ERDC/CERL TN-21-1
September 2021

Autonomous Transport Innovation: A Review of Enabling Technologies

by Matthew T. Gross and Julie L. Webster

PURPOSE: This document is the first of the technical note series under the Autonomous Transport Innovation (ATI) research program. The series intends to be an introduction on autonomous vehicles (AVs), their testing, and associated infrastructure. A review of technologies that enable vehicle autonomy is necessary to provide the basis for understanding vehicle performance in testing scenarios and in actual use.

INTRODUCTION: This report provides an overview of the wide variety of sensors and sensing applications that are currently in use in autonomous systems. Sensors are categorized in comparison tables according to the medium used (e.g., electromagnetism, acoustics, or tactility). The assortment of sensors covered in this report is non-exhaustive as there are limitless ways to sense the surrounding world. Additionally, the value of specific sensors to particular autonomous applications varies, and these caveats are noted throughout the report. This report also briefly covers AV technology trends, including developments with promise and deficiencies.

ONBOARD SENSORS: An AV needs a combination of varying sensors to effectively perceive its surroundings in diverse situations. This report covers four sensor types, including electromagnetic, acoustic, state, and tactile. Each type will be defined, contextualized for its use in autonomy, and evaluated in terms of its efficacy.

Electromagnetic Sensors. Electromagnetic (EM) sensors are the most commonly used sensor type because public roads are designed to give drivers visual cues that are communicated via the EM spectrum (Table 1)[1]. The EM spectrum is the light people see every day; it includes a wide range of frequencies. The colors in the spectrum shown in Figure 1 display how these frequencies are seen by humans. Understanding which EM frequencies will travel through Earth's atmosphere is important because AV sensors use these frequencies to transmit signals. The signals travel by air to objects in the environment, are reflected by these objects, and ideally are not absorbed.[2] One limitation of using EM as a sensing medium is weather conditions. The frequencies emitted are absorbed by precipitation (high humidity, fog, rain, sleet, hail, snow, etc.), blocking a large portion of the EM spectrum due to increased water content.

---

[1] An example of a nonvisual cue is the rumble strip, a series of grooves across a road or along its edge. Upon contact with the strip, vehicle tire noise changes and warns drivers of speed restrictions or road edges.

[2] With LiDAR, for example, it is desirable that a frequency get return signals from physical objects in the surrounding environment but not get returns from the atmosphere itself.

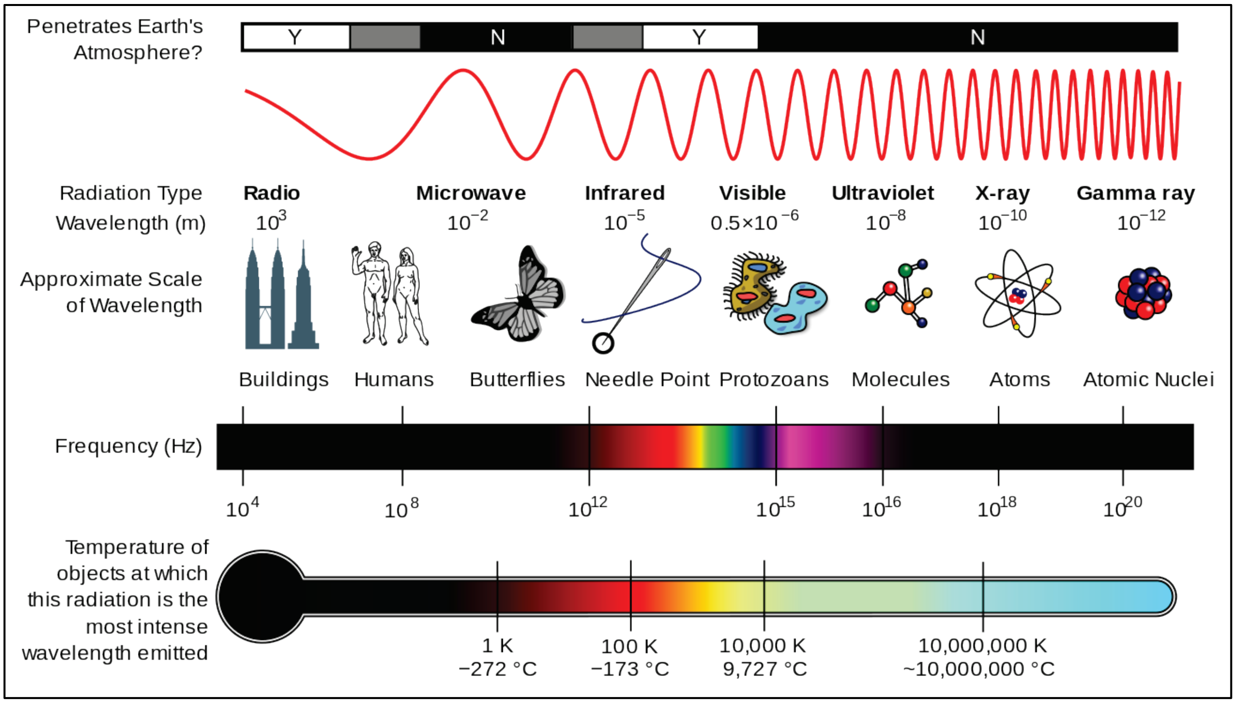

Figure 1. EM infographic (Impey 2020).

The most prevalent sensors found on AVs—Radio Detection and Ranging (RADAR), Light Detection and Ranging (LiDAR), and cameras—all work in the EM spectrum and rely on objects to reflect the same wavelength of light that is emitted.[1] Due to material makeup, the reflection property more effectively covers metal objects such as cars but may struggle detecting other materials.

RADAR: RADAR detection utilizes light (which has a constant speed) and metallic surfaces (which reflect radio waves like a mirror). By emitting a radio wave and measuring its return time and direction, it is possible to identify the distance and direction of objects. RADAR can provide depth perception and velocity detection for objects in inclement weather but delivers lower resolution than LiDAR, to be discussed in the following paragraph. To see how RADAR compares to other sensors in terms of acuity and range, see Figure 2.

LiDAR: LiDAR uses the same technique as RADAR, but instead of using radio waves, it uses infrared (IR) waves that will reflect off of nearly anything. One key difference between the two detection systems is accuracy. The radio waves used by RADAR detect objects at farther distances with larger margins of error, typically ranging between 1 and 2 m.[2] The infrared waves used by LiDAR detect objects at shorter distances with smaller margins of error, typically within centimeters. LiDAR can provide detailed depth perception, which is helpful for collision avoidance, yet, when met with inclement weather conditions, the performance decreases. Infrared

---

[1] An exception is passive cameras that do not emit light, such as a cell phone camera. A reflector is classified as an object that reflects back 80 percent or greater of the emitted light, and an absorber typically reflects back less than 10 percent of the emitted light

[2] For a full list of the spelled-out forms of the units of measure used in this document, please refer to US Government Publishing Office Style Manual, 31st ed. (Washington, DC: US Government Publishing Office 2016), 248-52, https://www.govinfo.gov/content/pkg/GPO-STYLEMANUAL-2016/pdf/GPO-STYLEMANUAL-2016.pdf.

waves are absorbed by water, making detection difficult in rain or snow, for example. To see how LiDAR compares to other sensors in terms of acuity and range, see Figure 2.

