Release Letter

| Product:            | Intelligent Video Analytics  
|                     | Essential Video Analytics |
| Version:            | 6.40                        |

1. General

Intelligent Video Analytics 6.40 and Essential Video Analytics 6.40 are the successors of Intelligent Video Analytics 6.30 and Essential Video Analytics 6.30, respectively.

Use Essential Video Analytics for business relevant applications including intruder detection indoors and in smaller, controlled outdoor environments, and Intelligent Video Analytics for mission-critical outdoor intrusion detection even in extreme weather conditions. Both can be used for people counting applications, where Intelligent Video Analytics can cope with twice as many people in the scene.

Intelligent Video Analytics is a continuously growing product, whereas the development of Essential Video Analytics is restricted to user interface improvements.

Bosch IP cameras are grouped by their common product platform (CPP) generation. Intelligent Video Analytics 6.40 is available on Bosch IP cameras of CPP4, CPP6, provided these cameras have an additional dedicated hardware accelerator unit (FPGA), and on all CPP7 and CPP 7.3 7000-9000 series cameras. Best performance, doubled detection range, ship tracking and compensation of shaking cameras are only available on CPP6, CPP7 & CPP7.3. Essential Video Analytics is available on all CPP7 and CPP7.3 4000-6000 series cameras.
2. New Features

With 6.40, cameras of platform CPP7.3 are supported. Intelligent Video Analytics is now available on all CPP7.3 cameras of range 7000-9000, and Essential Video Analytics is now available on all CPP7.3 cameras of range 4000-6000.

New features of Intelligent Video Analytics (CPP6/7) 6.40: on AUTODOME / MIC 7000-9000 of CPP7.3:
- Improved Intelligent Tracking. The newest generation of Intelligent Tracking ensures smoother camera motion for more comfortable viewing as well as more reliably tracking objects even under challenging scenes. The object is tracked reliably even if it stops moving, or when it moves through small groups of other moving objects. Geolocation of tracked objects is provided if the camera is calibrated and assigned a location in the corresponding coordinate system.
- Detect “object in field” even during camera motion. Outside of prepositions and even while the PTZ camera is moving, whether on guard tours or manually, Intelligent Video Analytics can now detect and alarm on “object in field”. These alarm fields are defined globally across all potential PTZ views, and any part of those global fields currently within the cameras view becomes active. Thus, guard tours sweeping across the monitored areas in order to provide an intruder less time to slip by can now automatically alarm on those intruders as well. Geolocation of the tracked objects is now provided if the camera is calibrated and assigned a location in the corresponding coordinate system.
  - Person height must be ≥ 12% of the image height (cars respectively depending on their aspect ratio), and all objects must move at least 1 VCA pixel per VCA frame to be detected. Camera pan / tilt speed ≤ 5° per second and similarly slow zoom recommended.

New features of both Intelligent Video Analytics 6.40 and Essential Video Analytics 6.40 on all supported platforms:
- Calibration simplification and rework:
  - Calibration has been completely moved to Camera -> Installer Menu -> Positioning in the device web page, and to Main -> Positioning in the Configuration Manager
  - Infographics explaining the calibration and geolocation configuration values have been added.
  - For DINION / FLEXIDOMES, calibration is now independent of the VCA profile. When upgrading from a previous firmware, the first calibration found in either VCA profile 1 or 2 will be taken. The assumption is that both calibrations should be the same, as the camera can only look at a single field of view. When downgrading, any available calibration will be put in both VCA profiles. Thus, performance changes when upgrading / downgrading are possible in theory, as standard tracking can behave a bit differently if calibration is available, but should not occur in practice.
  - For DINION / FLEXIDOMES, automatic setting of internally known calibration values is available via a new button “AutoSet”. In the device web page, this will open a light box showing and live-updating all internally known values, from where the values can be set into the calibration. In the Configuration Manager, the values will directly be put into the calibration.
  - For AUTODOME / MICs, it is now possible to calibrate the whole camera by just entering the height the camera is installed at. The assumption is that the AUTODOME / MICs are always...
perfectly “ceiling mounted”, and any derivation cannot be modeled in the global calibration. Once this global calibration is set, all preposition-related VCA profile calibrations are automatically computed via the internal knowledge of pan, tilt and zoom values. Furthermore, for the preposition-related VCA profiles, a dedicated calibration overwriting the global calibration can be defined, to cope with different ground planes at different prepositions, non-horizontal ground planes, and potential inaccuracies at high zoom values. Thus, upgrading and downgrading will not change calibration at preposition-related VCA profiles.

