Top Level Newsletter: Connected Vehicle

(Published by IEEE Orange County Section)

July 2023
Volume 41.0

Editor-in-Chief: Kay Das, IEEE Life Member,
Systems Research Development and Innovation

Vol 41: This issue includes articles relevant to the sciences of AI and deep learning in connected vehicles:

1. Software-Defined Imaging: A Survey,
2. Information-to-Model Matching for AI in Open Radio Access Networks (extension to article in v.40.0)
3. The Deep Learning Revolution (by Terrence Sejnowski): summary of a published book
4. LIDAR Object Perception Framework for Urban Autonomous Driving

General Notes
This series of newsletters is intended to provide the IEEE member with a top level briefing of the many different subjects relevant to the research, development and innovation of the connected vehicle.

The objective is to provide a platform for fast learning and quick overview so that the reader may be guided to the next levels of detail and gain insight into correlations between the entries to enable growth of the technology. Intended audiences are those that desire a quick introduction to the subject and who may wish to take it further and deepen their knowledge. This includes those in industry, academia or government and the public at large. Descriptions will include a range of flavors from technical detail to broad industry and administrative issues. A (soft) limit of 300 to 600 words is usually set for each entry, but not rigorously exercised.

As descriptions are not exhaustive, hyperlinks are occasionally provided to give the reader a first means of delving into the next level of detail. The reader is encouraged to develop a first level understanding of the topic in view. The emphasis is on brief, clear and contained text. There will be no diagrams in order to keep the publication concise and podcast-friendly. Related topics in the case of Connected Vehicle technology, such as 5G cellular and the Internet of Things will be included. The terms Connected Vehicle and Automated Driving will be used inter-changeably. Articles from other published sources than IEEE that add to the information value will occasionally be included.

This newsletter forms part of the regional Advanced Technology Initiative (ATI) of which connected vehicles form a constituent part. Technical articles solely from IEEE journals/magazines are referred to by their Digital Object Identifier (DOI) or corresponding https link. The link for each article is provided. Those readers who wish to delve further to the complete paper and have access to IEEE Explore (www.ieeexplore.ieee.org) may download complete articles of interest. Those who subscribe to the relevant IEEE society and receive the journal may already have physical or electronic copies. In case of difficulty please contact the editor at kaydas@mac.com. The objective is to provide top level guidance on the subject of interest. As this is a collection of summaries of already published articles and serves to further widen audiences for the benefit of each publication, no copyright issues are foreseen.

Readers are encouraged to develop their own onward sources of information, discover and draw inferences, join the dots, and further develop the technology. Entries in the newsletter are normally either editorials or summaries or abstracts of articles. Where a deepening of knowledge is desired, reading the full article is recommended.

1. Software-Defined Imaging: A Survey, by (Suren Jayasurya et al)

Published in: Proceedings of the IEEE (Volume: 111, Issue: 5, May 2023)
Page(s): 445 – 464

Abstract: Huge advancements have been made over the years in terms of modern image-sensing hardware and visual computing algorithms (e.g., computer vision, image processing, and computational photography). However, to this day, there still exists a current gap between the hardware and software design in an imaging system, which silos one research domain from another. Bridging this gap is the key to unlocking new visual computing capabilities for end applications in commercial photography, industrial inspection, and robotics. In this survey, we explore existing works in the literature that can be leveraged to replace conventional hardware components in an imaging system with software for enhanced reconfigurability. As a result, the user can program the image sensor in a way best suited to the end application. We refer to this as software-defined imaging (SDI), where image sensor behavior can be altered by the system software depending on the user's needs.

The scope of our survey covers imaging systems for single-image capture, multi-image, and burst photography, as well as video. We review works related to the sensor primitives, image signal processor (ISP) pipeline, computer architecture, and operating system elements of the SDI stack. Finally, we outline the infrastructure and resources for SDI systems, and we also discuss possible future research directions for the field.

Introduction

Image sensing has become ubiquitous in modern society, ranging from industrial uses in the workplace all the way to personal entertainment through the sharing of photos and videos via social media. Driven by the development of CMOS image sensors in the 1990s and 2000s, image sensing has become cheap, affordable, and, when integrated with smartphones, extremely portable and easy to use. Indeed, billions of photos are taken and uploaded on the internet each day.

