Automatic Navigation through a Single 2D Image using Vanishing Point

Geetha Kiran A¹, Murali S²

¹Computer Science and Engineering, Malnad College of Engineering, Hassan, Karnataka, India
²Computer Science and Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India

Abstract - Image based navigation paradigms have recently emerged as an interesting alternative to conventional methods. This paper focuses on the problem of automatic navigation through Road scenes that mainly consist of a single vanishing point. The algorithm infers frontier information directly from the image to navigate through Road images. The major cue to terminate the navigation is the vanishing point. The proposed algorithm has 3 major steps: First, the preprocessing techniques are applied to the given image to find the vanishing point. Second, compute the distance from the ground truth position to the vanishing point which is used as the termination point for navigation. Finally, create the navigation by cropping the image. Our approach is fully automatic, since it needs no human intervention. The approach finds applications, mainly in assisting autonomous cars, virtual walk through ancient time images and in forensics. Qualitative and quantitative experiments on nearly 150 Real-road images in different scenarios show that the proposed algorithm is more efficient and accurate.

Keywords : canny edge detector, hough transform, vanishing point, termination point, video generation

1. Introduction

Navigation through a single 2D image is a challenging problem due to limited information from the input image. In Imaging devices, there is a trade-off between the images (snapshots) and video because of the limitation in storage capacity. Video clips need more storage space compared to images. This motivated to navigate through a single 2D image rather than storing video clips. Humans analyze variety of single image cues and act accordingly, unlike robots. The work is an attempt to make robots analyze similar to humans using single 2D image.

The task of generating video from photographs is receiving increased attention in many of the applications mainly with Road images. The application domain where the method can be applied are forensics and to assist autonomous cars by generating video from a single 2D image and assessing in advance - how far there is a straight road? If there is any suspected person or item in our path of journey, it could be detected prior and necessary action can be taken. We are addressing the key case where dimension of the real world object or measurement of object dimension in 2D plane is unknown. However navigation through a single image with the above constraints is very difficult because of the perspective view. Alternatively, navigation on a single image using proper ground known i.e., vanishing point could be easier. The work is carried out on outdoor images, mainly for Road scenes. We describe a unified framework for navigation through a single 2D image in lesser time. The input image may be easily acquired since no calibration target is needed or we can download Road images from internet.

This paper focuses on the automatic navigation through Road scenes that mainly consist of single vanishing point. Vanishing point is the major cue for obtaining the termination point for video generation. The approach can be used in a variety of applications, including forensics, virtual drive through the ancient images, to assist autonomous cars. It provides users with the important details available in the image while navigating.

The paper is organized as follows. In section 2 a review on the related works is highlighted. Section 3 gives description of finding the vanishing point from single view scene constraints and computing the distance from the ground truth position to the vanishing point. This is followed by the method of video generation in section 4. Implementation details are presented in section 5. Finally, some of the experimental results are presented in section 6.

2. Related Work

It is observed that some methods have been developed for detecting vanishing point on a single image, few which are directly relevant to the work are highlighted here. Techniques for estimating vanishing points can be roughly
divided into two categories. One requiring the knowledge of
the internal parameters of the camera and the other operates
in an uncalibrated setting. A large literature exists on
automatic detection of vanishing points, after Barnard [1]
first introduced the use of the Gaussian Sphere as an
accumulation space. He suggested that the unbounded space
can be mapped into the bounded surface of the Gaussian
sphere. Tuytelaars et. al. [2] mapped points into different
bounded subspaces according to their coordinates. Rother
[3] pointed out these methods could not preserve the
original distances between lines and points. In this method,
the intersections of all pairs of non-collinear lines are
considered as accumulator cells instead of a parameter
space. But these accumulator cells are difficult to index,
searching for the maximal from the accumulator cells is
slow. The simple calculation of a weighted mean of
pairwise intersection is used by Caprile et. al.[4].
Researches [5-7] have used vanishing point as global
constraint for road, they compute the texture orientation for
each pixel and select the effective vote-points, then locate
the vanishing point by using a voting procedure. Hui Kong
et. al. [8-11] have proposed an adaptive soft voting scheme
which is based upon a local voting region using high-
confidence voters. However, there are some redundancies
during the voting process and the accuracy on updating
vanishing point. Murali S et. al. [12,13] have detected edges
using canny edge detector and then the created edge is
subjected to hough transform. The maximum votes of first
N number of cells in the hough space is used for computing
the vanishing point. We use the similar framework in [12-
13] in our work to decide the vanishing point.
A very few Researchers have proposed different methods
for navigation through a Single 2D image. Shuqiang jiang
et. al [14] have proposed a method to automatically transform static images to dynamic video clips in mobile
devices. Xian-sheng Hua et. al [15] developed a system
named photo2video to convert a photographic series into a
video by simulating camera motions. The camera motion
pattern (both the key-frame sequencing scheme and
trajectory/ speed control strategy) is selected for each
photograph to generate a corresponding motion photograph
clip. A region based method to generate a multiview video
from a conventional 2-dimensional video using color
information to segment an image has been proposed by
Yun-Ki-Baek et. al [16]. Na-Eun Yang et.al [17] have
proposed method to generate depth map using local depth
hypothesis and grouped regions for 2D-to-3D conversion.
The various methods of converting 2D to stereoscopic 3D
images involves the fundamental, underlying principle of
horizontal shifting of pixels to create a new image so that
there are horizontal disparities between the original image
and the new version. The extent of horizontal shift depends
on the distance of the feature of an object to the stereoscopic
camera that the pixel represents. It also depends on the inter-
lens separation to determine the new image viewpoint.
The methods proposed by the authors for detecting
vanishing points have made certain assumptions specific to
the application. These artifacts are not of much importance
in our work, this made us to propose a new method as
proposed in [12,13], which decides the vanishing point in
lesser time. Using the vanishing point, the distance from the
ground truth position to the vanishing point could be
computed. This helps in navigating through the single Road
image.