Between RADAR and LiDAR. Technology operating in an area between IR and microwave radiation on the EM spectrum (marked in gray for "Penetrates Earth's Atmosphere?" in Figure 1) is currently the subject of research. This region of the spectrum is of interest because signals operating in these frequencies can travel for more than 200 m one way before reflection and return a significant reading when reflected off of metal, a material that makes up the bulk of kinetic obstacles on roadways today. A limitation therein is that technology in this range does not detect pedestrians effectively. However, because this range has a wavelength of approximately 1 mm, it provides relatively high definition and equipment resolution.[1] Systems operating in this range are comparable to LiDAR and also function with higher performance during inclement weather than systems using other frequencies. The near-millimeter wavelength characteristic of systems operating in this range earns them the shortened name *mmWave RADAR*. For more information on mmWave RADAR, see page 10.

Standard Visible Light (red-green-blue [RGB]) Cameras. Depth, direction, and speed data from RADAR and LiDAR can be correlated with information obtained from passive visual sensors (e.g., RGB cameras) to identify detected objects. RGB cameras are those used for still images and videos in daily life. In the case for AVs, multiple RGB cameras are mounted onto the vehicle in locations that ideally provide a 360 deg view of surroundings. These cameras are used because roadways were designed for human sight, which functions in the RGB spectrum. RGB cameras allow the AV to sense, analyze, and respond to road markings and traffic signals. A single camera can provide detailed information for object identification, but multiple cameras offer the added benefit of depth perception. Regardless of number, camera performance degrades in low light and inclement weather. To see how these passive visual sensors compare to other sensors in terms of acuity and range, see Figure 2.

---

[1] Equipment resolution is a measurement specification related to the accuracy of a device in positioning an object in its field of view.

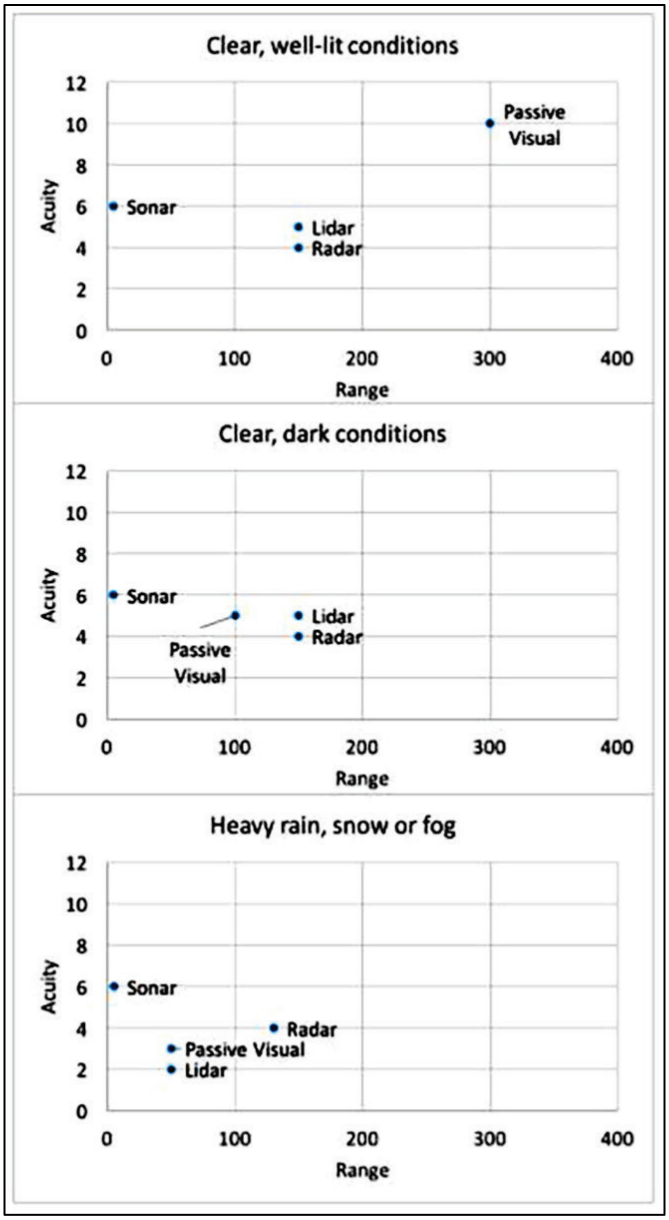

Figure 2. Circumstantial sensor acuity vs. range in meters (Barnard 2020).

Other Cameras. Thermal cameras work outside of the visible spectrum in the IR zone, which is why thermal cameras are often called IR cameras. These cameras take advantage of the "black body radiation" concept that assumes most objects creating heat also emit light. The ability to see heat by means of a "heat signature" is valuable to an AV because it can then detect pedestrians and make decisions to avoid impact. While thermal cameras reveal heat signatures, RGB cameras do not perceive them (Figure 3). For this reason, thermal cameras are sometimes used to augment RGB cameras in low-light conditions. This is done by attaching IR emitters to a thermal camera to convert a light source normally not viewable by humans into one that is viewable (i.e., night vision). To see how this type of passive visual sensor compares to other sensors in terms of acuity and range, see Figure 2.

Figure 3. Highlighting the differences in what a normal RGB camera can see versus the details that a thermal camera can detect (FLIR 2020).

The industrial light field camera, also called a plenoptic camera, features a grid of thousands of small lenses that sits between a conventional camera's sensor and its series of regular lenses. This grid allows the camera to take a picture of the back of the regular lens series. This picture can then be fed into a computer to mathematically rebuild the image with any desired focus point within the captured shot. The image rebuild provides dramatically increased clarity at the focus point (Figure 4). This capability may be useful in helping AVs with object identification. A single image with a predetermined focus point can be captured in approximately 200 ms (Perwass and Wietzke 2012). When applied to autonomous driving, extra computation time would likely be required to repeatedly refocus the camera on passing points of interest. However, the ability to refocus an image without retaking the image altogether reduces moving object ambiguity, increasing computational confidence in object detection. If an AV can more reliably detect objects—specifically humans—from greater distances, it can travel faster more safely, gaining utilization and time efficiencies. Although modern computers can handle the imaging algorithms required of industrial light field cameras, the large processing power demands and slow refresh rate make this technology impractical to most AV manufacturers. Other limiting factors are its high starting price (approximately $1,600) and size (larger than a standard RGB camera).

