

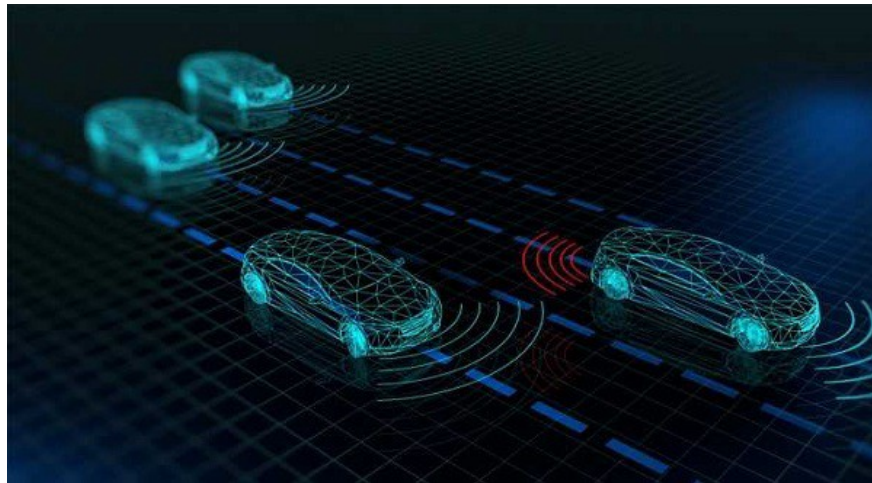
Multi-Sensor Data Fusion (MSDF) for Driverless Cars, An Essential Primer



Lance Eliot [Follow](#)

Apr 1 · 20 min read ★

Dr. Lance B. Eliot, AI Insider



Multiple sensors on self-driving cars requires proper sensor fusion

A crucial element of many AI systems is the capability to undertake Multi-Sensor Data Fusion (MSDF), consisting of collecting together and trying to reconcile, harmonize, integrate, and synthesize the data about the surroundings and environment in which the AI system is operating. Simple stated, the sensors of the AI system are the eyes, ears, and sensory input, while the AI must somehow interpret and assemble the sensory data into a cohesive and usable interpretation of the real world.

If the sensor fusion does a poor job of discerning what's out there, the AI is essentially blind or misled toward making life-or-death algorithmic decisions. Furthermore, the sensor fusion needs to be performed on a timely basis.

Humans do sensor fusion all the time, in our heads, though we often do not overtly put explicit thought towards our doing so. It just happens, naturally.

The other day, I was driving in the downtown Los Angeles area. There is always an abundance of traffic, including cars, bikes, motorcycles, scooters, and pedestrians that are prone to jaywalking. There is a lot to pay attention to. Is that bicyclist going to stay in the bike lane or decide to veer into the street?

I had my radio on, listening to the news reports, when I began to faintly hear the sound of a siren, seemingly off in the distance. As I strained to try and hear a siren, I also kept my eyes peeled, anticipating that if the siren was occurring nearby, there might be a police car or ambulance or fire truck that might soon go skyrocketing past me.

I decided that the siren was definitely getting more distinctive and pronounced. The echoes along the streets and buildings was creating some difficulty in deciding where the siren was coming from. I could not determine if the siren was behind me or somewhere in front of me.

At times like this, your need to do some sensor fusion is crucial.

Your eyes are looking for any telltale sign of an emergency vehicle. Maybe the flashing lights might be seen from a distance. Perhaps other traffic might start to make way for the emergency vehicle, and that's a visual clue that the vehicle is coming from a particular direction. Your ears are being used to do a bat-like echolocation of the emergency vehicle, using the sound to gauge the direction, speed, and placement of the speeding object.

I became quite aware of my having to merge together the sounds of the siren with my visual search of the traffic and streets. Each was feeding the other.

This is the crux of Multi-Sensor Data Fusion.

