www.prophesee.ai

WHITE PAPER

Event-based sensing enables a new generation of machine vision solutions

EVENT-BASED SENSING IS A NEW PARADIGM IN ARTIFICIAL VISION. INSPIRED BY HUMAN BIOLOGY, IT PROMISES TO ENABLE A SMARTER AND SAFER WORLD BY IMPROVING THE ABILITY OF MACHINES TO SENSE THEIR ENVIRONMENTS AND MAKE INTELLIGENT DECISIONS ABOUT WHAT THEY SEE.

This technology will take over from the frame-based approach used by traditional photography and video cameras in many machine-vision applications. Event-based solutions will make machine vision more accessible and more effective, enabling a world in which production lines run more quickly, humans and machines work together more safely, autonomous vehicles are accident-free, and drones intuitively avoid collision at high speeds.

AMONG THE ADVANTAGES OF EVENT-BASED VISION ARE:

- Affordable high-speed sensing
- Ability to operate in uncontrolled and variable lighting
- Reduced computational, memory and communications requirements
- Low energy consumption

THESE ADVANTAGES ENABLES EVENT-BASED SOLUTIONS FOR:

INDUSTRY 4.0

Fast object recognition enabling higher throughputs of production lines and a shift from discrete to continuous processes.

Fast visual feedback loops for robotics systems, to enable safer co-working between robots and humans.

Complex multi-object tracking that is resilient to occlusions and so enables improved operational safety and optimization.

MOBILE • WEARABLE • IoT

Low-power 'always on' features to enable next-generation user interfaces on smart mobile devices.

Low-power simultaneous localization and mapping for AR/VR devices.

Fast change detection that highlights regions of interest in surveillance and monitoring scenes, reducing the amount of data that need to be analysed further.

Computational imaging leveraging images and events combined for image deblurring or ultra-slow motion videos.

AUTONOMOUS DRIVING

Fast and robust detection for forward collision avoidance, enabling faster automatic emergency braking.

Adaptive sensing in rapidly changing lighting conditions and improved light flicker mitigation.

Lower costs and computational overhead will enable the use of more cameras for increased redundancy and safety.

THE STRENGTHS AND WEAKNESSES OF CONVENTIONAL IMAGING

WE HAVE BEEN USING CONVENTIONAL IMAGING IN MACHINE VISION FOR DECADES, SO WHY SHIFT TO AN EVENT-BASED APPROACH NOW?

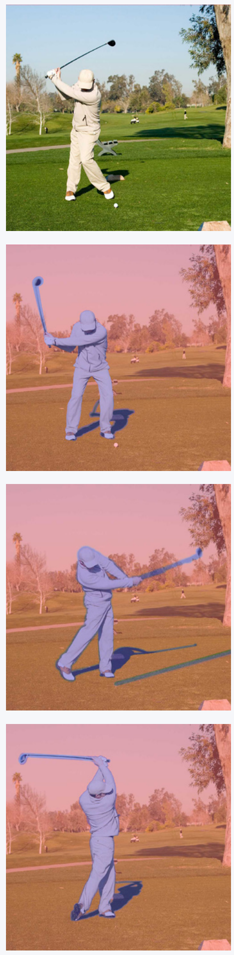
Conventional imaging uses a frame-based approach, in which all the pixels in a sensor measure the light falling upon them at pre-defined points in time, and report their values to the support circuitry in synchrony.

Do this once and, in the right lighting conditions, you get a good-quality still image. Do it more rapidly and you can fool the human brain into thinking that the sequence of still images with which it is presented is actually continuous movement.

This approach works well for humans, but is not ideal for machine-vision applications. One reason for this is that a conventional camera applies a common frame rate to the entire scene. This means moving elements are likely to be under sampled while stationary ones are extensively and needlessly replicated.

Machine-vision systems are therefore burdened with processing large amounts of useless or bad data, using expensive, power-hungry processors, high-bandwidth communications links and memory, to no useful effect.