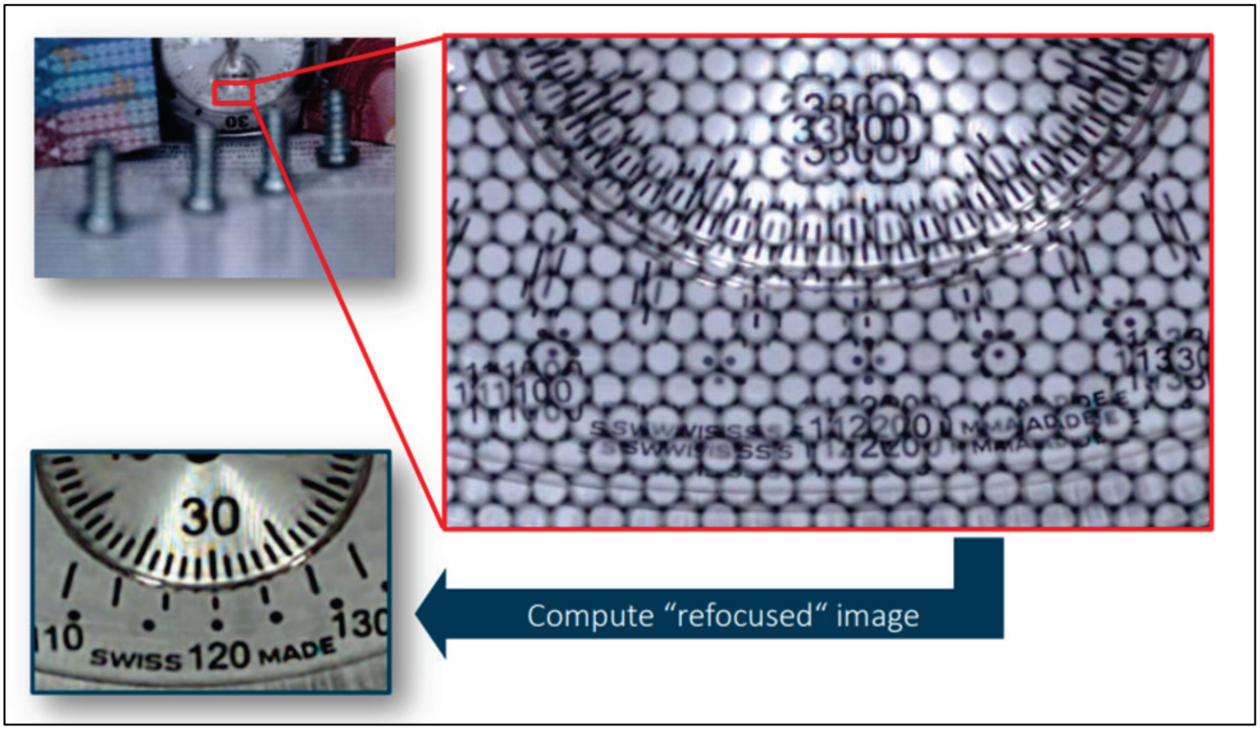

Figure 4. Raw versus computed image from a plenoptic camera showing the difference before and after computation (Perwass 2014).

The stereo camera is made up of two RGB cameras working together to measure the angle to an object from two points. Using this information, the distance to the object can be approximated. For delineating the edges of an object, a stereo camera has the same depth capabilities as a single camera. In AVs, stereo cameras are often overlooked because both LiDAR and RADAR work well at detecting depth. However, these cameras can also be particularly useful for long range path finding (Hadsell et al. 2009) or used with industrial light field cameras to provide active depth perception. Furthermore, stereo cameras use less computational power and cost less when compared to LiDAR, while increasing range.

Vertical Cavity Surface Emitting Laser. The vertical cavity surface emitting laser (VCSEL) is a semiconductor-based laser diode that emits light as a fine optical beam plane. While useful for extremely high distance resolution, VCSELs are rarely employed on AVs due to their narrow and short-range detection. These lasers have an extremely fine resolution but do not emit enough photons for the EM wave to have a strong return signal. Some users have mitigated the detection issue by combining complex algorithms, selective filtering, and single photon avalanche detectors (SPADs) to catch the return signal from over 200 m away. A SPAD is a type of transistor that permits a large flow of electrons through the transistor gate when hit by a single photon, which acts as the gate key. These devices recognize weak single-photon return signals, thus permitting high-definition detection.

Table 1. Electromagnetic (light) sensors.

Name	Description	Output	Measures of Accuracy	Operation Range	Limitations
Radio Detection and Ranging (RADAR) [type 1]	Emits EM radio wave, times response	Proximity	M	1-300 m	Interference[1]
Radio Detection and Ranging (RADAR) [type 2]	Emits EM radio wave, times response and records direction and speed	Obstacle ID	M m/s	1-160 m	Interference
Light Detection and Ranging (LiDAR)	Emits laser in circle/sphere and times response to give series of distances	2D/3D Point Cloud	cm	1-20 m	Weather, Interference, Obstruction
Standard Visible Light (RGB) Camera	Measures Red, Green, and Blue wavelengths of light under optics	Digital image	Resolution	N/A	Distance, Obstruction
Thermal (heat) Camera	Measures infrared wavelengths of light under optics	Heat signatures	Resolution, Temperature	N/A	Interference, Distance, Obstruction
Industrial Light Field Camera	Takes a picture of the back of the camera lens, allows post-focusing	Image at all depths	Focus	N/A	Glare, Distance
Stereo Cameras	Measures 2 RGB camera inputs and compares	Depth, image	Resolution, Distance	N/A	Glare, Distance
Vertical Cavity Surface Emitting LASER	Emits and receives laser EM wave for precise distance	Distance	mm	0-20 cm	Interference
Night Vision Camera	Emits infrared EM waves, measures IR wavelengths through optics	Image at night	Resolution	N/A	Interference, Distance, Obstruction

Acoustic Sensors. Non-visual technologies are also useful in autonomous systems. Acoustic sensors use sound waves to perceive the environment. They can attain high reflectivity and definition off of close-range objects. Their inherent deficiency in sensing distant objects makes them less popular than EM spectra sensors. Despite this deficiency—and as the autonomy market looks to off-road applications—innovation, development, and use of acoustic technologies will likely be needed. Table 2 introduces sensors with future potential for off-roading. Acoustic sensors are used today primarily for their high-resolution detection for a vehicle that is traveling at low speeds and measuring nearby objects. An example application is parking, as this action is done at a slow, controlled speed and with objects such as parking blocks or curbs nearby. RADAR and LiDAR are not used for such applications because they require a minimum distance between object and vehicle for detection.

---

[1] Interference is the act of broadcasting an additional EM signal at the same frequency as an initial signal but louder so the initial signal is overpowered.

Table 2. Acoustic sensors.

Name	Description	Output	Measures of Accuracy	Operation Range	Limitations
Ultrasonic Sensor	Echo, reads timing	Distance	cm	4-50 cm	Distance, Interference
Ultrasonic Profiling Sensor	Echo, reads reflection	Profile	dB	< 10 m	Distance, Interference
Microphone	Listens and gives a profile	Profile	dB	< 10 m	Distance, Interference

Vehicle State Sensors. Another sensing medium is the vehicle itself. This environment includes all sensors that are installed to measure forces the vehicle is receiving and exerting, as well as sensors for finding the vehicle position within a reference frame (Table 3). Associated sensor readings are combined to form what is called the *vehicle pose*, which includes its 3-axis position and 3-axis rotation. This is added to the odometry, vehicle speed, and acceleration in three axes to obtain the entire state of the vehicle. In an ideal world, this information, along with a known starting position, would be enough to precisely locate the vehicle after a given amount of time in operation. In reality, many of these sensors are prone to drift, developing a small amount of error that accumulates over time. The process of finding the current position from all past vehicle states is called *dead reckoning*, and this can be done using encoders, cameras, landmark tracking, and anything that is not Global Positioning System (GPS) or a GPS-like system (Vivacqua et al. 2017).

Two state sensors commonly used in AVs are GPS and Inertial Measurement Units (IMUs). GPS provides location data that can be correlated to a given map. IMUs can detect the path a vehicle is traveling by tracking the forces the vehicle body experiences. This is particularly useful when GPS signal is lost. As with most sensors, IMU path tracking is not error free, and the calculated and real position will diverge over time.

Table 3. State sensors.