- Drawing the calibration elements “vertical line”, “ground line” and “ground angle” for unknown calibration values is possible via “Sketch”, which is a button in the device web page and a tab in the Configuration Manager for all non-panoramic cameras. Instead of “fix”ing the values that should not be changed by this calculation, a “calculate” check box needs to be selected for values that should be calculated with this tool. For AUTODOME / MICS, this is possible for preposition-related VCA profiles, and the values from the global calibration act as default settings. If no VCA profile has been assigned to a preposition, then the button “Sketch” is inactive.

- Limitations:
  - No calibration support for AUTODOME / MIC inverted / canted modes
  - When using older configuration manager, VideoSDK or MPEG Active X versions, the configuration will say that the camera is not calibrated. Please update the corresponding software for a correct configuration.

- Geolocation on AUTODOME/MIC prepositions has been corrected to have the proper azimuth alignment
- The object bounding box visualization now includes a cross at its center.
- Device Web Page “Live” can now show the video analytics rule elements (fields, lines, routes) and their alarm status.
- JPEGS can now optionally include the information about alarm objects for
  - MJPEG streams (since FW 6.30)
  - Remote snapshot requests: http://192.168.0.1/snap.jpg?VCAOverlay=1 (since FW 6.30)
  - Alarm email snapshots (since FW 6.32)
- ONVIF support: Added object class and object shape polygon to the ONVIF metadata output.
Introduction

Intelligent Video Analytics and Essential Video Analytics are software algorithms that detects behavior of objects within an environment monitored by a video camera and generates alarm events that can be processed further in a CCTV system. They make it possible to capture and evaluate directional movement of objects, apply configured filter rules and to combine these rules, thereby largely preventing false alarms. Both algorithms adapts automatically to changing environmental conditions and are therefore non-sensitive to perturbing influences such as rain and small tree movements.

Intelligent Video Analytics is more robust than Essential Analytics and can cope with heavy rain and snow falls, strong wind moving vegetation or shaking the camera, and water surfaces and fountains in the background.

Both algorithms provide easy to use, intuitive configuration via graphical user interface as a part of an advanced wizard structure in the Task Manager. Improved with an intuitive configuration option it is possible to provide the complete property information (object type, size, speed, aspect ratio, direction, color) for an object just by clicking it in the live scene.

Intelligent Video Analytics has the auxiliary function “Flow”, which supports optical flow to detect the direction of objects independent from the structure of the background and the size of the objects. It also has the ability to detect frontal faces and upload the best face snapshot to an ftp server or Dropbox account.

The metadata generated by Intelligent Video Analytics and Essential Video Analytics, transmitted to live video or to storage, is used to display overlay graphics and allow retrospective forensic search. Bosch Video Client (BVC) and Bosch Video Management System (BVMS) support full forensic search. The camera web page, the Video Security Client (VSC) and the Video Security Apps support a limited forensic search called smart search.
3. Features

General
- Robust object detection and tracking for indoor and outdoor usage in sparsely populated scenes.
- People tracking and counting in well-populated scenes. Intelligent Video Analytics can cope with twice the amount of people than Essential Video Analytics.
- Built-in tamper monitoring detects camera hoooding/masking, blinding, defocusing, and repositioning. An indication is shown if the reference image check works and when the alarm will be triggered.
- Two different VCA configurations (profiles) per DINION / FLEXIDOME allow for different settings of different conditions (like day and night)
- 16 different VCA configurations (profiles) for AUTODOME / MIC permit to combine a dome scene position with a certain VCA configuration (profile)
- Intelligent Video Analytics also includes frontal face detection. The faces can be searched for using simple search, or snapshots of the best face image can automatically be uploaded onto an FTP or Dropbox account.