This brute-force approach works, within limits, for some current applications, but may not be suitable for new vision tasks, such as in automotive vehicles, that need to understand scenes in real time, or in environments with limited power, bandwidth and computing resources.



- Oversampled: Sky, grass and trees
- Undersampled: Motion of golfer, club and ball

Consider a scene with a fast-moving object in front of a static background, such as a golfer addressing the ball.

When acquiring such a scene with a conventional video camera, important information about the fast movement of the arm, the club and the ball may be lost (as in the third frame), because the scene is not being sampled quickly enough, while static parts of the image, such as the tree, are oversampled, recreating data that do not contain any new information.

The universal frame rate of conventional sensors means that information is often lost from areas of a scene that have been undersampled.

THE EVENT-BASED ALTERNATIVE

THE DEVELOPMENT OF EVENT-BASED SENSING HAS BEEN INSPIRED BY OUR EVOLVING UNDERSTANDING OF HOW THE HUMAN VISION SYSTEM WORKS.

Human vision gives us a huge evolutionary advantage, at the cost of sustaining a brain powerful enough to interpret the vast amount of data it produces. Evolution's frugal nature therefore led to the emergence of shortcuts in the visual-processing centres of our brains to cope with this data deluge.

The photoreceptors in our eyes only report back to the brain when they detect a change in some feature of the visual scene, such as its contrast or luminance. Evolutionarily, it is far more important for us to be able to concentrate on the movement of a predator within a scene than to take repeated, indiscriminate inventories of the scene's every detail.

Recent research on human's ability to recognise objects suggests that humans can gather useful data from a scene that is changing at rates of up to 1000 times a second – a far higher rate than the 24, 30 or 60 frame/s that we use to represent movement on television or in movies. A huge amount of useful information is encoded in these changes, which most fixed frame-rate cameras never even see due to their low sampling rates.

Event-based sensing doesn't use a fixed frame rate but instead relies upon each pixel to only report what it sees when it senses a significant change in its field of view. This approach reduces the amount of redundant data transmitted by the sensor, saving processing power, bandwidth, memory and energy. It enables sensors to be built with much higher dynamic ranges than is usually the case, because each pixel automatically adapts to the incident light. For this reason, event-based sensors aren't restricted by high contrast in the scene such as a car's headlights at night, in the way that a conventional sensor would be. And, event-based sensors allow to cost-effectively record events, that would otherwise require conventional cameras running at up to tens of thousands of frames per second.

The Prophesee event-based sensor has further advantages. Its output is a time-continuous data stream that represents a visual event as a sequence of addresses for each pixel that senses it. This spatio-temporal data stream provides a more direct way of representing dynamics and motion in the event sensor's field of view than inferring it from frame-to-frame processing of a standard sensor's output. These characteristics create opportunities to rethink today's imaging and machine-vision data processing, and to address emerging computer-vision strategies, such as machine learning, in a new way.

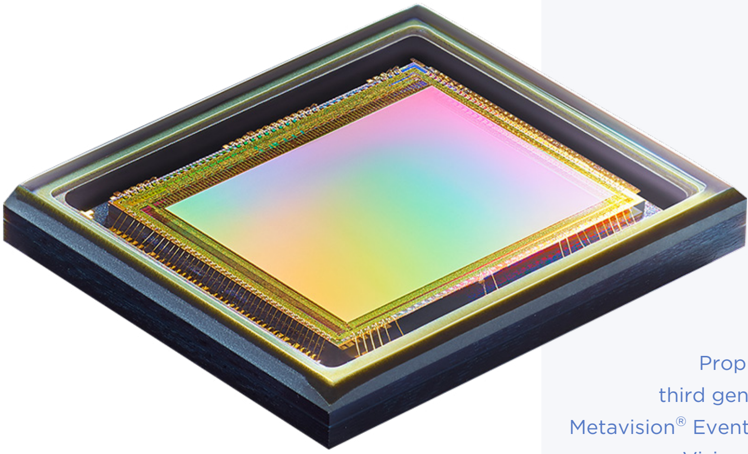
THE PROPHESEE ECOSYSTEM FOR EVENT-BASED SOLUTIONS

PROPHESEE HAS DEVELOPED AN EVENT-BASED ECOSYSTEM COMPRISING OF HARDWARE AND SOFTWARE OFFERINGS THAT ENABLES DEVELOPERS TO BRING VISION SOLUTIONS TO A WIDE RANGE OF MARKETS