I had one kind of sensor, my eyes, providing visual inputs to my brain. I had another kind of sensor, my ears, providing acoustical inputs to my brain. My brain managed to tie together the two kinds of inputs. Not only were the inputs brought together, they were used in a means of each aiding the other. My visual processing led me to listen toward the sound. The sound led me to look toward where the sound seemed to be coming from.

My mind, doing some action planning of how to drive the car, melded together the visual and the acoustic, using it to guide how I would drive the car. In this case, I pulled the car over and came to a near stop. I also continued to listen to the siren. Only once it had gone far enough away, along with my not being able to see the emergency vehicle anymore, did I decide to resume driving down the street.

This whole activity of my doing the sensor fusion was something that played out in just a handful of seconds.

Suppose though that I was wearing my ear buds and listening to loud music while driving (not a wise thing to do when driving a car, and usually illegal to do while driving), and did not hear the siren? I would have been solely dependent upon my sense of sight. Usually, it is better to have multiple sensors active and available when driving a car, giving you a more enriched texture of the traffic and the driving situation.

Notice too that the siren was hard to pin down in terms of where it was coming from, along with how far away it was. This highlights the aspect that the sensory data being collected might be only partially received or might otherwise be scant, or even faulty.

Another aspect involves attempting to square together the inputs from multiple sensors. Imagine if the siren was getting louder and louder, and yet I did not see any impact to the traffic situation, meaning that no other cars changed their behavior and those pedestrians kept jaywalking. That would have been confusing.

In this case, I was alone in my car. Only me and my own “sensors” were involved in this multi-sensor data fusion. You could have more such sensors, such as when having passengers that can aid you in the driving task.

Peering Into The Fog With Multiple Sensory Devices

I recall during my college days a rather harried driving occasion. While driving to a college basketball game, I managed to get into a thick bank of fog. Some of my buddies were in the car with me.

Here's what happened.

My buddy in the front passenger seat offered to intently watch for anything to my right. The two friends in the back seat were able to turnaround and look out the back window. I suddenly had the power of six additional eyeballs, all looking for any other traffic. They began each verbally reporting their respective status.

All in all, we made it to the basketball game without nary a nick. It was a bit alarming though and a situation that I will always remember. There we were, working as a team, with me as the driver at the wheel. I had to do some real sensor fusion. I was receiving data from my own eyes, along with hearing from my buddies, and having to mentally combine together what they were telling me with what I could actually see.

When you are driving a car, you often are doing Multi-Target Tracking (MTT). This involves identifying particular objects or “targets” that you are trying to keep an eye on. While driving in downtown Los Angeles, my “targets” included the many cars, bike rides, and pedestrians. While driving in the foggy evening, we had cars coming from the right and from behind.

Your Field of View (FOV) is another vital aspect of driving a car and using your sensory apparatus. During the fog, my own FOV was narrowed to what I could see on the driver’s side of the car, and I could not see anything from behind the car. Fortunately, my buddies provided additional FOV’s. My front passenger was able to augment my FOV by telling me what was seen to the right of the car. The two in the backseat had a FOV of what was behind the car.

Those two stories that I’ve told are indicative of how we humans do our sensor fusion while driving a car.

In the news recently there has been the story about the Boeing 737 MAX 8 airplane and in particular two horrific deadly crashes. Some believe that the sensors on the plane were a significant contributing factor to the crashes. Though the matters are still being investigated, it is a potential example of the importance of Multi-Sensor Data Fusion and has lessons that can be applied to driving a car and advanced automation used to do so.

Multi-Sensor Data Fusion for AI Self-Driving Cars

What does this have to do with AI self-driving driverless cars?

At the Cybernetic AI Self-Driving Car Institute, we are developing AI software for self-driving cars. One important aspect involves the design, development, testing, and fielding of the Multi-Sensor Data Fusion.

