White Paper

Artificial Intelligence from the Edge to Cloud

Utilize 4D Digital Twins in Any Industry with Intel® SceneScape

4D scene-based digital twins are poised to transform defense, manufacturing, smart cities, transportation, security, healthcare, retail, and more.

Executive Summary

The ability to monitor, track, and analyze different spaces has become essential to business processes in many industries. In fact, ninety-five percent of companies said that location intelligence is important in achieving desired business results today, and ninety-one percent said that it will be even more essential in three to five years.[1]

Location intelligence, or spatial awareness, refers to an individual or groups understanding of a physical space and the activities within that space. This can include machines, people, objects, and more. Spatial awareness is gained through analyzing various inputs, such as data from different sensors within the space being looked at. From there, individuals or groups can apply this understanding of their space, and the activities within it, to valuable use cases in any industry. Common use cases in this context include restricted access zone monitoring, people or queue counting, anomaly detection, and asset tracking.

This paper introduces the Intel® SceneScape software platform, for 4D digital twin creation in use cases related to spatial awareness.

Usage Category	Key Value Creation Metrics
Situational Monitoring	▪ Lower cost by using existing sensors / infrastructure ▪ Increased context through multiple camera / sensor inputs
Logistics and Tracking	▪ Improved customer experience ▪ Reduction in lost / stolen assets
Assets & Operations Optimization	▪ Reduction in resolution time / maintenance cost ▪ Increased throughput efficiency
Control Optimization and Autonomy	▪ Reduction in resolution time ▪ Reduction in injuries

Figure 1: Example of Intel® SceneScape Use Cases and Value Delivered

Challenges in the Current State of Scene Analytics

Today, the market is oftentimes served by traditional "unimodal" point solutions. This is where specific sensors are installed to enable a single, specific use case. The vendor installing the sensor, such as a camera or RFID anchor, directly consumes the data to enable a given use case, such as people counting or asset tracking. The application is responsible for sensor context, such as labelling cameras as "front entrance" and "side entrance," so appropriate logic can be applied when a person is detected in that camera. Unimodal point solutions do not allow users to track multiple use cases in a single space in a connected way since they are built to serve one specific use case.

An alternative to traditional unimodal point solutions is multimodal sensing. In the market today, where multimodal sensing has been developed, it is typically in fully proprietary environments where customers are locked in with a specific vendor. If the vendor changes, all sensors need to be replaced and business processes completely revamped. This is typically a slow and resource intensive process. Moreover, these systems are typically built for a specific application and do not generalize solutions to the broadest extent. This hampers the ability for the full value of multimodal technologies to be realized.

Introducing Intel® SceneScape

Intel® SceneScape is a software platform that reaches beyond vision-based AI to realize spatial awareness from sensor data. It transforms data from many sensors to create and provide live updates to a 4D digital twin of your physical space. Digital twins can be applied to your use cases to look at past analytics, track what is happening in the present, and make predictive decisions for the future.

Intel® SceneScape unlocks business applications from raw sensor data by providing a digital twin of each scene. Objects, people, and vehicles within the scene are represented as overlays on the dynamic structure of the digital twin. Applications, autonomous systems, and mobile systems securely access the digital twin to make decisions about the state of the scene, such as whether a person is in danger, a part is worn or broken, someone has been waiting in line too long, a product has been mis-shelved, or a child has called out for help.

Powerful AI algorithms and AI hardware crunch all available sensor data to maintain the 4D scene graph (3D plus time), as quickly and accurately as possible. This enables users to see what is happening in near real time. With the Intel® Distribution of OpenVINO™ toolkit, Intel® SceneScape is able to use raw sensor data to create the 4D semantic digital replica by ingesting detections from 2D cameras and mapping them into the digital twin. The Intel® Distribution of OpenVINO™ toolkit also helps to abstract the different types of Intel® hardware accelerators, including CPU, GPU, VPU, FPGA, and GNA, enabling developers to write code once and deploy it.

A scene is defined by a fixed local coordinate system against which all sensors are provisioned to provide spatial context to sensor data. While geographical coordinates may be suitable for many applications, scenes usually use a fixed local Cartesian coordinate system. That scene could be a building, a ship, an aircraft, or a campus that has the global geographical coordinate system as a parent. To provide this context to sensor data, Intel® SceneScape manages:

- A unique scene and scene coordinate system.
- Exactly one scene parent for each sensor at a given time.
- The location of cameras, microphones, thermometers, and all other sensors within the scene. Maintained in the SceneScape database, sensor position and orientation, along with characteristics such as camera and microphone polar patterns provide critical data to for determining context from raw sensor data.
- Enablement of existing scene graph technologies and standards such as glTF[2], X3D[3], and Open Geospatial Consortium (OGC) 3D Tiles[4].

Intel® SceneScape Capabilities

Enhance and extend tracking beyond vision-based AI in any industry by applying Intel® SceneScape to any use case. Some of the key benefits and capabilities delivered with Intel® SceneScape include:

- **Scene Context** - Scene and Sensor Management utilizes knowledge about sensors to apply scene context. For example, the position of a smart camera in a building allows for mapping the context of detected people from the camera view into building coordinates.
- **Multimodal detection** - Multimodal tracking allows users to decide what sensors best fit their operational needs. Intel SceneScape readily handles visual, infrared, radio frequency (RF), Intel® RealSense™ tracking and depth sensing cameras, or even other environmental sensors.
- **Visual-based analytics and modeling** - Base analytics tools enable users to create and modify regions of interest, tripwires over a scene and dwell time in a region.
- **Multi-sensor data fusion** - Multiple sensor fusion means Intel SceneScape will detect an object of interest at multiple angles in

different sensors but know to only display the object of interest once on the scene graph removing duplicates and reducing error.

- **Versatility** - Support for third party inferencing.
- **Realistic visualization** - Support for Intel® OpenVINO™ models that provide high-angle camera detection and human poses.

Intel® Geti™ Software Platform Compatibility Unlocks New Potential for AI Model Development

The Intel® Geti™ platform enables teams to build computer vision models rapidly and efficiently. By easing labor-intensive tasks in model development, such as data upload and labeling, model training, optimization, and retraining, the platform speeds up time to model. Professionals can build custom models for use cases across verticals and applications, including detecting defective parts in a production line, segmenting weeds in an agricultural field, automating inventory management, or other digitization and automation projects involving computer vision.

While tracking your space with Intel® SceneScape, you can take as little as 20-30 still frames from your camera sensors or a video and export those into the Intel® Geti™ platform to build an initial model. You can build a custom AI model with the Intel® Geti™ platform and then import that model into Intel® SceneScape for a monitoring experience that is custom-built for your specific needs.

Try Intel® SceneScape Today

Intel® SceneScape gives users the live tracking data needed to assess the people and objects within their spaces by creating 4D digital twins of their physical environment. This spatial awareness adds value to use cases including, but not limited to, restricted access zone monitoring, people or queue counting, anomaly detection, and asset tracking.

To learn more about Intel® SceneScape, contact us at IOTG-PublicSector@intel.com or visit us at intel.com/scenescape.

[1] S. Blumenthal, H. Bruyninckx, W. Nowak and E. Prassler, "A scene graph based shared 3D world model for robotic applications," 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 2013, pp. 453-460, doi: 10.1109/ICRA.2013.6630614.
[2] glTF – "The JPEG of 3D": https://www.khronos.org/gltf
[3] X3D standard: https://en.wikipedia.org/wiki/X3D
[4] OGC 3D Tiles: https://www.ogc.org/standards/3DTiles
[5] Intel® Geti™ Software Platform: geti.intel.com