Name	Description	Output	Measures of Accuracy	Limitations
Attitude and Heading Reference System	Combines IMU and GPS to give more accurate information	Localization	m	Drift, interference
IMU	Combines an inclinometer, accelerometer, and gyroscope to give more accurate information	Pose	m	Drift, interference
Inclinometer	Uses Earth's gravitational field to measure angle relative to gravity	Incline	degrees	N/A
Magnetometer	Electronic compass	North, south, east, west orientation	degrees	Interference
Gyroscope	Measures rotations on 3 axes	Orientation changes	degrees	Drift
Accelerometer	Measures acceleration on 3 axes	Velocity changes	m/s	Drift

GPS	Takes in timing of signals from satellites, compares to satellite almanac, operates anywhere in United States	Position	< 8 m	Interference, obstruction
Odometer/Wheel Encoder	Measures angle of wheel/track cog	Distance traveled	in.	Traction loss

Tactile Sensors. Tactile sensors measure pressure on a surface and require direct contact (Table 4). While these aspects make them unhelpful for crash avoidance, tactile sensors have utility when exploring off-road applications. The tactile medium is one that is used by few (if any) autonomy efforts outside of the Robust Autonomous Vehicle for Off-road Navigation (RAVON) project (Armbrust et al. 2011). This project made use of a tactile push-bar to feel around for impassable objects to avoid. The sense of touch offered by the push-bar allowed RAVON to respond to contact stimuli. This is particularly useful when a vehicle is accelerating unsuccessfully into a tree or teetering on a boulder after bottoming out on a rocky path, for example. A tactile sensor can provide definitive information on whether a vehicle has made physical contact, though contact can be reasonably assumed with data from other sensors types.

Table 1. Tactile sensors.

Name	Description	Output	Measures of Accuracy	Operation Range	Limitations
Pressure Sensor	Tells how hard a part is being pressed	Single axis pressure	Pascals	N/A	Distance
Contact Sensor	Tells if a part is being pressed	Single axis contact	Binary	on/off	Distance
Electro resistive Sensor	Reads resistance between two terminals dug in the soil	Resistance	Ohms	N/A	Noise, moisture

DATA FUSION AND ALGORITHMS: Once raw data are collected by sensors, it must be fed into a decision-making process. This process is vast and complex, and there are intricate problems left unsolved in the translation from theory to real-world application. Consequently, numerous approaches exist for autonomous decision-making. Also, computer system processing is bound by hardware capabilities, and thus system optimization[1] is necessary.

Full Machine Learning. Machine learning algorithms can be used as a control system, or feedback loop, between AV sensing and actuation. These algorithms can be difficult to properly execute, and harder to debug. This type of system is usually referred to as an End-to-End Machine Learning (E2EML) process. According to Kocić et al. (2019), in an E2EML process, the raw sensor data are fed into a computer. Then, the computer makes a decision in response to the entire process, from responding to incoming information to outputting the correct control signals for steering, braking, and acceleration. To train this machine-learning algorithm, it is fed a library of driving scenarios that the computer must use to decide on an action. The outcome is then graded

---

[1] In this context, system optimization is the process of reducing decisions made by a computer such that the least amount of computational work is necessary to complete a task. For example, if a computer system only acts when encountering three 1s in a row, then it is not necessary to see 32 or 64 digits of 1s and 0s at a time.

as either correct or incorrect. The computer is rewarded for correct responses and punished for incorrect responses. Eventually, the algorithm learns situational responses to information from the given group of sensors and their fused outputs (i.e., payload).

The next step in learning is usually accomplished by having a human drive around in a vehicle outfitted with the autonomous payload attached (as it would in full autonomy mode). The algorithm will continually convert payload output into actions it would execute if it had control of the vehicle. A secondary computer compares those actions to the human actions, using the human control as the *answer key*. The advantage to this process is that specific payloads are carefully calibrated for human impersonation. Unfortunately, the process is not modular, and the sensor payload cannot be modified in a piecemeal manner. Once a payload is implemented, any desired edits require complete retraining of the algorithm.

The E2EML process also requires a massive amount of computing power, and the debugging process is difficult because decisions made by the computer are not logically trackable. The computer can only record raw sensor data, not specific objects encountered. Even though a point array produced by a LiDAR unit is recordable, the computer's classification of objects is not reachable in an E2EML control system. For example, it is not possible to detect if the computer recognized a traffic light or its signal color. The ability to log specific objects and their traits could be useful in determining required changes, either to the test data or to the sensor payload for altering computer sensitivity to specific events or more general scenarios. Ultimately, the E2EML process may be considered inefficient as engineers must guess how to correct a failed system, and fixes are probabilistic rather than deterministic.

Modular Processing. Instead of an E2EML control system alone, sensor modules and sensor fusion can be introduced. These modify the end-to-end system by dividing up the work and allowing for results determined by one unit to be recorded as they are passed to other units. Modular sensor/sensor fusion setups consist of two hierarchal steps. The first step is using a module handling raw data from a sensor to return more decisive data. One example is a module that takes in data from a camera and returns images of objects overlaid with their type classification. The second step is sensor fusion, or the merging of data from multiple sensors to create new data for downstream processing units. This step can involve raw input from sensors or processed input from sensor modules (or both). Information from any and all sensors can be used as needed. An example is fusing raw sensor input from a LiDAR unit with a mapping module that cleans up the LiDAR point array and turns the recorded surroundings into a three-dimensional (3D) map that continuously updates using images from the RGB cameras.

The fusion example above could feed into another module made for obstacle avoidance. The obstacle avoidance module could take the 3D map and determine which paths are passable and which are obscured by objects that must be yielded to or avoided. The output from this module could feed into a subsequent pathfinding module, which intakes the current vehicle position and end-goal position on a map, to determine the best path forward. The pathfinding module could make amends between the predetermined route of a given map and the real-time obstacle data and use this information with traffic rules to execute the correct maneuvers.

Although a modular sensor/sensor fusion system is more complex than E2EML, it is more practical from an engineering perspective. The overall results are deterministic regardless of probabilistic machine learning algorithms used in the individual modules employed. An E2EML control system augmented with sensor fusion is more practical and allows for modular changes; the sensor payload can be changed without the need for changing the core decision algorithm (Shmoolik 2018).

Partial Machine Learning. Linearly coding a module means the outputs are deterministic and do not change unless the programmer intervenes and changes the code. Machine learning means the machine can learn for itself by running available inputs through test scenarios until the most successful result is achieved. This result is then used to alter machine outputs. The machine learning process usually takes place in a laboratory and is tested before use. Once the module is operating to a given standard, the machine output is locked in, and the machine can no longer learn. Machine learning and linear code can be used simultaneously in modular processing units, thus resulting in partial machine learning.

In an AV, the processing modules are any part of the autonomous control system that excludes the sensors and actuators themselves. Figure 5 shows a sample layout for autonomous decision-making using a variety of system designs (e.g., modular processing and partial machine learning). Layouts can be expanded or retracted as necessary to be more or less specific with regards to module task assignments, clustering of modules, number of sensors, and sensor types. For example, a thermal camera could be added to the layout in Figure 5 to assist in detecting people and animals.

