

Détection des quintes de toux et des voix stressées

Fonction d'analyse audio sur certaines caméras à micro
intégré

Mai 2024

Avant-propos

La *détection des quintes de toux et des voix stressées* est une application d'analyse audio qui détecte les incidents par l'écoute permanente des bruits environnants.

L'application est livrée préinstallée sur certaines caméras Axis à micro intégré. Elle consiste en deux algorithmes de détection distincts, que vous pouvez choisir d'utiliser séparément ou simultanément.

Le détecteur de toux détecte les toux uniques et les quintes de toux pour que le personnel puisse réagir rapidement auprès des personnes qui en ont besoin. Le détecteur de voix stressée identifie les schémas sonores associés à la contrainte, à la colère ou à la peur, d'où son intérêt pour dissuader la criminalité, limiter les violences ou détecter les personnes en détresse.

La confidentialité est préservée, car la *détection des quintes de toux et des voix stressées* n'a pas besoin de stocker de données audio pour fonctionner correctement. Le son n'est pas enregistré, sauf si vous avez explicitement activé l'enregistrement audio.

Vous pouvez configurer plusieurs paramètres pour optimiser le comportement de cette fonction d'analyse dans votre scénario. La *détection des quintes de toux et des voix stressées* effectue des contrôles d'intégrité réguliers pour confirmer qu'elle fonctionne correctement.

Table des matières

1	Introduction	4
2	Détecteur de toux et détecteur de voix stressée	4
3	Optimisation de la détection audio	5
4	Sensibilité aux autres capteurs	5
5	Incrustations	6
6	Types d'événement et état d'intégrité	6
7	Confidentialité	7

1 Introduction

Les fonctions de détection audio des caméras sont de précieux alliés en vidéosurveillance. Elles permettent une détection et une alerte précoces en cas d'incident potentiel, parfois même avant la vidéo.

La *détection des quintes de toux et des voix stressées* est une application d'analyse audio qui écoute en permanence l'environnement pour classer et filtrer les sons. Lorsque l'application détecte une quinte de toux ou une voix stressée, elle génère une alerte.

Ce livre blanc présente l'application de *détection des quintes de toux et des voix stressées*, en expliquant comment la configurer pour optimiser la détection.

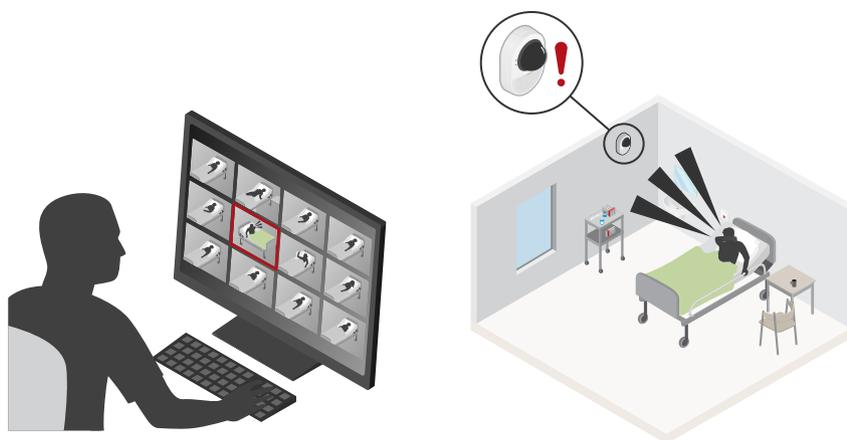
2 Détecteur de toux et détecteur de voix stressée

L'application de *détection des quintes de toux et des voix stressées* est livrée préinstallée sur certaines caméras Axis à micro intégré. Les détecteurs repèrent en temps réel les indicateurs audibles d'incident directement dans la caméra. Vous pouvez choisir d'utiliser un des détecteurs ou les deux.

Le détecteur de toux fonctionne en détectant les toux et en les comptabilisant dans un laps de temps donné. Il permet au personnel de réagir rapidement lorsqu'une personne tousse ou est prise d'une quinte de toux. Le détecteur peut détecter une quinte de toux ou une seule toux selon le paramétrage choisi.

Pour réduire le nombre de notifications dans un très court laps de temps, un bloc de temps de 5 secondes démarre dès que la première toux est comptabilisée. Si l'analyse est définie sur 3 toux en 30 secondes, elle ne comptabilisera la prochaine toux que si au moins 5 secondes se sont écoulées après la toux précédente. Les toux intermédiaires au sein du bloc de 5 secondes ne seront pas comptabilisées. Cela signifie qu'avec ces réglages, une alerte ne sera envoyée qu'après 3 toux espacées d'au moins 5 secondes chacune.

