COP507 - Computer Vision & Embedded Systems

B717426

29th December 2017

Contents

1 Introduction

Make and Model Recognition (VMMR) has been a subject area of interest within computer vision both in industry and academia. The importance of the ability to recognise a variety of vehicles spans across different disciplines and applications such as driver assistance, traffic monitoring, law enforcement and surveillance. As stated by Dehghan et al (2017), this categorisation task has been a problem for classical computing however, a result of great advances in artificial intelligence, this task has been made possible with accuracy rates above 90 percent. The task given was to design and implement a VMMR system with the capability to recognise the make and model of a vehicle based on only the front view (that is, headlights, grill and bumper). Based on the image of a vehicle fed as input, the output of the system should be the make and model of the vehicle. This report explains a practical use of the VMMR system that was implemented, furthermore, it explains the design of the system and any performance restrictions or performance enhancements that will be expected should the system implemented in hardware.

There some basic processes that are essential for any computer vision application. These include pre-processing the acquired images, feature detection and feature extraction. Pre-processing involves making the image data conducive for feature extraction. Some of these processes includes noise reduction, morphological operations, colour correction and image re-sampling. Feature extraction is the process by which key points are obtained from the image. These key points consist of blobs, edges and corners.

2 Preprocessing Tasks

This section discusses the pre-processing tasks that are required to implement the Vehicle Make and Model System in practical scenarios.

2.1 CCTV camera located at the side of a motor way

As a consequence of this system being implemented outdoors, there are certain factors that cannot be controlled prior to the CCTV footage being taken. These factors include illumination, motion blur, problems with focus and the camera being placed at an angle which affects processing of the image. The following subsections explain solutions to these problems found when processing the images taken by the camera.

2.1.1 Skewing & Rotation

Considering the camera is at the side of the motor way, one thing to consider is that the angle at which the camera is placed. Therefore, before performing any feature detection or extraction, the images taken from that camera must be set to the right orientation. This could be done by skewing the image or rotating the image as this will set the image to the right orientation. Traditional mathematical algorithms and formulas exist which can be used to correct the angle of rotation or the degree of skewing required. Considering that since an image can be represented as a matrix of pixel values, affine transformation matrices could be applied to the image to correct the problem (Gonzales & Woods, 2008).

Before these algorithms can be applied to the images, it must be detected whether there is a need to apply skew correction to the images captured. Patel et al. (2015) discuss an approach for skew angle detection and correction based on the radon transform. The radon transform was used because of its accuracy and low computational cost. The radon function calculates projections of the image matrix in particular directions (Patel et al, 2015). For an image matrix $f(x,y)$ a projection is considered as a collection of line integrals. The line integrals are calculated from various sources along beams, lateral paths, in a specific direction. The skew detection algorithm proposed by Patel et al (2015)

is as follows. First of all, the input image is binarised and morphological filling operations are applied. After that, edges are detected using the canny edge detector. This process is used for locating relevant transitions within the image. The radon transform is then applied on the output of the edges detected. Next, a maximum value is searched for within the radon transform coefficients. After that, a skew angle is found from the maximum intensity value that was found from the radon coefficients. The image is then rotated with the skew angle.

2.1.2 Brightness & Colour Correction

Another factor to consider is the lighting conditions. The reason being that, in the day time, pictures taken may or may not have the appropriate lighting. In the night time, the images taken may be very dark. Therefore, in order to make the images conducive for training, the images will need to be made brighter. A remedy to this problem could be applying a brightness adjustment algorithm. According to Szeliski (2010), this can be done by multiplying or adding the image by the constant.

$$g(x) = \alpha f(x) + \beta$$

Where α and β are parameters which influence the contrast and brightness of the image and where α is greater than zero.

Alternatively, colour balancing and correction algorithms can be used to improve the illumination and contrast of the image. Gasparini and Schettini (2004) introduce a method of unsupervised color correction for removing colour cast.

2.1.3 Motion blur removal

Motion blur is a phenomenon which occurs when there are inconsistencies between the speed of the moving object and the rate at which the camera captures frames. The problem with motion blur is that, it attenuates the image signal. Removing the motion blur from the image involves two processes. Firstly, estimating the amount by which the image has been blurred and thereafter, recovering a more realistic image using deconvolution (point-wise division in the frequency domain). Cho (2010) proposes a method of deconvolution referred to as iterative distributed reweighting (IDR) for removing motion blur. This

method enhances the visual quality of reconstructed images as compared with traditional deconvolution algorithms used for removing motion blur.

Cho (2010) suggests the estimation of the blur kernel by examining edges within the blurred image. The edges within the blurred image encode estimations of the blur kernel from which the blur can be removed using the inverse radon transform. Using the edge information in this case will yield desirable results because, the object of interest is a man made object and man made objects consist of many edges. As a result of the loss of high frequency information, prior information concerning natural images is used to fill out the missing information. This is done using content aware image priors rather than sparse priors because, content aware prior adjusts its features to local textures and hence, it increases the quality of textures within the reconstructed images (Cho, 2010). After this has been completed, a maximum a posteriori (MAP) estimator is used to restore the image using piecewise smooth features, however, since the MAP does not always accurately reconstruct rich textures, IDR is used to improve the depiction of rich textures.

2.1.4   Shake Removal

3   Reference List

Cho,T.,S. (2010). *Motion blur removal from photographs.* Submitted to the Department of Electrical Engineering and Computer Science in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering and Computer Science. Cambridge Massachusetts: Massachusetts Institute of Technology.

Gonzalez, C., R., Woods, E., R. (2008). *Digital Image Processing.* Upper Saddle River New Jersey: Pearson.

Patel, J., Shah, A., Patel, H. (2015). Skew Angle Detection and Correction using Radon Transform. *International Journal of Electronics, Electrical and Computational System.* [Online] Volume 4 (May) Available from: `http://academicscience.co.in/admin/resources/project/paper/f201505101431314896.pdf`. [Accessed: 2nd January 2018].

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications.* New York: Springer.