

Detecção de tosse e voz rouca

Análíticos de áudio em câmeras selecionadas com microfone integrado

Maio 2024

Resumo

A *Deteção de tosse e voz rouca* é um aplicativo de analíticos de áudio que detecta incidentes ouvindo o áudio ao redor 24 horas por dia, 7 dias por semana.

O aplicativo vem pré-instalado em câmeras Axis selecionadas com microfone integrado. Ele consiste em dois algoritmos de detecção separados e você pode optar por usar um deles ou ambos.

O detector de tosse detecta tosses únicas ou ataques de tosse, permitindo que o pessoal responda rapidamente às pessoas necessitadas. O detector de voz rouca identifica padrões sonoros associados à coação, raiva ou medo, tornando-o uma ferramenta eficaz para dissuasão de crimes, redução de agressões ou indicação de pessoas que precisam de ajuda.

A privacidade é protegida porque a *deteção de tosse e voz rouca* não precisa armazenar nenhum dado de áudio para funcionar corretamente. O áudio não é gravado, a menos que você tenha ativado explicitamente a gravação de áudio.

Você pode definir várias configurações para fazer com que os analíticos funcionem de maneira ideal para seu caso de uso. A *deteção de tosse e voz rouca* também realiza verificações contínuas de integridade para verificar o funcionamento adequado.

Sumário

1	Introdução	4
2	Detector de tosse e detector de voz rouca	4
3	Ativar detecção de áudio ideal	5
4	Reconhecimento multissensor	5
5	Sobreposições	6
6	Tipos de eventos e status de integridade	6
7	Privacidade	7

1 Introdução

Analíticos de detecção de áudio em uma câmera são um complemento poderoso para o videomonitoramento. Eles permitem a detecção precoce e alertas sobre possíveis incidentes, possivelmente antes de serem descobertos no vídeo.

A *Detecção de tosse e voz rouca* é um aplicativo de analíticos de áudio que escuta o ambiente 24 horas por dia, 7 dias por semana, para classificar e filtrar sons. Quando é detectado um ataque de tosse ou voz rouca, o aplicativo gera um alerta.

Este white paper apresenta a *detecção de tosse e voz rouca* e como ela é configurada para detecção ideal.

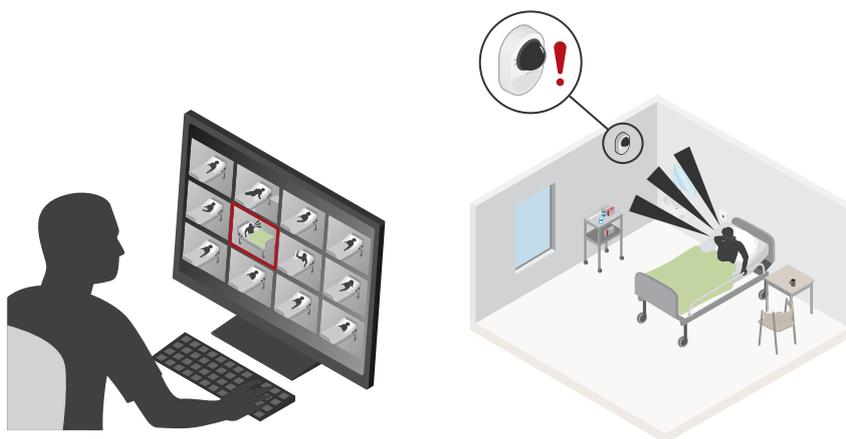
2 Detector de tosse e detector de voz rouca

A *detecção de tosse e voz rouca* vem pré-instalada em câmeras Axis selecionadas com microfone integrado. Os detectores captam indicadores sonoros de incidentes em tempo real, diretamente na câmera. Você pode optar por usar um dos detectores ou ambos.

O detector de tosse detecta e conta sons de tosse em um intervalo de tempo. Ele permite que o pessoal responda rapidamente quando alguém está tossindo ou tendo um ataque de tosse. O detector pode detectar ataques de tosse ou uma única tosse, dependendo de como você o configurou.

Para reduzir as várias notificações de eventos em um intervalo de tempo muito curto, assim que a primeira tosse é contada, começa um tempo de bloqueio de 5 segundos. Se o analítico for configurado para 3 tosses em 30 segundos, ele só contará a próxima tosse se houver um intervalo de pelo menos 5 segundos da tosse anterior. Tosses intermediárias dentro do espaço de bloqueio de 5 segundos não serão contadas. Isso significa que, com essas configurações, um alerta só será enviado após a contagem de 3 tosses com pelo menos 5 segundos de intervalo entre cada uma delas.