In accordance with the development of image sensor technology, the rise of image processing, computer vision, and computational photography has been similarly meteoric. Advances in algorithms, including state-of-the-art machine learning using deep neural networks, have enabled high-fidelity visual computing. However, while both image-sensing hardware and the software that implements and realizes the algorithms have grown exponentially, there is still relatively little codesign between the two domains. This is due to a number of factors: not only technical challenges of interfacing analog sensing components with digital computation but also social and cultural challenges of different communities of researchers and industries communicating with one another across the stack.

For example, the traditional image sensor faces major hurdles to allow for flexibility and reconfiguration in its operating modes. A recent study showed that changing image sensor resolution costs hundreds of milliseconds in latency, and this was mostly due to software operating system (OS) bottlenecks. Most image sensors have been slow to expose knobs to the software developer, such as resolution, exposure, and quantization; in contrast, most imaging and vision algorithms do not exploit these sensing mechanisms efficiently and adaptively at runtime. This leads to lost opportunities for improved reconfigurability of imaging systems in practice.

To address these issues, an interdisciplinary community of researchers and practitioners has embraced **software-defined imaging** (SDI) to improve the technology of visual computing systems. A software-defined image sensor offers several dimensions of sensing configurability along with system support and programming abstractions to support application-specific needs. Hardware parameters, such as exposure, resolution, and frame rate, are programmable in image sensors, and software-defined image sensors exploit this programmability alongside software algorithms to optimize certain imaging task metrics, such as energy, latency, and task performance. This field has a vertical-oriented mindset connecting knowledge of sensor physics and electronics, analog and mixed-signal circuits, digital systems and architectures, OSs and programming languages, and end applications of computer vision and computational photography.

In this article, we survey this emerging area of SDI, and highlight key works across the hardware and software stack in the literature. We point out how this area intersects other popular areas of research including embedded computer vision, deep learning, computational photography, and digital hardware acceleration for imaging applications. However, we also point out the unique nature of SDI research, cross-cutting traditional boundaries to allow synergy among the stack.
In addition, we begin to discuss the growing issue of how to quantify performance and evaluate quality in this cross-disciplinary area. We survey existing metrics, tools, and benchmarks that are open source and available to the public. These tools can be used to help improve basic research, generate training datasets for machine learning applications, and model/analyze system efficiency in

terms of latency and energy for the potential deployment of solutions. We highlight how this public infrastructure is distributed across the software–hardware stack from application programming interfaces (APIs) to image signal processors (ISPs) to hardware devices.

Why Software-Defined Imaging?

One of the major endeavors of this article is to coin the phrase "SDI," which we denote in this article by the acronym SDI, as an umbrella term and concept, which connects various strands of existing research. We do not seek to be overly prescriptive or limiting in our definition of SDI, but rather to coalesce around a common research agenda while having the flexibility to allow for offshoots and branches of interesting research potential. To help draw motivation for SDI, we actually first look at the research parallels in the communications community and how this enabled successful applications and high impact on those fields.

Software-defined radio (SDR) is now a mature field within communications. It was originally coined in the 1980s, but a bulk of the research came in the 1990s and 2000s along with the growth of the telecom industry. SDR centers around the idea of a radio communication system where traditional hardware components, such as filters, mixers, filters, amplifiers, modulators/demodulators, and detectors, are instead augmented or replaced with software and embedded processing. This allows radio systems to be flexible and reconfigurable and to maximize their usage of the spectrum at beneficial tradeoff points for system design. SDR has spurred a wide variety of work and has been implemented in many modern communication standards.

Similarly, our idea of SDI centers around the idea of the traditional hardware components and mechanisms of image sensors being replaced or augmented by software alternatives. The main reason for doing so is flexibility: allowing sensors to adaptively change their sensing behavior for new environments and application demands. For SDI, we focus on key sensor primitives, such as region-of-interest readout, exposure, frame rate, and quantization, since these are all programmable and implemented onboard the image sensor with dedicated hardware. SDI research aims to design the hardware–software stack from the sensor, ISP, compute architecture, and OS to allow for reconfigurability and programmability for these primitives, driven by the end applications. By doing so, we can actually achieve favorable tradeoffs between system efficiency and application-specific task performance for imaging and computer vision.