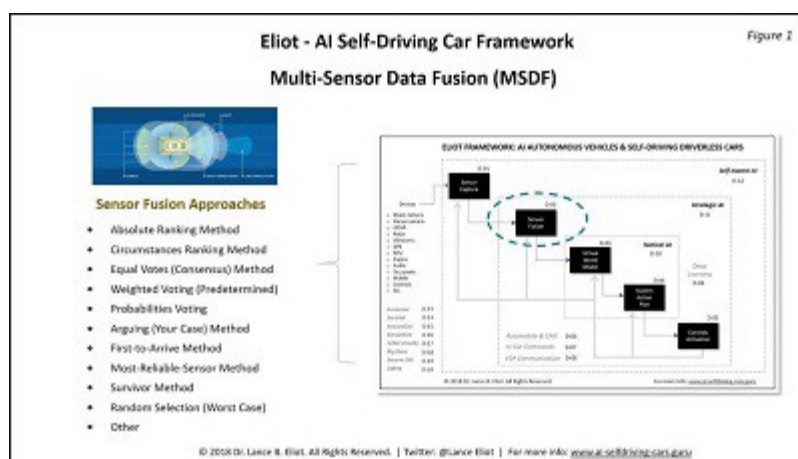
I'd like to first clarify and introduce the notion that there are varying levels of AI self-driving cars. The topmost level is considered Level 5. A Level 5 self-driving car is one that is being driven by the AI and there is no human driver involved.

For self-driving cars less than a Level 5, there must be a human driver present in the car. The human driver is currently considered the responsible party for the acts of the car. The AI and the human driver are co-sharing the driving task.

Another key aspect of AI self-driving cars is that they will be driving on our roadways in the midst of human driven cars too.

Returning to the topic of Multi-Sensor Data Fusion, let's walk through some of the key essentials of how AI self-driving cars undertake such efforts.

Take a look at Figure 1.



I've shown my overall framework about AI self-driving cars and highlighted the sensor fusion stage of processing.

Per my earlier remarks about the crucial nature of sensor fusion, consider that if the sensor fusion goes awry, it means that the stages downstream are going to be either without needed information or be using misleading information.

One of the major challenges for sensor fusion involves dealing with how to collectively stitch together the multitude of sensory data being collected.

You are going to have the visual data collected via the cameras, coming from likely numerous cameras mounted at the front, back, and sides of the self-driving car. There is the radar data collected by the multiple radar sensors mounted on the self-driving car. There are likely ultrasonic sensors. There could be LIDAR sensors, a special kind of sensor that combines light and radar.

Thus, you will have to stitch together sensor data from like-sensors, such as the data from the various cameras. Plus, you will have to stitch together the sensor data from unlike sensors, meaning that you want to do a kind of comparison and contrasting with the cameras, with the radar, with the LIDAR, with the ultrasonic, and so on.

Each different type or kind of sensor provides a different type or kind of potential indication about the real-world. They do not all perceive the world in the same way. This is both good and bad.

The good aspect is that you can potentially achieve a rounded balance by using differing kinds or types of sensors. Cameras and visual processing are usually not as adept at being indicative of the speed of an object as does the radar or the LIDAR. By exploiting the strengths of each kind of sensor, you are able to have a more enriched texturing of what the real-world consists of.

If the sensor fusion subsystem is poorly devised, it can undermine this complimentary triangulation that having differing kinds of sensors inherently provides.

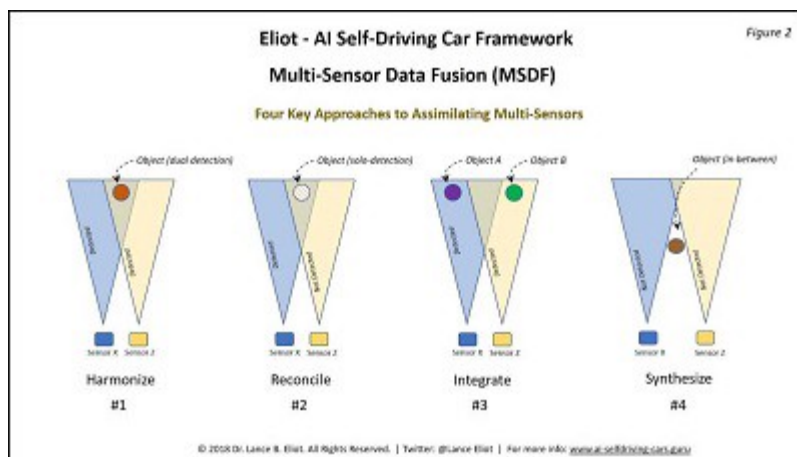
Let's though all acknowledge that the more processing you do of the multitude of sensors, the more computer processing you need, which then means that you have to place more computer processors and memory on-board the self-driving car. This adds cost, it adds weight to