ENABLING PARADIGM SHIFTS IN MACHINE VISION

Prophesee's event-based approach to vision sensing means that vision system designers who want to capture fast events no longer need to make a trade-off between running their cameras at high frame rates and dealing with large amounts of redundant data. The volume of data the sensor produces is now governed by the activity in its field of view, automatically adjusting as the scene conditions evolve.

Looking at a static scene will generate no events, but if there is a burst of action, the camera adapts automatically to capture it instantly. This makes it easier and more cost effective to acquire and analyse very fast motion, even if it is interleaved with times or areas in which motion is absent.



Prophesee's third generation Metavision® Event-Based Vision sensor.

The Prophesee sensor records a rotating robotic arm as a continuous stream of spatio-temporal events.

Each pixel provides information at the rate of change in its field of view, not at an arbitrary, pre-set and fixed frame rate.

An event-based approach also means that dynamic scenes can be analysed as a highly resolved sequence of events that form spatio-temporal patterns representing features such as the edges, trajectories, or velocities of objects. The figure above illustrates how the Prophesee sensor would record a rotating robotic arm.

The mathematics describing such features in space and time is simple and elegant and so yields efficient algorithms and computational rules.

In one comparison, Prophesee's event-based approach to dynamic sensing achieved temporal resolutions of tens of kHz where a frame-based approach struggled to reach 60Hz.

This is possible because each visual event is handled as an incremental change to a continuous signal, which can be analysed at low computational cost, when compared to the simultaneous analysis of all the pixels in many complete frames. An event-based approach also makes it easier to correlate multiple views of a scene. This eases tasks such as 3D depth reconstruction in multi-camera stereoscopy set-ups, because if two or more cameras sense an event at once, it is likely they are observing the same point.

Also, analyzing the way in which the illuminance of a single pixel changes over time enables the development of new ways to solve key vision challenges, such as object recognition, obstacle avoidance, and the simultaneous localization and mapping processes that are vital to enabling vehicle autonomy.

SCAN TO LEARN MORE

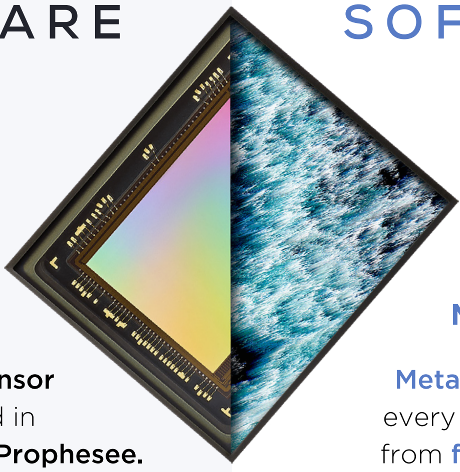
bit.ly/propheseeproducts

HARDWARE

SOFTWARE

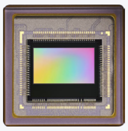
SCAN TO LEARN MORE

bit.ly/Metavisionintelligence



METAVISION® SENSING

Prophesee successfully **built 4 sensor generations.** The last one realized in collaboration between **Sony** and **Prophesee.**

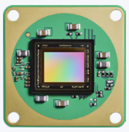


IMX636ES (HD)

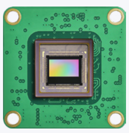
METAVISION® EVALUATION KIT 4 – HD

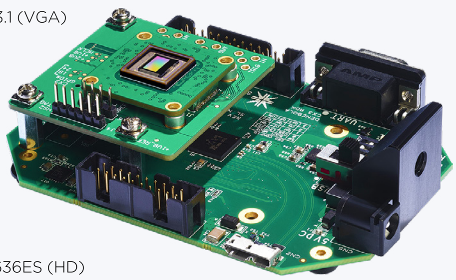
Ultra-light and compact HD Metavision® Evaluation Kit built to endure field testing conditions.

