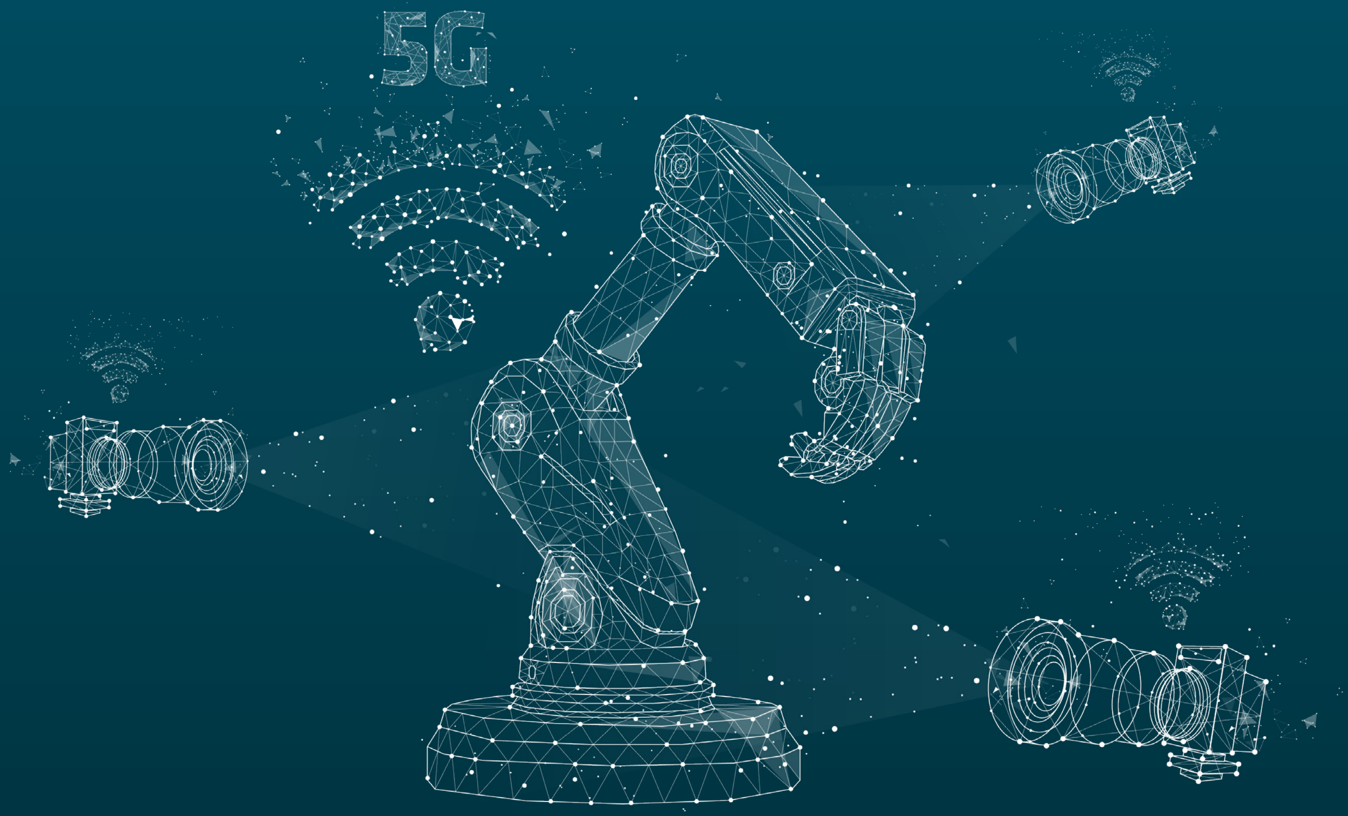


5G-ACIA White Paper

Architecture and Technology for Machine Vision in Industrial Applications with 5G

5G Alliance for Connected Industries and Automation

1 Executive Summary

Typical machine vision applications can be quite demanding both in terms of high availability and bandwidth. However, and this is one of the key learnings from this document, the convergence of vision sensor solutions with new technology like 5G or edge computing can provide solutions which satisfy the communications and control-loop characteristics needed for several use-cases in manufacturing. The ability to offload pre- and post-processing tasks to edge servers or even the cloud in non-real-time applications is a key element to more flexible and cost-effective deployments of machine vision in industrial automation. Finally, the manifold possibilities to reduce the bandwidth of image streams primarily by standard compression and encoding techniques may not come into play in applications that must rely on images of high quality. However, in monitoring, tracking, and remote-control use cases, uncompressed video data is of less importance.

Machine vision technology powered by 5G networks will highly depend on the ease of integration. Therefore, it is recommended to make the switch from wired to wireless solutions as easy as possible. The possibility to use the same transport layers based on existing machine vision standards as well as have nearly the same transmission performance available compared to wired interfaces will clearly drive the adoption of 5G in vision-assisted industrial automation and even allow for real-time applications.

About 5G-ACIA

The 5G Alliance for Connected Industries and Automation (5G-ACIA) was established to serve as the central and global forum for addressing, discussing, and evaluating relevant technical, regulatory, and business aspects with respect to 5G for the industrial domain. It reflects the entire ecosystem and all relevant stakeholder groups, including the operational technology (OT) industry (industrial automation companies, engineering companies, production system manufacturers, end users, etc.), the ICT industry (chip manufacturers, network infrastructure vendors, mobile network operators, etc.), academia, research institutes, and other relevant players. The paramount objective of 5G-ACIA is to ensure the best possible applicability of 5G technology and 5G networks to the industrial domain. 5G-ACIA's mission is to make sure the interests and needs of the industrial domain are adequately considered in 5G standardization and regulation and that ongoing 5G developments are understood by and effectively transferred to the industrial domain.

2 Table of Contents

3 Introduction

Machine vision has the potential to play an ever-increasing role in the industrial automation context:

a) the use of artificial intelligence increases the potential applications spectrum from simple vision sensor to vision-guided automation.
b) The relevant upcoming standards are done in an OSI-Layered manner (IEEE, IETF, OPC-F, etc.), keeping interoperability across layers. This allows the different component manufacturers to focus on their application, thus reducing the development costs while increasing the addressable market.

Therefore, it is important to understand which machine vision applications would benefit from 5G usage and how this can be implemented. Therefore, the purpose of this white paper is to provide guidance to 5G factory operators and OT engineers on which system functions are needed and their Key Performance Indicators, from both 5G and from vision control systems perspective, using the following methodology:

1. Select a set of representative machine vision use cases in industrial applications and corresponding industrial 5G requirements
2. Description of the different machine vision system technical and architectural options
3. Description of the different options to integrate a machine vision system in a 5G network

4 Use-cases and requirements

4.1 Introduction

Use cases documented and analysed in this document are described in a consistent manner.

It starts with a brief description which explains the industrial environment with its operational conditions, the systems that are involved and the tasks that are performed by these systems in the given environment. Furthermore, inter-system communication aspects, relevant system / service flow steps as well as postconditions and results of the use case are analysed.

Based on the use-case analysis and the derived system capabilities, technical requirements for the 5G Networks are reported.

4.2 Stationary Use-cases

4.2.1 Process Automation Plant Asset Monitoring

1. Description
To keep a plant up and running it is essential that assets such as pumps, valves, heaters, instruments, etc., are well maintained.

The assets are monitored for timely recognition of any degradation (e. g., leakage of pipes and valves, wear and tear of bearings) to support and plan maintenance work. Smart sensors (e. g., vibration/temperature/acoustic sensors, video camera, gas/flame detector) are used for detecting abnormal conditions of assets.

The operation for this use case can be in a wide service area, and interaction with the public network (e. g., service continuity, roaming) may be required.

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Process Automation Plant Asset Monitoring	Transport of images	99,99	< 100ms to 60 s	≤ 2Mbit/s	4 MByte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
100ms to 60 s	Stationary	10 000	≤ 100 Km^2	Periodic	

Source: 5G-ACIA / ZVEI e. V.

2. Pre-condition
Smart sensors providing relevant data for asset condition are distributed over the plant. All nodes are connected to the 5G system.

3. Service flow
Data from smart sensors are transmitted to storage within a defined time interval. In the case of an actual failure, an event is transmitted immediately.

4. Post-condition
Data and event information are available where needed for processing and displaying. Assets are maintained in an optimal manner.

5. Requirements
- Sensors are battery driven, so the communication module must be highly energy efficient
- Communication service must support high sensor density and provide low latency and high service availability, required to transfer media stream (e. g., video, image, audio, vibration) from smart sensors. Characteristic parameters, as well as influence quantities as defined by 3GPP for this use case, are reported in reference TS22.104, clause A.2.3.2.

4.2.2 AR based Product Quality & Augmented Field Procedure with Digital Twin

1. Description
This use case describes the support for Product Quality and Augmented Field Procedures using Digital Twins and Augmented Reality.

A digital twin is a digital and virtual model of a functional system, combined with real-time aspects of how the system operates. As such, the digital twin incorporates the up-to-date behavior of the system and not simply the design intention represented by a digital model.

Augmented reality (AR) refers to technology that can superimpose digital data and visual images onto streaming video of the physical world. Such superimposition of digital information directly on physical objects in a video stream allows humans to simultaneously process physical information and digital data, resulting in an improved ability to assimilate information, make decisions, and efficiently execute required tasks.

The use-case for product-quality envisions developing and incorporating complex digital models that represent the digital twin. The digital twin would support analyses that utilize streaming data from physical systems to compare the behavior of the physical system against the design intention represented by the digital twin. Such analysis would leverage machine learning to detect and adjust for factory-specific anomalies in the production environment.

By gaining knowledge of machine performance in situ, adaptive digital models and AR overlays help identify production issues to improve overall production quality. Such adaptive manufacturing can drive higher throughput, improved capital utilization, and longer machine life.

The use-case for AR-based field procedures drives enhanced safety and increased operational efficiency by making field procedures mobile and augmented. An example of an AR-based field procedure system would:

- Guide and assist the user step-by-step through the workflow and record all execution details
- Overlay field data as augmented context in the display of the physical area
- Enable accessible and up-to-date information on procedures and supporting information
- Guide manual inputs and trigger actions
- Capture photos and comments
- Support manual and automatic data capture

The combination of digital twin and AR technologies in an industrial or manufacturing setting results in several benefits; in particular, the combination:

- Enables near real-time monitoring and control of machines and processes using 5G and Edge Computing
- Enhances safety and increased operational efficiency by making field procedures mobile, augmented, and integrated
- Provides immersive experience on actual products using augmented reality during manufacturing to enhance quality
- Supports predictive simulation models to improve operations efficiency

For these use-cases, edge computing provides for low-latency AR overlays and incorporation of AI/ML techniques to identify manufacturing issues and improve product quality as well as to enable offline adjustments for optimization, adaptations, and preventive operations on the machines. Edge-compute based AR algorithms and edge-compute based AL/ML inferencing models leverage locality and site-specific parameters of the manufacturing facility.