The following object tracking modes can be selected:
- Standard tracking, e.g. for intrusion detection
- 3D tracking, e.g. for enhanced distances with Intelligent Video Analytics
- 3D people tracking, e.g. for indoor people counting
- Museum mode, e.g. for asset protection – don’t touch!
- Ship tracking (with Intelligent Video Analytics on CPP6/7 only)

The following tasks can be selected:
- Detect objects within, entering, or leaving an area
- Detect multiple line crossing from single line up to three lines combined in a logical row
- Detect objects traversing a route
- Detect loitering in an area related to radius and time
- Detect objects which are idle for a predefined time span
- Detect removed objects
- Detect objects who’s properties such as size, speed, direction, and aspect ratio change within a configured time span according to specification (for example something falling down)
- Count objects crossing a virtual line
- Count objects within an area and alarm if a predefined limit is reached.
- Detect a certain crowd level in a predefined field

The following filters can be applied to alarm objects:
- Object class (upright persons, bikes, cars and trucks)
- Object size
- Object speed
- Object direction (two different directions can be defined)
- Object aspect ratio
The following tasks can be selected for Intelligent Video Analytics “Flow”:

- Object color

- Detecting flows in a certain area and if needed filter for the main direction speed and activity of the moving objects:
  - The user may define up to two flow directions
  - In counter flow mode the user does not have to define a main flow direction. Instead, the algorithm automatically detects the direction of the main flow and triggers an alarm if objects move against this direction

- Detect a certain crowd level in a predefined field

**Easy configuration:**

- Wizard structure of task manager guides through the setup
- Broad range of predefined detection tasks available
- Up to eight independent tasks for alarm generation can be created per channel
- Filters for object type, size, speed, two-way direction, aspect ratio and color are available to create more specific detection rules for every task
- All spatial information like detector lines, detector fields, sensitive area, configured object size, object aspect ratio, object direction and more are graphically drawn into the scene and can be manipulated there for flexible and easy configuration
- Enriched with intuitive “Click-object-in-scene” configuration of object filters
- Graphical statistics for more transparency regarding alarm results
- Enhanced camera calibration with a choice of three calibration modes and direct feedback
- Task scripting offers the possibility to combine tasks and thus hugely extend the predefined tasks

**Features which need calibration:**

- Object filter for size and speed in metric or imperial system
- Object filter by type “upright person, bike, car, truck
- 3D tracking, which tracks objects on the ground plane
- 3D people tracking, which interprets everything as person and tracks these on the ground plane.

**Forensic Search** allows to define any task / object filter computation as well as the task combinations available via the task script language.

**Smart Search supports the following search tasks:**

- Any object
- Line crossing
- Object in field
- Any face
Differences between Intelligent Video Analytics and Essential Video Analytics

Essential Video Analytics currently has the full feature set of Intelligent Video Analytics with the following exceptions:

- No ship tracking
- No frontal face detection
- No “Flow” algorithm
- Less robustness for extreme weather conditions
- No compensation for shaking cameras
- No double resolution / detection distance
4. Specific Explanations and Limitations

General
- Intelligent Video Analytics is running on a separate hardware accelerator (FPGA) which means it is running totally independent from the H.264 encoding or other functions.
- Performance of Intelligent Video Analytics on CPP6/7 cameras will be better than on CPP4 due to a different processing unit with more processing power.
- You have to choose between Intelligent Video Analytics and Intelligent Video Analytics Flow. It is not possible to run both functions at the same time.
- Intelligent Video Analytics and Essential Video Analytics consist of two parts: The generation of metadata describing the detected objects, which will be stored together with the video, and live alarming or forensic search based on this metadata. In the configuration, all options in the tab “metadata generation” change the generated metadata. Therefore these options are not available for forensic search later on. All alarming as defined in “Tasks” as well as the inspection of metadata can be done both for live alarming and during forensic search.
- Only if Intelligent Video Analytics and Essential Video Analytics are active can they create metadata for live viewing, alarming and for storage. Forensic search therefore just can deliver results in a retrospective search for moving objects in the database for cameras that have had one of these algorithms enabled for the recording.
- The forensic search function for the feature “crowd detection” is only working on the crowd fields you defined before in the camera or in the encoder. Metadata for crowd will only be generated in this crowd fields and cannot be changed for a different crowd fields in Forensic Search.