the car, it consumes electrical power, it generates heat, and has other downsides.

Four Key Ways Approaches to MSDF Assimilation

Let's consider the fundamental ways that you assimilate together the sensory data from multiple sensors.

Take a look at Figure 2.



I'll briefly describe the four approaches, consisting of harmonize, reconcile, integrate, and synthesize.

- Harmonize

Assume that you have two different kinds of sensors, I'll call them sensor X and sensor Z. They each are able to sense the world outside of the self-driving car. We won't concern ourselves for the moment with their respective strengths and weaknesses, which I'll be covering later on herein.

There is an object in the real-world and the sensor X and the sensor Z are both able to detect the object. This could be a pedestrian in the street, or maybe a dog, or could be a car. In any case, I'm going to simplify the matter to considering the overall notion of detecting an object.

This dual detection means that both of the different kinds of sensors have something to report about the object. We have a dual detection of

the object. Now, we want to figure out how much more we can discern about the object because we have two perspectives about it.

This involves harmonizing the two reported detections. Let's pretend that both sensors detect the distance of the object. And, sensor X indicates the object is six feet tall and about two feet wide. Meanwhile, sensor Z is reporting that the object is moving toward the self-driving car, doing so at a speed of a certain number of feet per second N. We can combine together the two sensor reports and update the virtual world model that there is an object of six feet in height, two feet in width, moving toward the self-driving car at some speed N.

Suppose we only relied upon sensor X. Maybe because we only have sensor X and there is no sensor Z on this self-driving car. Or, sensor Z is broken. Or, sensor Z is temporarily out of commission because there is a bunch of mud sitting on top of the sensor. In this case, we would know only the height and weight and general position of the object, but not have a reading about its speed and direction of travel. That would mean that the AI action planner is not going to have as much a perspective on the object as might be desired.

As a quick aside, this also ties into ongoing debates about which sensors to have on AI self-driving cars. For example, one of the most acrimonious debates involves the choice by Tesla and Elon Musk to not put LIDAR onto the Tesla cars. Elon has stated that he doesn't believe LIDAR is needed to achieve a true AI Level 5 self-driving car via his Autopilot system, though he also acknowledges that he might ultimately be proven mistaken by this assumption.

Some would claim that the sensory input available via LIDAR cannot be otherwise fully devised via the other kinds of sensors, and so in that sense the Teslas are not going to have the same kind of complimentary or triangulation available that self-driving cars with LIDAR have. Those that are not enamored of LIDAR would claim that the LIDAR sensory data is not worth the added cost, nor worth the added processing effort, nor worth the added cognition time required for processing.

- Reconcile

I'd like to revisit the use of sensor X and sensor Z in terms of object detection.

Let's pretend that sensor X detects an object, and yet sensor Z does not, even though the sensor Z could have. In other words, the object is within the Field of View (FOV) of sensor Z, and yet sensor Z isn't detecting the object. Note that this is vastly different than if the object were entirely outside the FOV of sensor Z, in which case we would not have any expectation that sensor Z could detect the object.

We have a bit of a conundrum on our hands that needs reconciling.

Sensor X says the object is there in the FOV. Sensor Z says the object is not there in the same FOV intersection. Yikes! It could be that sensor X is correct and sensor Z is incorrect. Perhaps sensor Z is faulty, or obscured, or having some other difficulty. On the other hand, maybe sensor X is incorrect, namely that there isn't an object there, and the sensor is X is mistaken, reporting a "ghost" of sorts, something that is not really there, while sensor Z is correct in reporting that there isn't anything there.

