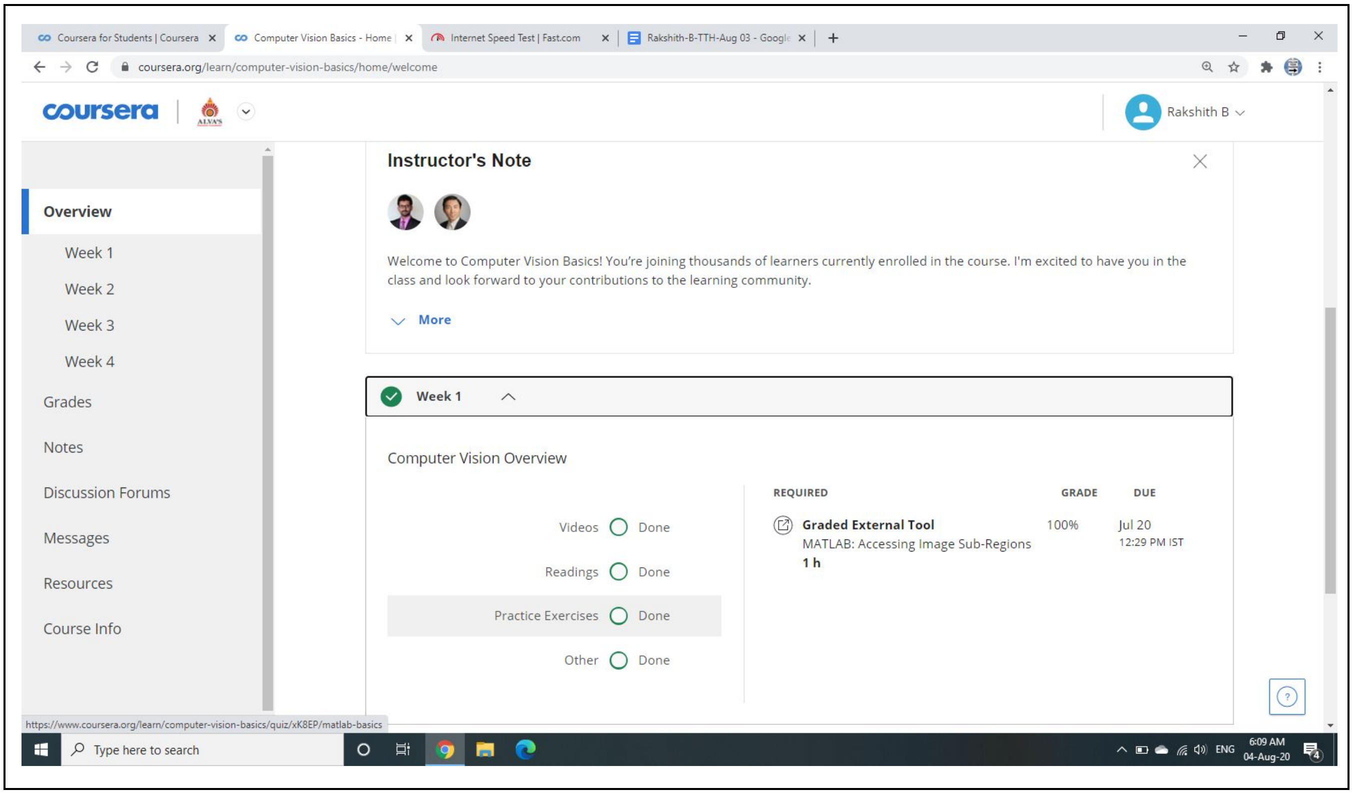
REPORT AUGUST 03

Date:	03 AUGUST 2020	Name:	Rakshith B
Course:	Coursera	USN:	4AL16EC409
Topic:	Computer Vision Basics	Semester & Section:	6th SEM B
Github Repository:	Rakshith-B		

