NEWS RELEASE

Visual Intelligence Will Define the Vehicle

2026-11-03

For the past decade, the automotive industry has been trying to build a better voice assistant.

The objective made sense. As vehicle interfaces became more complex and physical controls disappeared, drivers needed a safer, simpler way to access navigation, climate, media, and communications. Voice promised to replace menus with natural interaction.

Generative AI brings that promise closer. Today's systems can understand everyday language, retain context, coordinate multiple functions, and respond conversationally. But this progress also creates a strategic challenge: when every automaker can offer capable conversational AI, voice alone is no longer a meaningful differentiator.

The more important question is what happens after the vehicle understands us.

Can it determine what matters in the moment?

Can it combine information from the road, the cabin, the vehicle, and its occupants?

Can it communicate the answer without demanding unnecessary attention?

And can it turn intelligence into an experience that feels native to the vehicle and distinctive to the brand?

This is why the industry may be debating the wrong interface.

The future is not voice versus touch, or video replacing voice. It is the emergence of multimodal visual intelligence: AI that perceives context, reasons across multiple sources of information, and presents the right insight through

the right interface.

Voice will remain essential. But increasingly, voice will be the input. Visual intelligence will define the experience.

AI creates more language, but drivers have less attention to spare

Voice works well when a driver's hands are occupied and attention must remain on the road. It is a natural way to initiate a task, request information, or control a function.

Its limitation is that speech is sequential. Information must be delivered one word at a time.

The more capable generative AI becomes, the more visible this limitation becomes. AI can produce detailed explanations, recommendations, and alternatives, but a driver does not necessarily want a long, fluent answer. Inside a moving vehicle, the objective is not to maximize conversation. It is to minimize the effort required to understand and act.

Voice is efficient for expressing intent. Visuals are often more efficient for communicating the result.

- A route comparison is easier to interpret on a map than through a spoken list.
- A charging recommendation is more useful when location, availability, range, and expected arrival state appear together.
- An automated maneuver is easier to understand when the vehicle identifies the relevant cyclist, lane boundary, or road condition.

A 2023 study supported by the AAA Foundation for Traffic Safety examined how auditory and visual warning configurations affected takeover time, driving performance, eye-glance behavior, trust, and usability. The research reinforces that modality, timing, and message specificity influence driver behavior. The objective is not to select one universal interface, but to use the appropriate modality for the situation.

The display becomes the visible face of vehicle intelligence

Historically, displays have shown information generated by separate systems: speed from the instrument cluster, a route from navigation, media from infotainment, and warnings from ADAS.

AI changes the display's role.

It becomes the point where the vehicle's intelligence is interpreted for a human being.

A larger display presenting more information can create more complexity. The opportunity is to use AI to determine:

- what deserves attention
- which occupant should receive it
- where it should appear
- whether it should be visual, spoken, or haptic
- how much detail is appropriate
- when it should disappear

This is a more significant shift than the move from physical controls to touchscreens.

The interface is no longer static.

It can adapt to the user, the task, the environment, and the driving state.

At Visteon, we see this progression through the evolution of the digital cockpit. Our cockpit domain controller architecture consolidates functions such as the instrument cluster, infotainment, multiple displays, driver monitoring, and ADAS visualization onto a common computing foundation. SmartCore™ capabilities include multi-display support, sensor fusion, context-aware interaction, and adaptive user interfaces, for example.

This foundation is critical because a fragmented electronics architecture cannot easily deliver a coherent AI experience. When sensing, reasoning, graphics, driver information, and infotainment operate independently, the result feels like a collection of features. When they share context and computing resources, they can behave as one system.

The bigger opportunity is visual reasoning

The claim that video will matter more than voice captures part of the transition, but it can be interpreted too narrowly.

Video entertainment will become increasingly relevant, especially for passengers and while vehicles are stationary. But the more strategically important opportunity is video as an input to intelligence.

Cameras provide contextual information that a voice assistant cannot obtain from speech alone. They can support occupant and driver monitoring, scene understanding, gesture interpretation, personalization, surround vision, and environmental awareness.

When combined with speech and vehicle data, these inputs allow the system to understand what was said, who said it, what is happening around them, and which response is appropriate.

Visteon's high-performance compute tech reflects this direction. Its hybrid multimodal architecture combines speech, camera inputs, and vehicle data, with on-device AI, advanced graphics, multimedia capabilities, and support for driver and passenger interfaces.

More visual content is not the objective

The shift toward visual AI must be approached with discipline.

More information is not necessarily better information. More visual content is not automatically a better experience.

The goal, therefore, should not be to place more AI-generated material in front of the driver. It should be to use AI to reduce the information burden.

The same system might provide a brief visualization while the vehicle is moving, a spoken response when visual attention is constrained, richer information for a passenger, and an immersive video experience when the vehicle is parked. During a safety-critical event, it may combine visual, auditory, and haptic signals.

The winning interface will not prioritize one modality but instead select and coordinate them intelligently.

What this changes for automakers and suppliers

In the AI era, automakers will not be defined by the models they use. They will be defined by the experiences they own.

Conversational capabilities will quickly become a commodity. Real differentiation will come from how intelligence interacts with the vehicle, understands context, coordinates information across screens, and translates a brand's identity into behavior.

For suppliers, this changes the value equation. Multimodal AI demands expertise across compute, sensing, displays, graphics, audio, software, thermal management, safety, cybersecurity, and lifecycle operations.

The prize is enabling intelligence to work reliably at vehicle scale. That is a systems engineering challenge, and increasingly, the industry's most important differentiator.

That is where Visteon has a distinctive role. Displays, cockpit compute, camera integration, graphics, driver information, and automotive software are not adjacent technologies assembled for an AI demonstration. They are core elements of production cockpit systems. Bringing them together creates the foundation for an experience in which intelligence is not only available, but visible, useful, and trusted.

The new competitive battleground

Voice will become infrastructure. It will remain indispensable, but capable conversational systems will become widely available.

Visual intelligence will become differentiation. The ability to perceive context, explain vehicle behavior, coordinate displays, personalize experiences, and make intelligence understandable without creating distraction will define the next generation of the digital cockpit.

The automakers that lead this transition will be those whose vehicles know what to show, when to show it, and when not to show anything at all.