2. Pre-condition

A digital twin of the machine in its current state is available at the edge for overlaying with the physical machine. Low-latency and high bandwidth communication infrastructure to transmit all the required information of the physical system to the edge is available.

3. Service flow

The actual information from the physical system is transmitted to the edge. The actual physical system state is superimposed the digital twin and made available. In the case of any anomaly, an event is triggered.

4. Post-condition

An edge compute environment supplements the 5G network to provide low-latency access to the AR servers and Digital Twin models. The application components residing within the edge compute environment include the Digital Twin models, the AR rendering server, the AI/ML trained model to trigger the Closed-Loop Control module, and an MQTT Proxy that collects telemetry data from and submits control commands to the factory equipment.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Telemetry Data & Events	Transport of images	99,999	< 20ms	Dependent on specific video profile	4MB

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
< 20ms	Stationary	Dependent on specific application	10 000	Periodic	

Source: 5G-ACIA / ZVEI e. V.

4.2.3 Augmented Reality for Remote Operation

1. Description

During the running manufacturing process, if a machine or robot has an insufficient level of confidence in the immediately following action (e. g., the decision about how to grab a workpiece or how to insert and place a workpiece in a machine), the machine/robot asks a human, e. g. the human operator or machine owner, for help. At this time, production has already been started, i. e., the concerned machine/robot is blocked. Therefore, this task is time-critical with respect to machine outage times. This use case as defined in [1][2], is triggered when the human operator is physically away and needs to monitor or control the machine/robot.

It may play a crucial role for the following applications:

- Remote operation of the physical asset (e. g., robot arm mounted with video camera)
- Remote quality check (e. g., of the machine/robot or a workpiece)

2. Pre-condition

A digital twin of the robot and its surrounding environment in its current state is available for a remote expert. The environment data near to machine/robot needs to be available with low-latency and high bandwidth communication infrastructure.

3. Service flow

A remote expert uses AR device to view a live video feed coming from the physical asset (e. g., robot arm). 3D replicas of the physical asset as well as the surrounding items (e. g., workpiece) are displayed in addition to the live video of the real physical asset. The input device to AR device controls or moves the physical asset (robot arm). Models of the physical asset, as well as the surrounding items, are rendered directly on a local computational unit on the AR display up to a certain complexity. For more complex models, rendering needs to be offloaded to the edge.

4. Post-condition

The human operator experiences a real-time update of digital twin and its surrounding environment and can move physical asset via input device.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Live video from physical asset	Transport of images	99,999	≤ 10 ms	10Mbit/s	4 Mbyte
Remote control of physical asset	Control Data	99,9999	≤ 10 ms	100 kbit/s	50 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m²]	Communication Attributes	Remarks
≤ 16 ms	Stationary	Dependent on specific application	10 000	Periodic	
≤ 10 ms	Stationary	Dependent on specific application	10 000	Deterministic	

Source: 5G-ACIA / ZVEI e. V.

- depending on the application area and field, the head mounted device (HMD) might have to be adjusted. E. g., the HMD must be attached to a helmet on construction sites, to ensure the adherence of health and safety regulations. If worn outside, an optical filter must be attached to prevent AR image quality loss.
- very stringent requirements in terms of latency and bandwidth must be considered. E. g., high bandwidth and ultra-low latency for camera video stream/rendered image to remote expert wearing AR device, ultra-low latency for interactive communication between input device and rendered robot. Sufficient FPS transmission together with low latency is required to avoid display stutters that can lead to motion sickness, headaches, etc. Following are the requirements on network characteristics of a 5G device:
  - Guaranteed bandwidth of approx. 50 Mbit/s from camera of machine/robot to AR server/device
  - Guaranteed bandwidth of approx. 10 Mbit/s from AR server/device to machine/robot
  - Network latency ≤ 10 ms
  - Jitter to avoid stuttering ≤ 0.5 ms
  - 60 FPS for AR device

4.2.4 Inline Multi-camera Manufacturing

1. Description

This use case relates to a camera-assisted production cell, where a robot arm performs sorting or palletizing operations.

A set of image sensors with wireless and/or wired connectivity can be placed around a production cell or be integrated into machines and robots. Image analysis allows 3D image reconstruction, or scene analysis, and can be utilized for the control of industrial machines, AGVs or robots. Based on the image analysis results, a controller unit controls the robot arm position to pick up and place items into the correct package or onto a pallet.

Depending on the hardware available and the processing power on the controlling devices, centralized or distributed 3D image analysis (and then control) can be performed. Practical applications are: Pick-and-Place/Sorting, Inspection, Palletizing, Functional safety.

2. Pre-condition

All image sensors, actuators (for the arm) and the motion controller are switched on and connected to the network.

3. Service flow

A set of fixed image sensors (placed around the production cell) sends images to a central controller that combines them and performs image reconstruction and analysis. Based on the analysis, the controller unit controls the robotic arm. An embedded high-definition image sensor, mounted on the robot arm, allows the controller unit to perform measurements for fine guiding of the robot arm during picking and placing operations.

The image analysis allows detecting unexpected presence of a person or a machine (e. g., AGV) near the robot arm and adjusts the speed of the robotic to prevent any collision.

4. Post-condition

The robotic arm operates safely and collected image are stored for offline quality inspection.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Inline Multi-camera Manufacturing	Transport of images	99,9999	≤ 10ms	2560MBit/s	4MB
Inline Multi-camera Manufacturing	Control Data	99,999999	<1-10ms	100 kbit/s - 1 Mbit/s	50 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m²]	Communication Attributes	Remarks
≤ 10 ms	Stationary	Dependent on specific application	10 000	Periodic	
<1-10ms	Stationary	Dependent on specific application	10 000	Deterministic	

Source: 5G-ACIA / ZVEI e. V.

- synchronization of the several fixed and mobile image sensors. The synchronization accuracy ranges from 10 µs to 1 µs

- the controller unit requires images transmission under guaranteed timing conditions. The associated transmission latency should not exceed 10 ms
- a transfer interval ranging from 1 ms to 10 ms is needed for the control unit to control the robot arm position that picks up mobile objects
- a target service availability of 99.999 999 % is needed for the control unit to control the robot arm for functional safety purpose. The end-to-end latency should not exceed 8 ms
- the size of the images delivered by the image sensors ranges from 640 p to 4 K.

4.2.5 Inline Multi-Camera Quality Inspection

1. Description
Multiple 5G enabled 2D cameras are placed around an inspection chamber to inspect a semi-finished product at once and from all sides. All cameras are triggered simultaneously via 5G and the acquired images are transmitted to an edge server that performs the following tasks:

- Aggregation of images
- Reconstruction of geometry
- Feature extraction and measurements
- Matching features and dimensions against digital twin
- Status message to control level (ok / not ok)
- Storage of data for retro-active inspection

2. Pre-conditions
All cameras are powered on and connected to the 5G system. There is enough light to acquire images. Uplinks and downlinks are available to trigger the cameras simultaneously and transmit the images being acquired. A physical or virtualized PLC is ready to stop a conveyor belt if needed for image acquisition or human intervention. An edge server is prepared to perform image processing tasks in time and to distribute results via 5G.

3. User-system & system-system interactions (flow of events)
A semi-finished product enters the inspection chamber. Light barriers or positioning sensors verify the item's target position, status update information sent by a PLC reaches an edge server dedicated to the image processing, which subsequently triggers all cameras at once. Images are then transmitted to the edge server and fed into the image processing pipeline. Afterwards, status update information of the image analysis is distributed from the edge server to the control level and PLCs receive commands to act according to the quality check status.

4. Post-conditions
The semi-finished product is quality-tested successfully, and results are documented for statistical evaluations. Image data is stored for retroactive inspection.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Multi-camera inspection	Transport of images	99,9999	>>1000ms	2340MBit/s (all cameras)	13MB (image)
Multi-camera inspection	Triggering of cameras	99,9999	<100ms	<0,5MBit/s	64Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
~4000ms	Not relevant	Up to 90	20m^2	Deterministic	
~4000ms	Not relevant	Up to 90	20m^2	Deterministic	

Source: 5G-ACIA / ZVEI e. V.

4.2.6 Bin Picking

1. Description
A smart 3D sensor is mounted on a robot's arm that is supposed to pick randomly organized items from a bin. The sensor detects and locates the items in the bin and calculates grasp or suction points based on the acquired 3D data. The pose data of the next logical or graspable item is transferred to the robot controller via 5G. The robot, no matter if it is stationary or installed on a mobile platform, then performs the grasp according to the planned trajectory. The smart sensor constantly transmits picking status messages per item to the factory's control level.

FW updates of the smart sensor allow to improve the system's capabilities or to change the program and adapt to several picking tasks.

2. Pre-conditions
All 3D sensors as well as robots are powered on and connected to the 5G system. The environmental conditions are controlled and in favor of the 3D method being used to acquire depth data. 3D sensors mounted on the arm of stationary or mobile robots are equipped with a pre-trained software stack and can interchange data with the robots' operating system. Robot controllers are ready to receive object pose data information from 3D sensors via 5G. There is a constant flow of items landing in the bins of interest.

3. User-system & system-system interactions (flow of events)
Items are constantly fed into a bin via a robot or a conveyor belt. A smart 3D sensor detects the objects and estimates their poses based on acquired 3D vision data. Object pose data reflecting estimated grasp points are transmitted from the 3D sensor to a 5G enabled PLC or directly to the robot controller, which ultimately starts the picking task. Afterwards, either the 3D sensor detecting the new object in the drop area or the robot controller analyzing the gripper's contact pressure or vacuum data sends a status message about the successful drop or place of the item to the system's HMI or a central server on the control level. If the robot system is supposed to pick new or other items from a bin in a flexible production line scenario, configuration data is transmitted from the HMI or a central control system to the smart 3D sensor, that is ultimately switching over to the right application program. Updates to the sensor's FW or application software, such as an update on neural network parameters,

are distributed from a central control server amongst all 3D sensors in the network.

4. Post-conditions

Items from the bin are picked successfully without failure or downtimes. Field and control room operators receive constant status updates about the picking progress. The robot system can adapt flexibly to new picking tasks.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Bin Picking	Update neural network parameters (FW update)	99%	>>1000 ms	<100 MBit/s (per device)	~100 MB (FW + network)
Bin Picking	Transfer of object pose data	99,9999%	<1000 ms	<0.2 Mbit/s	<512 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m²]	Communication Attributes	Remarks
Not relevant (twice a day)	Not relevant	5-100	>10.000 m²	Not relevant	
33ms	Not relevant	5-100	>10.000 m²	Periodic	

Source: 5G-ACIA / ZVEI e. V.