Le détecteur de voix stressée identifie les schémas sonores associés à la contrainte, à la colère ou à la peur dans la voix d'une personne. Dès la détection, le système envoie une notification automatique au personnel, sous la forme d'une alerte visuelle ou du déclenchement d'une alarme. Cet avertissement précoce permet au personnel de réagir rapidement. Il peut offrir son aide à une personne qui en a besoin ou éviter une escalade susceptible de conduire à une agression physique.



Détection des quintes de toux et des voix stressées, utilisée dans un environnement hospitalier.

3 Optimisation de la détection audio

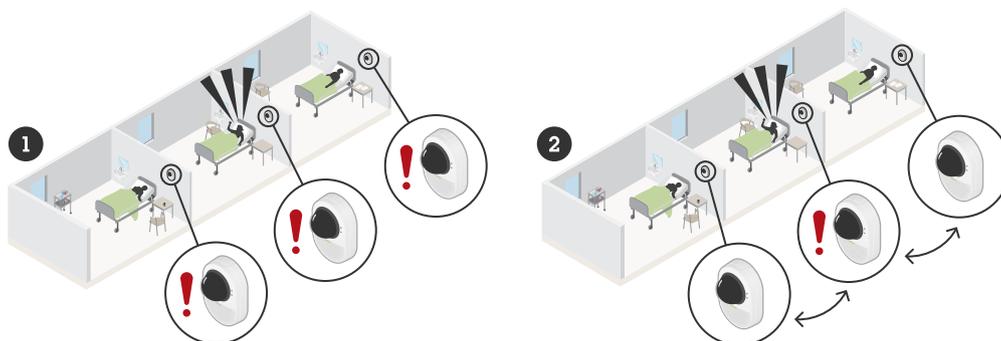
- **Positionnement des caméras.** La caméra qui dispose de la fonction d'analyse doit être placée à 1,5 m au moins des sources de bruits parasites, telles que climatiseurs, systèmes de sonorisation publique, haut-parleurs ou portes qui claquent. De préférence, la caméra doit aussi être dirigée sur la zone où détecter les bruits. Même si cette orientation n'est pas une exigence absolue, elle peut favoriser la précision de la détection. En effet, les bruits peuvent être altérés lorsqu'ils contournent des coins ou des obstacles. En particulier, les fréquences sonores ne se propagent pas toutes de la même façon.
- **Sensibilité.** Le système de détection peut être affiné avec les paramètres de sensibilité. Une sensibilité supérieure donne lieu à un plus grand nombre de détections. Elle accroît également le risque de détections indésirables (fausses alertes), qui peut être voulu lorsqu'il est essentiel de ne jamais manquer une détection. Avec une sensibilité plus basse, les détections sont signalées uniquement lorsque le bruit est correctement classé avec une certitude élevée. Un tel réglage accroît le risque de manquer des incidents potentiels, mais une sensibilité basse peut être nécessaire si elle évite de nombreuses fausses alertes.
- **Mode collecte de données.** Vous pouvez utiliser le mode collecte de données pendant un certain temps après l'installation pour recueillir des informations sur les types de bruit détectés. Les résultats et leur analyse peuvent fournir des précisions sur le niveau de sensibilité optimal pour l'installation en question.
- **Seuil de détection de toux.** Vous pouvez définir le seuil du nombre de toux nécessaires. Une alarme se déclenche alors uniquement lorsque le nombre de toux atteint le seuil pendant la période définie.
- **Paramètres avancés.** Les paramètres avancés sont réservés aux utilisateurs experts. Leur modification peut engendrer des détections incorrectes ou l'absence totale de détection. Néanmoins, dans certains scénarios, le changement de ces paramètres peut se révéler nécessaire. Il doit être réalisé uniquement après consultation d'un expert système.

4 Sensibilité aux autres capteurs

Lorsque les caméras sont installées à proximité les unes des autres, par exemple dans des salles adjacentes, elles peuvent toutes détecter le même incident audio. Cette redondance peut compliquer la localisation de l'incident.

Dans ces cas-là, la fonction de sensibilité aux autres capteurs peut être utile dans la détection des voix stressées. Lorsqu'elle est activée, si plusieurs caméras détectent la même voix stressée, la caméra qui l'a

détectée en premier sera celle qui déclenche une notification. De cette manière, les caméras à proximité interagissent pour réduire les faux positifs et éviter les notifications en double pour le même événement.



