UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

Corso di Laurea Magistrale in Informatica

Vision system for autonomously driven vehicles

Summary

Relatore: Prof. Domenico G. Sorrenti
Correlatore: Dr. Axel Furlan
Correlatore: Dr. Daniele Marzorati

Candidato: Roberto FERMI
Matricola: 048091

Anno Accademico 2009 – 2010

Summary

The aim of this thesis is to realize a vision system that can be applied to autonomously driven vehicles, to determine the vehicle/robot position in the world and recognize the presence of obstacle in its field of view.

Nowadays the most promising approach for computer vision is the use of cameras. Information generated by cameras images permit to have a global vision of world objects and permits to distinguish, after the application of specific image processing techniques, various typologies of objects, like people, cars, trees, etc. On the opposite side, as the real world points are projected onto a bi-dimensional plane, there's a loose of one dimension which is depth. When a camera image is captured, indeed, we're not able any longer to determine at which distance the objects represented in the image were. If we knew the object dimension, this aspect wouldn't be a problem, but this is not conceivable in this application, because of the big differences between different the objects in the scene.

To solve this problem trials to simulate what normally happens in human sight, which uses two images of the same scene taken from two different points of view are performed. In computer vision this is achieved using two cameras. The two images taken from both cameras must be referred to the same time and thus the shutter command must be properly synchronized, to avoid watching two scenes having nothing in common. Both cameras must also be calibrated: in other words, it is necessary to identify how they are placed one compared to the other. The calibration is used to make it possible to determine, for the first camera, the correspondence to the points in the second camera.

Before doing any high level image elaboration (like, for example, object recognizing and categorizing), using images captured from cameras, it is necessary to have enough image quality. Different response to light intensity in various acquiring conditions, electrical noise, dynamic range width and contrast are just a few parameters

that characterize the sensors and images viewed in this thesis. To compute mathematical analysis it is necessary, indeed, that in every acquisition condition, objects are well contrasted and defined, which means they should be well distinguishable the one from the others and also from the background. If in a controlled ambient like a laboratory or a well illuminated room this can be easily achieved, in real conditions it is really difficult. For example, let's think what happens when driving a car on a street in a sunny day, while approaching a tunnel. In this condition, it is sometimes very difficult also for human eyes to adapt to see correctly the details exposed to direct sunlight and the details in the dark, inside the tunnel at the same time. This is a situation of high dynamic range, where the dynamic is the image luminance. For state of the art cameras, in these conditions it is very difficult to obtain good quality images and it is rarely possible to have usable images. If it is not possible to have them, it will not be possible to determine obstacles that precede the vehicle and it will not be possible to understand what happens in the world surrounding the vehicle itself. For this reason, it is extremely important to have a good output from both cameras.

One interesting aspect in computer vision applications is the choice to adopt camera fixed parameters (like, for example, exposure time, aperture, digital gain, etc.) or to make them variable during the time, in relation to the specific surrounding conditions of the ambient where we're moving. The first approach has the main advantage of avoiding background computation scene analysis, but, generally, it has big limits in its usage. In this thesis, it will be shown how the use of particular digital sensor technologies, combined with image processing techniques, has made possible this scenario, allowing the use of camera fixed parameters and, at the same time, to have an excellent image contrast in every exposure condition, from the darkness of the night to the heavy backlight of a sunny day.

We will also show how it is possible to modify just one camera parameter (the exposure time) with background analyses, which take into account, besides particular scene illumination conditions, also specific indicators that permit to define how much motion blur is present in the images. This permits to determine if the exposure time is too high for what is happening in the world observed by the cameras. Indeed, in these circumstances it is important not to have a blurred image of the scene by compensating the lack of exposure time (with low exposure times images are darker) with post processing techniques capable of increasing image brightness.

In this thesis it will be presented the CPBD metric (Cumulative Probability of Blur Detection) that is based on the idea of Just Noticeable Blur (JBN), which is the minimum relievable difference from human between two brightness levels in the image. This metric permits to do psychometric evaluations about the amount of image blurriness, allowing a supervised control of cameras exposure time.