4.2.7 Collision Warning System

1. Description

Multiple 5G enabled 2D smart sensors are placed around a potentially dangerous area within the factory such that all angles of view within the area are covered. The sensors continuously monitor the area by performing the following tasks:

- Object detection
- Tracking detected objects
- Homography transformation and velocity estimation of tracked objects
- Transmission of tracklet data to an edge server
- Transmission of an encoded on-demand video stream

Based on the tracklet data from all sensors, the edge server performs the following tasks:

- Fusion of tracklets
- Collision prediction of the tracked objects
- Distribution of messages to the people (via smartphone) and vehicles at risk or to signaling devices in the area
- Storage of tracklet data for incremental improvement of fusion and prediction algorithms

2. Pre-conditions

All 2D smart sensors are powered on and connected to the 5G system. There is enough light to acquire images. An edge server is prepared to receive tracklet data and perform collision prediction in time and to distribute warning messages via 5G to the people and vehicles at risk or to signaling devices within the area.

3. User-system & system-system interactions (flow of events)

Multiple 2D smart sensors continuously monitor an area to detect objects. The sensor's internal image processing pipeline tracks the detected objects and calculates movement directions and velocities. Tracklet data, such as an object ID, bounding boxes, motion vectors, velocity information of each object is then transmitted to an edge server. The server fuses the tracklet data of all 2D sensors and predicts the probability of collisions between several objects. When a certain probability threshold is reached, the edge server sends out warning messages to the people and vehicles at risk or to signaling devices within the area. In addition, AGVs and/or AMRs in the area are automatically slowed down or stopped completely by the fleet management system. In the event of a potential collision, the 2D smart sensor covering the respective area is automatically triggered by the edge server and thereby starts streaming encoded video to the control room via 5G. Tracklet data is stored centrally for incremental improvements of fusion and association of tracklets as well as collision prediction algorithms.

4. Post-conditions

A dangerous situation is detected successfully by the tracking system and all workers and vehicles in the area are warned in time to prevent an accident. Tracklet data is stored for incremental improvements of fusion and association of tracklets as well as collision prediction algorithms.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Collision prediction	Tracklet data	99,9999%	<33ms	1 Mbit/s	<100 Byte
Collision prediction	Warning messages	99,9999%	<33ms	1 Mbit/s	<100 Byte
Collision prediction	Continuous encoded video stream	99,9999%	33ms	10 Mbit/s	<1 Mbyte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
33ms	Not relevant	5-1000	100 - >10.000 m^2	Periodic	
33ms	Not relevant	5-1000	100 - >10.000 m^2	Periodic	
33ms	Not relevant	5-1000	100 - >10.000 m^2	Periodic	

Source: 5G-ACIA / ZVEI e. V.

4.3 Mobile Use-cases

4.3.1 Mobile Robots and Automated Guided Vehicles (AGV)

1. Description

In the factories of the future there will be a coexistence between Automated Guided Vehicles (AGV) and humans. The introduction of AGVs will allow the transportation of products, pieces of product, tools, and raw materials across the factory according to logistic needs between storage areas and production lines. To execute these complex tasks, AGVs are to be mobile robots with the capacity to follow information flows on inventory and others, capacity for handling materials, monitoring and control, image processing, recognition, etc. Implementing edge computing features in a 5G network will increase the performance of AGVs use cases.

2. Pre-conditions

The 5G network provides low latency, high reliability, high network availability and high precision location for the remote control of unmanned vehicles [3GPP TR22.804 section 5.3.7].

In the mix of indoor / outdoor scenario for the network deployment context, a specific consideration is requested in terms of frequency band, penetration loses, handover processes and intra-band interference.

AGVs have a high level of safety to enable the AGV for shared human/moving machines areas. Emergency stop procedures possible by the AGV itself as well as by the central control unit.

3. Service flow

From the network deployment point of view, it can be a mix of indoor/outdoor or purely indoor. There are two options for control:

a) AGVs controlled by a centralized automatic controller, where AGVs are automatically steered to move materials in a restricted facility [3GPP TR22.804, section 5.3.7]. It requires live monitoring and remote-control applications.
b) AGV controlled by a human, where an additionally augmented reality application is required.

4. Post-conditions

AGV's are successfully remote controlled using live video monitoring, without downtimes and collisions, the direct human control is also possible. The factory logistic needs are satisfied by the AGV's and operators receive constant status updates about the status.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Mobile Robots and Automated Guided Vehicles (AGV)	Continuous encoded video stream	99,999	5ms	<10 Mbit/s (per stream)	4 MByte
Mobile Robots and Automated Guided Vehicles (AGV)	Control data	99,999	5ms	100 kbit/s - 1 Mbit/s	50 Byte
Mobile Robots and Automated Guided Vehicles (AGV)	Safety Data	99999	5ms	100 kbit/s - 1 Mbit/s	50 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
33ms	<2,5m/s	5-100	100 - >10000	Periodic	
<1-10ms	<2,5m/s	5-100	100 - >10000	Deterministic	
<1-10ms	<2,5m/s	5-100	100 - >10000	Deterministic	

Source: 5G-ACIA / ZVEI e. V.

4.3.2 Autonomous Navigation

1. Description

Several AMRS are deployed in a factory or a warehouse. They are all equipped with lidar as well as stereo vision RGB-D sensors for environmental sensing and a central processing unit that aggregates the sensor data. PLCs on board of the AMRs are controlling the movements and assure functional safety by acting according to zone control functions of the lidar sensors. The central processing unit fuses the depth data (X,Y,Z and intensity values) as well RGB images coming from the lidar and RGB-D sensors and transmits the information together with extrinsic calibration and IMU data to an edge sever via 5G. The edge server calculates point clouds from the depth data and assigns RGB-data to each point. The colorized point clouds are then used to generate or update a map of the AMRs' environment and to run different localization, object detection and recognition as well as classification tasks based on conventional and AI algorithms. The central navigation system, that also runs on edge servers, receives the resulting information and constantly transmits control data to all AMRs. This way, all AMRs are directed centrally via 5G to the destinations coming from the fleet management system.

2. Pre-conditions

AMRs are powered-on and connected to the 5G system. All sensors are powered by the AMR and are ready to acquire RGB, respectively depth data. There are prioritized uplinks available to transmit RGB and depth data to the edge server which is ready to receive and process the data accordingly. Downlinks are available to transmit time-sensitive control data back to autonomous robots that is subsequently translated in movements by the PLC.

3. User-system & system-system interactions (flow of events)

RGB-D and Lidar sensors stream data to an AMR's central processing unit in free run mode. The processing unit captures the sensor data, fuses it, and sends it to an edge server via 5G. The edge server receives the data and performs different image pre-processing and post-processing task, such as the filling of missing data, filtering, object detection and recognition as well as localization and mapping. Based on the localization functions, a fleet management system receives position data for live tracking and tracing of all robots as well as to perform path planning and traffic control management tasks. As soon as a new task with a navigation goal is scheduled by the fleet management system, the navigation stack - often based on ROS (Robot Operating System) - sends control data back to the AMR, that is subsequently fed into the robot's control loop, respectively a motion controller that controls the robot's drive motors.

4. Post-conditions

AMRs reach the navigation goals that are set by the fleet management system and navigates safely and autonomously through the factory or warehouse. Control room operators can rely on real-time position data for live tracking and tracing.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Navigation	Control data	99,9999	7ms	<1Mbit/s	50 Byte
Navigation	Fused 3D and RGB data (container)	99,9999	33ms	300 - 500MBit/s	1,6-2MB (RGB + depth)

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m^2]	Communication Attributes	Remarks
7ms	<2,5m/s	5-200	100 - >10000m^2	Periodic	
33ms	<2,5m/s	5-100	100->10000m^2	Periodic	

Source: 5G-ACIA / ZVEI e. V.

4.3.3 Obstacle Detection and Code Reading in AS/RSs

1. Description

Several robotic shuttles run on a rack structure at very high speeds, each operating on one level of the rack. The shuttles inject and extract loads, such as bin, totes, or pallets into and from the racks. 5G enabled smart 3D sensors that are mounted at the front and the back of the robotic shuttles detect obstacles in the aisle and make the shuttle stop in time. In parallel, the smart sensor sends event messages to the central execution system via 5G, and thereby control room operators are informed about the incident. In addition, shuttles are equipped with 5G enabled 2D code readers that constantly identify the items that are supposed to be extracted from the rack. The codes are compared with inventory data in the execution system and in case of inventory discrepancies, operators are notified to resolve the issue.

2. Pre-conditions

All sensors are powered on and connected to the 5G system. The environmental conditions are controlled and in favor of the 3D method being used to acquire depth data and images. Uplinks are available to transmit event messages and code information to the execution system.

3. User-system & system-system interactions (flow of events)

Robotic shuttles run on a rack structure following task planning instructions from the fleet management, respectively the warehouse executions system. Smart 3D sensors detect an obstacle in front of the shuttle and inform the shuttle's onboard controller to make it stop. In parallel an event message reflecting the incident is sent via 5G to the execution system by the smart sensor. The execution system cancels the task and notifies an operator about the obstacle in the aisle.

A smart 5G code reader constantly identifies the loads in the rack and transmits the information to the warehouse executions system. In case of a mismatch between the inventory data in the system and the actual data from the code reader, the scheduled extraction task is cancelled, and an operator is notified to check the load in the rack. Also, inventory data is updated according to code reader information.

4. Post-conditions

Robotic shuttles do not get stuck on the racks. Inventory data is up to date and the right loads are extracted from the racks.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Obstacle detection	Event messages	98-99 %	<1000ms	1 Mbit/s	100 Byte
Code reading	Decoded data	98-99 %	66ms	0,2 Mbit/s	<50 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m²]	Communication Attributes	Remarks
<1000ms	6m/s	10-100	1.000 - >10.000m²	Periodic	
66ms	6m/s	10-100	1.000 - >10.000m²	Periodic	

Source: 5G-ACIA / ZVEI e. V.

4.3.4 AMR Remote Operation

1. Description
Several AMRS are deployed in a factory or a warehouse. The robots are equipped with multiple 2D or RGB-D sensors for navigation and inspection purposes. The robots are travelling autonomously throughout a factory or warehouse environment, and they rely on their perception and navigation system to reach a given destination goal without incidents. However, in some cases a robot may get stuck and cannot free itself alone. The fleet management system recognizes that the AMR is not moving forward on its planned route and notifies a remote-control operator, who takes control of the AMR. The operator uses multiple camera views provided by the AMR's on-board 360° perception system to navigate the robot remotely out of the difficult situation. Afterwards, the remote-control session is ended by the operator and the AMR resumes its planned route automatically.

