-1C, they vary from 0 to 255. Each image is a 512×512 image. Since the original images possess poor intensity values, enhanced images of the originals are shown in [Figs.1 and 4] here for better clarity and understanding. Fig.1 shows enhanced images of the Kolkata city taken by IRS-1A and Fig.4 shows the enhanced image of a suburb named Barrackpore in the northern part of Kolkata, taken by IRS-1C.

The values considered for the length of the window(w) are 16 and 8. Minimum number of pixels on a line segment is considered to be 14 when $w = 16$, and is considered to be 7 when $w = 8$. The threshold value δ is taken to be a small value (*i.e.*, δ is $0.8,\ 1.2$ or 1.6) when the spatial resolution is coarser (*i.e.*, the image taken from IRS-1A), and the threshold is taken to be a larger value (*i.e.*, δ is $16, 24$ or 32) when the resolution is finer.

6. Analysis and Conclusion

In this article we proposed a method of extraction of regions from color images using HT. Definitions of uniformity of a region are provided. A region is defined by the union of several line segments. The method has been able to detect an arbitrary shape like the Hooghly river in the Kolkata images.

The results provided by the proposed algorithm are compared with the results obtained using *K-means* algorithm. Output of the *K-means* algorithm depends upon the initial partition of the data set, and thus the results may be different for different initial partitions. Here *K-means* algorithm is implemented taking five different initial partitions. The result which is closest to the ground truth is included here for each of the images considered. *K-means* algorithm is chosen for comparison because both the proposed algorithm and *K-means* with $K = 2$, result in classifying of pixels in two different clusters. The proposed method has been implemented with several sets of parameters as mentioned earlier on both the images. After the careful observation of the results obtained with these parameter values Fig.2 and Fig. 5 are found to be closest to the ground truth. The parameter set used to obtained these two results are ($w = 16,\ l = 14, \delta = 1.2$) and ($w = 16,\ l = 14, \delta = 24$) respectively. The results with other sets of parameters are not included here due to lack of space. If we observe the regions extracted in Kolkata images the river Hooghly is separated from the rest clearly by the proposed method. In Fig. 3(result found from *K-means* algorithm), river Hooghly has not been clearly separated from the concrete structure area present on both sides of the river. This is especially true for the southern part of river Hooghly [Fig.3]. The two bridges (Ravindra Setu and Bally Bridge) on the river are also not clearly distinguishable [Fig. 3], where as the proposed algorithm provides them [Fig.2 and 5].

The definition of homogeneity (definition 3) of a set of pixels is given with three different criteria and it is experimentally seen that all the three criteria give almost equivalent results visually. The results of using these homogeneity criteria are not shown here in the form of figures due to lack of space. Note that, with the criterion (iii), the algorithm needs less human intervention because it requires less number of parameters to be supplied. Criteria (i) and (ii) are strict in the sense that if the inequality fails for any one of the components, the set of pixels is declared as not homogeneous. In many situations, if the inequality fails for one or two components and if the over all variation is within a limit, the set of pixels can be said to be homogeneous. This is taken care of in criterion (iii) and thus it makes the algorithm robust. Thus the criterion (iii) in definition 3 is used to generate the results provided here.

Validity of a **PSLS** is controlled by tuning the parameter δ. Eigen values of the variance covariance matrix gives the variation present in the corresponding principal axes of the color vector. Thus, Trace(Σ) is the overall variation of color present in the set of pixels. This implies that the less

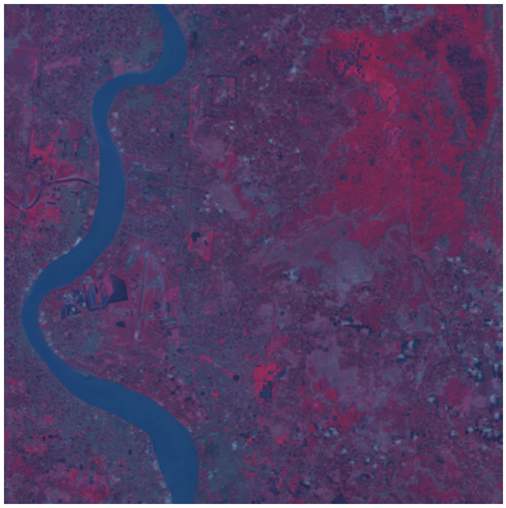

Figure 4. Original Barrackpore Image taken from IRS-1C

the value of Trace(Σ), less is the variation and more is the uniformity. In this way, by varying the value of δ, the validity condition (i.e a **PSLS** is a valid line segment) can be relaxed or tightened. Thus, by increasing the value of δ we are relaxing the validity condition for a line segment and increasing the chance of a line to be valid line segment. This justification is evident if we observe the results of the same original image with same w and l but different δ's. It is observed that by increasing the value of δ, the number of regions and the number of pixels in each region are increasing.