- **Supported Sensors:** IMX636ES (HD)
- **High Quality Aluminum alloy** casing
- **Ultra-compact and light:** 30x30x36mm, weighing just 40g
- **C/CS mount**
- **USB 3.0** Type-C connectivity
- **Free access to 5x award-winning Event-Based vision software suite:** Metavision Intelligence 2.3 onward, **2H** premium support, **Knowledge Center** access.



Gen3.1 (VGA)





IMX636ES (HD)

METAVISION® EVALUATION KIT 3 – VGA/HD

Board level VGA or HD Metavision® Evaluation Kit, ideal for first hands-on, cost-efficient evaluation.

- **Supported Sensors:** IMX636ES (HD), Gen3.1 (VGA)
- **USB 3.0** Micro-B connectivity
- **CS / S mount** versions available
- **Free access to 5x award-winning Event-Based vision software suite:** Metavision Intelligence 2.3 onward, **2H** premium support, **Knowledge Center** access.

METAVISION® INTELLIGENCE

Metavision® Intelligence Suite, covers every step of the development process, from **first discovery** to fast **prototyping** to **end-application development.**

METAVISION INTELLIGENCE SUITE

THE MOST COMPREHENSIVE EVENT-BASED VISION SOFTWARE SUITE

95 algorithms, 67 code samples and 11 ready-to-use applications

OpenEB

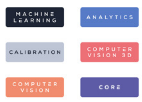
OPEN SOURCE ARCHITECTURE

Unlocking full interoperability between our software and hardware devices

NEURAL INFORMATION PROCESSING SYSTEMS

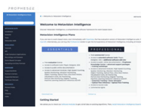
LEADING EVENT-BASED ML TOOLKIT

Leverage the most performant object detector to date spotlighted at NeurIPS 2020



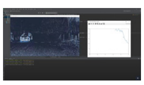
6 EXTENSIVE MODULE FAMILIES

Composed of 6 main module families covering a wide range of machine vision fields



EXTENSIVE DOCUMENTATION

270+ pages of regularly updated content on docs.prophesee.ai, more than 20 Jupyter notebooks, reference data and extensive guidelines



GET RESULTS IN MINUTES

The largest collection of pre-built pipelines, extensive datasets, code samples, GUI tools and more

A MORE EFFICIENT COMPUTER VISION

TRACKING

The sparsity of the data generated by an event-based camera directly improves the efficiency of some computer-vision algorithms, such as tracking objects in a videosurveillance setup with a static camera where events are generated only on pixels of interest.

Conversely, the almost continuous sampling of the visual scene enabled by the asynchronous behaviour of the sensor brings new ways of optimizing computer vision. Tracking algorithms can afford to only consider pixels neighbouring an object being tracked, instead of a large region of the image, because an object is now guaranteed to trigger an event at every pixel it moves through.