There are various means to try and reconcile these seemingly contradictory reports. I'll be getting to those methods shortly herein.

- Integrate

Let's suppose we have two objects. One of those objects is in the FOV of sensor X. The other object is within the FOV of sensor Z. Sensor X is not able to directly detect the object that sensor Z has detected, rightfully so because the object is not inside the FOV of sensor X. Sensor Z is not able to directly detect the object that sensor X has detected, rightfully so because the object is not inside the FOV of sensor Z.

Everything is okay in that the sensor X and sensor Z are both working as expected.

What we would like to do is see if we can integrate together the reporting of sensor X and sensor Z. They each are finding objects in their respective FOV. It could be that the object in the FOV of sensor Z is heading toward the FOV of sensor X, and thus it might be possible to inform sensor X to especially be on the watch for the object. Likewise, the same could be said about the object that sensor X currently has detected and might forewarn sensor Z.

My story about driving in the fog is a similar example of integrating together sensory data.

- Synthesize

In the fourth kind of approach about assimilating together the sensory data, you can have a situation whereby neither sensor X and nor sensor Z has an object within their respective FOV's. In this case, the assumption would be that neither one even knows that the object exists.

You sometimes have a chance at guessing about objects that aren't in the FOV's of the sensors by interpreting and interpolating whatever you do know about the objects within the FOV's of the sensors. This is referred to as synthesis or synthesizing of sensor fusion.

Remember how I mentioned that I saw other cars moving over when I was hearing the sounds of a siren. I could not see the emergency vehicle. Luckily, I had a clue about the emergency vehicle because I could hear it. Erase the hearing aspects and pretend that all that you had was the visual indication that other cars were moving over to the side of the road.

Within your FOV, you have something happening that gives you a clue about what is not within your FOV. You are able to synthesize what you do know and use that to try and predict what you don't know. It seems like a reasonable guess that if cars around you are pulling over, it suggests an emergency vehicle is coming. I guess it could mean that aliens from Mars have landed and you didn't notice it because you were strictly looking at the other cars, but I doubt that possibility of those alien creatures landing here.

So, you can use the sensory data to try and indirectly figure out what might be happening in FOV's that are outside of your purview. Keeping in mind that this is real-time system and that the self-driving car is in motion, it could be that within moments the thing that you guessed might be in the outside of scope FOV will come within the scope of your FOV, and hopefully you'll have gotten ready for it. Just as I did about the ambulance that zipped past me.

Voting Methods of Multi-Sensor Data Fusion

When you have multiple sensors and you want to bring together in some cohesive manner their respective reporting, there are a variety of methods you can use.

Take a look at Figure 1 again.

I'll briefly describe each of the voting methods.

- Absolute Ranking Method

In this method, you beforehand decide a ranking of sensors. You might declare that the cameras are higher ranked than the radar. The radar you might decide is higher ranked than the LIDAR. And so on. During sensor fusion, the subsystem uses that predetermined ranking.

For example, suppose you get into a situation of reconciliation, such as the instance I described earlier involving sensor X detecting an object in its FOV but that sensor Z in the intersecting FOV did not detect. If sensor X is the camera, while sensor Z is the LIDAR, you might simply use the pre-determined ranking and the algorithm assumes that since the camera is higher ranking it is “okay” that the sensor Z does not detect the object.

There are trade-offs to this approach. It tends to be fast, easy to implement, and simple. Yet it tends toward doing the kind of “tossing out” that I forewarned is not usually advantageous overall.

- Circumstances Ranking Method

This is similar to the Absolute Ranking Method but differs because the ranking is changeable depending upon the circumstance in-hand. For example, we might have setup that if there is rainy weather, the camera is no longer the top dog and instead the radar gets the topmost ranking, due to its less likelihood of being adversely impacted by the rain.

