

Detección de ataques de tos y voz estresada

Analítica de audio en algunas cámaras con micrófono integrado

Mayo 2024

Resumen

Detección de ataques de tos y voz estresada es una aplicación de analítica de audio que detecta incidentes escuchando lo que ocurre en las inmediaciones durante las 24 horas del día.

Esta aplicación viene preinstalada en algunas cámaras Axis con micrófono integrado. Incluye dos algoritmos de detección diferentes y puede usar ambos o solo uno.

El detector de tos detecta tosidos aislados o ataques de tos para alertar al personal cuando alguien necesita ayuda. El detector de voz estresada identifica patrones de sonido asociados a la coacción, el enfado o el miedo, por lo que resulta útil para frenar actividades delictivas, reducir los asaltos o saber cuándo alguien necesita ayuda.

La privacidad está garantizada porque *Detección de ataques de tos y voz estresada* no necesita guardar datos de audio para ofrecer los resultados esperados. El audio no se graba a menos que active expresamente la grabación de audio.

Es posible configurar varios ajustes para adaptar la analítica a sus necesidades concretas. *Detección de ataques de tos y voz estresada* revisa continuamente el estado de la aplicación para garantizar su correcto funcionamiento.

Índice

1	Introducción	4
2	Detector de tos y detector de voz estresada	4
3	Claves para una detección óptima del audio	5
4	Función multisensor	5
5	Superposiciones	6
6	Tipos de eventos y estado	6
7	Privacidad	7

1 Introducción

La analítica de detección de audio para cámaras es un complemento muy interesante a la videovigilancia. Facilita la detección temprana y el envío de alertas sobre posibles incidentes, a menudo antes de que se identifiquen en el vídeo.

Detección de ataques de tos y voz estresada es una aplicación de analítica de audio que escucha lo que ocurre en las inmediaciones durante las 24 horas del día para clasificar y filtrar los sonidos. Al detectar un ataque de tos o una voz estresada, la aplicación genera una alerta.

Este documento técnico presenta la aplicación *Coughing fit and stressed voice detection* y explica cómo configurarla para una detección óptima.

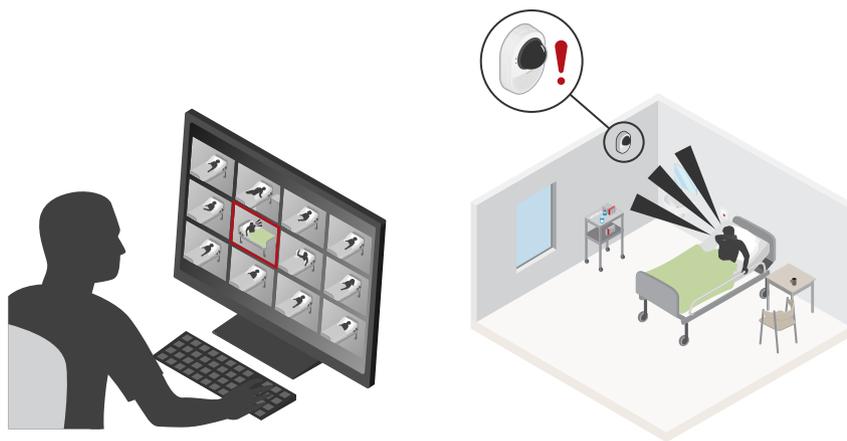
2 Detector de tos y detector de voz estresada

Coughing fit and stressed voice detection viene preinstalado en algunas cámaras Axis con micrófono integrado. Los detectores capturan indicadores audibles de incidentes en tiempo real, directamente en la cámara. Tiene la opción de usar solo un detector o ambos.

El detector de tos detecta los tosidos y los cuenta dentro de un intervalo de tiempo. Este sistema permite al personal responder enseguida cuando alguien tose o tiene un ataque de tos. Puede detectar un ataque de tos o un tosido aislado, según como lo configure.

Para evitar un gran número de notificaciones de evento en poco tiempo, se inicia un bloque de tiempo de 5 segundos nada más contabilizarse el primer tosido. Si la analítica está configurada en 3 tosidos cada 30 segundos, solo contará el siguiente tosido si transcurren por lo menos 5 segundos desde el anterior. Los tosidos que se produzcan dentro del bloque de 5 segundos no se contabilizarán. Por tanto, con estos ajustes solo se enviará una alerta una vez que se hayan contado 3 tosidos, con un mínimo de 5 segundos entre cada tosido.