2. Pre-conditions
On-board 2D or RGB-D sensors are powered-on, and the AMR is connected to the 5G system. There are uplinks available to transmit several video streams via 5G to the remote-control room's monitoring system. If the control room is outside the factory or warehouse, also public network links must be available for video data transmission.

3. User-system & system-system interactions (flow of events)
An AMR get stuck in a difficult situation. The fleet management system recognizes the incident, and a remote-control operator is notified to take over control of the robot. The operator starts a remote-control session and gets access to the robot's motion control system. By starting the session, the AMR's perception system is triggered to stream multiple encoded 2D video streams via 5G to the remote-control room's monitoring system. With the help of the live 360° view of the AMRs surrounding, the operator can free the robot and gets it back on track. Afterwards, the operator ends the session, and the fleet management system receives a status update message. The AMR resumes the pending task and starts navigating autonomously to reach the destination goal.

4. Post-conditions
The AMR is freed from the difficult situation remotely and can resume its task. No operation personnel must be onsite to resolve the issue.

5. Requirements

System Task	Communication Stream	Communication Service Availability [%]	End-to-end latency [ms]	Average Service Data Rate [Mbit/s]	Message Size (Payload) [byte]
Remote operation	Continuous encoded video stream	98-99%	16ms	10 Mbit/s (per stream)	<1 Mbyte
Remote operation	Control data	98-99%	<50ms	<1Mbit/s	50 Byte

Source: 5G-ACIA / ZVEI e. V.

Transfer interval, target value [ms]	End device speed [m/s]	# End Devices	Service Area [m²]	Communication Attributes	Remarks
16ms	<2,5m/s	5-100	100 - >10000m²	Periodic	
50ms	<2,5m/s	5-100	100 - >10000m²	Periodic	

Source: 5G-ACIA / ZVEI e. V.

5 Architectural and Technology options

5.1 Introduction

In the following chapter, the different components and architectural options of a machine vision system are presented. The goal is to explain what a Machine Vision system is, the different options of building such a system and finally, the different interfaces to integrate machine vision system in an automation system. It enumerates the most used technologies today and current standardization efforts for upcoming interface technologies.

5.2 Machine vision system components and architecture

The typical machine vision camera is based on an image sensor that converts the incoming light into electrical impulses, a lens that focuses the light of the scene and projects it sharply on the sensor surface, and a processor that pre-processes the data coming from the sensor. The pre-processed image data is then transported via a wired interface.

The same interface that is used for image transport is also used to configure and control the camera via an API or to supply the device with power (Power over Ethernet). Additional IO connectors allow, for instance, synchronization with external lighting components or with other cameras (multi-camera operation) as well as triggering the sensor with low latency.

In addition to the camera, a machine vision system includes other components such as integrated or external lighting, which illuminates the scene either permanently or sequentially, and a host system, which receives and processes the data coming from the camera. In general, there are three main machine vision system approaches:

1. **Machine Vision Camera System**
   Modular combination of one or more standard machine vison cameras, lenses, cabling and lighting components and a separate industrial image processing PC running the application specific image processing software.

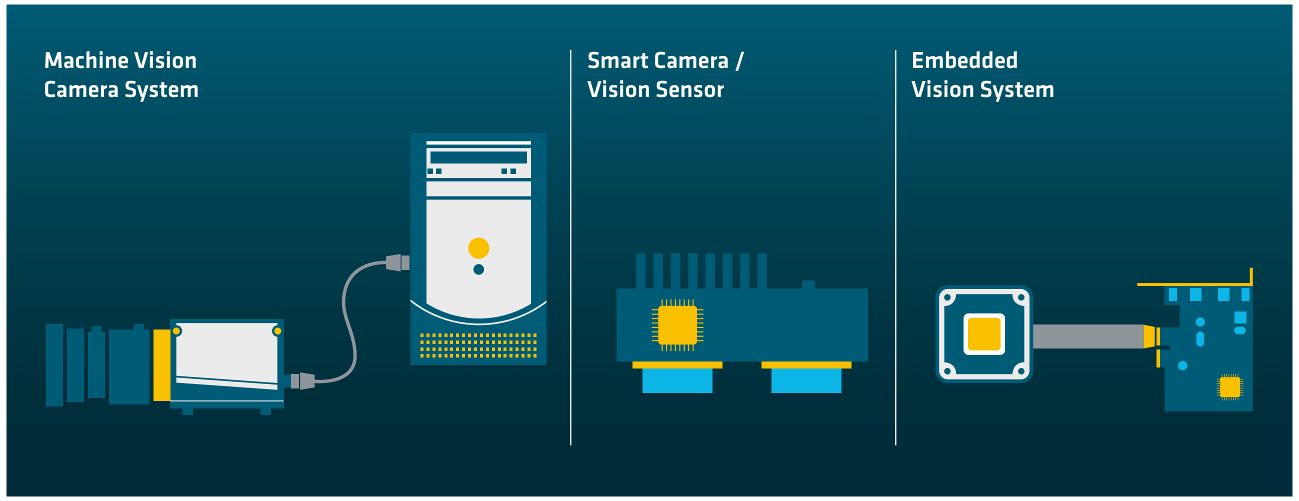
2. **Smart Camera / Vision Sensor**
   Integrated camera device that combines a camera, lens, and application processing in a single intelligent device. Some smart cameras are also equipped with lighting components. While vision sensors are dedicated to one specific applications task, smart cameras do host a complete library of vision algorithms and a development environment in which users can flexibly program an image processing pipeline to solve a given problem.

3. **Embedded Vision System**
   Embedded Vision Systems make use of embedded computing hardware such as "System on Chip" and "System on Module" components and their heterogeneous processing architectures based on CPUs, GPUs, NPUs, ISPs and DSPs to deliver outstanding performance at very low cost. This system approach uses low-level interfaces, such as MIPI CSI-2 or GMSL to connect single board camera modules and transport acquired image data directly into the processor's memory. Depending on the application and the performance of the embedded computer, the image processing is performed on the embedded computing device directly or offloaded to an edge server.

As the eyes of a machine or a robot, cameras have become one of the most important sensors to help control a machine or monitor a process in industrial automation. In this capacity, machine vision systems are connected to the DCS (Distributed Control System) of a factory by interfacing with PLCs or other types of controllers, such as a robot system control unit. In terms of communication, machine vision systems mainly communicate either via TCP/IP or different Ethernet-based fieldbus protocols, such as Ethernet/IP, EtherCAT or ProfiNet with other participants in the network.

The design of a vision system is very much dependent on the system requirements of the overall system in which the vision system shall be integrated. Besides the optical setup with major criteria such as field of view or working distance, the object and environment conditions are important factors that influence the design. Equally important, however, are the interface and processing requirements. Cable length, image acquisition speed (frames per second), interface throughput, as well the as the application specific processing time based on conventional or AI based image processing algorithms, are decisive criteria to consider when designing a machine vision system. 5G can be a game changer in these aspects of system design by providing high-bandwidth and low-latency data transmission capabilities, thereby increasing installation and deployment flexibility while reducing the spatial footprint of vision systems at the same time.

Figure 1: Main machine vision systems approaches



Source: 5G-ACIA / ZVEI e. V.

5.3 Technology options

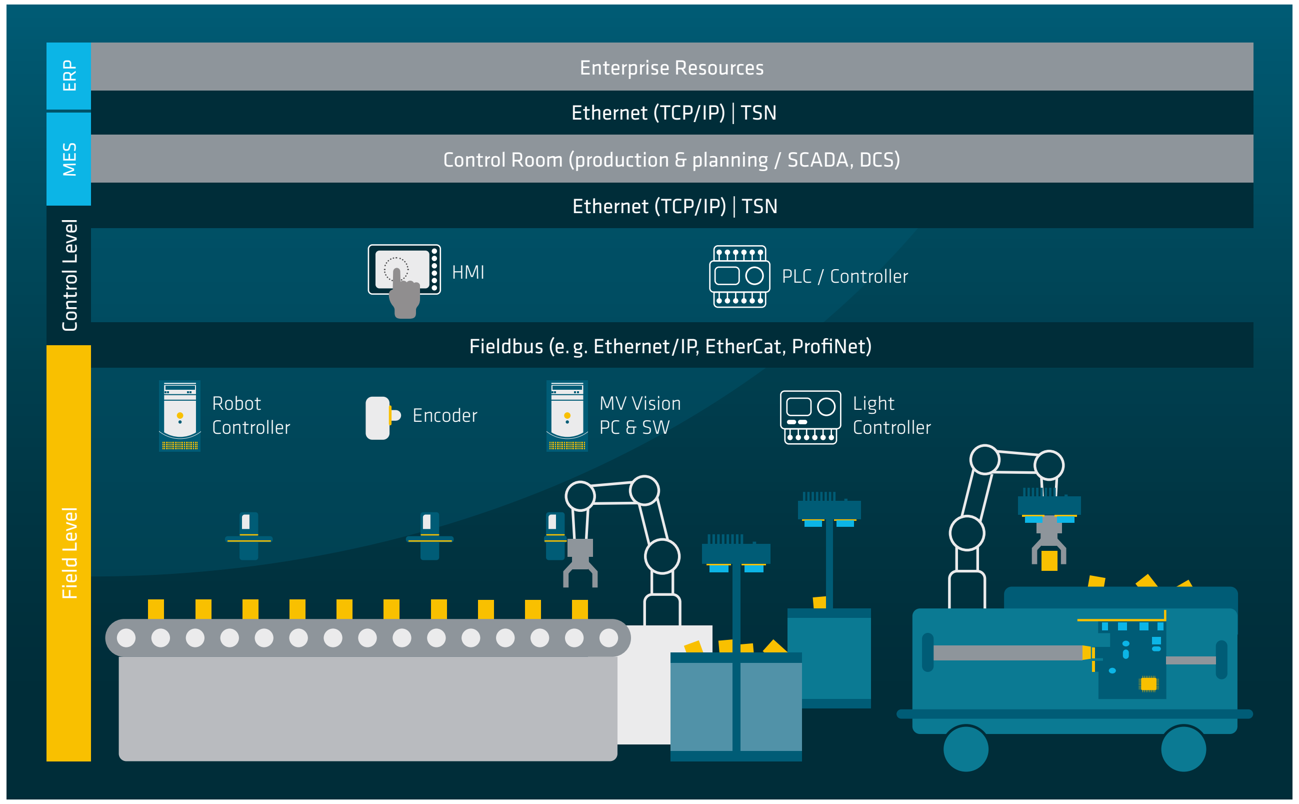
5.3.1 Vision sensor interface

Machine vision system architectures do not only rely on high-level plug & play interfaces. In the recent years an increasing number of embedded vision applications have appeared on the market which make use of low-level interfaces such as LVDS (Low Voltage Differential Signaling) or MIPI CSI-2 which is hosted by the MIPI Alliance – or Mobile Industry Processor Interface Alliance. Well-known MIPI standards include the Display Serial Interface (DSI) and the Camera Serial Interface (CSI) specifications.