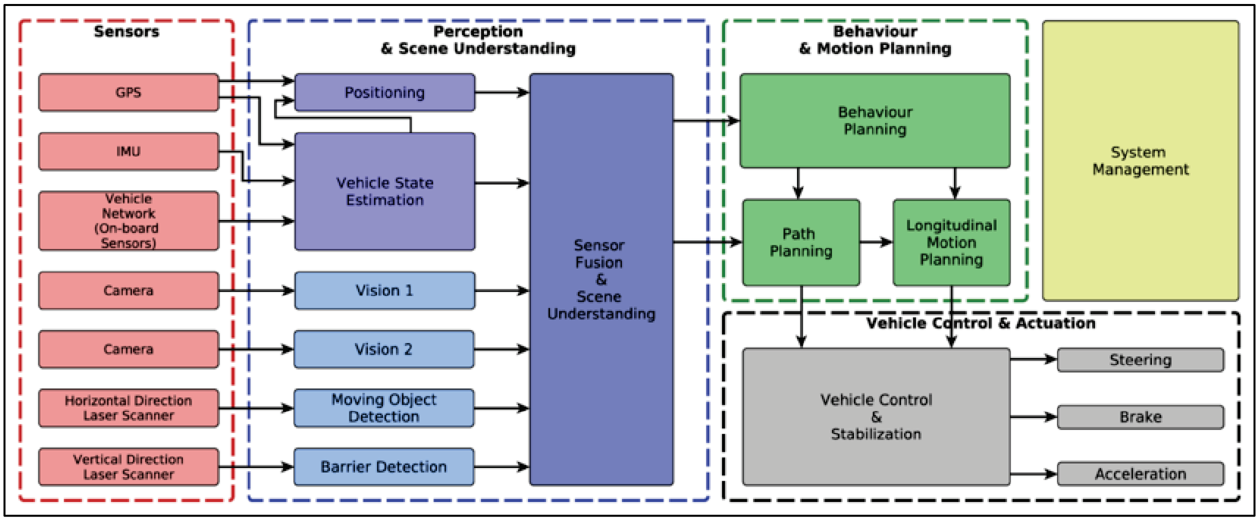

Figure 5. Sample modular layout of an autonomous control system (Kuhnt 2016).

Edge Cases. Despite improvements in interoperability, sensor modularity and sensor fusion have their drawbacks. The main limitation of modularity (separating the process into divisible portions of work) is that raw sensor data can be lost along the way. This data loss creates vulnerabilities in rare situations known as edge cases. For example, the camera module might be caught in the glare of a sunrise and fail to detect pedestrians, or the LiDAR unit might be impaired by rain diffusing its laser pulses before they can reach their intended targets in the physical surroundings. The results are gaps in information being conveyed to the core control algorithm and a decline in sensor performance.

TECHNOLOGY TRENDS: AVs are currently in the transportation spotlight. With help from automobile manufacturers worldwide, and a decade of work behind them, autonomous innovation today is on the verge of making significant breakthroughs on a variety of enabling technologies. To close remaining technology gaps, researchers are attempting to use human detection, sensor validation and verification, road drivability assessments, and vegetation detection to overcome challenges for both on- and off-road autonomous driving.

Promising Technology. Much of the current research and development into autonomous systems focuses on communication methods. This is due to the importance of communication in differentiating manual driving from autonomous driving. The automated exchange of information increases transportation safety while freeing people from their driving responsibilities so they can be more productive with their time or enjoy leisure activities while on the move.

Wireless Technologies. Active research is underway to optimize the use of standards-based wireless technologies (e.g., Bluetooth, Long-Term Evolution (LTE), Fourth Generation (4G), and Fifth Generation (5G)) in autonomous applications. These systems receive immediate attention because they are well known, and most are time tested. Each wireless technology has its preferred uses, advantages, and drawbacks. The primary differences between wireless frequencies are the range (coverage) and bandwidth (speed) that the frequency bands provide. To date, Bluetooth and 4G have shown little utility. Bluetooth, for example, is currently not desirable due to its short range and the tendency of its signal to become muddled over a short distance due to background noise (this is measured as signal-to-noise ratio). 4G has a longer range but is too slow for most autonomous uses, and it consumes substantial bandwidth. LTE has a longer range, is too slow, and uses considerable bandwidth, but its ability to handle the combination of video, LiDAR, and (vehicle-to-everything) V2X communications is promising.

The longer-range, lower-frequency communication methods above correlate to slow transfer of information, much like that experienced when the deep bass sound of a rock band has a delayed arrival at the back of a concert venue. To increase data transfer speeds and range distances (desirable in autonomy), a middle-ground technology between Bluetooth and 4G is needed. That technology is 5G. 5G networks are being implemented with frequencies in the 2.4 GHz to 8 GHz range, which approximately covers the same transmitting frequencies of home Wi-Fi systems (Seker 2018). Lower bands (e.g., 2.4 GHz) provide coverage at a longer range but transmit data at slower speeds; higher bands (e.g., 8 GHz) provide less coverage but transmit data at faster speeds.

One area in which LTE and 5G excel is map sharing. Map sharing between communication platforms supports the creation of a cohesive picture of the general environment and immediate surroundings for all platforms in a given area. For example, in a dense urban environment, autonomous platforms must respond to traffic alerts, address route planning, communicate with other vehicles, and may interact with intelligent infrastructure.

Dedicated Short-Range Communications (DSRC). There are a few specific frequencies not readily absorbed by Earth's atmosphere. Communication devices operating in these frequencies can maintain their power over longer distances since air does not absorb power from their signal. To fully exploit these frequencies, professional bodies Institute of Electrical and Electronics Engineers and Society of Automotive Engineers partnered to establish the DSRC standards that set a common communication frequency for use by automotive manufacturers

(Kenney 2011). According to these standards, DSRC infrastructure for improving traffic flow, traffic safety, and other intelligent transportation service applications can be built around the 5.9 GHz transmission frequency at a bandwidth of 500K bps. An example of such infrastructure is V2X communication points that can be installed along roadways to pass information from a vehicle to any entity that may affect the vehicle.

mmWave. Operating in the near-millimeter wavelength dimension earns these systems the shortened name *mmWave RADAR.*[1] Some larger wavelength systems do not give information about the direction of an object, but mmWave RADAR technologies allow light beams to be emitted and sensed at a desired angle to give information about the direction of a reflected signal. If the direction of a reflected EM wave can be determined, then that direction can be related to detection of nearby objects. This tactic is useful in AVs for mapping trackable nearby obstacles.

DSRC claims it operates in the mmWave spectrum but scientifically speaking, it does not. DSRC operates at 5.9 GHz with a 5 cm wavelength, while true millimeter wave (mmWave) technology is located higher into the frequency spectrum at 30 GHz through 300 GHz. These higher frequencies cannot penetrate inclement weather as well as DSRC systems because they lose power quickly. This is due to water molecules more readily absorbing the energy from higher wavelengths (Figure 6).

Unlike larger wavelength systems that require a separate antenna, the small wavelength of mmWave systems allows their antennae to be incorporated into microchips, thus reducing manufacturing costs. They can also be incorporated into small quarter-wave antennae used to build beamforming radar systems. These systems focus a wireless signal towards a specific receiver rather than having the signal broadcast in all directions. The result is a faster and more reliable signal. Beamforming is useful to AVs as it allows RADAR to operate with higher definition when penetrating through inclement weather that would blind most LiDAR systems.