3. Vanishing Point Detection

Images considered for modeling are perspective. In a
perspective image, lines parallel in the world space appear
to meet at a point called Vanishing point. Vanishing points
provide a strong geometric cue for inferring information
about 3 dimensional structure of a scene in almost all kinds
of man-made environment. There are methods available for
detecting vanishing points with known camera parameters
and also with uncalibrated setting. The method described in
this section requires no knowledge of the camera parameters
and proceeds directly from geometric relationships. The
steps involves detecting edges using canny edge technique
to identify the straight lines depending upon the threshold
fixed by the hough transform, compute the vanishing point
using the intersection points of the lines. The above steps
have been explained in the subsequent sections.

3.1 Line Determination

The given color image is converted to gray. Lines are edges
of the objects and environment present in an image. These
lines may or may not contribute to form the actual vanishing
point. The existence of the lines are obtained by applying
the canny edge detection algorithm. The versatility of the
canny algorithm is to adapt to various parameters like the
sizes of the Gaussian filter and the thresholds. This will
allow it to be used in detecting edges of differing
characteristics.

For an image as in Figure 1(a), after converting it to gray
(Figure 1(b)), the edges are detected by applying Canny
dge detection algorithm. A set of white pixels containing
edges are obtained and the rest of the contents of the image
are removed. A Canny edge detected image is shown in
Figure 1(c). This image contains pixels contributing to
straight lines and also other miscellaneous edges. Considering all these pixels of the edges contributing to the
straight lines, Hough transformation is applied on the image.
The result is shown in Figure 1(d).
As the outcome of the Hough transformation, a large number of straight lines are detected. These straight lines depend upon the threshold fixed up for the Hough transformation. Points belonging to the same straight line in the image plane have corresponding sinusoids which intersect in a single point in the polar space (Figure 1(d)). The need for calculating the number of straight lines is that there could be several straight lines in the image which intersects each other at different points in the image plane.

In such case there arises a situation that more than one peak value in the polar space is obtained. Thus by selecting the number of peak values (in descending order of their votes) equal to the number of straight lines ‘N’ present in the image we restrict the unwanted lines which may not contribute to the real vanishing point. This reduces the computational complexity of vanishing point detection to only the number of straight lines contributing the possible vanishing point.

### 3.2 Intersection Point of any Two Lines

Lines drawn by Hough transformation are on edges of the object and environment in an image. These lines may or may not contribute to form the actual vanishing point. Depending upon the number of lines present in the image, the number of peaks in the Hough space is fixed up in a descending order of their occurrences. Each peak in the hough space signifies the existence of a longer edge in the image than any other points in the Hough space and hence a peak is formed. These peaks of the voted points of the hough space are calculated to find the intersection between two lines to calculate the vanishing point. Finding the intersection points for all combination of lines selecting two at a time, corresponding one (x,y) pair is obtained. The number of pairs of x and y values obtained for all combinations is given by the relation

\[
Nc_2 = \frac{N!}{((N-2)! \times 2!)}
\]

where N is the number of peaks selected. These (x,y) pairs are the probable vanishing points. All of them are within the vicinity of the actual vanishing point. We have taken the mean of the probable vanishing points (Figure 2(a)-blue color), since they are within the vicinity of the actual vanishing point. In our work vanishing point is used to find the distance from the ground truth position to the detected vanishing point position. The distance obtained is used in the next section to facilitate the termination point for the navigation.