Limitations of Intelligent Video Analytics and Essential Video Analytics object tracking:
- Due to reflections, objects or motion might not be reliably detected or too many objects or motions might be detected. False alarms might occur due to:
  - reflective background
  - glass (glazed building frontages)
  - water as a background
  - cones of light moving in the dark
- Sudden appearance of spotlights, moving headlights or torch cones etc. are lightening up an area that might be detected as an “object”.
- Large areas of reflected light can also cause spurious motion detection. However, light reflections caused by falling raindrops, for example, are small enough to be ignored for statistical purposes and owing to the uniform nature of their motion.
- Motion of vegetation due to wind is covered for slow, continuous and uniform wind. If this movement overlaps with objects, false as well as missed detections are possible. To avoid this, adjusting the camera position.
- Strong wind, storms and heavy peak blasts from different directions, especially in the foreground of a scene, might trigger false alarms.
- Suddenly appearing sharp shadows of clouds, trees and buildings can be mistake as objects. Soft shadows are covered by the algorithm.
An object in strong sunlight with crisp shadow may therefore be registered within its outlines including this shadow. This has to be taken into account for aspect ratio and object size configuration. Soft shadows are covered by the algorithm.

A constant background is necessary in order to detect motion reliably and to assign that motion to a certain object. The more the background moves, the harder it is to distinguish moving objects from it. For instance, a person walking in front of a hedge that is moving heavily in the wind will very probably not be detected properly.

Merging effects (outline bubble over more than one object) might occur with objects overlapping or passing-by close to each other. This means the occurrence of a new (bigger) object in the scene and the loss of the former already detected and tracked object IDs with all the consequences to the selected detection tasks. The same applies when these object separate again. To avoid this, please check your scene and camera position to the best and take it into account when setting up the configuration.

If the image consists to a certain extent of nothing but moving objects – in other words if objects cannot be distinguished from each other or from the background – the motion of an individual object cannot be detected (e.g. individuals in a large crowd, idle object in a large crowd).

The detection and analysis of objects entering the image will be delayed until significant size and motion have been observed. To avoid this, center all evaluations in the image. With Intelligent Video Analytics on CPP6/7, the sensitivity parameter allows you to additionally choose a trade-off between fast object detection and less false alerts due to insignificant motion.

“Click-object-in-scene” configuration: The quality of the metric results provided (size, speed, aspect ratio) is very much depending on the correct calibration. Furthermore it has to be mentioned that the color filter used in the “click-object-in-scene” function is of course related to the outlined area of an object. In most of the cases, this outlines include additional surroundings like background (e.g.: asphalt). To concentrate on the real object specification, it is recommended to delete these unwanted colors from the histogram using the “Clear” button.

**Minimal object size & differences in processing resolution**

Intelligent Video Analytics and Essential Video Analytics typically use less resolution than available, and different processing resolutions on different devices and for different image aspect ratios. Here the processing resolutions for different video aspect ratios

- **Essential Video Analytics**
  - 4:3 - 320x240
  - 16:9 - 320x180

- **Intelligent Video Analytics (CPP6/7), 3D tracking on, noise suppression OFF / MEDIUM, for moving / started / stopped objects**
  - 1:1 - 640x640
  - 4:3 - 640x480
  - 16:9 - 640x360

- **Intelligent Video Analytics (CPP6/7), 3D tracking off or noise suppression STRONG or placed/taken objects**
  - 1:1 - 320x320
  - 4:3 - 320x240
  - 16:9 - 320x180
Intelligent Video Analytics (CPP4)
  o 5:4 - 255x204
  o 4:3 - 240x180
  o 16:9 - 320x180

In corridor mode, the height and width are switched. Intelligent Video Analytics and Essential Video Analytics can reliably detect objects that have at least 20 square pixel in this internal resolution, e.g. 3x8 pixel for an upright person.

Limitations of automatic calibration:
  • Calibration can only be done for a single, flat, horizontal ground plane.
  • For MIC and AUTODOMES, a proper vertical mount is assumed, otherwise the global calibration will not work.
  • The more horizontally the camera is looking, the more accurate calibration needs to be, and automatic calibration may not be accurate enough.
  • The larger the focal length and covered distance, the more accurate calibration needs to be, and automatic calibration may not be accurate enough.