---

[1] An mmWave technology known as Simultaneous Wireless Information and Power Transfer (SWIPT) aimed to transmit power over the air to improve electric vehicle range when supplemental power was provided via remote infrastructure. Patent No. US20140292090A1 for this technology was filed in 2011 but has since been abandoned. The loss of power from transmitting through the air over larger distances was found to be inefficient. Regardless of beam angle, the system was wasting more energy than it provided (Ponnimbaduge Perera et al. 2017; Cordiero 2011).

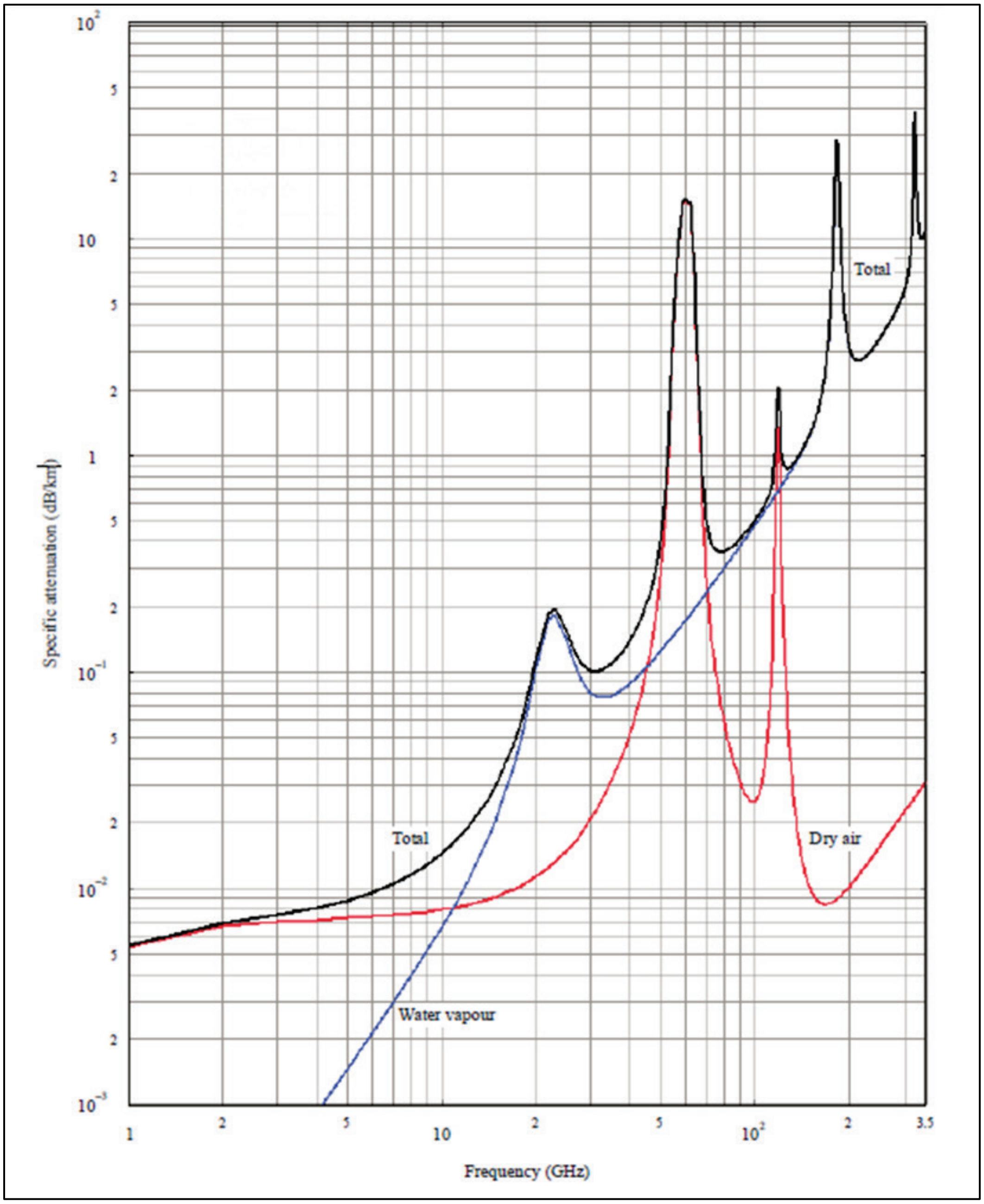

Figure 6. Absorption of light at mmWave frequencies; more attenuation of transmission signals means more power loss[1]; attenuation due to water vapor translates to impassability of signals during inclement weather (Tynan 2012).

---

[1] In Figure 6, 1 dB attenuation equates to 20% signal power loss, 3 dB attenuation equates to 50% signal power loss, and 10 dB attenuation equates to 90% signal power loss.

Technology Gaps. Research under the ATI program is concerned with applications of vehicle autonomy in both the civilian and military domains. Technology gaps remain in both realms, and developers will need to address these shortcomings in order to field safe self-driving vehicles.

On-Road Autonomy Shortcomings. On the civilian side, cheaper and more effective human detection is needed. Car manufacturers are frugal with their use of thermal cameras because they are expensive in comparison to other sensors. However, these cameras are the most effective means of directly assessing the presence of humans on and near a roadway (Figure 7). Cost-benefit analyses into the use of thermal cameras in autonomy could prove favorable. Work is ongoing to find algorithms for addressing the blind spots and edge cases that have allowed pedestrian fatalities, such as the March 2018 Uber incident (Harvey 2018).