OPTICAL FLOW

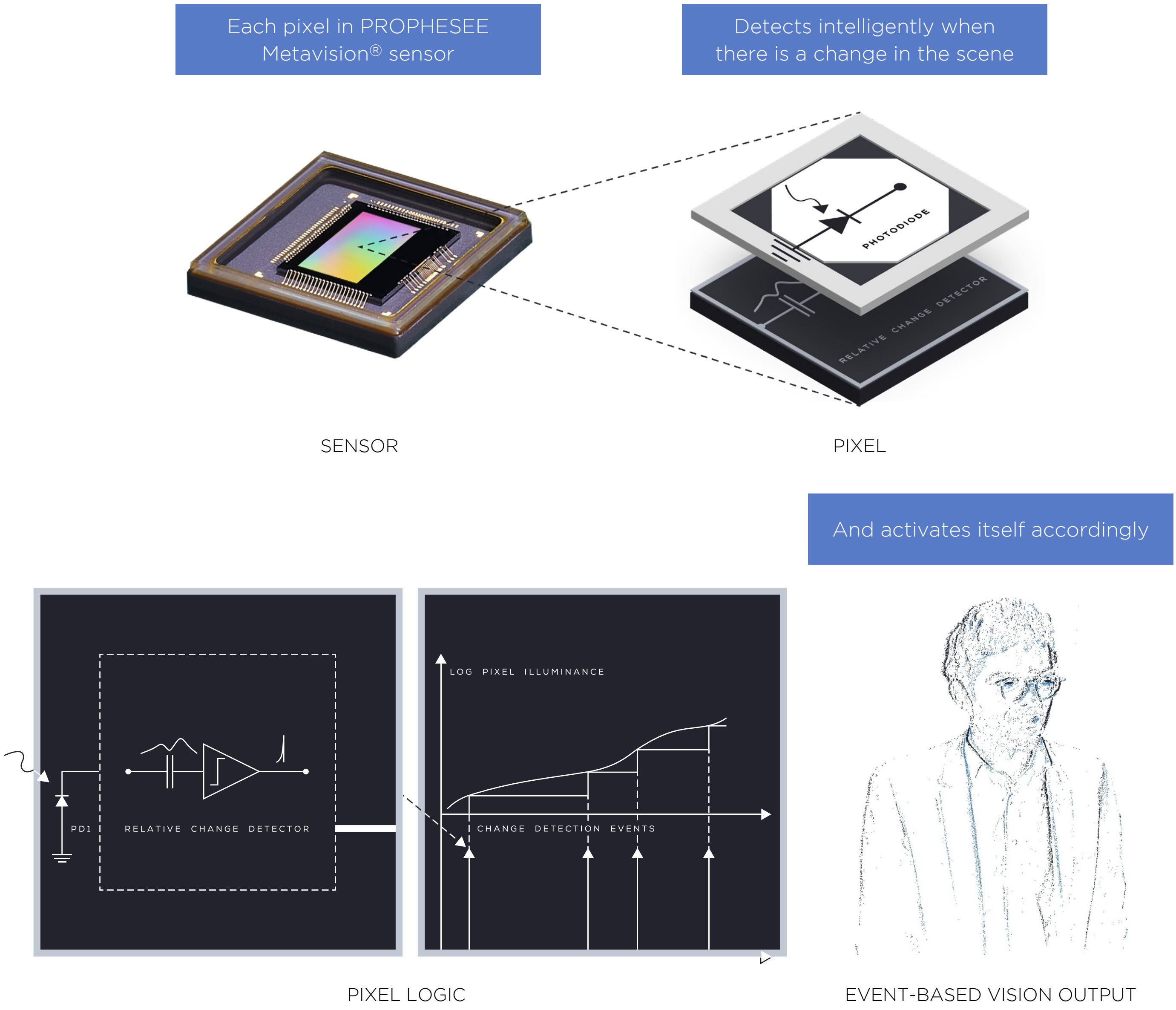
Following the same idea, computation of visual flow can also be greatly simplified. When a moving object is captured by an event-based camera, its edges move through every pixel of the sensor, generating a surface in space-time. If it moves slowly, it takes a lot of time to go from pixel to pixel and generates a steep surface.

If it moves fast, the surface becomes almost horizontal since the contour takes very little time to go from one pixel to the next. Visual flow is then reduced to simply computing the slope of spatio-temporal surfaces, an operation much less computationally intensive than its counterpart in frame-based vision, where an object or a feature needs to be tracked or matched between frames, after it has moved over a large number of pixels.

STEREO

The large temporal resolution of the data produced by an event-based camera also simplifies the process of resolving depth data from 2D images of a moving scene. If two cameras viewing the same scene each report an event at exactly the same time, it's possible to assume the event happened to the same part of the image. This enables 3D positioning by triangulation.