Limitations of automatic classification of object type:
  • Camera must be calibrated and 3D tracking must be active
  • No differentiation of crawling or rolling persons from animals. Only upright walking or standing persons are classified as persons.
  • Persons and bikes seen from the front are easily confused. A bike from the front will only be classified as such if it is fast enough, otherwise it will be classified as a person.
  • No differentiation between bicycle and motorbike.
  • Small objects with only a few pixel can be confused (for example, objects far away from the camera).
  • Object class may change over time.
  • All objects start as an unknown object. They are only classified over the time if the object class can be determined sufficiently reliable.

Limitations of color configuration:
  • An object is almost never displayed in a consistent color in the image data. Pixels on the outer edge of a detected object in particular often contain the color information of the background and not the object. Objects such as automobiles comprise a variety of parts (body, windows, tires). Each individual part of the object is displayed in a different color, for example the mudguards in red and the tires in black.
  • The color properties of an object depend upon the lighting conditions. If the lighting conditions in a captured image changes, then the captured color of the object also changes. Objects on a street appear in different hues depending on the time of day and weather conditions.
  • An object that changes its position or direction of motion may then appear with different color properties. For example, automobiles are often marked on the side in color but not on the back. When people are seen from the front, the hue of the face determines the color impression; however, if the person turns around, the color properties are then defined by the hair or headdress.
- Having a lot of different colors of the spectrum selected and low precision set nearly the entire color spectrum is selected to be detected. This means nearly all colors might trigger unwanted alarms. It is suggested to be more selective and/or precise in these cases.
- To define a specific color nuance, use the slider for saturation to select from more colors out of the spectrum. Up to five colors can be selected for one object. The importance of the colors in the search is from left to right: 25%; 20%; 15%; 10%; 5%. The reason to start with 25% is that objects normally consist of several colors, e.g. a car out of windshields (white or mirror effect), tires (black or dark grey), bumpers (black or dark grey) and finally the car paint we are actually looking for. The pure car paint might cover just 25% of the object therefore the algorithm starts with 25% as for the main color to guarantee realistic search results.
- With the slider precision you define the accuracy of the color match in alarm detection. With the slider being at the far left side (meaning “0”) the selected color is not detected at all. With the slider at the far right (“full”), the color has to match exactly to be detected. Be aware that this “selected range of accuracy” is shared across all several selected colors. This means: one color could take all range of accuracy for itself and the other colors hence have to match exactly or all of the colors share less accuracy for each of them.

Limitations of Intelligent Video Analytics Flow:
- Minimum object size for a reliable detection in Flow must be at least eight blocks, each with a size of 8 x 8 (QCIF-pixel).
- Objects too fast which cross the camera view in less than two seconds, might not be detected by Flow.
- Flow is only capable to detect objects that have an almost constant velocity and linear movement.
- Heavy camera movement like camera pans and camera vibration can cause false alarms and leads to objects not being detected with Flow.
- The higher the sensitivity value, the more tolerant is Flow in detecting objects that are temporarily covered.
- The range of the main flow must at least have 45° to detect even objects that move slowly.

Limitations of “3D tracking”:
- To activate “3D tracking”, a calibration of the scene is required. If the scene is not well calibrated the tracking may deliver wrong or no results. Both tracking modes assume that all objects move on a single flat ground plane. Multiple floors, staircases, vertical object motion etc. may lead to wrong results.
- “3D tracking” requires a camera height of more than 2.5 meter (above 3 meter recommended) to work robustly. By using more elaborate algorithms and prior real-world knowledge, the 3D tracking mode can improve the object tracking robustness.
- In “3D tracking” mode, objects that are located entirely above the horizon (e.g. flying objects) are not detected, since object motion here is restricted to the ground plane.

Limitations of 3D people tracking:
- For a top-down perspective: Heights of the camera more than 3 meter (recommended 4 meter). Lens to be chosen that head diameter of heads of persons is optimally between 7%
and 14% of the screen width and 8% and 16% of the screen heights.