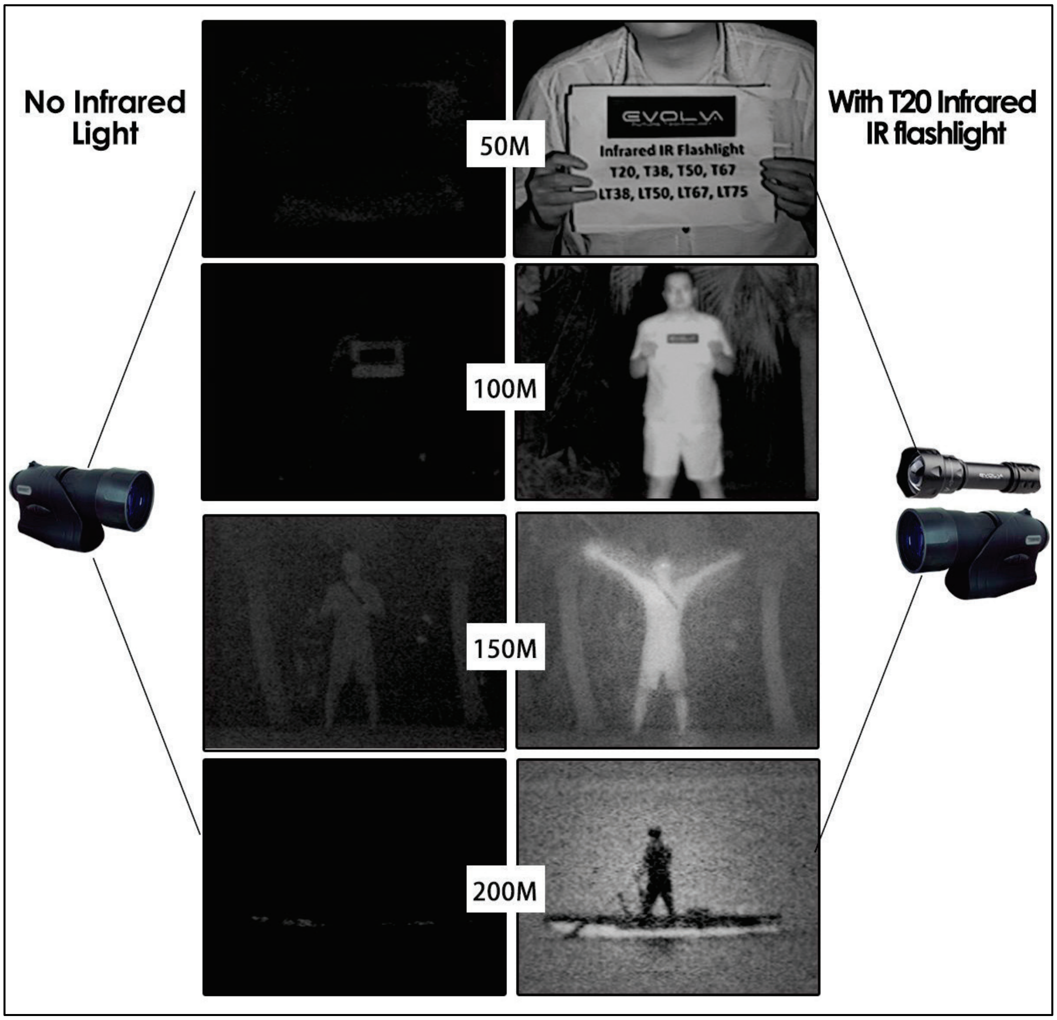

Figure 7. IR/thermal vision with and without infrared illumination over several distances (Brock Industries 2017).

Another inadequacy in the autonomy realm is sensor validation. Customary validation is accomplished in two ways, either by redundancy or cross-validation. System redundancy works at the sensor level and, for example, could involve multiple cameras covering the same detection area. Cross-validation involves higher levels of processing to resolve multiple inputs. For example, RGB cameras may detect images of a person while a LiDAR unit also detects the outline of a person and a thermal camera simultaneously detects the heat signature of a person. All three inputs considered collectively would conclude there is a human detected.

Alternatively, sensor validation can be conducted from within an E2EML control system. Machine learning algorithms are normally programmed for forward processing, meaning sensor data go in and control signals come out. To verify sensors, these algorithms can be written to reverse the processing order. Specifically, control signals can be generated by comparing a multitude of possible sensor input scenarios to expected control signal output scenarios. In this way, finding when and where a vehicle would accelerate, brake, and turn can be individually assessed and corrected within the control system before allowing the algorithm to continue developing during its training period. While reverse algorithms are more complex, they are more easily diagnosable, and they provide a unique perspective that can expose system flaws.

Off-Road Autonomy Shortcomings. On the tactical side, advances are needed in determining path drivability. Civilian autonomous transportation works on the assumption that all roads and paths are drivable and that all ground is relatively flat. These assumptions may be overly optimistic in some civilian cases but are generally unrealistic in off-road tactical scenarios. Improvements in road-drivability determination may come from acoustic or tactile sensor innovation because those mediums are largely untapped due to current range limitations.

Other areas for improvement include finding more effective ways to conduct sensor data fusion between modules and providing increased sensor redundancy to allow for higher speeds of travel. With high-speed operations comes the problem of excessive vibration. People, mechanical components, and electronics can only withstand so much vibrational energy before being damaged. Therefore, advances in terrain roughness estimation are needed. This capability remains underdeveloped because on-road autonomous driving dominates the marketplace. In the 2005 Defense Advanced Research Projects Agency Grand Challenge race, the Stanford University robot *Stanley* encountered vibration problems and was subsequently forced to run at reduced speeds (Thrun et al. 2006). After the race, the Stanford team was able to generate a roughness estimation of the ground up to 20 m ahead of the vehicle using several LiDAR units. This estimation allowed for subsequent higher-speed operation (Stavens and Thrun 2012). Fourteen years have passed with few advancements addressing terrain roughness and the excessive vibrations it creates, and yet technology now exists that could mitigate their effects.

Vegetation detection is another off-road technology gap. If an AV is to effectively operate in vegetated environments, then it must be able to determine the density and possibility of success for driving through vegetation. One approach for designing a vegetation detection module might combine ray tracing algorithms (that reliably approximate the density of sparse trees) with wavelengths of light (Figure 8) to classify vegetation. Light wavelengths can be used to detect the ground beneath the vegetation and determine the appearance of the vegetation. When this information is combined with camera input, and heat signatures from thermal cameras, improvements can be made in vegetation detection that increase driving confidence and thus operating speeds.

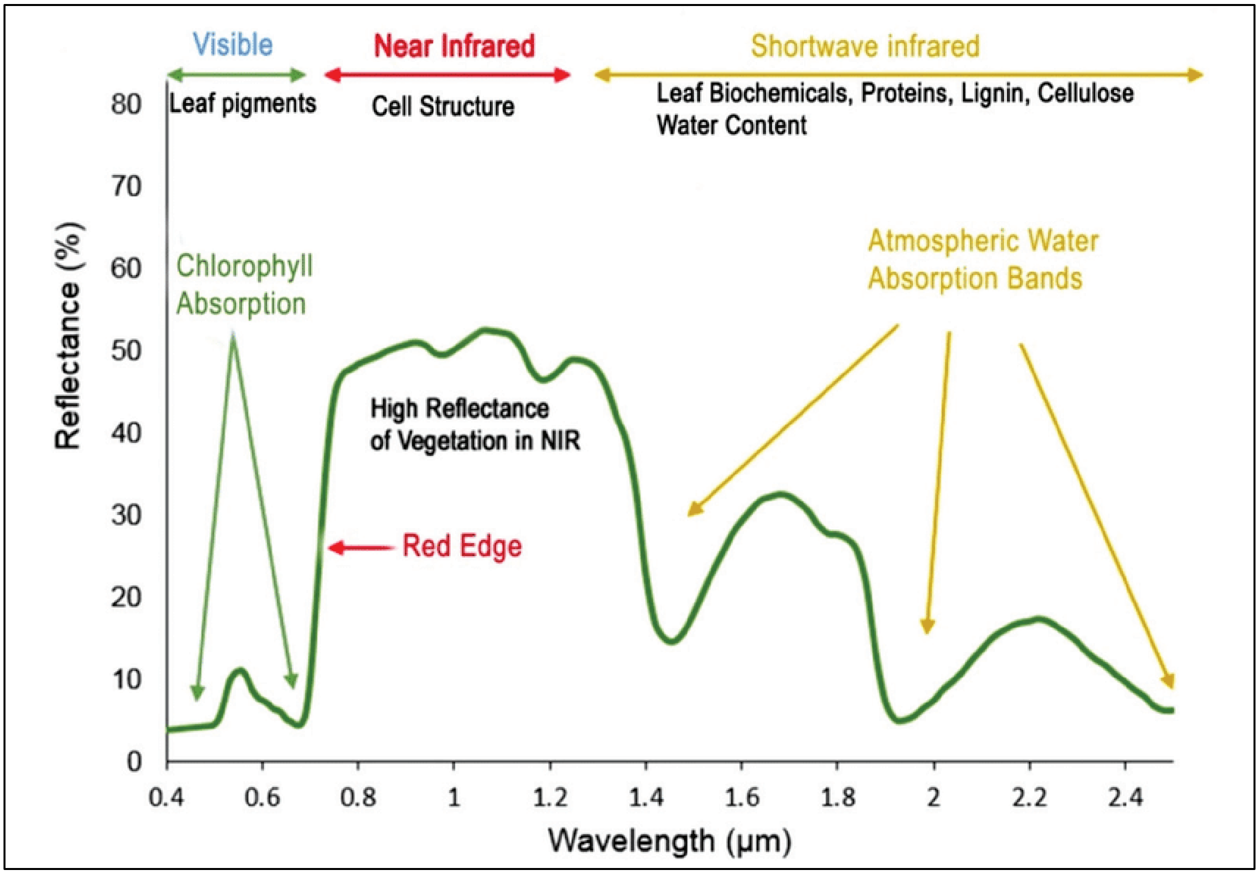

Figure 8. Spectral reflectance curve of vegetation showing use of wavelengths in vegetation classification (Roman and Ursu 2016).

