



# The Future of Autonomous Vehicles

Think Like a Robot, Perceive Like a Human

By Dr. James Doty and Blair LaCorte

FUT1WP\_21\_0226



1 Park Place, Suite 200, Dublin, CA 94568 • 925.400.4366

© 2021 AEye, Inc. All rights reserved.

## Introduction

*By James Doty, MD*

For over three decades, I've studied and performed surgery on the human brain. I have always been fascinated by the power, plasticity and adaptability of the brain, and by how much of its amazing capacity is dedicated to processing and interpreting data we receive from our senses. With the rapid ascension of Artificial Intelligence (AI), I began to wonder how developers would integrate the complex, multi-layers of human perception to enhance AI's capabilities. I have been especially interested in how this integration would be applied to meet the demands of the entire continuum of autonomy: from ADAS to mobility, as well as automated and partially automated applications in trucking, transit, construction, rail, intelligent traffic systems (ITS), aerospace and defense. To me, it is clear that the artificial intelligence needed to drive these applications will require artificial perception modeled after the greatest perception engine on the planet—the human visual cortex. These vehicles will need to think like a robot, but perceive like a human.

To learn more and to better understand how this level of artificial perception will be created, I became an advisor to [AEye](#), a company developing cutting edge artificial perception and self-driving technologies, helping them use knowledge of the human brain to better inform their systems. This is known as biomimicry: the concept of learning from and replicating natural strategies from living systems and beings (plants, animals, humans, etc.) to create more responsive and intelligent technologies and products. Essentially, biomimicry allows us to fit into our existing environment and evolve in the way life has successfully done for billions of years. But why is incorporating biomimicry and aspects of human perception integral to the development and success of autonomous vehicles?

Because nothing can take in more information and process it faster and more accurately than the human visual cortex. Humans classify complex objects at speeds up to 27Hz, with the brain processing 580 megapixels of data in as little as 13 milliseconds. If we continue using conventional sensor data collection methods, we are more than 25 years away from having AI achieve the capabilities of the human brain in robots and autonomous vehicles. Therefore, to enable self-driving cars to safely move independently in crowded urban environments or at highway speeds, we must develop new approaches and technologies to meet or exceed the performance of the human brain. The next question is: *how?*



## Not All Objects Are Created Equal (See everything, and focus on what is important)

Humans continuously analyze their environment, always scanning for new objects, then in parallel (and as appropriate) focus in on elements that are either interesting or potentially pose a threat. We process at the visual cortex very fast, with incredible accuracy, using very little of the brain's immense processing power. If a human brain functioned as autonomous vehicles do today, we would not have survived as a species.

In his book *The Power of Fifty Bits*, Bob Nease writes of the ten million bits of information the human brain processes each second, but how only fifty bits are devoted to conscious thought. This is due to multiple evolutionary factors, including our adaptation to ignore autonomic processes like our heart beating, or our visual cortex screening out less relevant information in our surroundings not needed for survival.

This is the nature of our intelligent vision. While our eyes are always scanning and searching to identify new objects entering a scene, we focus our attention on objects that matter as they move into areas of concern, allowing us to determine an appropriate response. In short, we *search* a scene, consciously *acquire* the objects that matter, and *act* on them as required.

Current autonomous vehicle sensor configurations utilize a combination of LiDAR, cameras, ultrasonics, and radar as their “senses” that collect serially (one way) and are limited to fixed patterns of search. These “senses” collect as much data as possible, which is then aligned, processed, and analyzed long after the fact. This post-processing is slow and does not allow for situational changes to how sensory data is captured in real time. Because these sensors don't intelligently interrogate, up to 90% of the sensory data collected is thrown out as it is either irrelevant or redundant by the time it is processed. This act of triage also comes with a latency penalty. At highway speeds, this latency results in a car moving more than 20 feet before the sensor data has been fully processed. Throwing away data you don't need with the goal of being efficient is inefficient. A better approach exists.

The overwhelming task of sifting through this data—every tree, curb, parked vehicle, the road, and other static objects—also requires immense power and data processing resources, which slows down the entire system significantly, and introduces risk. These systems' goal is to focus on everything and then try to analyze each item in their environment, after the fact, at the expense of timely action. This is the exact opposite of how humans process spatial and temporal data in situations that we associate with driving.

**AEye's intelligent LiDAR sensor system, iDAR™ (Intelligent Detection and Ranging)**, enables autonomous vehicles to “search, acquire, and act” as humans do. It does this by defining new data and sensor types that more efficiently communicate actionable information while maintaining the intelligence to analyze this data as quickly and accurately as possible. It achieves this through its innovative high-performance, solid-state active LiDAR.