Programmable Sensor Mechanisms

Since the invention of CCD and CMOS image sensors in the late 20th century, there has been an increased emphasis on programmability in their design to enable user flexibility during capture. In this section, we will curate and discuss research works that highlight the development of programmable camera sensors. In particular, we will delve into detailed discussions on frame rate, exposure, region of interest (ROI), and quantization, which are the most widely configured parameters in the image sensor. In addition to describing the hardware advances for these sensor mechanisms, we will also examine algorithms that exploit programmable sensor knobs for end-applications in imaging and machine vision.

Note that we analyze imaging systems for single capture (e.g., conventional photography), multiframe or burst photography (e.g., high dynamic range (HDR) mode and panorama stitching), and video (e.g., continuous shooting mode). We do not differentiate between these modes as modern systems typically can perform capture in all of these modalities. However, we do highlight when research has been designed and targeted for a specific application in mind. (1299 words)

DOI: 10.1109/JPROC.2023.3266736

2. Information-to-Model Matching for AI in Open Radio Access Networks (by Jorge Martin Perez et al)

Published in: IEEE Communications Magazine (Volume: 61, Issue: 4, April 2023) **Page(s):** 58 - 63

Abstract: Open Radio Access Network (O-RAN) is an emerging paradigm, whereby virtualized network infrastructure elements from different vendors communicate via open, standardized interfaces. A key element therein is the RAN Intelligent Controller (RIC), an Artificial Intelligence (AI)-based controller. Traditionally, all data available in the network has been used to train a single AI model to be used at the RIC. This article introduces, discusses, and evaluates the creation of multiple AI model instances at different RICs, leveraging information from some (or all) locations for their training. This brings about a flexible relationship between gNBs. In 5G, a gNodeB is responsible for radio communication with UEs in its coverage area, known as a cell. A gNodeB, (new name for a base station for 5G) may be a physical entity, such as a tower, or the AI models used to control them, and the data such models are trained with. Experiments with real-world traces show how using multiple AI model instances that choose training data from specific locations improve the performance of traditional approaches following the hoarding strategy.

Introduction: Virtual Radio Access Network (vRAN) is arguably one of the most exciting recent innovations of the networking ecosystem. It is enabled by the Software-Defined Networking (SDN) approach, and allows the functions traditionally performed by base stations (currently gNBs) to become *virtualized* and *split* across multiple network nodes, including newly-introduced entities called Central Units (CUs), Distributed Units (DUs), and Radio Units (RUs). Such a functional split allows different decisions to be made at different nodes *and* with different time scales. For example, RUs can perform real-time radio management, while CUs can adjust higher-level resource allocation at longer time scales. The different CU, DU, RU units corresponding to different gNBs can now be hosted in edge or cloud servers, sharing location in some cases and reducing costs for the operators through the remote management of the components thanks to its virtualized nature. The promising results of vRAN gave rise to initiatives, such as Open Radio Access Network (O-RAN) or Cisco's Open vRAN Ecosystem Group, aiming at creating an open and interoperable RAN ecosystem where open APIs and interfaces can be integrated connecting different vendors components. O-RAN has been so far the vRAN initiative receiving more attention, also thanks to the open-source community created around it.

Conclusion:

We have proposed and analyzed a new approach to the integration of AI in O-RAN scenarios, allowing to assign different model instances to each gNB of the network, independently *choose* the data each instance is trained on. Our approach deviates from the state of the art in that it does not seek to train one model instance for the whole network and to train it using all available data; giving more flexibility than fully-centralized and fully-distributed approaches.