- Calibration of the scene is required, and tracking should be set to “3D people tracking”.
- Other moving objects, reflections on the ground, blinking lights, changing light conditions, shadows, trolleys or persons carrying bags or umbrellas can lead to wrong counting results.
- Children close to other persons may not be detected.
- The algorithm is not suitable to count persons in crowded areas, or to count persons that stand still for a very long time.
- The amount of people that can be tracked in real time is
  - ~ 10 for Intelligent Video Analytics on CPP4
  - ~ 20 for Intelligent Video Analytics on CPP6/7
  - ~ 10 for Essential Video Analytics

If more than this amount of people is in the scene, there will occur more and more frames without metadata. The tracking will continue correctly nonetheless as long as possible.

- Line counting at the edge of the field of view may not work.
- 3D people tracking will not work in low lighting conditions.

**Limitations of counting:**

- In addition to the limitations of 3D people tracking, partly concealed objects, objects which cannot be tracked well, e.g. because of speed, background conditions or size will lead to wrong counting results.
- The counter can be reset by reloading the configuration or via a RCP+ command.

**Limitations of water vehicle tracking:**

- Only available in Intelligent Video Analytics on CPP6/7.
- The tracking mode “water vehicle tracking” needs to be enabled.
- This tracking mode should not be used to track people at a shore.
- Applications like a lock where water rushes in white foam into the lock once the doors are opened are not supported.
- False alerts may occur due to high waves.
- Two boats following each other closely will be detected as one boat only, as the wave of the first boat and thus the second boat as well will be suppressed.

**Limitations on panoramic cameras:**

- The algorithms run on the warped image only, though the results can be transformed into dewarped image views.
- The calibration wizards are not available as they are based on straight lines not available in the warped image.

**Limitations of geolocation:**

- The camera needs to be calibrated and given a geolocation itself. Tracking is only possible on a single ground plane.
- Performance will be better if 3D tracking mode is enabled.
- A viewer to show the geolocations on a map is needed.
- On CPP4 AUTODOME / MICs, no geolocation is available during Intelligent Tracking.
Limitations of idle/removed object detection:
- Placed / taken object detection is most robust if the object placed or taken is much smaller than the object which handles it.
- If a person places or removes a bike, the bike can be detected as placed / taken or as started / stopped as the bike is similar in size to the person. Therefore check for all idle / removed objects if this is of interest.
- Cars should always be detected as started / stopped objects as they are much larger then persons entering or leaving them.
- With Intelligent Video Analytics on CPP6/7, in regions with much background movement only stopped objects will be detected. Activation of the detection all other idle / removed object types in these regions is only possible by deactivating the noise suppression.

Limitations of face detection:
- Only frontal faces with deviation of up to 20% from frontal can be detected
- On CPP4, faces will not be detected in 3D people tracking
5. System Requirements

- Configuration Manager 5.50 or newer
- Web Browsers:
  - Microsoft Internet Explorer 11.0 or higher
  - Mozilla Firefox using MPEG-ActiveX 6.10 or newer

Additional information:
- The software functionality is part of the firmware release 6.40 and higher, for the products listed to support it.
- All installed Intelligent Video Analytics and Essential Video Analytics 6.30 licenses are upgraded automatically when the device is upgraded to firmware version 6.40.
- Intelligent Video Analytics 6.40 is part of the following cameras with firmware release 6.40 or higher
  - every 7000-9000 series CPP7 and CPP7.3 camera
  - every IVA CPP6 camera
  - every IVA Enabled CPP4 camera
  - every IVA Ready CPP4 camera with IVA license enabled
- Essential Video Analytics 6.40 is part of every 4000-6000 series CPP7 and CPP 7.3 camera with FW 6.40 or higher.
- Set-up of Intelligent Video Analytics and Essential Video Analytics 6.40 is achieved using either the camera/encoders web page setup or the configuration manager, which is available for download via the Bosch Website.
- Intelligent Video Analytics and Essential Video Analytics consume CPU power – please check chapter 4 (Specific Explanations and Limitations) for details.
- Forensic Search is a system feature of Bosch Video Client (BVC) and Bosch Video Management System (BVMS).