Whereas LVDS is used in FPGA based systems, MIPI CSI-2 has become the de facto standard for transferring image and configuration data between low-cost camera modules and embedded computing devices. The MIPI CSI-2 specification describes the physical layer of the signal transfer (D-PHY or C-PHY) as well as the CSI-2 protocol for image data transfers. The standard also specifies an interface for camera configuration via I^{2}C, namely CCI (Camera Control Interface). The MIPI Alliance has also recently approved the use of MIPI CSI-2 over Wi-Fi.

The MIPI Alliance recently specified how to map CSI-2 to IEEE 1722 AVTP (Audio Video Transport Protocol). It is expected that this work can be leveraged for an Ethernet network,

Figure 2: Simplified automation pyramid



Source: 5G-ACIA / ZVEI e. V.

since the IEEE 1722 is a common Audio and Video Transport for all IEEE 802 network types (including wired IEEE 802.3 and wireless IEEE 802.11), and thus including Time Sensitive Network Applications in Bridged Local Area Networks.

5.3.2 Transport layer options

Low latency values are predominately important for real-time applications in which speed is crucial or

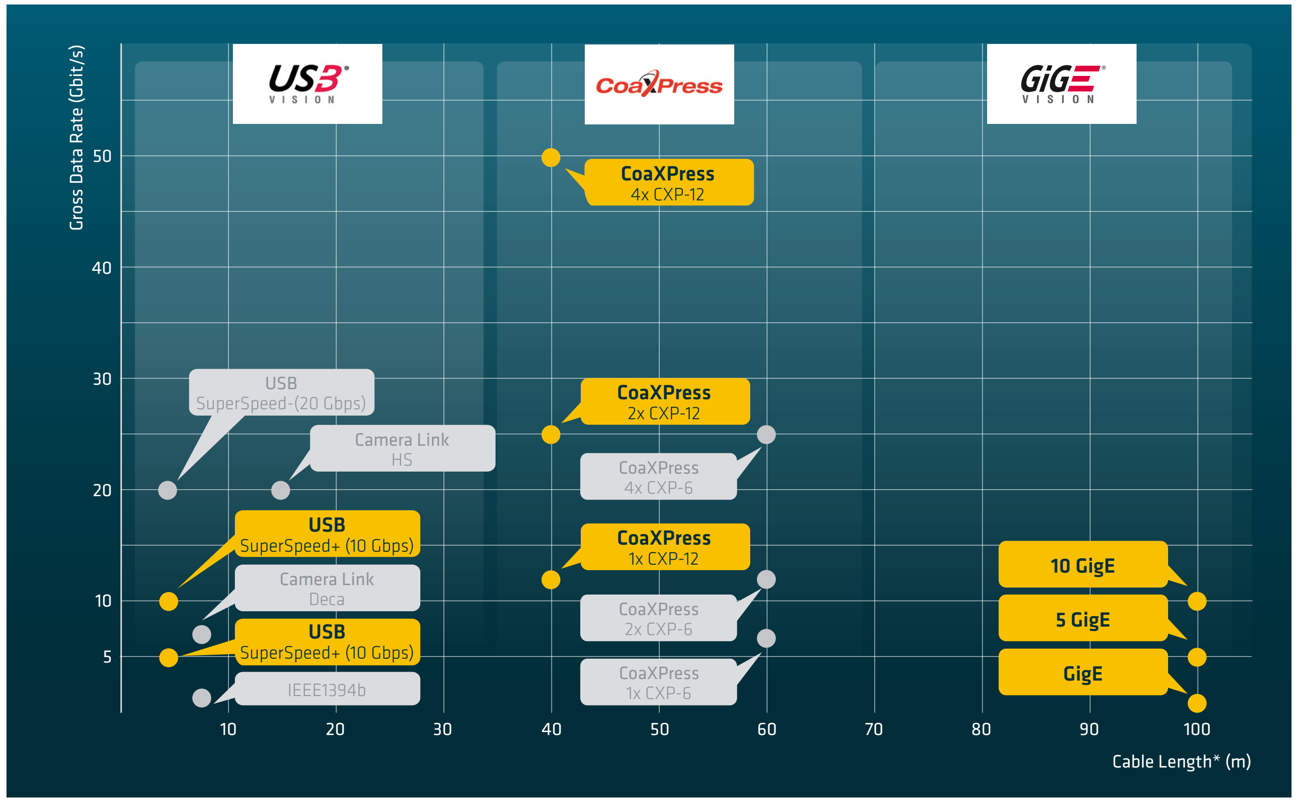
cameras must be synchronized with other system components, such as lighting.

Depending on the form of triggering, latency and jitter values can vary considerably. While a hardware trigger uses a dedicated wire or electronic signal which is sent directly to an input pin of the camera, software trigger sends a command to the camera's configuration channel which is less responsive and thus adds to overall latency. Depending on the quality of the implementation of the host software and camera, this delay ranges between 80 and 500 µs for a typical vision system. A HW trigger on the other hand, allows for more real-time in the magnitude of 10 µs.

The main camera interfaces for industrial imaging that are available today are hosted by the A3, Association for Advancing Automation (GigE Vision®, Camera Link®, Camera Link HS® and USB3 Vision®) and the JIIA, Japan Industrial Imaging Association (CoaXPress). GigE Vision was originally introduced in 2006 by major players of the machine vision industry and advanced to version 2.2. in June 2022.

Due to its technical flexibility in terms of bandwidth, cable length and multi-camera functionality, GigE Vison has become the most used interface in the machine vision industry.

Figure 3: Most important machine vision interfaces



Source: 5G-ACIA / ZVEI e. V.

Furthermore, the possibility to obtain power over the data lines by using Power over Ethernet (PoE) and high compatibility with existing network infrastructure makes GigE Vision the interface of choice for many users.

When it comes to transferring a large amount of data over greater distances CoaXPress comes into play.

The standard was officially released in 2011 and became established in industrial image processing, eventually advancing to CXP 2.0 in 2021. CoaXPress is particularly popular in the semiconductor industry.

In 3D AOI systems (Automated Optical Inspection), for example, large amounts of data must be realized at high resolution without significant latency. Other applications include print inspection, food inspection, intelligent traffic solutions (ITS), and medicine.

Besides the previous enumerated transport technologies, there is further technology that may serve machine vision for industrial use-cases, this is specified in the IEEE Std 1722™2021[3] The IEEE Std.1722™-1-2013 was defined by the AVB task group, which meanwhile has changed its name to the Time Sensitive Networking (TSN) task group to highlight that its applicability goes beyond just audio/video bridging. Due to the native TSN nature of IEEE Std 1722™2021, it is expectable a broad acceptance from the industry, both wired and wireless.

IEEE Std 1722™2021 defines a standard way to transport media streams over a LAN, connecting IEEE Std 1722-2021 streaming data providers and one or multiple stream data consumers. This standard enables interoperable streaming by defining media formats and encapsulations, media synchronization mechanisms and multicast address assignment.

The so-called Audio/video bridging/time sensitive networking discovery, enumeration, connection management, and control (ATDECC) define the communication protocols between ATDECC Entities, thus enabling discovering other ATDECC Entities, identify their capabilities, manage audio/video bridging (AVB) streams, control media parameters, and allows ATDECC Entities to report health and diagnostic status to ATDECC Controllers. These steps can be used together to form a system of end stations that interoperate with each other in a standards compliant way. The application that will use these individual steps is called an ATDECC Controller and is the third actor in the ATDECC Talker, ATDECC Listener and ATDECC Controller device relationship.

As mentioned in the previous chapter, there is at the time of the writing of this document a specification proposed by Mi-Pi-Alliance to the IEEE TSN Task Group, defining the mapping of CSI-2 to IEEE 1722 AVTP (Audio Video Transport Protocol). However, there is no decision on the timing or content of an IEEE 1722 release supporting this mapping.

5.3.3 Application Layer Interface

All already standardized machine vision interfaces rely on GenICam™, a generic programming interface for all kinds of cameras and a standard of its own. GenICam™ describes the features supported by industrial digital cameras, a transport layer interface for grabbing images, and a generic end-to-end configuration interface. By GenICam™, machine vision cameras become interchangeable no matter what interface technology is used. The standard is hosted by the European Machine Vision Association (EMVA).

6 Machine vision systems with 5G

6.1 Introduction

Machine vision use cases, as described in section 4, can benefit from integration with a non-public 5G network in two aspects: use cases which are mobile by nature benefit from having no communication wires, which could obstruct the movements of devices and in many cases limit the physical space where assets may roam considerably; use cases of stationary nature may also benefit from the 5G integration, as a wireless communication enables frequent reconfiguration when required, as well as expanding brown field production systems with smart sensors and accessories.

5G non-public networks enable a flexible configuration of different communication bearer types ranging from ultra-low latency to very high bandwidth capabilities, supporting a large variety of machine vision use cases.

6.1.1 Key-features and capabilities of 5G supporting machine vision applications.

User data in camera-based systems often has extremely high bandwidth requirements, especially when raw vision data must be transferred over the radio interface. Edge cloud technologies may be applied in some cases to limit that bandwidth, e. g. by applying smart image pre-processing or AI-based processing. Simple cameras (e. g. camera heads) are limited regarding this type of processing due to size and cost constraints. Due to these varying demands, a non-public 5G network must cope with connections of extreme high to low bandwidth for the same or of different vision systems camping on the network. Apart from bandwidth flexibility, machine vision system also benefits from the superior systems reliability of non-public 5G network, compared to other wireless technologies such as Wi-Fi.

Machine vision use cases, like many other industrial use cases, have often significant security and privacy demands that can be fulfilled by a non-public 5G network, as outlined in ref [6].

Some of the key feature of the 5G technology, which in particular machine vision use cases benefit from, are listed below.

- *Multiple parallel data flows for one device*
  Applying different QoS profiles to different data flows, enables the users to establish e. g. low-latency and average-latency communication bearers, in parallel. Following the same QoS profile principles, bearers with other attributes (e. g. bandwidth) can be established. One such paramount attribute apart from latency is bandwidth. QoS 5G principles are detailed in ref [5]

- *Integration with ERP systems*
  ERPs integrated with non-public 5G networks via exposure APIs allows the factory operator to modify data flows attributes. That modification allows the adaptation of 5G resources for many use cases and assets, optimizing scarce resources such a radio spectrum. Exposure services which are offered by 5G networks are described in ref [6].

- *Flexible coverage area for radio connectivity*
  5G coverage spans, by nature, from a large geographical area – e. g. an industry or processing campus - to a restricted to very small locations – e. g. the size of a pico-cell covering one production cell only.