El detector de voz estresada identifica patrones asociados a la coacción, el enfado o el miedo en la voz de una persona. Cuando reconoce un patrón, el sistema envía una notificación automática al personal con una alerta visual o mediante la activación de una alarma. Esta advertencia precoz permite responder con más rapidez: ayudar a una persona que lo necesite o frenar un enfrentamiento que podría terminar en una agresión física.



Uso de la detección de ataque de tos y voz estresada en un entorno sanitario.

3 Claves para una detección óptima del audio

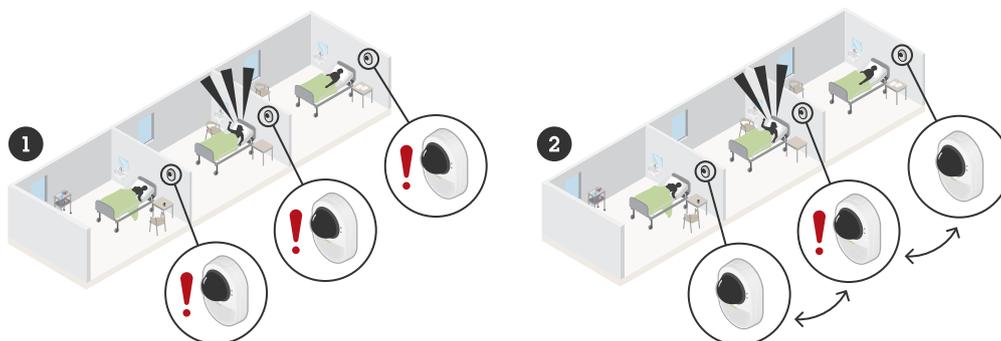
- **Colocación de la cámara.** La cámara con analítica debe estar a una distancia mínima de 1,5 metros de posibles interferencias, como sistemas de climatización, sistemas de megafonía o altavoces y puertas que producen ruido al cerrarse. Idealmente tiene que estar también en la línea de visión del sitio donde quiere detectar el audio. Aunque esta última condición no es imprescindible, sí es importante para una detección más precisa porque cuando el sonido tiene que superar esquinas u obstáculos puede afectar el resultado. Por ejemplo, no todas las frecuencias tienen el mismo grado de curvatura.
- **Sensibilidad.** Los ajustes de sensibilidad permiten personalizar el sistema de detección. Cuando la sensibilidad es más alta, aumentan las detecciones. Crece el riesgo de detecciones no deseadas (falsas alarmas), pero a veces puede ser necesario si hay que detectarlo absolutamente todo. Con una sensibilidad más baja, solo se notifican las detecciones cuando no hay dudas sobre la clasificación del sonido. Esta configuración aumenta el riesgo de no registrar incidentes potenciales, pero puede ser la mejor opción para evitar un exceso de falsas alarmas.
- **Modo de recopilación de datos.** Puede usar el modo de recopilación de datos durante un tiempo tras la instalación para hacerse una idea del tipo de sonidos detectados. Los resultados y su análisis pueden darle una información muy útil para saber cuál es el nivel de sensibilidad óptimo para esa instalación concreta.
- **Umbral de detección de tos.** Puede configurar el número de tosidos necesarios. La alarma se generará solo cuando se llegue al umbral configurado para el periodo de tiempo asignado.
- **Configuración avanzada.** La configuración avanzada está pensada solo para los usuarios expertos. Los cambios pueden generar detecciones incorrectas o una ausencia de detecciones. Sin embargo, en algunas situaciones concretas pueden ser necesarios. En este caso, hay que realizarlos con la ayuda de un experto en el sistema o siguiendo sus indicaciones.

4 Función multisensor

Cuando hay varias cámaras colocadas cerca, por ejemplos en habitaciones contiguas, varias cámaras pueden detectar el mismo incidente de audio y eso puede dificultar saber dónde se ha producido el incidente.

En la detección de voz estresada, la función multisensor resulta muy útil en estos casos. Cuando está activada y varias cámaras captan la misma voz estresada, solo generará una notificación la cámara que la

captó primero. De este modo, las cámaras cercanas suman esfuerzos para reducir los falsos eventos y las notificaciones duplicadas por el mismo evento.



