BriefCam

TRANSFORMING VIDEO SURVEILLANCE INTO ACTIONABLE INTELLIGENCE

# Video Characteristics for Best Video Results


July 2021

# Contents

# Introduction

The BriefCam processing engine scans the video using advanced computer vision techniques and algorithms, essentially extracting moving objects, separating the background, and storing the objects and background samples. BriefCam's VIDEO SYNOPSIS® is subsequently generated using the artifacts stored in the database.

Some videos are better suited for this type of processing than others and camera deployment can be optimized for more effective use of the product. This document describes the characteristics of a video that would lead to the best possible quality of video analytics.

# Camera Characteristics

There are several camera characteristics that can affect the performance of the video analytics engine.

## Camera Types

The type of cameras that are installed affect the performance of the object extraction, classification, and face recognition.

As a general note, it is recommended that color cameras be used wherever possible.

### Fisheye Cameras (Distorting Lenses)

Fisheye cameras, which provide a 360, 180 and even 120 degree view of the scene, produce lower quality results than regular videos (some objects may seem upside-down, some will be distorted). This is especially true when calculating proximities, size, and speed as the geometry of the scene will be heavily distorted and will not be extracted by BriefCam's algorithms – making proximity information unavailable.

### Thermal and Infrared Cameras

Thermal and infrared (IR) cameras, which are useful in poorly lit areas or at night, can detect people, but with lower quality results than with "regular" cameras, especially when extracting features such as detailed classification (e.g. trucks, vans), person attributes including gender and, of course, color extraction.

Both thermal and infrared cameras might not detect vehicles well. For thermal videos, this is because the vehicles' heat-radiating areas might cause them to appear differently than vehicles in "normal" videos. For infrared video (which implies that the video was shot at night), detection of vehicles at night is problematic because the vehicles' headlights might dazzle the camera and make detection and classification harder. This might also be the case for other objects that reflect infrared lights – as infrared video is often shot using IR active illumination.

In thermal videos, other filters, such as Person Attributes, will not operate well because the image of a person will not have any details. In infrared videos, some Person Attributes might work, for example, a handheld bag will still look like a handheld bag.

In both infrared and thermal videos, Color filters will not work because the image is usually grayscale.

### Surveillance Cameras

Some surveillance (CCTV) cameras produce smaller and low quality face images, resulting in poorer results for face recognition.

### PTZ Cameras

A PTZ camera, which constantly moves or scans the scene, would provide a video that is unusable for BriefCam, due to the inability to identify the background which is changing constantly. However, the video from a PTZ camera that is stationary most of the time, but occasionally move to another position, can be processed. Note that processing results of the video footage recorded during the position change are not usable, but the processing recuperates when the camera stabilizes in its new position (after 5 minutes).

## Camera Placement

CCTV cameras are usually placed at a high vantage point (near the ceiling or on a pole) in order to gain an overview of the scene. This allows to detect and track objects in the scene with minimum occlusions.

### Optimal Proximity Accuracy

For optimal proximity accuracy, it is recommended to place the camera so that there is a view of a single ground plane. In order for the proximity functionality to work, the camera should be pointing 10-80 degrees downwards (not horizontal or vertical) with an optimal angle of between 30-45 degrees. The camera should be placed 2.5.-3 meters/8.2-9.8 feet above the ground. If the camera is indoors, it should be placed close to the ceiling. In addition, a fisheye camera (or other distorting lenses) should not be used and the resolution needs to be good enough for human pose detection.

### Optimal Face Recognition Accuracy

For face recognition, a good recommendation when using BriefCam is to place cameras in all entrances and exits, and to position the cameras in such a way that the faces are seen clearly and as close to eye-level as possible with a vertical angle of approximately 30 degrees or less. In addition, it's best to have the scene set up, when possible, where the people are walking in and out individually and not in groups, such as through a turnstile.

In general, face recognition produces the best results when:

- The cameras are positioned at eye level.

- The lighting lights the front of objects sufficiently, resulting in a quick shutter speed that produces a crisp image.
- The lighting gives good contrast and the faces are not lit from behind.
- The focus of the camera is on the area where you expect faces to appear.
- The camera is steady to prevent image smearing.

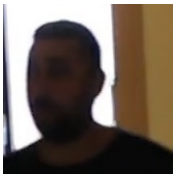
## Optimal Face Mask Detection Accuracy

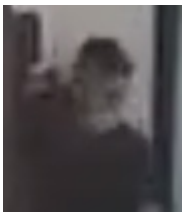
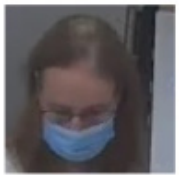
Like in Face Recognition, the best results for face mask detection are achieved when the camera is at eye-level, at a shallow angle, and with no occlusions and good lighting.

Face mask detection requires a captured face quality of at least 1-star. This also means that the face resolution should be more than 24x24 pixels across the face.