- *Network slicing*
  The technology to subdivide a larger 5G networks into many logically separated virtual 5G networks is called network slicing. Such slices, or logically separated virtual 5G, can be offered to different enterprises, guaranteeing resource separation as well as data privacy.

- *Positioning*
  5G offers the function to locate the position of a 5G device, eliminating the need to install another positioning solution

- *Integration with existing industrial communication systems*
  5G can be integrated with existing wired communication networks on Ethernet- or on IP-level. A 5G network in that respect appears to the industrial communication systems as a L3 IP-router or as a L2 Ethernet-bridge. For deterministic communications 5G also supports DetNet and TSN technologies, as defined in ref [8] and ref [7].

- *5G deployed with Edge Computing*
  5G NPNs can be, as an option, be deployed with edge computing capabilities, as outlined in ref [11]. Applying edge computing capabilities may provide significantly lower latency communication, avoiding high bandwidth data tromboning to machine vision systems, as well as to ensure data privacy.

6.1.2 5G features mapping with use-cases and their requirements

There are three major communication types with respective technology features and key performance indicators (KPIs) define for 5G. These communication types can be configured and deployed in non-public 5G networks for different machine vision devices, depending on the envisioned use case and respective communications requirements.

The three communication types, their key characteristics as well as their applicability for machine vision use-cases are described below.

eMBB: enhanced Mobile Broadband:
Following features are supported
  - Peak data rate of 10 Gb/s

mMTC : massive Machine-Type Communications
Following features are supported
  - High IoT device density of one million/km2
  - Low-rate data (1 to 100 kb/s)
  - Optimal energy consumption of 10% for LTE network

URLLC: Ultra-Reliable Low-latency Communications
Following features are supported
  - Latency below 1 ms
  - Availability of 99.9999 %
  - Supports high speed mobility

6.2 Integrating 5G and machine vision systems

6.2.1 Machine vision building blocks and options

Machine vision systems consist of three major building blocks, as outlined in chapter 5 of this document.

- A vision controller system, which host an industrial vision application controller. That application uses an industrial protocol (e. g. OPC-UA, Profinet, EtherCAT, etc.) to connect either via L2 (industrial Ethernet) or L3 (IP) to
- A vision embedded compute system, which terminates the industrial protocol towards the vision controller system and connect e. g. via Ethernet to one or more cameras.
- The camera which captures pictures and send pre-processed images them to the vision embedded compute system for processing, analysis, and event triggering.

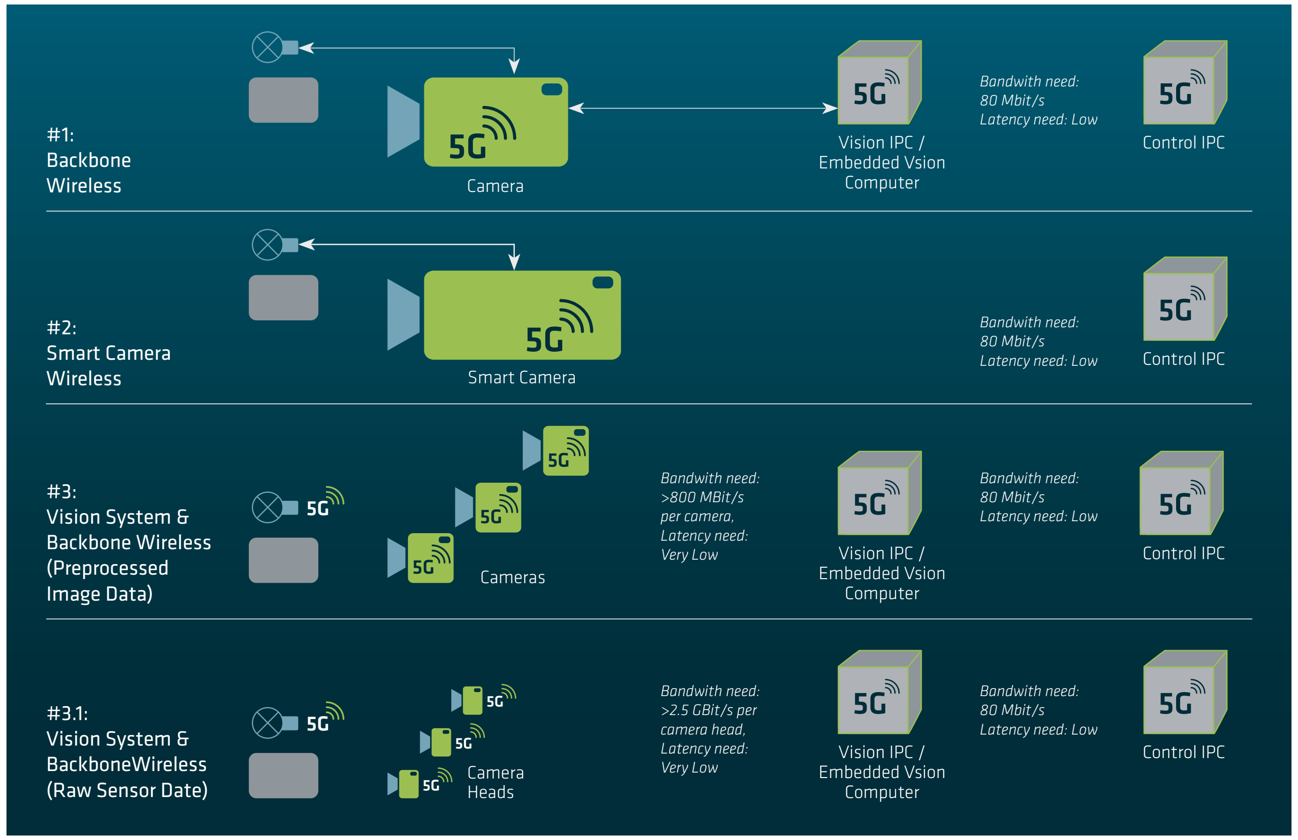
As an option, the cameras and the vision embedded controller can be combined to form a smart camera. Yet, as an additional option, a camera can consist of a simpler camera head only which captures images and uses e. g. MIPI to transfer raw image data to an embedded vision controller.

These different deployment options are described in chapter 5. In that chapter, also the expected bandwidth demands are shown, where the simple camera head solution produces raw data with a demanded bandwidth of more than 2.5 Gbit/s, while a camera's demands are in the range of 800 Mbit/s and above.

Integration of these different vision solution options with a 5G system are depicted in fig. xx. It should be noted that the bandwidth demands and the radio integration points for the different options will have a tremendous impact on the 5G system capacity needs, and in some cases may even go beyond the capabilities of commercial private 5G networks deployments.

An industrial vision system as used today with wired industrial communication networks, does not have any 5G radio capabilities as per default. In the following section we analyze and propose how the different industrial machine vision systems can be integrated with 5G NPNs [9] and what equipment and capabilities are needed.

Figure 4: Vision System Scenarios with 5G



Source: 5G-ACIA / ZVEI e. V.

6.2.2 5G systems building blocks and interconnected OT networks

5G system consist of two major building blocks, the 5G network and the 5G UE (User Equipment), as depicted in figure 5 below.

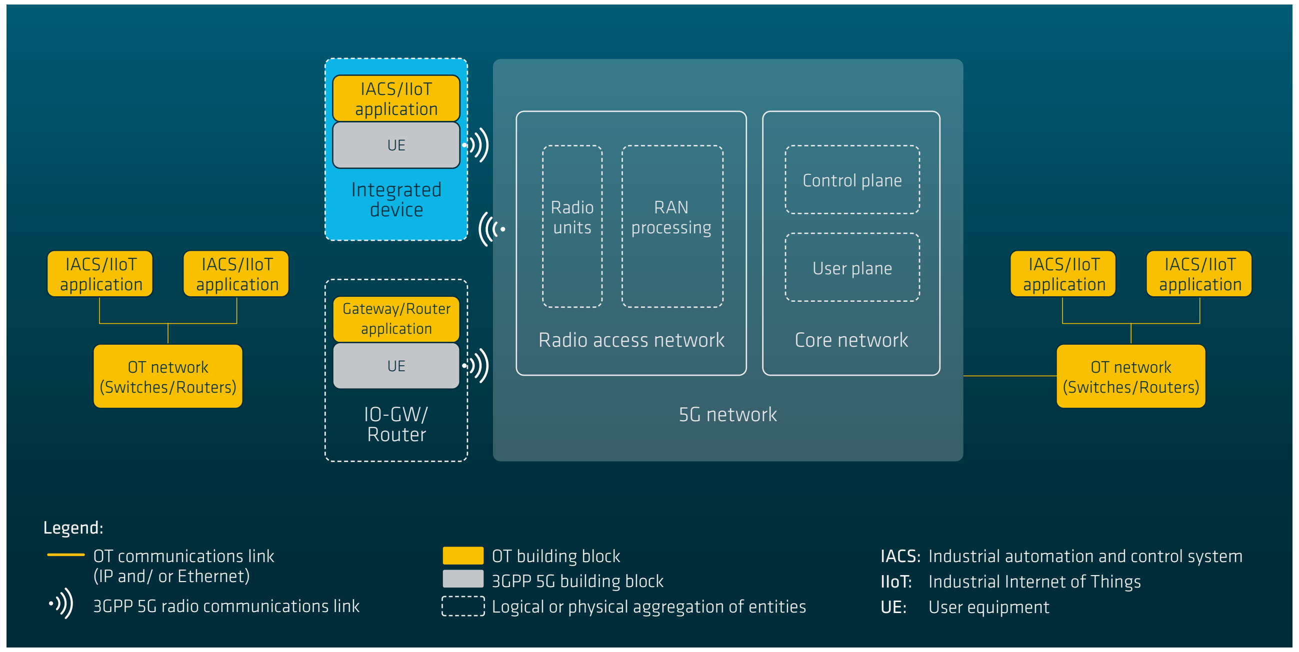
The 5G networks consists of a core network and of the radio access network. It connects the radio interface to the 5G UE, which consists of a 5G module with an embedded chipset – that module is responsible for the radio stack. Details of the UE and of industrial device decompositions are detailed in the white paper (ref [10]). On a genera level, industrial devices can be integrated, i.e. the UE and the sensor/actuator or controller in an integrated, self-contained, entity.

On the other hand, an IO-GW/router type of devices contains a UE and contains also L2 and or L3 communication infrastructure. These types of devices are used to connect separate sensors/actuators or controllers, which may be unaware of the 5G connectivity provided by the IO-GW/router.

Regarding the ingress/egress at the core network side, external OT industrial (wired) communication networks are integrated using standard L2 and L3 interfaces. Sensors/actuator or controllers connected to these OT networks and communicating with radio-connects sensors/actuator or controller may be also unaware of the of the 5G connectivity provided by the 5G network.