There are many post-processing techniques and the most appropriate for our project is the pixel-binning. This consists in grouping (or more precisely summing) the energy stored inside each contiguous photosensitive element in camera sensors, to obtain a bigger light intensity in poorly illuminated scenes, at the expense of image resolution. Another very useful technique diffused also in other subjects, other than computer vision, is the histogram equalization that allows to redistribute equally light intensity levels in all available range. This has the advantage of increasing the image contrast. In this thesis it will be shown a particular variant of the histogram equalization technique that allows to obtain better results by saturating a predefined amount of image data. This predefined value has been experimentally determined.

This work has focused on the following steps:

- **State of the art analysis**. This first part of the project has required the research and the analysis of available technologies to identify the most suitable for us.
- **Stereo head creation**. During this part of the project we have dealt with the physical stereo head realization. It has been necessary to identify which mounting position would have been the most appropriate. We also have had to find a way to

make cameras firmly established with USAD structure. Next step has been the realization of the stereo head.

- **LinLog technology analysis**. It has been necessary to realize a specific program to read the private parameters inside the camera about LinLog technology settings. Then we have performed some exhaustive experimental tests to identify which parameters gave the better results.
- **Low level image processing techniques identification**. In this part of the project, we have chosen which techniques to apply to captured images. We opted for pixel binning and histogram stretching with stretchlim.
- **C-language implementation**. The next step has been the writing of C code to implement the techniques shown before. Then, it has been necessary to integrate all this code with DAFNE library developed by D.R. Marzorati that includes, inter alia, the disparity map computation.
- **Identification of fixed camera parameters.** After the implementative part, we went through the experimental identification to understand which parameters would perform the best in fixed parameters camera settings.
- **Experimental verification of the results.** After the integration with DAFNE, we proceeded to an exhaustive verification process of the obtained results. Tests have been performed in different acquisition conditions, in closed spaces and in open spaces. We chose to test the system affordability also in different weather conditions, like snow, night and during the sunset with heavy backlight.
- **Blur detection techniques analysis**. Although good results were obtained in fixed parameters camera settings, we've done an in depth research about the state of the art concerning the image blur detection techniques.
- **Choosing and testing blur detection metric.** During this part of the project we have evaluated the effectiveness of the different blur detection metrics previously analyzed and we have chosen the CPBD metric, developing a first version using Matlab

- **Experimental evaluation of blur detection metric.** We performed tests based on previously captured dataset to identify the effectiveness of CPBM metric in relieving and identifying the motion blur.

At the end of this project, experimental results have been analyzed.

LinLog technology by PhotonFocus is, indeed, very promising and, under some circumstances, the results obtained have been impressive.

However, in order to use it correctly and also to be able to identify a set of parameters that allowed to have good results, required both a big research effort and a big practical effort for testing this technology in every condition.

To allow the best results from captured images, we choose to sacrifice image resolution using the pixel binning technique, in order to have bright images also in very dark capture conditions.

Stretchlim function for histogram equalization brought well contrasted images even when the traditional histogram equalization technique was unable to modify the original images.

Concerning the ambitious project of having a fixed parameter camera setting for each scene illuminating condition, it is possible to say that we reached excellent results, despite having accepted a loss in image resolution. However, the chosen exposure time could not be appropriate for higher speed than the one reachable from USAD, or in high movement scenes.

Thus we proceeded to test the images motion blur detection metric. Specific tests were performed on images got from USAD stereo-head. We realized that CPBD metrics permits to identify in a very precise way the level of blurriness of an image.

This project has a large number of further developments and the main ones are the following:

- Pixel binning technique can be brought to be dynamic by using it, for example, in a less strong way (using a 3x3 or a 2x2 binning) when not needed. This should be done to obtain a higher image resolution when the scene illumination is enough and on the other hand to decrease the resolution, using a more aggressive pixel binning, when the light is too low.
- Algorithmic optimization of disparity map computation can be improved, trying to find techniques that allow for a reduced computational time
- It is necessary to write in C-code the algorithm of the blur-detection metric, presented in this thesis
- It is necessary to create a background process that performs the feedback on the camera exposure time, basing the analysis on the mean illumination intensity and on the blur detection metric
- The information obtained with the disparity map and the free space map, can be used to drive in an autonomous way, interacting with ROS and the other navigation sensors.