Accuracy should be above 90% in scenes with good conditions (Good lighting and resolution. Small pitch and yaw of the faces).

### Examples of Various Face Mask Conditions

| | |
|---|---|
|  | bad lighting and resolution |
|  | bad angle (yaw) but good quality |
|  | bad image quality and angle |
|  | bad resolution, quality, and illumination (and a person with stubble) |

| | |
|---|---|
|  | non-face recognition quality |
|  | good quality (some tough angles) |
|  | |
|  | |

# Video Parameters

The following are some technical video parameters and the respective recommendations for effective video analytics.

## Frame Rates

Officially, BriefCam strongly recommends a minimum frame rate of 8 frames per second (FPS). A frame rate lower than 8 can also be used, but the quality of the detection, tracking, and classification is affected.

The maximum recommended frame rate is up to 15 frames per second. If a camera is configured for a higher number of frames per second, BriefCam's processing algorithm lowers the frame rate by skipping some of the frames. In the table below you can see how many frames BriefCam processes and how many frames are skipped.

| FPS | Processed Frames | Details |
|---|---|---|
| 6 | 6 | Video is processed as is, with performance degradation (i.e. reduced event detection rate and potential misses to be expected). |
| 15 | 15 | Video is processed as is, at the input frame rate (i.e. each individual frame is processed, as opposed to alternating frames). |
| 16 | 8 (16÷2=8) | BriefCam processes alternating frames, effectively reducing the number of frames to be processed by half. |
| 25 | 12 (25÷2=12.5) | |
| 60 | 15 (60÷2=30, then 30÷2=15) | |

## Detection Performance

Detection of objects is subject to the following criteria:

- **Minimum duration** – Any object spanning a time period longer than one second is detected (sometimes even 5 frames are enough to start detecting)
- **Minimum object size** – The minimum size of objects that can be detected depends on the contrast between specific objects and the background. Typical minimum object size is approximately 10 x 10 pixels (100 pixel area) for a 1MP video. Note that for videos of larger resolutions, the computer vision engine may process in a lesser resolution for better processing efficiency. Thus, an original video of 5MP processed at 1MP will allow to detect objects of approximately 22 x 22 pixels (500 pixel area) or larger (the minimum area size for detection is proportional to the original video resolution). In order to exploit the full resolution, and be able, for example, to detect 10 x 10 objects in a 5MP video, the processing resolution may be increased to 5MP, but more processing time is needed to process this high resolution video.
- **Minimum contrast** – The minimum contrast of objects that can be detected depends on the size of specific objects. Typical minimum object contrast is of approximately 15 grey levels.

# Resolution

The resolution needs to match the desired analytics goals and the camera's field of view. The minimum detection and classification resolutions need to be met, where the higher the resolution the better the analytics accuracy, particularly for face recognition where the more facial features that are seen, the better the accuracy. For example, an outdoor camera covering an entire intersection should be 4K to allow faces to be efficiently recognized.

Different resolutions between archive and real-time multi-streams can be different – as long as they both have the same aspect ratio.

Variable-resolution cameras that change resolutions cannot be used with BriefCam.

Videos that are recorded with a variable-FPS rate might cause the tracking and detection mechanisms to behave sub-optimally in cases of counting. This is because during BriefCam engine's 5-minute warmup, the engine determines the rate of FPS and uses this number to sample 3 frames per second for detection. For example, if 15 FPS is used, every 5th frame is used for detection. For variable-rate FPS, the number selected in the 5-minute warmup is probably incorrect for different parts of the movie, which leads to jittery movement as perceived by the algorithm and unpredictable detection and tracking accuracy. It is recommended to not use variable-rate FPS and to set the camera to fixed-framerate.

## Face Recognition Required Resolution

The minimum resolution required for successful face recognition is 24x24 pixels across the person's face or 12 pixels between the eyes.

You can use this size to calculate the distance at which BriefCam can perform face recognition by taking into consideration the lens FOV (field of view) angle and the number of pixels in the camera sensor.

The steps for calculating the distance are:

1. Calculate the width (or height) in meters of the FOV at a certain distance using the horizontal – or vertical FOV angle (and the tangent function).
2. See how many pixels there are per this width (or height) – by checking the sensor spec.
3. Calculate the width (and height) in meters per pixel (by dividing the value from step 1 by the value in step 2).

The table below is based on this minimum resolution limit and approximates the maximum operational distance of BriefCam's face recognition engine per different camera resolutions and horizontal field of view (in degrees).

**Face recognition - Approximate maximum distance**

| Resolution | HFOV*=40deg | HFOV=60deg | HFOV=90deg |
|---|---|---|---|
| 720p | 13.5 m/44.3 feet | 9.4 m/30.8 feet | 6.8 m/22.3 feet |
| 1080p | 20.2 m/66.3 feet | 14.1 m/46.2 feet | 10.2 m/33.5 feet |

*HFOV = horizontal field of view

BriefCam

www.BriefCam.com

*Visit us on social media*