Image of the Session



Today is an exciting time to work on computer vision, as it is becoming an integral technology in our daily lives. Applications of computer vision, range from tasks such as Industrial Vision Systems, say inspect bottles painting by on a production line to robots that can comprehend the world around them. Computer Vision is outperforming humans on certain real-world tasks such as circuit board inspections and face recognition under controlled conditions. There has also been great progress in traditional application areas like multimedia, robotics, and medical imaging. Moreover, new application areas keep arising such as augmented reality, autonomous driving, Internet of things, human-computer interaction and vision for the blind. There are growing opportunities for Computer Vision to provide outreach to non-traditional areas such as astronomy, nanotechnology, novel brain imaging techniques, scientific analysis and many more. In health care, Computer Vision has the potential to read real value. While computers won't completely replace healthcare personnel, there is a good possibility they will compliment routine diagnostics that require a lot of time and expertise of human physicians. In reality, recent advance s in the field of Computer Vision are enabling vision algorithms to surpass human vision capabilities. There are many kinds of computer vision systems. Nevertheless, all of them contain these basic elements. A power source, at least one camera, processor, as well as, control and communication cables or some wireless interconnection mechanism. In addition, a practical vision system contains configurable software, as well as, a display in order to monitor the system. A class of moving cameras are egocentric vision systems composed of a wearable cameras that automatically takes pictures from the first-person perspective like the GoPro cameras. The demand for fast and robust processing of vision task is so high that vision processing units are emerging as a new class of processors to complement CPUs and GPUs which are Graphics Processing Units. One of the key applications of Computer Vision is visual surveillance. Human supervision simply cannot scale up to the needs of visual surveillance. There are too many objects and events to keep track of. Visual surveillance is an active area of research to make it work robustly on an energy constraints system like a drone. The next major Computer Vision application area is biometrics. The most widely used biometrics application is fingerprint-based identification and authentication. Shown in many movies, iris face recognition is a popular biometric application. Iris recognition was used in 2002 to identify the woman in this Emblematic National Geographic cover image from 1985. Face recognition is also a widely use computer vision based biometric application. Most of the smartphones these days have face unlock feature. Alright, so, computer vision in multimedia and entertainment. If you have ever use face filters or played augmented reality based games, you have experienced Computer Vision powered applications in first-hand. Speaking of multimedia applications, majority of the image processing applications have a computer vision aspect to them. To fit an image into screens of different aspect ratios is not a trivial problem because cropping and scaling wouldn't work. Image re-targeting is one of the sophisticated image processing applications which uses scene carving to retain the salient regions in the image. Of course, visual effects is one of the areas where Computer Vision is widely used. Computer Vision is indispensable when it comes to navigation. Stereo vision and depth sensor based vision are widely used for robot navigation systems. One of the popular features of Google Navigation App is Google Street View created by using panorama stitching.The properties and characteristics of the human visual system often give inspiration to engineers who are designing computer vision systems. Just to give you a glimpse of the capabilities of human vision, let me show you an image for a couple of seconds, and ask you a couple of questions.

Although the image is highly pixellated and blurred, you were able to infer the context and key information from it.

The objective of high-level vision is to infer the semantics, for example, object recognition and scene understanding. A challenging question for many decades is how do you achieve invariant recognition? That is, how do you recognize 3D objects from different view directions? High-level vision works for image understanding and video understanding. It works on answering questions like, is there a car in the image? or is the person in the video jumping? Based on the answers of these questions we should be able to fulfill different tasks in intelligent human-computer interaction, intelligent robots, smart environments, and content-based multimedia. Let us look at the high-level vision concepts in more detail. Visual recognition applications if implemented robustly can open up limitless opportunities. These applications can serve as a repository for visual knowledge and can assist humans using them. So why is visual recognition such a hard problem? Before we delve into that let us look at the visual recognition techniques that are out there.

Although the image is highly pixellated and blurred, you were able to infer the context and key information from it.

The objective of high-level vision is to infer the semantics, for example, object recognition and scene understanding. A challenging question for many decades is how do you achieve invariant recognition?

That is, how do you recognize 3D objects from different view directions? High-level vision works for image understanding and video understanding. It works on answering questions like, is there a car in the image? or is the person in the video jumping? Based on the answers of these questions we should be able to fulfill different tasks in intelligent human-computer interaction, intelligent robots, smart environments, and content-based multimedia. Let us look at the high-level vision concepts in more detail. Visual recognition applications if implemented robustly can open up limitless opportunities. These applications can serve as a repository for visual knowledge and can assist humans using them. So why is visual recognition such a hard problem? Before we delve into that let us look at the visual recognition techniques that are out there.