Our performance evaluation, leveraging real-world traces, shows how our approach yields very attractive trade-offs between training time and learning effectiveness, by combining data from different sources in a flexible manner. Future research directions stemming from our work include characterizing *a priori* the usefulness of data for AI training, trade-offs between data transfer delays and AI training time, and the impact of AI accuracy over the performance of concrete applications. (527 words)

DOI: https://doi.org/10.1109/MCOM.003.2200401

3. The Deep Learning Revolution (book by Terrence Sejnowski)

ISBN 9780262038034, Published: October 23, 2018. Publisher: The MIT Press, 352 pages

Abstract:: How deep learning—from Google Translate to driverless cars to personal cognitive assistants—is changing our lives and transforming every sector of the economy**.**

The deep learning revolution has brought us driverless cars, the greatly improved Google Translate, fluent conversations with Siri and Alexa, and enormous profits from automated trading on the New York Stock Exchange. Deep learning networks can play poker better than professional poker players and defeat a world champion at Go. In this book, Terry Sejnowski explains how deep learning went from being an arcane academic field to a disruptive technology in the information economy.

Sejnowski played an important role in the founding of deep learning, as one of a small group of researchers in the 1980s who challenged the prevailing logic-and-symbol based version of AI. The new version of AI Sejnowski and others developed, which became deep learning, is fueled instead by data. Deep networks learn

from data in the same way that babies experience the world, starting with fresh eyes and gradually acquiring the skills needed to navigate novel environments. Learning algorithms extract information from raw data; information can be used to create knowledge; knowledge underlies understanding; understanding leads to wisdom. Someday a driverless car will know the road better than you do and drive with more skill; a deep learning network will diagnose your illness; a personal cognitive assistant will augment your puny human brain. It took nature many millions of years to evolve human intelligence; AI is on a trajectory measured in decades. Sejnowski prepares us for a deep learning future. (259 words)

4. LIDAR Object Perception Framework for Urban Autonomous Driving (by Jongho Kim et al)

Published in: IEEE Vehicular Technology Magazine (Volume: 18, Issue: 2, June 2023)

Abstract: This article describes the development and implementation of a 3D lidar perception framework to guarantee the precise cognition of the surrounding environment for urban autonomous driving. The proposed framework consists of two different detection modules operating in parallel: a deep learning-based and a geometric model-free cluster-based method. The first module utilizes the convolutional gated recurrent unit (ConvGRU)-based residual network (CGRN). The module aims to repredict 3D objects based on results from a continuous single-frame detection network. A vision-fusion methodology based on 2D projection is adopted for postprocessing in the first module. The second module utilizes geometric model-free area (GMFA) cluster detection and is designed to cope with false-negative cases of unclassified objects from the prior module. For the second module, a cluster variance-based ground removal is conducted to prevent false-positive cases. A kinematic model-based particle filter (PF) is then applied to estimate the dynamic states of detection. The suggested framework has been developed with real-time operation in mind, to be implemented in autonomous vehicles equipped with automotive lidars and low-cost cameras. The test results show that the framework with CGRN and GMFA successfully improved the surrounding object detection and state estimation accuracy in urban autonomous driving.

Advancements in the field of autonomous driving has led to a shift in driving scenarios, from highways to urban roads, to accommodate level 4 autonomous driving. Unlike highways, urban roads with substantial traffic require much more information regarding the surroundings to ensure that safe decisions are made in autonomous driving systems. Consequently, the lidar sensor, which utilizes reflections to obtain point clouds with high positional accuracy, is utilized by most autonomous vehicles for recognizing short- to medium-ranged objects in urban scenarios.
Lidar object perception for modern autonomous driving systems generally consists of detection and tracking. The detection module distinguishes and categorizes the 3D box from the point clouds of surrounding objects. Furthermore, according to the tracking-by-detection paradigm, the tracking module tracks IDs and estimates

the dynamic states of the moving objects from the continuous detection sequence. Although many researchers have studied perception modules individually for several years, the entire perception framework as a whole has not yet been fully discussed and still requires the evaluation and validation of real time and robustness for autonomous vehicle implementation. Therefore, this article proposes a lidar object perception framework, including detection and tracking and several postprocessing methods. Focusing on the application to urban autonomous driving, the proposed framework is designed to confirm the real-time capability and robustness against perception failure…..

In future studies, our research team plans to develop auxiliary networks to infer the driver intentions, such as lane changes and yields from 3D object sequence data, by following the concept of CGRN. Furthermore, we plan to design a network that can integrate the results of multiple domains, such as radar, V2X, and AVM cameras, apart from lidars and cameras. (473 words)

DOI: https://doi.org/10.1109/MVT.2023.3236480