O detector de voz rouca identifica padrões sonoros associados à coação, raiva ou medo na voz de uma pessoa. Após o reconhecimento, o sistema envia uma notificação automática aos funcionários por meio de um alerta visual ou acionando um alarme. O aviso prévio permite que o pessoal responda rapidamente. Eles podem oferecer ajuda a uma pessoa necessitada ou prevenir uma escalada que poderia levar à agressão física.



Detecção de tosse e voz rouca, utilizada em ambiente de saúde.

3 Ativar detecção de áudio ideal

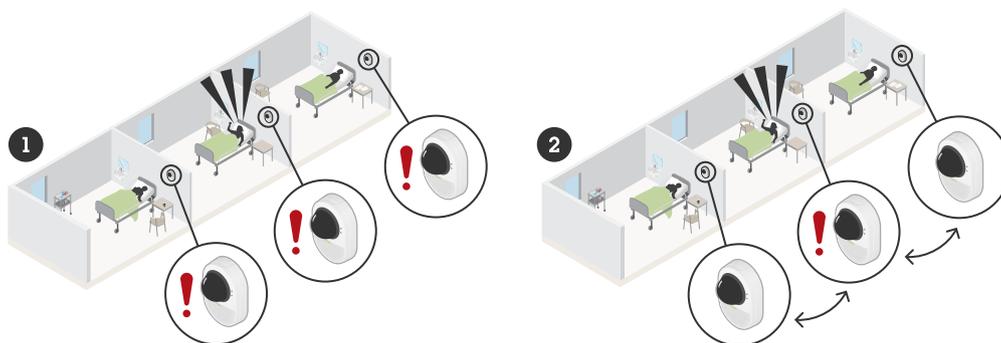
- **Posicionamento da câmera.** A câmera com analíticos deve ser colocada a pelo menos 1,5 metro (5 pés) de distância de fontes de ruído interferentes, como sistemas HVAC, sistemas de PA ou alto-falantes, e portas batendo. Além disso, a câmera deve ser preferencialmente colocada na linha de visão da área onde você deseja detectar o áudio. Embora a linha de visão não seja um requisito estrito, ela pode permitir uma detecção mais precisa. Isso ocorre porque os sons podem ser afetados quando eles contornam cantos ou obstáculos. Por exemplo, nem todas as frequências se curvam no mesmo grau.
- **Sensibilidade.** O sistema de detecção pode ser ajustado por meio das configurações de sensibilidade. Uma sensibilidade mais alta produzirá mais detecções. Ela aumenta o risco de detecções indesejadas (alarmes falsos), mas pode ser necessária quando é crucial nunca perder uma detecção. Com uma sensibilidade mais baixa, as detecções serão reportadas apenas quando houver certeza de que o som está classificado corretamente. Isso aumenta o risco de perder potenciais incidentes, mas pode ser necessária uma sensibilidade baixa quando, de outra forma, existiriam muitos alarmes falsos.
- **Modo de coleta de dados.** Você pode usar o modo de coleta de dados por um período de tempo após a instalação para obter informações sobre quais tipos de áudio são detectados. Os resultados e a análise podem fornecer informações sobre qual nível de sensibilidade é ideal para uma instalação específica.
- **Limite de detecção de tosse.** Você pode definir o limite de quantas tosses serão necessárias. Um alarme será acionado somente quando o número de tosses atingir o limite no período de tempo atribuído.
- **Configurações avançadas.** As configurações avançadas são apenas para usuários experientes. As alterações podem levar a detecções incorretas ou a nenhuma detecção. No entanto, para cenários específicos, pode ser necessário alterar essas configurações. Isso deve ser feito somente quando aconselhado ou em consulta com um especialista em sistemas.

4 Reconhecimento multissensor

Quando as câmeras são colocadas próximas umas das outras, por exemplo, em salas adjacentes, o mesmo incidente de áudio pode ser detectado por múltiplas câmeras. Isso pode tornar mais difícil identificar onde o incidente está ocorrendo.

Para detecções de voz rouca, o recurso de reconhecimento multissensor pode ser útil nesses casos. Quando estiver ativado e várias câmeras captarem a mesma voz rouca, apenas a câmera que captou primeiro vai

acionar uma notificação. Dessa forma, câmeras próximas trabalham juntas para reduzir eventos falsos e notificações duplicadas para o mesmo evento.