The objective of high-level vision is to infer the semantics, for example, object recognition and scene understanding. A challenging question for many decades is how do you achieve invariant recognition? That is, how do you recognize 3D objects from different view directions? High-level vision works for image understanding and video understanding. It works on answering questions like, is there a car in the image? or is the person in the video jumping? Based on the answers of these questions we should be able to fulfill different tasks in intelligent human-computer interaction, intelligent robots, smart environments, and content-based multimedia. Let us look at the high-level vision concepts in more detail. Visual recognition applications if implemented robustly can open up limitless opportunities. These applications can serve as a repository for visual knowledge and can assist humans using them. So why is visual recognition such a hard problem? Before we delve into that let us look at the visual recognition techniques that are out there.

The objective of computer vision is to make computers see and interpret the world like humans and possibly even better than us. Human vision performs multiple visual tasks quite effortlessly and effectively. How is visual information processed and understood in biological systems?

What is the nature of computation involved in visual tasks? And how might we build machines that can see? Partial answers to these questions have been offered over several decades by researchers in the fields of biology, neuroscience, and computer science. Let's say someone across the room throws a ball at you and you catch it.

This appears to be a simple task, but in reality this is one of the most complex processes to comprehend, let alone recreate. Let us try to analyze this task step by step. First, the light rays of the ball pass through both eyes and strike on their respective retinas. The retinas do some preliminary processing before sending the visual responses through optical nerves to the brain,

with the real world better, so vision is a very important task there. So, myself also partially work on this in the sense that we want to human hands to help you to interact in your real virtual scenario. So, we wanted a camera when you put your hand on top of the camera, or the camera should understand what is your hands' motion. For example, if even you manipulate objects there, it can do that.

Vision is a very important technique for augmented reality. Augmented reality, for example, if you apply with Microsoft HoloLens. So for example, it can give you, on top of the real thing you see, it can give you a virtual information stuff. For example, if you wear the augmented reality glass, the Google Glass lens. If you've heard about Google Glass, of course, now it disappeared, but if you've had a Google Glass before, that's the concept for augmented reality. So vision is important because if you want to put anything on that makes sense to the real thing, you have to understand it first. For example, if you take the, with your glass, see a building, if you want to make an augmented reality so that you can attach information of that building. So this is a building built 100 years ago, and has a very nice restaurant there, so on and so forth with information. We have to first understand, identify which building it is. So, you may rely on computer vision, object recognition, detection, segmentation to know what what you see as a first step. Whenever you know, computer vision can tell you what you see through the camera then it can augment a reality can function by putting even more information there or help you interact

I'm Jeff Bier. Am the President of BDTI, which is a specialized consulting firm, and founder of the Embedded Vision Alliance, which is an industry association. My background is my degrees are in electrical engineering. Early in my career, my focus was on Embedded Digital Signal Processing, figuring out how to squeeze demanding digital signal processing algorithms into the small cost in power budgets available in portable battery-powered consumer products. Then around the year 2010, I started to get interested in computer vision and got the idea that we were approaching the point where it would become possible to put computer vision into these same kinds of cost and power constraint systems, and shifted the focus of my consulting business at that time to computer vision and figuring out how to help people deploy computer vision in real-world products. Also around that same time, I started the Embedded Vision Alliance. It's a partnership that's working to inspire and empower all kinds of product developers to incorporate practical computer vision into their products.

where the visual cortex does the heavy lifting of thorough analysis. The brain taps into its knowledge base, classifies the object and dimensions, and having predicted its path, decides to act on it by sending signals to move the hand and catch the ball. This takes place in a tiny fraction of second without almost no conscious effort and almost never fails depending upon how much prior catching practice you've had. Recreating human vision isn't just a hard problem, it's a set of them, each of which relies on the other. More formally, computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding.

It is an interdisciplinary field that works for high-level understanding from digital images or videos. Computer vision has a dual goal.

From the biological science point of view, computer vision aims to come up with computational models for human visual system.

From the engineering point of view, computer vision aims to build autonomous systems to perform some of the tasks which the human visual system can perform and even surpass it in many cases. Many vision tasks are related to extraction of 3D and temporal information from time varying 2D data, or in other words, videos. Of course, the two goals are intimately related.

Now pay high attention to this picture I'm going to show only for a fraction of second. Are you ready? Although the image has been flashed for a split second, you were again able to establish the context. If I show you the image for a longer duration, you can make more observations for further analysis. Historically, computer vision as a field of research has been notoriously difficult.

One main reason for this complexity is that the human visual system is simply too good for many visual task. The computer vision system suffer by comparison. For instance, a human can recognize faces under all kind of variations and illumination viewpoint expression, etc.