Stereoscopic survey-Application to traffic

***Abstract*—Stereo vision is one of the main researched domains of computer vision, and it can be used for different applications, among them, we can find extraction of the depth of a scene, estimating the distance between the camera and other objects and then estimate the distance between objects which we are going to see in this article.**

Stereoscopy is a computer vision discipline used for recording and representing stereoscopic images (3D) . It can create an illusion of depth using two pictures taken from slightly different angles. There are actually two possible ways to take stereoscopic pictures : first, by using special two-lens stereo cameras or systems with two single-lens cameras joined together, The 3D camera consists of two cameras of parallel optical axes and separated horizontally from each other by a small distance and these two cameras are combined together in a single frame in order to get a 3D image. The 3D camera is used to produce two stereoscopic pictures for a given object. The distance between the cameras and the object can be measured depending on the distance between the positions of the objects in both pictures, the focal lengths of both cameras and the distance as well. In this article, we provide a comprehensive survey of this continuously growing field of research, summarize the foremost commonly uses of stereoscopic cameras especially its applications to traffic, and discuss their benefits and limitations. In retrospect of what has been achieved so far.

***Index Terms*—Distance, stereo vision, template matching, stero cameras Depth estimation.**

I. Introduction

About ninety percent of human perception of its surroundings is done through the visual sense. Since that the amount of information is sent to the brain through the eyes and processed are several times the volume of information that reaches the brain through other sensory organs.

Stereo vision is suitable for many applications for various reasons. The most important one is the similarity to binocular vision in humans and animals. Stereo vision consists of two identical cameras, which allow obtaining pictures from two distinctive viewpoints. Depth is determined by finding the disparity in the images of the same 3D points.

Applications of stereo vision are numerous. They range from entertainment (videos, games) to more specialized applications such as the educational ones and medical applications like body exploration, therapeutic purposes, traffic application, and so forth.

One of the most important stereoscopy tasks is definition of the 3D coordinates of a randomly chosen point on the observation scene of a stereoscopic system. In order to get reconstruct 3D images from stereo images, an accurate calibration is crucial.

The paper outline is as follows: After an overview of the differences between monocular and stereoscopic cameras systesms , calibration techniques are introduced. Consequently, analysis of the stereo vision applications is conducted. The paper closes with a dissucssion followed by a conclusion.

II. Comparison of monocular and stereo-cameras systems

A. Monocular vision

Mono vision systems are based on the classification of objects, meaning that they must be “trained” to detect how certain objects look. Those objects are classified as “classes”. Classes may include pedestrians, cyclists, passenger cars, trucks, etc. each in its own separate class. For each captured image, an algorithm tries to match one or more of the “classes” in its library. If the system finds a high accurate match, it can estimate the distance to the known object based on the previous knowledge it has of how those objects are supposed to appear for each distance. If it doesn’t find any match, the mono vision system will proceed as if there were no objects at all.

This is the main drawback in mono vision system technology. Sure, it does wonders and looks like magic thanks to AI and deep learning advancements, but the moment it runs into an unclassified object, it’s practically blind and acts as if there is nothing. This was demonstrated numerous times in reported accidents of semi-autonomous vehicles using monocular cameras to avoid obstacles or to alert the driver of an upcoming obstacle,so the impossibility to obtain an accurate distance between the camera and other objects.

For this reason, mono vision systems cannot be used as the sole technology sensor in future autonomous vehicles and estimating the real distance between objects. These systems must be combined with other technologies such as LiDARs and radars.

Fig. 1: Monocular camera example

B. Stereoscopic vision

Stereo vision is a technique for building a three dimensional description of a scene observed from several viewpoints. It is considered passive if no additional lighting of the scene, for instance by laser beam, is required. So defined, passive stereovision happens to be very attractive for many applications in robotics, including 3-D object recognition and localization as well as 3-D navigation of mobile robots.

Most of the research on passive stereo vision has been devoted to binocular vision for which two cameras are observing the same scene from two slightly different viewpoints. As soon as two image points are matched, i.e., identified as corresponding to the same physical point, it is possible to compute the three-dimensional coordinates of this physical point.