The iDAR platform can be broken down into four simple levels, each designed to meet the needs of specific use cases or applications in mobility, ADAS, trucking, transit, construction, rail, intelligent traffic systems (ITS), aerospace, and beyond. These four levels are: *iDAR at Design*, *Triggered iDAR*, *Responsive iDAR*, and *Predictive iDAR*. *iDAR at Design* allows for the creation of a single, deterministic scan pattern to deliver optimal information for use cases like powerline inspection, while with *Triggered iDAR*, a library of



deterministic patterns can be triggered by external input, such as maps, speed, and weather. I'll circle back to *Predictive iDAR* later on in this paper, but first, I'd like to discuss, in depth, *Responsive iDAR*.

In this level, the entire iDAR platform is completely situationally aware, adjusting, in real time, how it scans the scene, and where to apply density and extra power. Feedback loops (which we will discuss later) and other sensors, such as camera and radar, inform the LiDAR to focus on specific points of interest. The system is intelligent, proactively understanding and interrogating a scene, and perpetually optimizing its own scan patterns and data collection to focus on the information that matters most. This is akin to human perception.

Therefore, unlike standard LiDAR, AEye's active LiDAR is situationally adaptive, modifying scan patterns and trade resources such as power, update rate, resolution, returns, and range. This enables iDAR to dynamically utilize sensor resources to optimally search a scene, efficiently identify and acquire critical objects, such as a child walking into the street or a car entering an intersection, and determine the appropriate course of action. Doing this in real time is the difference between a safe journey and an avoidable tragedy.

Ultimately, iDAR is the only artificial perception platform for ADAS, mobility, and beyond that is truly software-configurable, enabling intelligent and active sensing. By using active LiDAR to intelligently collect data at the sensor level, iDAR is designed to adapt to new technologies and algorithms, continually evolving to minimize cost and maximize performance.

The iDAR platform enables customizable data collection based on the customer or host's applications, as well as offers a myriad of adjustable and customizable configuration settings to meet the demands of autonomous applications.

## Humans Learn Intuitively (Feedback loops enable intelligence)

As previously mentioned, the human visual cortex can scan at 27Hz, while current sensors on autonomous vehicles average around 10Hz. The brain naturally gathers information from the visual cortex, creating feedback loops that help make each step in the perception process more efficient. The brain then provides context that directs the eyes to search and focus on certain objects, to identify and prioritize them, and decide on the most effective course of action, while largely ignoring other objects of less importance. This prioritization allows for greater efficiency and increases temporal and spatial sampling, not only scanning smarter, but scanning better.

Try it yourself. Look around, and notice how there is a multitude of depths, colors, shadows, and other information to capture with your eyes—and then there is motion. Then, consider what you know from experience about what you are seeing: *Is a certain object capable of movement or is it likely to remain static? Could it behave predictably or erratically? Do you find value in the object, or do you consider it disposable?* While you don't consciously make these observations, your brain does.

Current sensor systems on autonomous or semi-autonomous vehicles are optimized for "search" which is then reported back to a central processor. The search is done with individual passive sensors which apply the same power, intensity, and search pattern everywhere at every time—regardless of changes in the environment.



Even more limiting: data flows only one way—from the passive sensors to the central processor—with no ability to actively adapt or adjust its collection. All intelligence is added after it has been fused and decimated, with up to 95% thrown out, when it is too late to learn and adjust in real time.

AEye's iDAR is an active, multidimensional system that relies on feedback loops to efficiently and effectively cycle information to modify reactions appropriately, and in real time, just like in humans. The camera can talk to the LiDAR and the sensor system can talk to the path planning system simultaneously, in real time.

In addition to improving response time, these feedback loops enable artificial intelligence to be more effectively integrated with artificial perception. Today's sensor systems passively return the same type of data no matter the situation. Pushing sensory data capture and processing to the sensor, rather than the centralized processor, enables faster integrated feedback loops to inform and queue actions. In this way, the iDAR system is able to continually learn so that over time, it can become even more efficient at identifying and tracking objects and situations that could threaten the safety of the autonomous vehicle, its passengers, other drivers, and pedestrians.

This is what AEye calls *Predictive iDAR*. Just like a human, *Predictive iDAR* understands the motion of everything it sees, which enables the system to deliver more information with less data, focusing its energy on the most important objects in a scene while paying attention to everything else in its periphery. The end result of Predictive iDAR is Motion Forecasting through neural networks. Like human intuition, Predictive iDAR can “sense” (i.e., predict) where an object will be at different times in the future, helping the vehicle to assess the risk of collision and charter a safe course).

## Orthogonal Data Matters (Creating an advanced, multidimensional data type)

Orthogonal data refers to complimentary data sets which ultimately give us more quality information about an object or situation than each would alone. This allows us to identify what in our world is critically important, and what is not. Orthogonality concepts for high information quality are well understood and rooted in disciplines such as quantum physics where linear algebra is employed and orthogonal basis sets are the minimum pieces of information one needs to represent more complex states without redundancy.