There are trade-offs to this approach too. It tends to be relatively fast, easy to implement, and simple. Yet it once again tends toward doing the kind of “tossing out” that I forewarned is not usually advantageous overall.

- Equal Votes (Consensus) Method

In this approach, you allow each sensor to have a vote. They are all considered equal in their voting capacity. You then use a counting algorithm that might go with a consensus vote. If some threshold of the sensors all agrees about an object, while some do not, you allow the consensus to decide what the AI system is going to be led to believe.

Like the other methods, there are trade-offs in doing things this way.

- Weighted Voting (Predetermined)

Somewhat similar to the Equal Votes approach, this approach adds a twist and opts to assume that some of the voters are more important than the others. We might have a tendency to believe that the camera is more dependable than the radar, so we give the camera a higher weighted factor. And so on.

Like the other methods, there are trade-offs in doing things this way.

- Probabilities Voting

You could introduce the use of probabilities into what the sensors are reporting. How certain is the sensor? It might have its own controlling subsystem that can ascertain whether the sensor has gotten bona fide readings or maybe has not been able to do so. The probabilities are then encompassed into the voting method of the multiple sensors.

Like the other methods, there are take-offs in doing things this way.

- Arguing (Your Case) Method

A novel approach involves having each of the sensors argue for why their reporting is the appropriate one to use. It's an intriguing notion. We'll have to see whether this can demonstrate sufficient value to warrant being used actively. Research and experimentation are ongoing.

Like the other methods, there are trade-offs in doing things this way.

- First-to-Arrive Method

This approach involves declaring a kind of winner as to the first sensor that provides its reporting is the one that you'll go with. The advantage

is that for timing purposes, you presumably won't wait for the other sensors to report, which then speeds up the sensor fusion effort. On the other hand, you don't know if a split second later one of the other sensors might report something of a contrary nature or that might be an indication of imminent danger that the first sensor did not detect.

Like the other methods, there are trade-offs in doing things this way.

- Most-Reliable Method

In this approach, you keep track of the reliability of the myriad of sensors on the self-driving car. The sensor that is most reliable will then get the nod when there is a sensor related data dispute.

Like the other methods, there are trade-offs in doing things this way.

- Survivor Method

It could be that the AI self-driving car is having troubles with the sensors. Maybe the self-driving car is driving in a storm. Several of the sensors might not be doing any viable reporting. Or, perhaps the self-driving car has gotten sideswiped by another car, damaging many of the sensors. This approach involves selecting the sensors based on their survivor-ship.

Like the other methods, there are trade-offs in doing things this way.

- Random Selection (Worst Case)

One approach that is obviously controversial involves merely choosing among the sensor fusion choice by random selection, doing so if there seems to not be any other more systemic way to choose between multiple sensors if they are in disagreement about what they have or have not detected.

Like the other methods, there are trade-offs in doing things this way.

- Other

You can use several of these methods in your sensor fusion subsystem. They can each come to play when the subsystem determines that one approach might be better than the other.

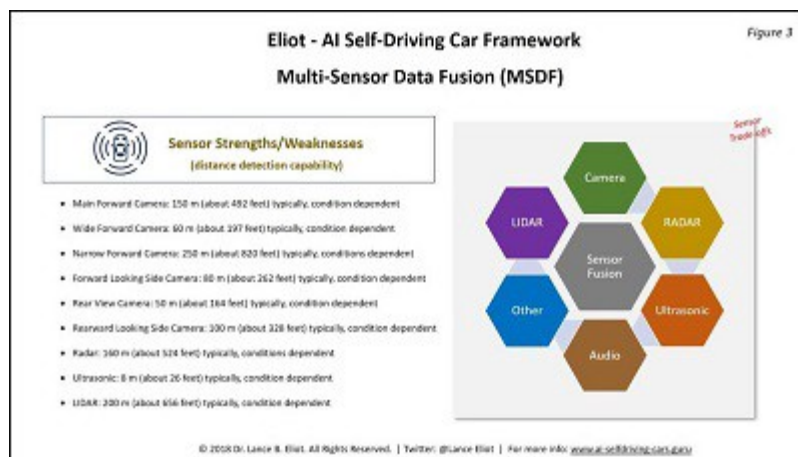
There are other ways that the sensor fusion voting can also be arranged.