The certain object is selected on the left picture, while the same object on the right picture is automatically detected by optimisation algorithms techniques which searches for minimum difference between both pictures. The calculation of object's position can be calculated by doing some geometrical derivations. The accuracy of the position depends on picture resolution, optical distortions and the distance between the two cameras.

In stereo video/images you have more information per frame/image allowing for creating a 3D presentation of the image/video signal (depth). You may create mono images/video from the 3D. You are not able to create a 3D presentation from a mono image/video signal. That the reason that old movies, that were not captured with two cameras from two slightly different direction are not convertible into 3D

Fig. 2: Stereo cameras examples

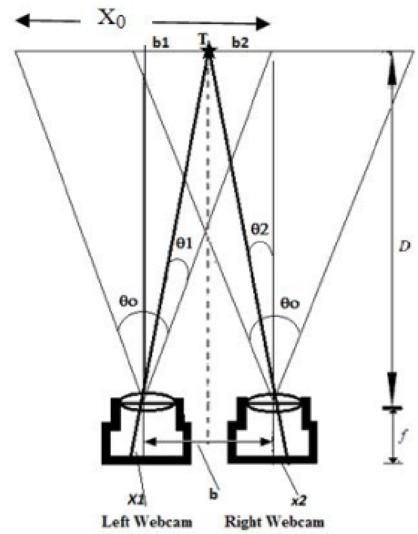

Fig. 3: Schematic diagram of an object pictured by two horizontally aligned cameras separated by small distance (stereoscopy).

The diagram of fig. (1) shows two cameras, left camera and right camera aligned so that their optical axes parallel and at a distance b from each other. The two webcams have the same parameters, i.e. the same **focal length f** and the same view angle θ_0 image of the target T will be at distance x1 in the left camera and at a distance x_2 in the right camera. Applying basic triangulation will lead to the following equations:

$$\frac{b_1}{D} = \frac{-x1}{f}$$

,

$$\frac{b_2}{D} = \frac{x2}{f}$$

Since $b = b1 + b2$, then,

$$b = \frac{D}{f}(x_2 - x_1)$$

$$D = \frac{bf}{x_2 - x_1}$$

$$\tan\frac{\theta_0}{2} = \frac{x_0}{D} = \frac{x_1}{f}$$

$$f = \frac{x_0}{2\tan(\frac{\theta_0}{2})}$$

$$D = \frac{bx_0}{2\tan(\frac{\theta_0}{2})(x_2 - x_1)}$$

In this equations above, x_0 is the width of the image in pixels, $x_2 - x_1$ is the disparity between the two images in pixels. Since the distance between both webcams b, the webcam angle view θ and image width are constant for 3D webcam, then it seems that the distance D is inversely proportional to disparity $x_2 - x_1$.

III. Stereoscopic cameras calibration

The calibration in general is a process that can determine the relationship between the values of the measured quantity indicated by the measuring instrument and therefore the corresponding values of physical quantities. The calibration is extremely important and is required for several activities and tasks. Without the process of calibration, it is impossible to obtain a reference to actual values and correct interpretation of the results. For the same reason, the cameras require a calibration process.

The stereoscopic cameras calibration is the process, which provides a multiple parameters that define and relate to the specific characteristics of the camera . These are critical during extracting 3D information from 2D images, measure object size in global units, visual-odometer process, reconstruct a 3D scene and many applications in computer vision.

A. *State of the art of calibration methods*

In the last years, scientists and researchers have studied various methods of calibrating cameras. This section provides an overview and in-depth analysis of some of these methods. In Fig 2, we categorize the most crucial methods due to their limitations.

This section presents the state of art of calibration methods, which use a calibration object in order to work. Many computer vision applications use them. This kind of group is popular when engineers assume that the camera parameters are fixed. Various assumptions are obligatory depending on a specific method. Every calibration method requires well-synchronized images provided by cameras.