- 1 *Sensibilité aux autres capteurs désactivée : les caméras des salles adjacentes détectent le même incident de voix stressée et produisent plusieurs alarmes.*
- 2 *Lorsque la sensibilité aux autres capteurs est activée, seule la caméra la plus proche signale une détection.*

Lorsque la sensibilité aux autres capteurs est activée, vous créez des groupes de dispositifs associés, qui réunissent les caméras à portée de détection audio les unes des autres. Quelques restrictions sont à prendre en compte :

- Tous les dispositifs associés doivent être configurés pour utiliser la synchronisation de l'heure NTP.
- Tous les dispositifs associés doivent exécuter la même version de l'application de *détection des quintes de toux et des voix stressées*.
- Tous les dispositifs associés doivent pouvoir intercommuniquer sur le réseau.

Si l'une des conditions ci-dessus n'est pas satisfaite, le dispositif associé repasse en mode autonome et se place dans un état dégradé.

5 Incrustations

Un spectrogramme en temps réel et les notifications de l'application peuvent être incrustés par-dessus le flux vidéo. Vous pouvez personnaliser la taille des incrustations, les déplacer à la position souhaitée et régler leur opacité au moyen d'un curseur.

Les notifications de l'application affichent les événements détectés par la caméra et le statut de l'application.

Le spectrogramme donne une représentation visuelle du son. L'écoute du son et la vue de sa représentation visuelle en simultanée peuvent vous aider à déterminer rapidement la gravité d'un incident.

6 Types d'événement et état d'intégrité

Les événements générés par le détecteur de toux et le détecteur de voix stressée sont *sans état* (« stateless »). Ce sont des occurrences momentanées déclenchées par une détection. Une fois la durée du bloc événement écoulée (cinq secondes, configurable), une nouvelle détection génère un nouvel événement.

Le statut d'intégrité de la *détection des quintes de toux et des voix stressées* est accessible par l'utilisation d'*événements avec état*. Dans les événements avec état (« stateful »), l'état de l'événement reste actif tant que la condition est satisfaite et change uniquement lorsque la condition est résolue.

Des contrôles d'intégrité sont intégrés pour vérifier le fonctionnement correct de la fonction et avertir en cas d'écart. On peut distinguer trois états d'intégrité pour l'application de *détection des quintes de toux et des voix stressées* :

- État sain : fonctionnement normal. La détection est possible.
- État dégradé : fonctionnement en mode dégradé. Cet état est généralement causé par des facteurs temporaires, tels que la perte d'une caméra associée, un son haché en raison d'un volume très élevé ou la saturation de la mémoire tampon pour l'audio. Dans l'état dégradé, les détections sont possibles, mais la probabilité de faux positifs et de détections manquées peut être supérieure. L'état dégradé se résout habituellement de lui-même.
- État dysfonctionnel : absence de fonctionnement. Aucune détection n'est possible. Cet état est généralement causé par des facteurs qui ne disparaissent pas d'eux-mêmes, par exemple la prise en charge audio désactivée dans les paramètres du dispositif ou la mise en sourdine du gain d'entrée audio.

Les états dégradé et dysfonctionnel sont indiqués dans le panneau d'information et dans l'incrustation de texte (si elle est activée). L'opérateur est donc informé de la situation de l'application.

Un signal de pulsation se déclenche toutes les 60 secondes (lorsque la fonction est activée, délai configurable). Il peut servir côté réception à vérifier si la fonction d'analyse est opérationnelle et déclencher une alerte en cas d'absence de réception de pulsation. Les événements de pulsation ne sont pas envoyés lorsque l'état dysfonctionnel est actif.

7 Confidentialité

Comme les données audio sont traitées et analysées dans la caméra, il est inutile de les stocker pour que la *détection des quintes de toux et des voix stressées* fonctionne correctement. L'enregistrement du son pendant les événements n'est effectif que lorsqu'il est explicitement activé. L'enregistrement peut être utile pour mener les investigations après un incident, résoudre des problèmes de faux positifs ou réécouter des incidents si le système de gestion vidéo ne le permet pas.

À propos d'Axis Communications

En concevant des solutions qui améliorent la sécurité et les performances de l'entreprise, Axis crée un monde plus clairvoyant et plus sûr. En tant qu'entreprise de technologie de réseau et leader de l'industrie, Axis propose des solutions de vidéosurveillance, de contrôle d'accès, d'interphonie et de systèmes audio. Les performances de ces solutions sont améliorées grâce à des applications d'analyse intelligentes et une formation de haute qualité.

Axis emploie près de 4 000 personnes dans plus de 50 pays et collabore avec des partenaires technologiques et d'intégration de systèmes dans le monde entier pour fournir des solutions clients adaptées. Axis a été fondée en 1984 et le siège social se trouve à Lund, en Suède.