ARTICLE IN PRESS

INTEGRATION the VLSI journal xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

INTEGRATION, the VLSI journal

journal homepage: www.elsevier.com/locate/vlsi

Algorithm and hardware implementation for visual perception system in autonomous vehicle: A survey

Weijing Shi[a], Mohamed Baker Alawieh[a], Xin Li[b,c,*], Huafeng Yu[d]

[a] Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[b] Electrical and Computer Engineering Department, Duke University, Durham, NC 27708, USA
[c] Institute of Applied Physical Sciences and Engineering, Duke Kunshan University, Kunshan, Jiangsu 215316, China
[d] Boeing Research and Technology, Huntsville, AL 35758, USA

ARTICLE INFO

Keywords:
Algorithm
Hardware
Autonomous vehicle
Visual perception

ABSTRACT

This paper briefly surveys the recent progress on visual perception algorithms and their corresponding hardware implementations for the emerging application of autonomous driving. In particular, vehicle and pedestrian detection, lane detection and drivable surface detection are presented as three important applications for visual perception. On the other hand, CPU, GPU, FPGA and ASIC are discussed as the major components to form an efficient hardware platform for real-time operation. Finally, several technical challenges are presented to motivate future research and development in the field.

1. Introduction

The last decade has witnessed tremendous development of autonomous and intelligent systems: satellites in space, drones in air, autonomous vehicles on road, and autonomous vessels in water. These autonomous systems aim at progressively taking over the repeated, tedious and dangerous operations by human, especially in an extreme environment. With the Grand Challenge and Urban Challenge of autonomous vehicles, organized by the Defense Advanced Research Projects Agency (DARPA) in 2004 and 2007 respectively [1], autonomous vehicles and their enabling technologies have received broad interests as well as investments from both academia and industry. After these challenges, major developments have been quickly switched from academic research to industrial commercialization. Automotive Original Equipment Manufacturers (OEMs) such as GM, BMW, Tesla, Daimler and Nissan, tier-one suppliers such as Bosch, Denso and Delphi, as well as software companies such as Google, Uber and Baidu, have progressively joined the global competition for self-driving technology. Many have already noticed Google self-driving cars in Mountain View, CA and Austin, TX, and Uber cars in Pittsburg, PA for road testing.

The revolution of autonomous driving brings up a lot of discussions on various issues related to society, policy, legislation, insurance, etc. For instance, how would the society accept autonomous vehicles when their behavior is still unknown and/or unpredictable? How should the policy be made to accelerate the development and deployment of autonomous vehicles? What laws should be built to regulate autonomous vehicles for their integration into our society? How do we handle the accidents with insurance involving autonomous vehicles? The recent reports on artificial intelligence [2,3] may be good references for thinking and addressing these questions.

Different from conventional vehicles, autonomous vehicles are equipped with new electrical and mechanical devices for environment perception, communication, localization and computing. These devices include radar, LIDAR, ultra-sonic, GPS, camera, GPU, FPGA, etc. They also integrate new information processing algorithms of machine learning, signal processing, encryption/decryption and decision making. The autonomy level [4] of these autonomous vehicles is finally determined by the combination of all devices and algorithms at different levels of maturity.

In the literature, a number of these new devices and technologies were first integrated into vehicles as the enablers for Advanced Driver Assistant System (ADAS). However, ADAS only provides simple and partial autonomous features at low levels of autonomy. Yet, ADAS has been proved to be valuable in improving vehicle safety. Examples of ADAS include lane departure warning system, adaptive cruise control, blind spot monitor, automatic parking, etc. These systems generally work with the conventional vehicle E/E (Electrical/Electronic) architecture, and do not require any major modification of vehicle architecture. ADAS has been extensively adopted for today's commercial vehicles with low cost.

On the other hand, an increasing number of companies are extremely interested in research and development of high-level

* Corresponding author at: Electrical and Computer Engineering Department, Duke University, Durham, NC 27708, USA.
E-mail addresses: weijings@cmu.edu (W. Shi), malawieh@cmu.edu (M.B. Alawieh), xinli.ece@duke.edu (X. Li), huafeng.yu@boeing.com (H. Yu).

http://dx.doi.org/10.1016/j.vlsi.2017.07.007
Received 20 May 2017; Received in revised form 24 July 2017; Accepted 26 July 2017
0167-9260/

autonomy. Namely, an autonomous car can drive itself, instead of providing driver assistant only. At this level of autonomy, it requires vehicles to sense the surrounding environment like humans, such as distance of obstacle, signalization, location, moving pedestrians, etc., as well as making decisions like humans. These requirements lead to the adoption and integration of a large set of new sensing devices, information processing algorithms, hardware computing units, and then lead to new automotive E/E architecture design where safety, security, performance, power consumption, cost, etc., must be carefully considered.

In spite of all the technical and social challenges for adopting autonomous vehicles, autonomy technologies are being developed with significant investment and at a fast rate [5,6]. Among them, visual perception is one of the most critical technologies, as all important decisions made by autonomous vehicles rely on visual perception of the surrounding environment. Without correct perception, any decision made to control a vehicle is not safe. In this paper, we present a brief survey on various perception algorithms and the underlying hardware platforms to execute these algorithms for real-time operation. In particular, machine learning and computer vison algorithms are often used to process the sensing data and derive an accurate understanding of the surrounding environment, including vehicle and pedestrian detection, lane detection, drivable surface detection, etc. Based upon the perception outcome, an intelligent system can further make decisions to control and manipulate the vehicle.