A very important aspect, is that the scenery must have suficient light, in order to detect many features in the images so the callibration resultat can be more accurate. Another very important aspect is the overlapping view of both cameras. Both cameras must simultaneously observe the larger part of the image (scene).

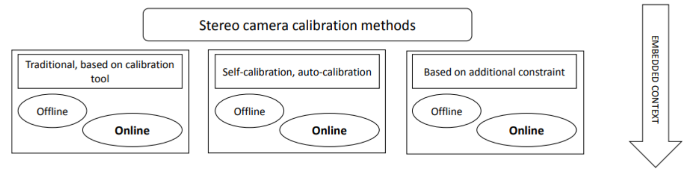

Fig. 4: Global calibration methods characterization

B. *Traditional Camera Calibration*

In this section, I'm going to introduce one of the most popular approaches of the traditional camera calibration methods. These methods require a special calibration tool or pattern to work .Fig. 3 shows us the four different shapes of this particular tool.

The pattern must be clearly visible (the whole tool) in the camera view from many different positions in relation to the static position of the camera. There is a reverse option in which the calibration tool has a stable position and it is observed from multiple camera positions (orientation). Depending on the different methods with a particular pattern, a various number of images from many perspectives is required. The traditional camera calibration method usually must stop application, system or current task to realize this procedure. Then, it must end and detect a calibration pattern, from many different views. Therefore, the traditional calibration method does not aspirate to have a potential to run in real time (during task). The procedure has to know the size of the calibration pattern. Each traditional method extracts clearly the coordinates of the calibration object.

Finally, this method provided in this section usually estimates all camera (intrinsic and extrinsic) parameters. Linking the image views obtained from cameras with the particular calibration patterns can provide very accurate results not achievable by other methods.

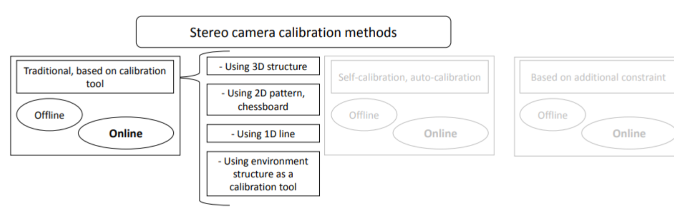

Fig. 5: Traditional calibration methods characterization.

In the market, we can find some applications to calibrate stereo cameras. One example is the Matlab stereo camera calibration toolbox .However, the license of such software is paid . The other one, the OpenCV is open source software, recommended by . Both applications have implemented the Zhengyou Zhang method. There is much more application available on the market. However, in this survey, we use these two as the reference methods.

The stereo cameras could be calibrated based on the traditional method. However, this kind of approach is not universal and not perfect due to fact that the calibration pattern is always required. which is not the case for every scene.

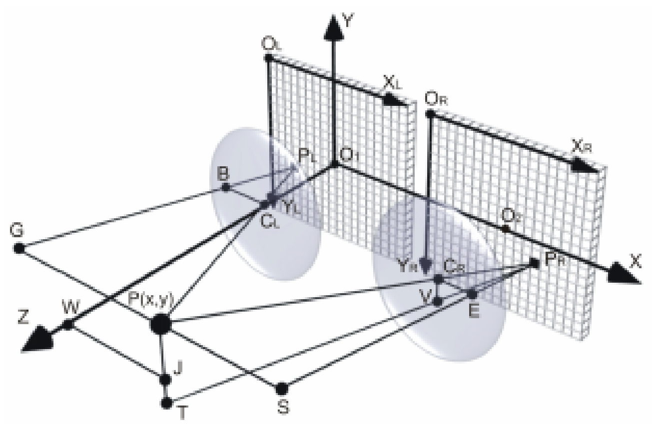

Fig. 6: 3D-view of geometric scheme of the stereoscopic system.

C. Self-Calibration methods

Self-calibration or auto-calibration method is the second callibration methods that we will discuss in this section,the procedures to use self-calibration methods do not make usage of any particular calibration object unlike the first method. the researchers try to get rid of all calibration patterns employed in the traditional callibration methods.