The technique can also be used to resolve two similar looking-objects in a scene, such as the windows of a building as a vehicle passes. Although two adjacent windows will look the same in a static image, their different relative positions to a moving event-based camera will create two separate streams of visual events whose signature can be used to distinguish the two windows.

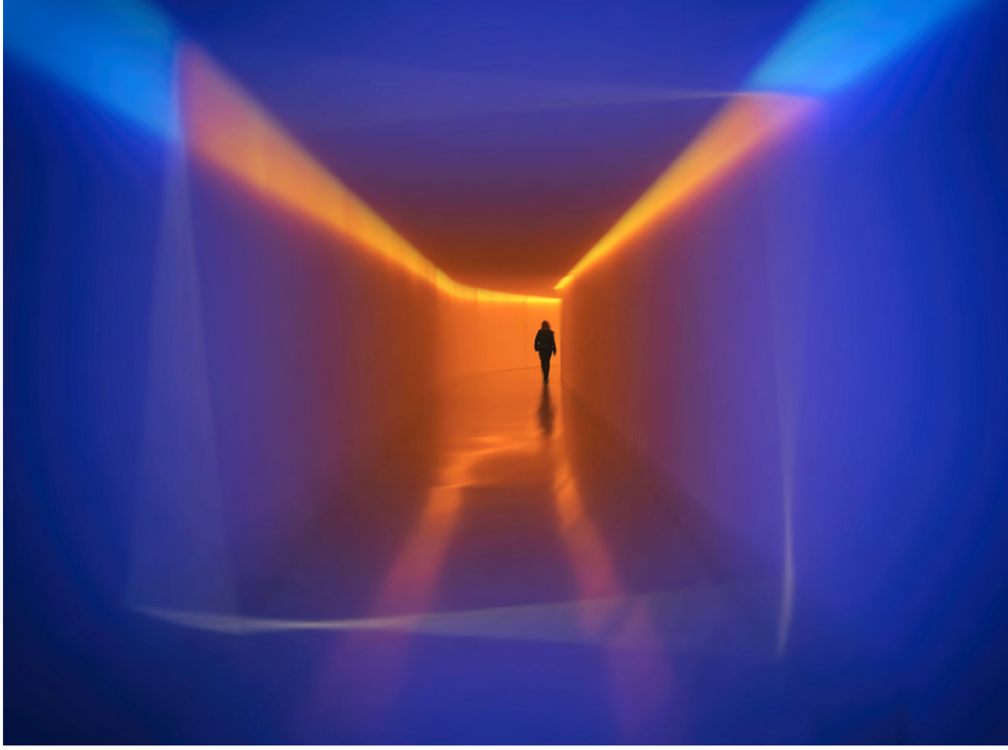


EVENT-BASED ACQUISITION OF DYNAMIC VISUAL INFORMATION

Implementing asynchronous delta modulation-based level crossing sampling, each pixel defines the timing of its own sampling points in response to its input by reacting to changes in incident light.

Pixel sampling events are read out as they happen. The output generated by an event camera hence is no t a sequence of images but a time-continuous stream of individual pixel data, generated and transmitted conditionally, based on changes happening in the scene.

Only pixels that are stimulated by these changres produce output while the other pixels in the array remain silent. Positive and negative changes are discriminated and marked (shown as black and blue dots in the accumulated event map image shown on the right).



A NEW DIMENSION IN MACHINE LEARNING

Event-based sensing enables new approaches to machine learning. An object recognition or detection algorithm that, until now, could only use the spatial information from a frame can now access another dimension: time.

For example, in a frame-based representation, a plastic bag and a rock lying in the road in front of a car may look a lot like each other but their dynamics, as captured by an event-based sensor, would clearly distinguish them.

Event-based sensors can also provide information that is difficult or impossible to derive using frames. Where a frame-based camera only sees a snapshot of the world, an event-based camera can see the oscillations of the arms and legs of a pedestrian, the rapidly rotating wheels of a car or a cyclist pedalling her bicycle as a distinct signal. With GPUs and other machine learning platforms focusing on processing growing volumes of data to increase their performance, the temporal information provided by event-based sensors could enable a complete shift in the inner properties used to recognise objects. This may either, enable better generalisation of learning (i.e. dramatically reducing the size of the datasets required for efficient machine learning) or, enable the capture of more subtle aspects of a scene such as the intent of a pedestrian waiting at a crossing, which can be read from slight changes in posture that a frame-based camera would not see.