When it comes to perception of moving objects, two types of critical orthogonal data sets are often required: spatial and temporal. Spatial data specifies where an object exists in the world, while temporal is where an object exists in time. By integrating these data sets along with others, such as color, temperature, sound, and smell, our brains generate a real-time model of the world around us, defining how we experience it.

The human brain takes in all kinds of orthogonal data naturally, decoupling and reassembling information instantaneously, without us even realizing it. For example, if you see that a baseball is flying through the air towards you, your brain is gathering all types of information about it, such as spatial (the direction of where the ball is headed) and temporal (how fast it's moving). While this data is being processed by your visual cortex “in the background,” all you're ultimately aware of is the *action* you need to take, which might be to *duck*. The AI perception technology that is able to successfully adopt the manner by which the human brain captures and processes these types of data sets will dominate the market.



Existing robotic sensor systems have focused only on single sensor modalities (camera, LiDAR, or radar) and only with fixed scan patterns and intensity. Unlike humans, these systems have not learned (nor have the ability) to efficiently process and optimize 2D and 3D data in real time while both the sensor and detected objects are in motion. To put it simply, they cannot use real-time orthogonal data to learn, prioritize, and focus. Effectively replicating the multidimensional sensory processing power of the human visual cortex requires a new approach to thinking about how to capture and process sensory data.

AEye's intelligent approach to sensing gives them the unique ability to capture "multiple senses" and gain a broader understanding of a vehicle's surroundings. By physically fusing their solid-state, active LiDAR with a hi-res camera, AEye has created a new data type called a Dynamic Vixel, which captures both camera pixels and 3D LiDAR voxels. Capturing both RGB and XYZ data allows vehicles to visualize like (if not better than) humans. Not only does color dominate our driving infrastructure (such as signage or traffic lights), it's one of the main drivers of the visual cortex. And Dynamic Vixels were expressly created to biomimic the data structure of the human visual cortex.

Like the human visual cortex, the intelligence inherent in Dynamic Vixels is then integrated in the central perception engine and motion planning system (which is the functional brain of the vehicle). They are dynamic because they have the ability to adjust to changing conditions, such as increasing the power level of the sensor to cut through rain, or revisiting suspect objects in the same frame to identify obstacles. Better data drives more actionable information.

## Looking Beyond Perception

It is exciting to witness how visionaries are empowering machines to more closely perceive the environment as a human would: evaluating risk, accurately gathering information, and responding and adapting to constantly changing conditions. This kind of autonomous vehicle and artificial intelligence integration could all but eliminate the woes of modern car culture.

What does it mean to think like a robot, but perceive like a human? It means logically mitigating human foibles like aggressive behavior, and avoiding the risks of fatigue, distraction, or alcohol. All this while striving to meet and exceed the capabilities of the human visual cortex and brain, the most powerful perception engine ever created. By doing this, we will ultimately save time, money, reduce stress, and improve our safety.

For autonomous vehicles, biomimicry informs us that artificial perception should put more processing at the sensor in order to function efficiently. By doing just that, AEye's iDAR sensor system has changed the way robotic vision is created and has defined new benchmarks for performance. iDAR sensors are the only ones that can successfully handle the industry's most challenging edge cases, and while achieving a scan rate in excess of 200Hz (6x human vision) with a detection range of one kilometer (3-5x current LiDAR sensors)—shattering records by going further and faster than any other sensor system and "driving" the promise of safe vehicle autonomy.



## About the Authors

**James R. Doty, MD**, is a clinical professor in the Department of Neurosurgery at Stanford University School of Medicine. He is also the founder and director of the Center for Compassion and Altruism Research and Education at Stanford University. He works with scientists from a number of disciplines examining the neural bases for compassion and altruism. He holds multiple patents and is the former CEO of Accuray. Dr. Doty is the New York Times bestselling author of **Into the Magic Shop: A Neurosurgeon's Quest to Discover the Mysteries of the Brain and the Secrets of the Heart**. He is also the senior editor of the recently released **Oxford Handbook of Compassion Science**.

**Blair LaCorte** is CEO of AEye and on the Board of the Positive Coaching Alliance, Kairos Foundation for Entrepreneurship, and has formerly served as a fellow at the Digital Strategy Center at the Tuck School at Dartmouth College; Executive Director of the Strategic Council on Security Technology; and a member of the Senate High Tech Advisory Board. Mr. LaCorte has had a lifelong passion for innovation, has garnered numerous patents across several domains and received numerous industry accolades including "Top 10 Marketer of the Year" by Ad Age and Business Marketing; "Innovator of the Year" by NASA; and "Product of the Year" by Industry Week. He has authored numerous business school case studies that are currently taught at the top fifty business schools and is the co-author of an upcoming book on relevancy and perception: **Relevancy...Rules**.