However, controllers are recommended to be aware of the 5G connectivity to manage bandwidth, latency and possible other 5G capabilities via the 5G system provided exposure interface, as explain in the white paper ref [6].

Figure 5: 5G system building block and interconnect with industrial OT networks and applications



Source: 5G-ACIA / ZVEI e. V.

6.3 Integration of industrial vision with 5G systems

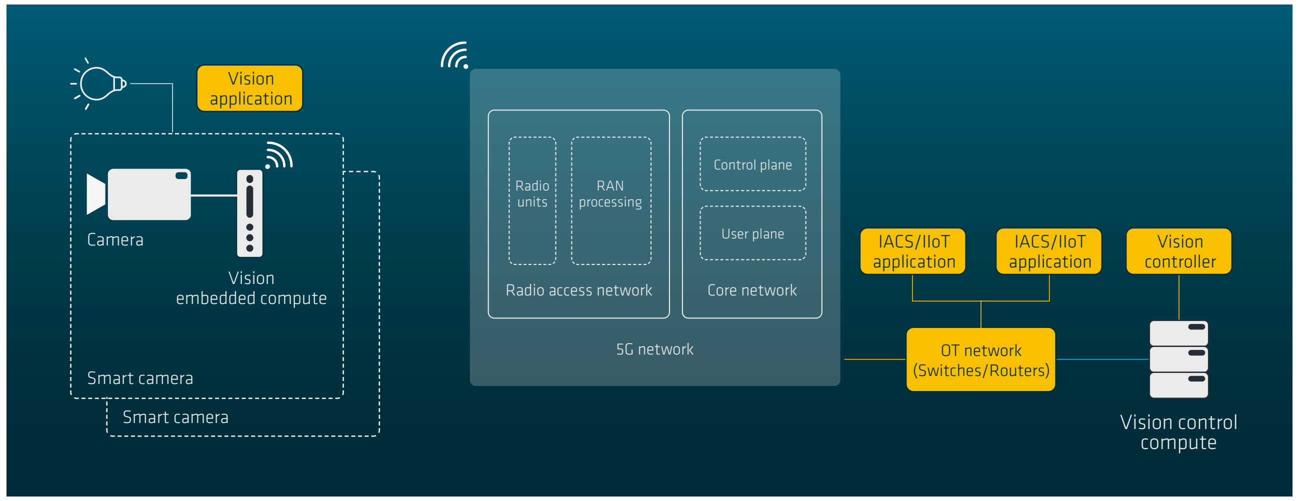
6.3.1 Integration of smart camera systems with 5G (case I)

The most commonly deployed vision system is the smart camera option. With this option, the embedded vision compute is integrated with the camera and uses locally stored algorithms to analyze and process the camera images. Based on that analysis, the controller will send events (triggers) to the vision compute, whose application logic will then steer the industrial manufacturing process. The 5G radio bandwidth demands for this integration case is rather modest and ranges from 10-100 kbit/s, sufficient for event trigger messages. These messages must be transferred (both up- and downlink) with short latencies in the range of 10-100 ms – exact values depend on the actual manufacturing process. That bandwidth may be higher, when the application also triggers the embedded controller to send images, e.g. for process documentation or additional quality checks. However, for image transfer the latency demands are much lower as no real-time processing is expected by the vision controller.

In this integration case I, the embedded vision compute must be enriched with a UE to form an integrated 5G device. That UE connects to the 5G radio ingress/egress point and communicates with the vision controller connected to 5G core network egress/ingress point. This integration option I is depicted in figure 6 below.

In that figure, the yellow blocks are industrial applications, which may or may not be 5G aware. The vision controller may be connected to the OT network with L2 or L3 protocols, depending on the industrial control protocol used.

Figure 6: Integration configuration I: Smart camera



Source: 5G-ACIA / ZVEI e. V.

In the figure 7 below, the protocol stack view is shown for that configuration option, assuming that the industrial control protocol is using L3 IP (TCP/UPD) at the reference point B.

At the reference point C, two data streams are shown: a data stream for low bandwidth, low latency trigger event message is shown; a second data stream with higher bandwidth but relaxed latency is also shown that can be used for e.g. on-demand or cyclic image transfer. These two data streams can be mapped to 2 QoS flows with different QoS attributes (ref [5]) which are defined in the vision controller or in an ERP-system managing the industrial assets, including the 5G connectivity. In this example a 5G-aware application is superior to a non 5G-aware application, as different communication demands can enhance system performance and can make efficient used of scarce radio resources.

Finally, the reference point A is the actual radio interface (Uu interface in 3GPP terms), which is managed by the embedded UE and the 5G network.

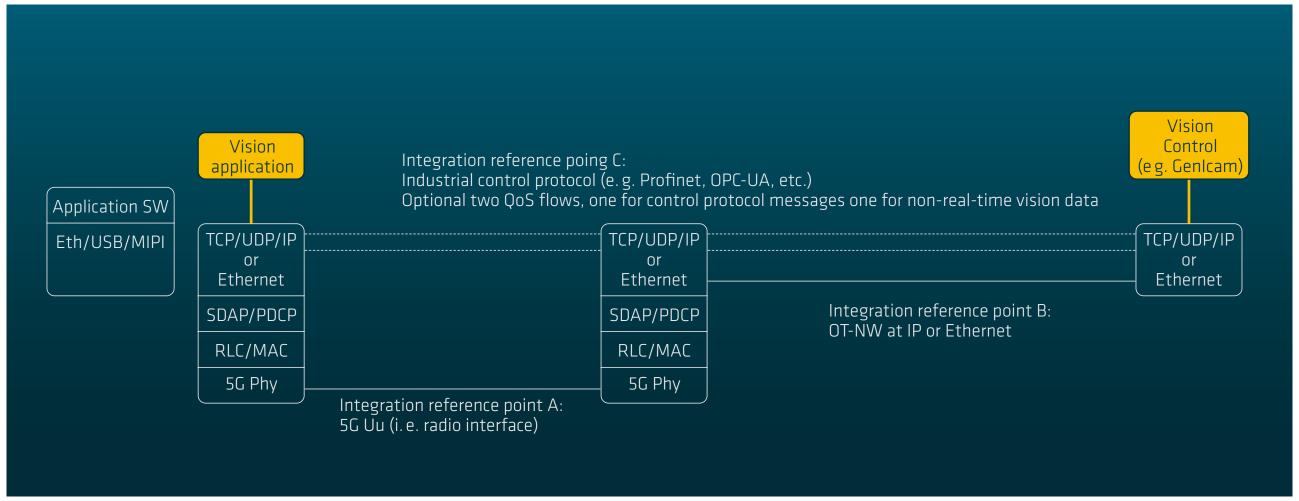
For this integration option I it is also possible to use a non-integrated device option, i.e. the UE may be not embedded with the smart camera (resp. with the embedded vision compute). Instead, the smart camera can be connected to an IO-GW/router, which then connects to the 5G network. Bandwidth and latency demands are not impacted by that IO-GW/router. An IO-GW/router may be commercially beneficial when e.g. many smart cameras are to be connected or when HW formfactors don't provide for an integrated UE solution.

This integration method (case I) can be applied to all use cases listed in section 4, assuming that smart cameras are deployed, which can be integrated on IP-level with the 5G network and with the industrial communication system.

6.3.2 Integration of camera systems with 5G (case II)

The major difference with case II is that the smart camera is by default separated, i.e. it consists of a camera and a remote embedded vision compute - that remote vision compute is deployed behind the 5G network. It may even not be an embedded system, instead a cloud application deployed in a data center in the OT network domain. Case II is depicted in fix z.1 below, where the camera is either enriched with a UE (forming an integrated 5G device), or the camera is connected

Figure 7: Integration configuration I: Smart camera, protocol stack view



Source: 5G-ACIA / ZVEI e. V.

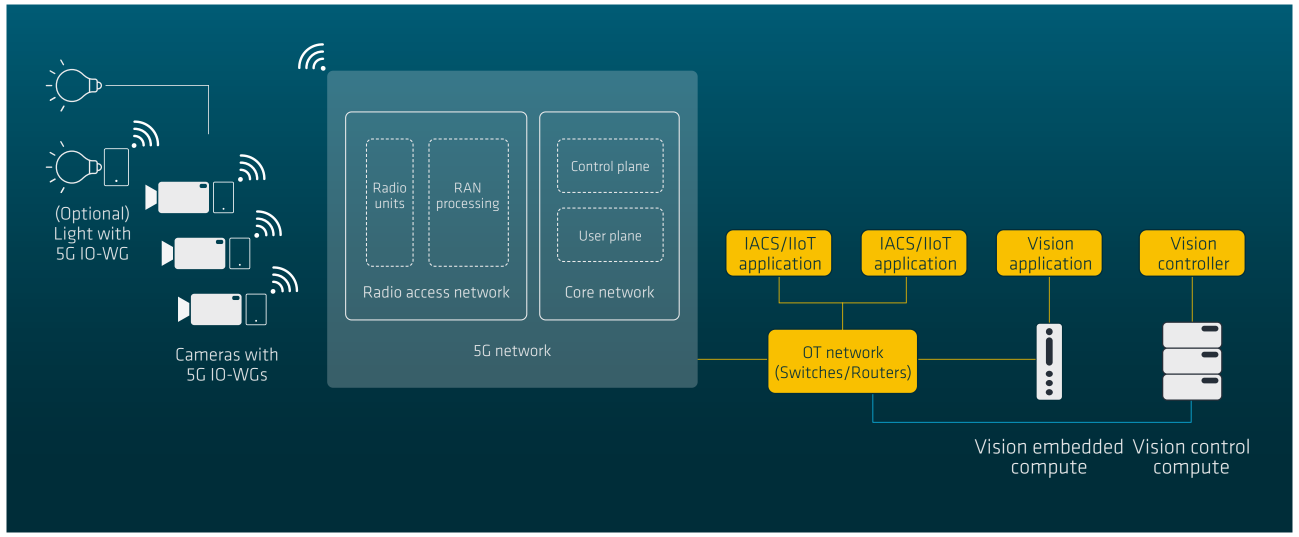
to an IO-GW/router instead, when a UE integration is neither commercially nor from a form factor's perspective beneficial.

Bandwidth, as well as latency demands, are significantly higher, as the camera(s) only pre-process and forward images over e. g. GigE standard L2 communication to the vision compute for analysis and processing. The vision controller is integrated with the OT network as in case I.

Furthermore, in this case II, the benefits of a 5G-aware vision controller or ERP-system also apply, as for case I above.