How Multiple Sensors Differ is Quite Important

Your hearing is not the same as your vision. When I heard a siren, I was using one of my senses, my ears. They are unlike my eyes. My eyes cannot hear, at least I don't believe they can. This highlights that there are going to be sensors of different kinds.

An overarching goal or structure of the Multi-Sensor Data Fusion involves trying to leverage the strengths of each sensor type, while also minimizing or mitigating the weaknesses of each type of sensor.

Take a look at Figure 3.



One significant characteristic of each type of sensor will be the distance at which it can potentially detect objects. This is one of the many crucial characteristics about sensors.

The further out that the sensor can detect, the more lead time and advantage goes to the AI driving task. Unfortunately, often the further reach also comes with caveats, such as the data at the far ends might be lackluster or suspect. The sensor fusion needs to be established as to the strengths and weaknesses based on the distances involved.

Here's the typical distances for contemporary sensors, though keep in mind that daily improvements are being made in the sensor technology and these numbers are rapidly changing accordingly.

- Main Forward Camera: 150 m (about 492 feet) typically, condition dependent
- Wide Forward Camera: 60 m (about 197 feet) typically, condition dependent
- Narrow Forward Camera: 250 m (about 820 feet) typically, conditions dependent
- Forward Looking Side Camera: 80 m (about 262 feet) typically, condition dependent
- Rear View Camera: 50 m (about 164 feet) typically, condition dependent
- Rearward Looking Side Camera: 100 m (about 328 feet) typically, condition dependent
- Radar: 160 m (about 524 feet) typically, conditions dependent
- Ultrasonic: 8 m (about 26 feet) typically, condition dependent
- LIDAR: 200 m (about 656 feet) typically, condition dependent

There are a number of charts that attempt to depict the strengths and weaknesses when comparing the various sensor types. I suggest you interpret any such chart with a grain of salt. I've seen many such charts that made generalizations that are either untrue or at best misleading.

Also, the number of criteria that can be used to compare sensors is actually quite extensive, and yet the typical comparison chart only picks a few of the criteria. Once again, use caution in interpreting those kinds of short shrift charts.

Take a look at Figure 4 for an indication about the myriad of factors involved in comparing different types of sensors.

As shown, the list consists of:

- Object detection
- Object distinction
- Object classification
- Object shape
- Object edges
- Object speed
- Object direction
- Object granularity
- Maximum Range
- Close-in proximity
- Width of detection
- Speed of detection
- Nighttime impact
- Brightness impact
- Size of sensor
- Placement of sensor
- Cost of sensor

- Reliability of sensor
- Resiliency of sensor
- Snow impact
- Rain impact
- Fog impact
- Software for sensor
- Complexity of use
- Maintainability
- Repairability
- Replaceability
- Etc.

Conclusion

It seems that the sensors on AI self-driving cars get most of the glory in terms of technological 'wizardry and attention. The need for savvy and robust Multi-Sensor Data Fusion does not get much airplay. As I hope you have now discovered, there is an entire and complex effort involved in doing sensor fusion.

Humans appear to easily do sensor fusion. When you dig into the details of how we do so, there is a tremendous amount of cognitive effort involved. For AI self-driving cars, we need to continue to press forward on ways to further enhance Multi-Sensor Data Fusion. The future of AI self-driving cars and the safety of those that use them are dependent upon MSDF. That's a fact.

For free podcast of this story, visit: <http://ai-selfdriving-cars.libsyn.com/website>

The podcasts are also available on Spotify, iTunes, iHeartRadio, etc.

More info about AI self-driving cars, see: www.ai-selfdriving-cars.guru

To follow Lance Eliot on Twitter: @LanceEliot

Copyright 2019 Dr. Lance Eliot