To meet the competitive requirements on computing for real-time operation, special hardware platforms have been designed and implemented. Note that machine learning and computer vision algorithms are often computationally expensive and, therefore, require a powerful computing platform to process the data in a timely manner. On the other hand, a commercially competitive system must be energy-efficient and with low cost. In this paper, a number of different possible choices for hardware implementation are briefly reviewed, including CPU, GPU, FPGA, ASIC, etc.

The reminder of this paper is organized as follows. Section 2 overviews autonomous vehicles and several visual perception algorithms are summarized in Section 3. Important hardware platforms for implementing perception algorithms are discussed in Section 4. Finally, we conclude in Section 5.

2. Autonomous vehicles

As an intelligent system, autonomous car must automatically sense the surrounding environment and make correct driving decisions by itself. In general, the function components of an autonomous driving system can be classified into three categories: (i) perception, (ii) decision and control, and (iii) vehicle platform manipulation [7].

The perception system of an autonomous vehicle perceives the environment and its interaction with the vehicle. Usually, it covers sensing, sensor fusion, localization, etc. By integrating all these tasks, we generate an understanding of the external world based on sensor data. Given the perception information, a driving system must make appropriate decisions to control the vehicle. The objective is to navigate a vehicle by following a planned route to the destination while avoiding collisions with any static or dynamic obstacle. To achieve this goal, the decision and control functions compute the global route based on a map in its database, constantly plan the correct motion and generate local trajectories to avoid obstacles.

Once the driving decision is made, the components for vehicle platform manipulation execute the decision and ensure the vehicle to act in an appropriate manner. They generate control signals for propulsion, steering and braking. Since most traditional vehicles have already adopted the electrical controlling architecture, manipulation units usually do not require any major modification of the architecture. Additionally, vehicle platform manipulation may cover emergency safety operations in case of system failure.

As the interface between the real world and the vehicle, an accurate perception system is extremely critical. If inaccurate perception information is used to guide the decision and control system, an autonomous car may make incorrect decisions, resulting in poor driving efficiency or worse, an accident. For example, if the traffic sign detection system misses a STOP sign, the vehicle may not make the correct decision to stop, thereby leading to an accident.

Among all perception functions, visual perception is one of the most important components. It interprets visual data from multiple cameras and performs critical tasks such as vehicle and pedestrian detection. Although an autonomous driving system usually has other non-visual sensors, cameras are essential because they mimic human eyes and most traffic rules are designed by assuming the ability of visual perception. For example, many traffic signs share similar physical shapes and they are differentiated by their colored patterns that can only be captured by a visual perception system. In the next section, we review several important applications for visual perception and highlight the corresponding algorithms.

3. Visual perception algorithms

Visual perception is mainly used for detecting obstacles that can be either dynamic (e.g., vehicle and pedestrian) or static (e.g., road curve and lane marker). Different obstacles may have dramatically different behaviors or represent different driving rules. For example, a road curve defines the strict boundary of the road and exceeding the boundary must be avoided. However, a lane marker defines the “soft” boundary of driving lane which vehicles may break if necessary. Therefore, it is not sufficient to detect obstacles only. A visual perception algorithm must accurately recognize the obstacles of interest. In addition to obstacle detection, visual perception is also used for drivable surface detection, where an autonomous vehicle needs to detect possible drivable space even when it is off-road (e.g. in a parking lot) or when the road is not clearly defined by road markers (e.g. on a forest road). Over the past several decades, a large body of perception algorithms have been developed. However, due to the page limit, we will only review a small number of most representative algorithms in this paper.

3.1. Vehicle and pedestrian detection

Detecting vehicles and pedestrians lies in the center of driving safety. Tremendous research efforts have been devoted to develop accurate, robust and fast detection algorithms. Most traditional detection methods are composed of two steps. First, important features are extracted from a raw image. A feature is an abstraction of image pixels such as the gradient of pixels or the similarity between a local image and the designed patterns. They can be considered as the low-level understanding of a raw image. A good feature can efficiently represent the valuable information required by detection while robustly tolerating the distortions such as rotation of image, variation of illumination condition, scaling of object, etc. Next, once the features are available, a learning algorithm is applied to further inspect the feature values and recognize the scene represented by the image. By adopting an appropriate algorithm for feature selection (e.g., AdaBoost [8,9]), a small number of important features are often chosen from a large set of candidates to build an efficient classifier.

Histogram of oriented gradients (HoG) [10] is one of the most widely adopted features for object detection. When calculating the HoG feature, an image is divided into a grid of cells and carefully normalized on local area. The histogram of the image gradients in a local area forms a feature vector. The HoG feature is carefully hand-crafted and can achieve high accuracy in pedestrian and vehicle detection. It also carries relatively low computational cost, which makes it popular in real-time applications such as autonomous driving. However, the design of hand-crafted features such as HoG requires extensive