SUMMARY: Today's AV platforms typically employ RADAR, LiDAR, and RGB cameras in some combination to sense and react to the world around them. This study reveals that other specialty technologies either exist or will soon mature that could improve current AV operability or produce wholly new capabilities that allow AVs to operate in edge cases. The US military has an active interest in these developments for its strategic application.

ADDITIONAL INFORMATION: This Construction Engineering Research Laboratory (CERL) technical note was prepared as part of the US Army Corps of Engineers, ATI program, by Annette Stumpf, Matthew Gross, Julie Webster, US Army Engineer Research and Development Center, CERL, Champaign, IL. Questions pertaining to this CERL technical note may be directed to Matthew T. Gross, Research Electrical Engineer, [Matthew.T.Gross@usace.army.mil.](mailto:Matthew.T.Gross@usace.army.mil)

This ERDC/CERL TN-21-1 should be cited as follows:

> Gross, Matthew T., and Julie L. Webster. 2021. *Autonomous Transport Innovation: A Review of Enabling Technologies.* ERDC/CERL TN-21-1. Champaign, IL: US Army Engineer Research and Development Center. http://dx.doi.org/10.21079/11681/42028

REFERENCES

Armbrust, Christopher, Martin Proetzsch, and Karsten Berns. 2011. "RAVON: Behaviour-Based Off-Road Robot Navigation." *KI - Künstliche Intelligenz* May: 155–160. doi:10.1007/s13218-011-0090-2.

Barnard, Michael. 2020. "Tesla & Google Disagree about LIDAR — which is Right?" Accessed Feb 24. *https://cleantechnica.com/2016/07/29/tesla-google-disagree-lidar-right/*

Brock Industries. 2017. *Night Vision IR Illumination: What Is It and Why You Need It*. *https://www.ultrafire.com/image/cache/catalog/9241/837f54bb2d211601dea1983bd872d23a-1200x1200.jpg*

Cordiero C. 2011. US20140292090A1, "Implementing Wireless Power Transfer with 60 GHz mmWave Communication." *Google Patents*. Accessed 20 February 2020. *https://patents.google.com/patent/US20140292090A1/en*

FLIR. 2020. *Voyager III Visible and Thermal Comparison*. *https://www.netzerotools.com/flir-voyager-iii-thermal-imaging-night-vision-infrared-camera*

Hadsell, Raia, Pierre Sermanet, Jan Ben, Ayse Erkan, Marco Scoffier, Koray Kavukcuoglu, Urs Muller, and Yann LeCun. 2009. "Learning Long-range Vision for Autonomous Off-road Driving." *Journal of Field Robotics* 26(2): 120–144. doi:10.1002/rob.20276. *https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.20276*

Harvey T. 2018. "Human Error and Advanced Technology: An Analysis of an NTSB News Release on Self-Driving Vehicle Fatality." *Professional Safety* 63(11): 55.

Impey, C. 2020. *Properties of Electromagnetic Radiation*. *https://www.teachastronomy.com/textbook/Detecting-Radiation-from-Space/Properties-of-Electromagnetic-Radiation/*

Kenney, J. B. 2011. "Dedicated Short-Range Communications (DSRC) Standards in the United States." *Proceedings of the IEEE* 99(7). *https://ieeexplore.ieee.org/document/5888501*

Kocić, J., N. Jovičić, and V. Drndarević. 2019. "An End-to-End Deep Neural Network for Autonomous Driving Designed for Embedded Automotive Platforms." *Sensors*. (Basel, Switzerland) May 3,2019. *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6539483/#*

Kuhnt F. 2016." Functional System Architectures towards Fully Automated Driving." *2016 IEEE Intelligent Vehicles Symposium (IV)*. *https://www.researchgate.net/figure/The-information-flow-between-the-modules-of-the-automated-vehicle-A1-the-vehicle-that_fig2_304657131*

Perwass, C. 2014. *Light Field Cameras*. Raytrix. *https://raytrix.de/wp-content/uploads/software/Raytrix_LightFieldCamera_2015_print1.pdf*

Perwass, C., and L. Wietzke. 2012. "Single Lens 3D-Camera with Extended Depth-of-Field." doi:10.1117/12.909882. http://dx.doi.org/10.1117/12.909882.

Ponnimbaduge Perera, Tharindu D., Dushantha Nalin K. Jayakody, Shree Krishna Sharma, Symeon Chatzinotas, and Jun Li. 2015. "Simultaneous Wireless Information and Power Transfer (SWIPT): Recent Advances and Future Challenges." *IEEE Communications Surveys & Tutorials* 20(1): 264–302. *https://ieeexplore.ieee.org/document/8214104*

Roman, A., and T. Ursu. 2016. "Multispectral Satellite Imagery and Airborne Laser Scanning Techniques for the Detection of Archaeological Vegetation Marks." *Landscape Archaeology on the Northern Frontier of the Roman Empire at Porolissum—An Interdisciplinary Research Project*. *https://www.researchgate.net/figure/The-spectral-reflectance-curve-of-vegetation-The-major-absorption-and-reflectance_fig4_315797574*

Seker, C. 2018. “A Review of Millimeter Wave Communication for 5G.” *2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. 19–21 October 2018. https://ieeexplore.ieee.org/document/8567053

Shmoolik, Mangan. 2018. *Detecting the Unexpected: The Path to Road Obstacles Prevention in Autonomous Driving*. PDF of a presentation given by VAYAVISION for the AutonomousTech in TLV, Israel. http://www.autonomoustechtlv.com/Portals/90/Shmoolik_1.pdf

Stavens, David, and Sebastian Thrun. 2012. "A Self-Supervised Terrain Roughness Estimator for Off-Road Autonomous Driving." https://arxiv.org/abs/1206.6872

Thrun, Sebastian, Mike Montemerlo, Hendrik Dahlkamp, David Stavens, Andrei Aron, James Diebel, and Philip Fong. 2006. "Stanley: The Robot that Won the DARPA Grand Challenge." *Journal of Field Robotics* 23 (9): 661–692. doi:10.1002/rob.20147. https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.20147

Tynan, P. 2012. “Atmospheric Attenuation.” *Earthshine Enterprises*. http://propagation.ece.gatech.edu/ECE6390/project/Fall2012/Team06/Webpage%20Folder/Webpage%20Folder/atmosphere.html

Vivacqua, Rafael, Raquel Vassallo, and Felipe Martins. 2017. "A Low Cost Sensors Approach for Accurate Vehicle Localization and Autonomous Driving Application." *Sensors (Basel, Switzerland)* 17(10): 2359. doi:10.3390/s17102359. https://www.ncbi.nlm.nih.gov/pubmed/29035334

NOTE: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.