The distance of the Road identified could directly be used to decide the number of frames to be generated, generally 1:2 depends on the length and it can be varied with requirements.
Figure 2. (a) Lines detected (white/green color), Probable vanishing points (black/blue color) (b) Vanishing point (black/pink color)

4. Navigation through a Single Image

The information obtained from the previous section is used to navigate through a single Road image. The input for the navigation are - single 2D image, computed termination point based on the distance from the ground truth position to the detected vanishing point. Based on this strategy, the frames for navigation are generated by cropping the image based on the size of the image up to the computed distance. The input image is considered as the first frame and the image is cropped based on the size of the predefined rectangle. Then the cropped image is resized to the original image and stored in an array of images. An appropriate set of key-frames are determined for each image based on the distance computed by using vanishing point. The images obtained after cropping is given in Figure 3.

Figure 3. (a) 1st Frame (b) 40th Frame (c) 80th Frame

Further navigation is based on the key frames stored in the array by writing the frames to the video file. This method provides vivid dynamic effect from global view to local details.

5. Implementation

Navigation through a single 2D Road image is a challenging task due to unknown z-coordinate. We have attempted to
automatically navigate through a single 2D Road images based on geometric relationships. The navigation process is decomposed into two parts as follows:

- To find the distance from the ground truth position to the detected vanishing point.
- To navigate through a single Road image based on the determined path.

The entire algorithm of the proposed method to generate video is as below:

1. **Step 1**: Canny edge detector is applied to the normalized image which yields a binary image with only the edge information.
2. **Step 2**: The obtained image is then subjected to Hough transform.
3. **Step 3**: Select only the maximum votes of first N number of cells in the Hough space for further calculation.
4. **Step 4**: Calculate the intersection points for all combination of lines selecting two at a time. These are the probable vanishing points.
5. **Step 5**: Take the mean of the probable vanishing points.
6. **Step 6**: Find the distance from the ground truth position to the detected Vanishing point position.
7. **Step 7**: Crop the image based on the size of the image.
8. **Step 8**: Cropped image is resized to the original image and stored in an array of images.
9. **Step 9**: The steps 7 and 8 are repeated until the termination point as computed by using vanishing point.

### 6. Experimental Results

The algorithm is applied to a test set of nearly 150 images obtained from different real-road images in different scenarios that mainly consist of single vanishing point, we have observed that the results presented are indicative of the algorithm performance. The images used in the experimentation are downloaded from internet and few of them are self captured. The steps involves detecting edges using canny edge technique, to identify the straight lines, compute the vanishing point using the intersection points of the lines. All of them are within the vicinity of the actual vanishing point. Based on the ground truth position, we compute the distance from the ground truth position to the computed vanishing point. We also have evaluated the algorithm by manually detecting the distance from the ground truth value to the vanishing point and compared it with the distance generated by our method. Experimented images have an error distance not more than 8 pixels as shown in Chart 1.

[![Chart 1. Comparison of vanishing point accuracy.](chart1.png)](chart1.png)

The first, intermediate and final frame generated by the method after finding the termination point is shown in Figure 4. We can observe the finer details in the intermediate and final frames that could be used in various applications including virtual walk through ancient time images, in forensics and in automated vehicle. The painting follows the geometric rules and also have color variations and therefore we can apply the methods developed here to have a virtual walk in the imaginary world.

### 7. Conclusion

An algorithm to navigate through a single 2D image is proposed and experimented for only Road Scenes. The work is experimented on nearly 150 images in difficult scenarios. This paper provides a solution to transform static single 2D image into video clips. It not only helps the users to enjoy the important details of the image but also provides a vivid viewing manner. The experimental results show that the algorithm is performing well on a number of outdoor scenes with Road. Further work may be extended to include investigating on more reliable Region Of Interest (ROI) detection techniques. Even finer details can be obtained from the key frames used in video generation. The work is done in view of assisting the automated vehicle at low cost.
References


Figure 4. (a) First Frame (b) Intermediate Frame (c) Last Frame