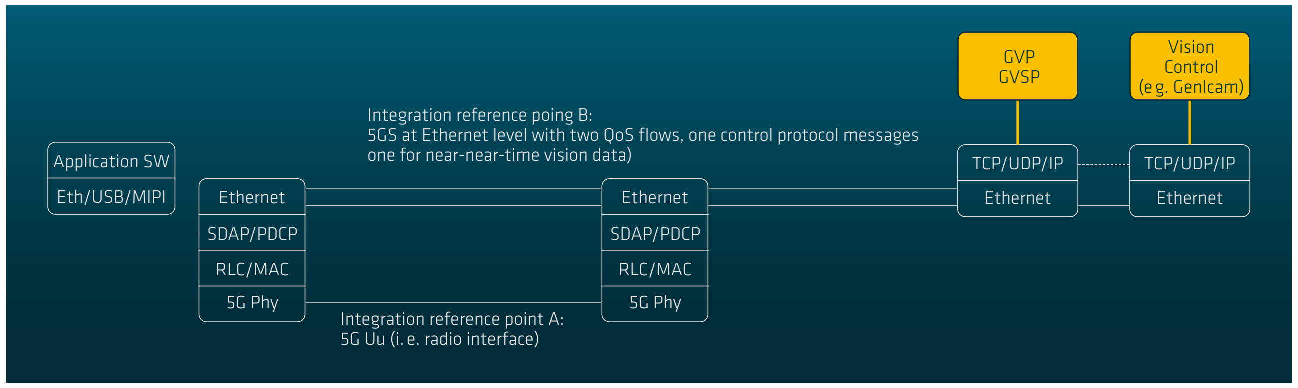
As the images from the camera are transported as L2 frames, the UE (or IO-GW/router) must support the '5G Ethernet-PDU session type' feature, to transport an Ethernet frame across the 5G network. The 5G network must support that feature, additionally both the network and the UE may support relevant TSN features to manage QoS L2 streams as described in ref [7].

Figure 8: Integration configuration II: Camera



Source: 5G-ACIA / ZVEI e. V.

Figure 9: Integration configuration II: Camera, protocol stack view



Source: 5G-ACIA / ZVEI e. V.

This integration method (case II) can be applied to those use cases listed in section 4, which use less capable camera systems as compared to smart cameras. Use cases applicable for this integration method are those demanding many cameras such as 'Inline Multi-camera Manufacturing', 'Inline Multi-Camera Quality Inspection' or 'Collision Warning System'.

6.3.3 Integration of camera systems with 5G (case III)

Case III constitutes the most bandwidth demanding integration alternative, because all raw image data must be transferred from the camera head(s) to the vision compute system deployed remotely, behind the 5G network. In this case an integration with a UE is most likely not possible due to the small formfactor of the camera heads. Therefore, we depict in figure 10 below the integration of the heads with IO-GW/routers. An IO-GW/router may connect many camera heads, if the capacity of that GW device and its embedded UE is sufficient.

In this configuration, the IO-GW/router must support an additional function, namely the conversion between MIPI and Ethernet. MIPI frames cannot be sent via a 5G system or any other L2/L3 compliant communication system.

Given the extreme bandwidth demands and complex integration, this case III may not have many industrial use cases examples that would justify its integration and deployment.

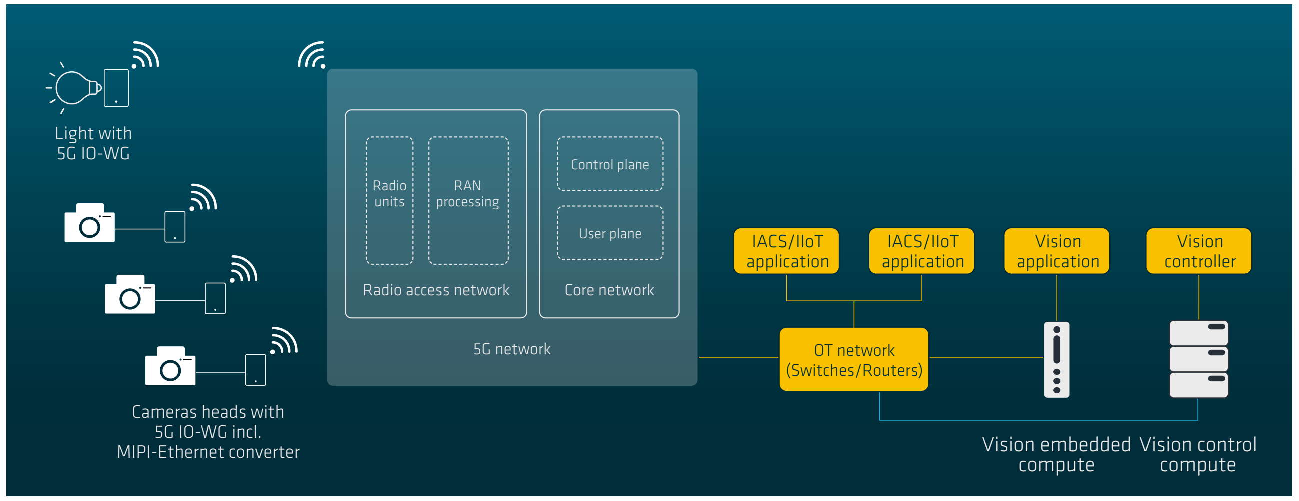
The protocol stack view for this case III is identical to the case II above and not repeated in this sub-section.

Again, as for the case I and II above, the benefits of a 5G-aware vision controller or ERP-system apply.

Given the extreme bandwidth demands and complex integration, this case III may not have many industrial use cases examples that would justify its integration and deployment.

This integration method (case III) can be applied for all use cases where simple, inexpensive, or physically confined camera heads are applied. The bandwidth demands with this method are very high, as raw image data is transferred over the 5G network. Special care must be taken (e. g. by the ERP system) to manage the scarce radio resources. Radio bearers and associated user data flows assigned to camera heads, which are not used at a certain time in the production process, should be dynamically modified to low-priority bearers and changed back to high-priority when used. That enables the smart and optimized usage of the limited networking resources amongst various use cases and various production cells.

Figure 10: Integration configuration III: Camera head



Source: 5G-ACIA / ZVEI e. V.

7 Conclusion

This document describes eleven use-cases of machine vision in industrial applications. These use-cases, covering a wide spectrum of mobile and stationary industrial automation applications, provide a good overview of the required 5G key performance indicators.

This document also demonstrates that 5G is a flexible technology, which can support the described machine vision use-cases. This flexibility is possible through the configuration of the 5G network Key performance Indicators.

This flexibility is further enhanced through the existing and upcoming standardized interfacing technologies, thus leveraging the integration of machine vision systems in 5G.

Therefore, this document clearly highlights the importance of bringing both 5G and TSN technologies closer together for advanced industrial automation.

3GPP
The 3rd Generation Partnership Project (3GPP) is a collaborative project that brings together standardization organizations from around the world to create globally accepted specifications for mobile networks. As its name implies, it was first created to establish such specifications for the third generation (3G) of mobile communication systems. It has continued its work for subsequent generations, including the one considered here, the fifth generation (5G).

5G-ACIA
5G-ACIA is the globally leading organization for shaping and promoting Industrial 5G.

8 References

[1] Deliverable D2.1 Use case specification & Open Call definition (first version) of EU-project IntellIoT (https://intelliot.eu/)
[2] Deliverable D2.2 Technology Analysis & Requirements Specification (first version) of EU-project IntellIoT (https://intelliot.eu/)
[3] IEEE Std 1722.1™2021 IEEE Standard for Device Discovery, Connection Management, and Control Protocol for Time-Sensitive Networking System
[4] 5G-ACIA white paper “*Security Aspects of 5G for Industrial Networks”*
[5] 5G-ACIA white paper *5G QoS for Industrial Automation - 5G-ACIA”*
[6] 5G-ACIA white paper “*Exposure of 5G Capabilities for Connected Industries and Automation Applications”*
[7] 5G-ACIA white paper “*Integration of 5G with Time-Sensitive Networking for Industrial Communications - 5G-ACIA”*
[8] 5G-ACIA white paper “*Integration of OPC UA in 5G Networks* “
[9] 5G-ACIA white paper “*NPNs for Industrial Scenarios”*
[10] 5G-ACIA white paper “*Industrial 5G Devices – Architecture and Capabilities”*
[11] 5G-ACIA white paper *“Industrial 5G Edge Computing – Use Cases, Architecture and Deployment”*

5G-ACIA White Paper
Architecture and Technology for Machine Vision
in Industrial Applications with 5G

Contact
5G-ACIA
Amelia-Mary-Earhart-Str. 12
60549 Frankfurt am Main
Germany
Phone: +49 69 6302-209
Email: info@5g-acia.org
www.5g-acia.org

Published by
ZVEI – German Electro and Digital Industry Association,
5G-ACIA – 5G Alliance for Connected Industries and
Automation, a Working Party of ZVEI
www.zvei.org

Published in March 2025

Design: COBRAND

9 5G-ACIA Members

As of Februrary 2025

- AALBORG UNIVERSITY DENMARK
- ABB
- Accelleran
- AILINK
- Airgain
- ANALOG DEVICES
- ASE
- ASKEY
- ASOCS
- ATAYA
- ATHONET
- BAYFU
- BECKHOFF
- BOSCH
- brown-iposs
- CAMPUS GENIUS
- Capgemini engineering
- cellnex
- celona
- 中国移动 China Mobile
- CTONE
- DEMAG
- Deutsche Messe
- dfki Deutsches Forschungszentrum für Künstliche Intelligenz German Research Center for Artificial Intelligence
- NTT docomo
- Druid
- EMERSON
- Endress+Hauser
- ERICSSON
- equinor
- FAU Friedrich-Alexander-Universität Erlangen-Nürnberg
- FESTO
- firecell
- Fraunhofer
- GHMT
- GXC
- HIRSCHMANN A BELDEN BRAND
- HMS
- HUAWEI
- IDLab INTERNET & DATA LAB
- ifak
- ifm
- infineon
- inIT
- ITRI Industrial Technology Research Institute
- KETI
- KEYSIGHT
- kontron
- MITSUBISHI ELECTRIC
- MOXA
- MUGLER TELCO NETWORKS
- NOKIA
- NXP
- OMRON
- orange
- PEPPERL+FUCHS
- PHOENIX CONTACT
- Qualcomm
- RELIABLE REAL-TIME RADIO
- Radisys
- Rockwell Automation
- ROHDE&SCHWARZ
- SAL SILICON AUSTRIA LABS
- salzburgresearch
- Schneider Electric
- SIEMENS
- SINTEF
- SoftBank
- SONY
- SPIE
- ST life.augmented
- T
- telent
- TNO
- TRUMPF
- 財團法人電信技術中心 TELECOM TECHNOLOGY CENTER
- u-blox
- UNIBERG
- University of Sheffield AMRC North West
- UPLINK
- verizon
- VOLKSWAGEN AKTIENGESELLSCHAFT
- WAGO
- Weidmüller
- YOKOGAWA

5G-ACIA.org