IMPROVED ROBUSTNESS

In parallel with the advantages brought by the asynchrony and temporal resolution of the sensor, its large dynamic range and logarithmic sampling of the scene make many computer vision algorithms more robust. One of the big challenges of computer vision is the uncontrolled nature of outdoor lighting.

Algorithms that perform very well in lab conditions may be much less effective when blinded by the sun or missing details hidden in shadows. In such conditions, the event-based sensor's sensitivity to changes in illuminance means that the way it perceives events does not change in low light, bright light, or very high dynamic range scenes.

This means event-based sensors are well suited to making seamless transitions from indoor to outdoor conditions a particularly useful facility in, for example, autonomous driving systems that have to cope with the vehicles to which they are fitted moving into or out of tunnels.

THE FUNCTIONAL ADVANTAGES OF EVENT-BASED SENSING

THE COMBINATION OF AN EVENT-BASED APPROACH, THE SENSOR WHICH IMPLEMENTS IT, A HARDWARE REFERENCE DESIGN AND A SOFTWARE STACK, WILL PROVIDE DEVELOPERS WITH A FAST PATH TO EVENT-BASED SOLUTIONS THAT SET NEW STANDARDS IN MACHINE VISION, INCLUDING:

- **Efficient streaming of visual data for real-time communication and surveillance**, thanks to the data compression inherent in not repeatedly reporting data from static parts of an image.

- **High-speed tracking of fast-moving objects**, thanks to the fact that the data stream from the sensor can be directly processed to measure object speed and predict trajectories.

- **Detection and real-time analysis of high-speed transient visual events**, thanks to each pixel's ability to dynamically adjust its sampling rate to the scene's needs.

- **Efficient stereoscopy**, through event-based correlation between multiple cameras simultaneously reporting a change in their fields of view when looking at the same 3D object.

- **Higher temporal-resolution and lower latency inputs** for real-time control systems.

- **Enhanced machine-learning strategies**, because event-based sensing adds time as another dimension of information to vision analysis tasks.

CONCLUSION

EVENT-BASED SENSING IS INSPIRED BY THE ELEGANT WAY IN WHICH NATURE HANDLES VISION TASKS IN THE HUMAN BRAIN, AS A SERIES OF TIME-CONTINUOUS EVENTS.

Prophesee's implementation of event-based sensing makes the technique widely available, setting new standards in machine vision in its current applications and widening its applicability to other markets.

Event-based sensing also offers a new paradigm for machine learning in image processing, shifting the emphasis from a frame-based analysis to a timed-event approach. This may make it possible to address current machine-vision challenges more efficiently, and create new ways to apply machine learning strategies.

ABOUT PROPHESEE

Prophesee is the inventor of the world's most advanced neuromorphic vision systems.

The company developed a breakthrough Event-Based Vision approach to machine vision. This new vision category allows for significant reductions of power, latency and data processing requirements to reveal what was invisible to traditional frame-based sensors until now. Prophesee's patented Metavision® sensors and algorithms mimic how the human eye and brain work to dramatically improve efficiency in areas such as autonomous vehicles, industrial automation, IoT, mobile and AR/VR.

Prophesee is based in Paris, with local offices in Grenoble, Shanghai, Tokyo and Silicon Valley. The company is driven by a team of more than 100 visionary engineers, holds more than 50 international patents and is backed by leading international equity and corporate investors including 360 Capital Partners, European Investment Bank, iBionext, Inno-Chip, Intel Capital, Renault Group, Robert Bosch Venture Capital, Sinovation, Supernova Invest, Xiaomi.

LEARN MORE

PARIS • GRENOBLE • SHANGHAI • TOKYO • SAN FRANCISCO