- 1 Sin función multisensor: cámaras situadas en espacios contiguos detectan el mismo incidente de voz estresada y crean varias alarmas.
- 2 Con función multisensor: solo la cámara situada más cerca notifica la detección.

La función multisensor permite crear grupos con las cámaras situadas cerca que pueden captar el mismo audio. Hay algunas restricciones:

- Todas las cámaras del mismo grupo deben tener configurada la sincronización de hora con NTP.
- Todas las cámaras agrupadas deben usar la misma versión de *Detección de ataques de tos y voz estresada*.
- Todas las cámaras agrupadas deben estar conectadas a través de la red.

Si no se cumple alguna de las condiciones anteriores, la cámara funcionará en modo independiente y aparecerá con el estado deficiente.

5 Superposiciones

Es posible mostrar como superposición en el vídeo un espectrograma en tiempo real y notificaciones de la aplicación. Además, se puede personalizar su tamaño y también su posición. Un regulador permite ajustar la opacidad.

Las notificaciones de la aplicación muestran los eventos detectados por la cámara y el estado de la aplicación.

El espectrograma es una representación visual del audio. La posibilidad de oír el audio y ver al mismo tiempo su representación visual es muy interesante para calibrar más rápidamente la gravedad de un incidente.

6 Tipos de eventos y estado

Los eventos generados por el detector de tos y el detector de voz estresada *no tienen estado*. Son incidentes puntuales activados por una detección. Una vez superado el tiempo asignado al bloque del evento (cinco segundos, configurable), se genera un nuevo evento si se detecta algo.

El estado de *Detección de ataques de tos y voz estresada* se refleja mediante *eventos con estado*. Con los eventos con estado, dicho estado sigue activo mientras o la situación y cambia solo cuando se ha solucionado.

Las comprobaciones de estado están diseñadas para saber que todo funciona bien y alertar en caso de errores. Hay tres estados posibles para *Detección de ataques de tos y voz estresada*:

- Estado correcto: funcionamiento normal. Se pueden llevar a cabo detecciones.
- Estado deficiente: la aplicación funciona pero en un modo deficiente. En general, se debe a factores temporales, como la pérdida de una cámara asociada, cortes en el audio por sonidos con un volumen muy alto o capacidad insuficiente de la memoria de audio temporal. En este estado, es posible detectar incidentes, pero son más habituales las falsas alarmas o los incidentes no detectados. Suele resolverse solo.
- Funcionamiento incorrecto: el sistema no está operativo. No se pueden llevar a cabo detecciones. Por lo general, este estado se produce por factores que no se resuelven solos, por ejemplo, porque se ha desactivado el audio en los ajustes del dispositivo o se ha silenciado la ganancia de entrada de audio.

Los estados deficiente e incorrecto se muestran en el panel de información y también con una superposición de texto (si está activada esta opción) para que el operador sepa que la aplicación tiene un funcionamiento deficiente o incorrecto.

Cada 60 segundos se activa un evento de latido (si está activado, ajuste de tiempo configurable). Sirve para verificar que la analítica está funcionando y alertar en caso de que no se reciban latidos. Los eventos de latido no se envían en caso de funcionamiento incorrecto.

7 Privacidad

Los datos de audio se procesan y analizan en la cámara. No es necesario guardar ningún tipo de dato de audio para que *Detección de ataques de tos y voz estresada* funcione correctamente. Solo cuando se activa expresamente la grabación se graba el audio de un evento. Esta opción puede ser útil al investigar incidentes, analizar la causa de falsos positivos o para volver a escuchar incidentes en el caso de que el sistema de gestión de vídeo no ofrezca esta posibilidad.

Acerca de Axis Communications

Axis contribuye a crear un mundo más inteligente y seguro a través de soluciones para mejorar la seguridad y el rendimiento empresarial. Como empresa de tecnología de red y líder del sector, Axis ofrece soluciones de videovigilancia, control de acceso y sistemas de audio e intercomunicación. Se ven reforzadas por aplicaciones de análisis inteligentes y respaldadas por formación de alta calidad.

Axis tiene alrededor de 4000 empleados dedicados en más de 50 países y colabora con socios de integración de sistemas y tecnología en todo el mundo para ofrecer soluciones personalizadas. Axis se fundó en 1984 y la sede está en Lund, Suecia