Connectedness and homogeneity of a set of pixels are the two properties, which are taken care of while deciding a set of pixels to be a region. Both of these properties do not depend on the maximum and minimum grey values in each of the components. Homogeneity condition is verified by checking the variation of color among the concerned set of pixels. Homogeneity condition only depends on the threshold value δ. The present work does not give a procedure to choose δ. This needs further research. The values of δ taken here are 1.2 for the IRS-1A images, and 24 for IRS-1C image. These values of δ are chosen experimentally. The value of Trace(Σ) of Kolkata and Barrackpore images are 29.756752 and 610.656483 respectively. These variances are computed over the whole image. It may be noted here that the values of Trace(Σ) in IRS-1A and IRS-1C images are significantly different. This may be the reason that the values of δ for Kolkata (IRS-1A) image is significantly smaller than the values of δ for Barrackpore(IRS-1C) image. The values of w are taken to be 16 and 8 because $w > 16$ may be too big whereas $w < 8$ may be too small for extracting regions. Value of l is taken to be 7 when $w = 8$ and l is taken as 14 when $w = 16$. The same value of these parameters are chosen also by Uma Shankar *et al*[10]. The value of w is taken to be a power of 2 though, in practice, w

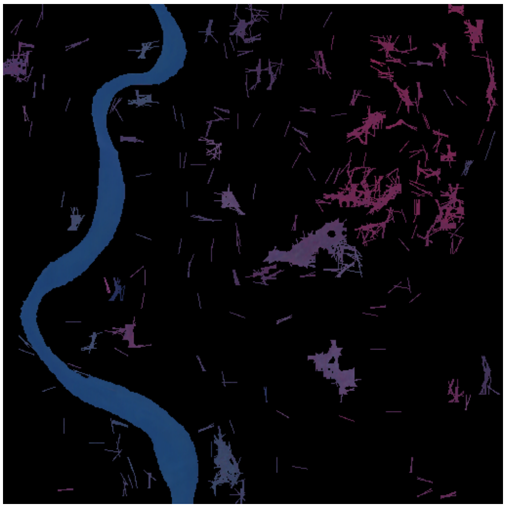

Figure 5. Extracted regions in Barrackpore image(IRS-1C) with w= 16, l= 14, δ = 24

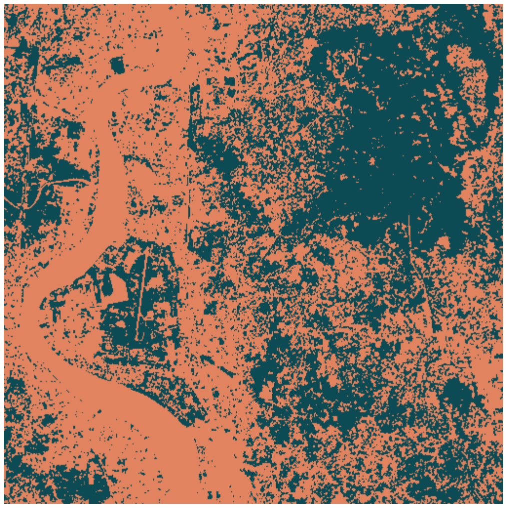

Figure 6. Extracted regions in Barrackpore image(IRS-1C) by *K-means* algorithm with $K = 2$.

can be any even number.

The proposed method fails to find an object if the object region contains many small uniform regions. Naked eye can find structures in an image whereas the proposed algorithm need not find them. Note that, the hexagonal portion of Kolkata image [Fig. 1] is not classified as a region by the proposed method. It is taken to belong to a single cluster by *K-means* algorithm [Fig. 3]. An increase in δ value of the proposed method would provide the hexagonal structure. But, it may be noted that an increase in δ would entail several other pixels to become part of a region, which need not be advisable always. Fine tuning of w, l and δ is a problem to be attempted in near future.

This article describes an approach to use HT to color images. Results are encouraging. Here line segments are obtained in color images using a particular definition of the word “Homogeneous”. Probably there may be other ways of defining homogeneity. In addition, given the concept of homogeneity, there may be other ways in which “Hough Transform” for binary/grey level images may be extended to color images. The approach described in this work need not be unique.

References

[1] R. Duda and P. Hart. Use of the Hough transform to detect lines and curves in pictures. *Commun. of the ACM*, 15(1):11–15, 1972 1972.

[2] J. Illingworth and J. Kittler. A survey of the Hough transform. *Comp. Vis., Grapics, and Image Processing*, 44(1):87–116, Oct. 1988.

[3] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall of India, 5th edition, 1999.

[4] A. Kesidis and N. Papamarkos. A window-based inverse Hough transform. *Patt. Recog.*, 33:1105–1117, 2000.

[5] A. L. Kesidis and N. Papamarkod. On the gray-scale inverse houg transform. *Image and Vis. Comput.*, 18(8):607–618, August 2000.

[6] V. Leavers. Which Hough transform ? *Comp. Vis. and Image Understanding*, 58(2):250–264, 1993.

[7] V. F. Leavers. The dynamic generalized Hough transform: Its relationship to the probabilistic Hough transforms and an application to the concurrent detection of circles and ellipses. *CVGIP: Image Understanding*, 56(3):381–398, Nov. 1992.

[8] R. Lo and W. Tsai. Gray-scale Hough transform for thick line detection in gray-scale images. *Patt. Recog.*, 28(5):647–661, 1995.

[9] C. R. Rao. *Linear statistical inference and its application*. John Elley & Sons, second edition edition, 1973.

[10] B. U. Shankar, C. A. Murthy, and S. K. Pal. A new gray Hough transform for region extraction from irs images. *Patt. Recog. Lett.*, 19(2):197–204, Feb. 1998.

[11] V. Shapiro. On the Hough transform of multi-level pictures. *Patt. Recog.*, 29(4):589–602., 1996.