Fig 4 presents a number of different methods that can compute the camera parameters with computationally fexible approach and different constraints. in addition, it is possible to use those methods online, while performing other tasks. The most popular strategies such as bundle adjustment (BA) optimization or epipolar constraint are detailed explain in the following sections.

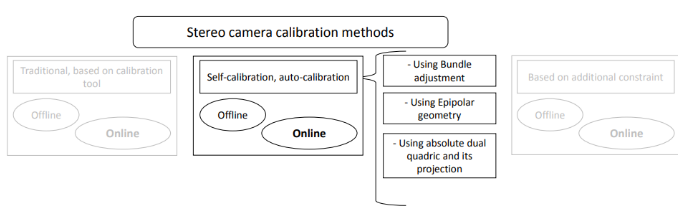

Fig. 7: Self-calibration methods characterization.

1) Bundle adjustment camera calibration methods : The camera calibration can be achieved by optimization of POI position from different camera poses. The bundle adjustment (BA) method is the one of procedure which can realize it. This technique is adapted in the eld of computer vision, where it calculates and optimizes the positions of multiple 3-dimensional POI from dierent view of the observer. Globally, the BA refers to a visual reconstruction where it creates the optimal 3-dimensional structure of the scene geometry. In addition, it can estimate the relative motion and vision parameters (camera position and/or calibration).

For the multidimensional optimization problem, the BA nds a set of camera parameters by minimizing the projection error between the measurements (all POI in each frame) and the predicted 3-dimensional position of observed points. The main diculty is to nd optimal parameters by minimizing cost of functions due to the scale problem. More precise camera parameters allow to minimize the error of POI positions and estimate the position of future frames and compare it with future measurements.

IV. Stereoscopic cameras applications

Stereo vision technology is used in a variety of applications, including people tracking, mobile robotics navigation, and mining. It is also used in industrial automation and 3D machine vision applications to perform tasks such as bin picking, volume measurement, automotive part measurement and 3D object location and identification and calculate the distance between the camera and objects so the possibility to figure out the distance between two objects by using some basic trinogometrical properties .

Stereo vision is also becoming the technology of choice for range sensing in mobile robotics navigation. The Bumblebee2 was selected as the primary range sensing method for the Princeton vehicle, “Prowler”, in the DARPA Urban Challenge, an autonomous vehicle research and development program that has robotic vehicles competing in a race in a complex urban environment.

A. Mobile robot

Automotive production is one of the largest market values in the world. Its economic potential constantly drives and enforces change. Continuous optimization improves many features of new car models. Road safety is a very important automotive aspect. According to World Health Organization (WHO), traffic accidents are one of the main causes of death in the world, because of this, there is a huge need to improve the car safety and reduce this statistic. Advanced driver-assistance system (ADAS) can help to solve this problem. It supports behavior on the road and improve driver safety. It can perform many activities such as human and road signs detection; calculate automatic maneuvers to generate collision-free trajectories, emergency stop systems to avoid collisions.

The 1rst models of vehicles equipped with a stereo camera sensor already appear in the automotive market. Usually the current stereo camera exists as one sensor (one box), which makes it impossible to get a wide base line between the cameras. The most common location of such a sensor is under the front mirrors on the windshield. Nowadays, this sensor provides only additional data to support more complex ADAS.

For these reasons, the automotive application is dierent from other CPS and the problem of the online stereo camera calibration is mostly overlooked because the cameras do not have the right to change their location in the expensive, rigid and complex stereo camera system. Nowadays ADAS try to propose a news solution, for example mount the cameras on the side mirrors. This solution increases a baseline between cameras, thus depth and field of view. In such case, it is mandatory to look at the online camera calibration because this type of installation requires continuous monitoring of the camera's position, because the mirrors are movable.