- 1 Sem reconhecimento multissensor: câmeras em salas adjacentes detectam o mesmo incidente de voz rouca e criam vários alarmes.
- 2 Com reconhecimento multissensor, apenas a câmera mais próxima relata uma detecção.

Com o reconhecimento multissensor, você cria grupos de peers para agrupar câmeras próximas que estejam dentro do alcance de captação de áudio umas das outras. Algumas restrições se aplicam:

- Todos os peers devem ser configurados para usar a sincronização de horário NTP.
- Todos os peers devem executar a mesma versão da *detecção de tosse e voz rouca*.
- Todos os peers devem ser capazes de se comunicar por meio da rede.

Se alguma das opções acima falhar, o peer voltará ao modo autônomo e se marcará como degradado.

5 Sobreposições

Um espectrograma ao vivo e notificações de aplicativos podem ser sobrepostos ao feed de vídeo. Você pode personalizar o tamanho das sobreposições e arrastá-las para a posição desejada. Você pode ajustar a opacidade da sobreposição usando um controle deslizante.

As notificações do aplicativo mostrarão os eventos detectados pela câmera e qual é o status do aplicativo.

O espectrograma fornece uma representação visual do áudio. Ouvir o áudio e ver simultaneamente sua representação visual pode ajudar você a determinar rapidamente a gravidade de um incidente.

6 Tipos de eventos e status de integridade

Os eventos gerados pelo detector de tosse e de voz rouca são *sem estado*. São ocorrências momentâneas que são desencadeadas por uma detecção. Depois que o tempo de bloqueio do evento (cinco segundos, configurável) expirar, uma detecção vai gerar um novo evento.

O status de integridade da *detecção de tosse e voz rouca* é refletido pelo uso de *eventos com estado*. Com eventos com estado, o estado do evento permanece ativo enquanto a condição ocorrer e alterna somente quando a condição for resolvida.

As verificações de integridade são integradas para verificar a operação adequada e alertar quando algo está errado. Três estados de integridade podem ser distinguidos para a *detecção de tosse e voz rouca*:

- Estado saudável: operação normal. As detecções são possíveis.
- Estado degradado: a operação está sendo executada em modo degradado. Isso normalmente é causado por fatores temporários, como perda de uma câmera peer, corte de áudio devido a sons muito altos ou saturação do buffer de áudio. No estado degradado, as detecções são possíveis, mas pode haver mais detecções falsas ou detecções perdidas. O estado degradado normalmente se resolve sozinho.
- Estado de mau funcionamento: sem operação. Nenhuma detecção é possível. Isso geralmente é causado por fatores que não se resolvem sozinhos, como a desativação do suporte de áudio nas configurações do dispositivo ou o silenciamento do ganho de entrada de áudio.

O estado degradado e de mau funcionamento será exibido no painel de informações e também na sobreposição de texto (se ativado), para que o operador saiba que o aplicativo está sendo executado com integridade degradada ou mau funcionamento detectado.

Um evento de pulsação é acionado a cada 60 segundos (quando ativado, configuração de tempo). Ele pode ser usado no receptor para verificar se os analíticos estão funcionando e alertar se nenhuma pulsação for recebida. Os eventos de pulsação não são enviados enquanto o estado de mau funcionamento estiver ativo.

7 Privacidade

Os dados de áudio são processados e analisados na câmera, e nenhum armazenamento de dados de áudio é necessário para que a *detecção de tosse e voz rouca* funcione corretamente. Somente quando ativado explicitamente, serão feitas gravações de áudio durante os eventos. Isso pode ser útil para análise forense ao investigar incidentes, para solução de problemas quando falsos positivos são relatados ou para ouvir incidentes caso isso não seja compatível com o sistema de gerenciamento de vídeo.

Sobre a Axis Communications

A Axis torna possível um mundo mais inteligente e seguro criando soluções para melhorar a segurança e o desempenho dos negócios. Como empresa de tecnologia de rede e líder do setor, a Axis oferece soluções em videomonitoramento, controle de acesso, intercomunicação e áudio. Nossas soluções são aprimoradas por aplicativos de análise inteligentes e apoiados por treinamento de alta qualidade.

A Axis tem cerca de 4.000 funcionários dedicados em mais de 50 países e colabora com parceiros de tecnologia e integração de sistemas em todo o mundo para fornecer soluções aos clientes. A Axis foi fundada em 1984 e tem sede em Lund, Suécia