Fig. 8: : Different uses of stereo cameras

B. Virtual and augmented reality devices

The virtual and augmented reality devices collect, process and control the data from sensors which provide such informations, in order to add the 2D, 3D figgures or artifcial information generated by computers to the display with real environments. Fig 8 presents the special devices like portable smart glasses or helmets. Those are able to change the optical properties of the environment around us in the real time so we can observe things like they are real. This area of research is quickly gaining popularity within the last years.

The first successes of virtual reality motion sensor, which try to capture the information, received in the real world and pass it to the computers was Microsoft's Kinect. It entered into the Guinness Book of Records as "the fastest-selling consumer electronics device" after more than 8 million copies were sold in the first 2 months. It shows a huge potential and great perspective in the future for this kind of devices. The Kinect is equipped with two cameras. The first one is a standard RGB video camera provides video images with a resolution of 640×480. The second camera is part of the whole sensor subsystem, which returns information about the depth of the local environment. It uses the infrared illuminator, which displays a cloud of points in front of the camera. The infrared camera sees its positions and size. Thanks to this data, it can calculate the distance to the local environment.

Fig. 9: Augmented and virtual reality devices

V. CONCULUSION

This paper provides a comprehensive survey of the uses of stereoscopic vision, the diffrence between mono camera and stereo camera and how stereo vision works basically, calibration methods of a stereoscopic camera and some examples of its applications

Finally, there are several related topics that have not been covered in this survey given the time constraint and the fact that i didn't work in group throughout this scientific project. This topic include image-based 3D object reconstruction using deep learning and video-based depth estimation and application of stereo scopic vision in traffic.

REFERECES

- [1] California State University "**DISTANCE ESTIMATION FROM STEREO VISION: REVIEW AND RESULTS**" **.**

- [2] Romuald Ginhoux Jens-Steffen Gutmann " **Model-Based Object Tracking Using Stereo Vision**" **.**

- [3] Manaf A. Mahammed, Amera I. Melhum, Faris A. Kochery "**Object Distance Measurement by Stereo VISION**" **.**

- [4] Mohammad Ebrahim Ashoori, Mahmoud Mahlouji "**Measuring the Distance between the Two Vehicles Using Stereo Vision with Optical Axes Cross** " **.**

- [5] Denis Grießbach,Martin Bauer, Andreas Hermerschmidt "**Stereo camera calibration with diffractive optical elements**" **.**

- [6] Jernej Mrovlje, Damir Vrančić " **Distance measuring based on stereoscopic pictures** " **.**

- [7] Gregory D. Hager, Wen-Chung Chang, and Stephen A. Morse. " **Robot hand-eye coordination based on stereo vision** " **.**

- [8] Yongjie Yan, Qidan Zhu, Zhuang Lin, and Quanfu Chen " **Camera calibration in binocular stereo vision of moving robot. In Intelligent Control and Automation** " **.**

- [9] Pathum Rathnayaka. Seung-Hae Baek. and Soon-Yong Park " **An effcient calibration method for a**

stereo camera system with heterogeneous lenses using an embedded checkerboard pattern. ” **.**

- [10] M. Alemán-Flores, L. Alvarez, L. Gomez, P. Henriquez, and L. Mazorra “ **Camera calibration in sport event scenarios** ” **.**

- [11] M. Bjorkman and J. O. Eklundh. “ **. Real-time epipolar geometry estimation of binocular stereo heads.** ” **.**

- [12] JRandolph Blake and Hugh Wilson “ **Binocular vision** ” **.**

- [13] T. Dang and C. Homann “ **Stereo calibration in vehicles. In Intelligent Vehicles Symposium** ” **.**

- [14] T. Dang, C. Homann, and C. Stiller “ **. Continuous stereo self-calibration by camera parameter tracking. Image Processing** ” **.**

- [15] Antonio Guiducci “ **Camera calibration for road applications. Computer Vision and Image Understanding** ” **.**

- [16] E. E. Hemayed “ **A survey of camera self-calibration. In Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance,** ” **.**

- [17] L. Grammatikopoulos, G. Karras, and E. Pets “ **An au-tomatic approach for camera calibration from vanishingpoints** ” **.**