

Audioanalysefunktionen in ausgewählten Kameras mit integriertem Mikrofon

September 2025

Zusammenfassung

Erfassung von Hustenanfällen und schreiender Stimme ist eine Audioanalyseanwendung, die Ereignisse erfasst, indem sie rund um die Uhr in Umgebung überwacht.

Die Anwendung ist in ausgewählten Axis Kameras mit integriertem Mikrofon vorinstalliert. Sie besteht aus zwei gesonderten Erfassungsalgorithmen, und Sie können wählen, einen davon oder beide zu nutzen.

Der Hustenmelder erfasst einzelne Hustenereignisse oder Hustenanfälle, so dass Mitarbeiter rasch auf Personen reagieren können, die Hilfe benötigen. Der Melder für schreiende Stimmen ist ein wirksames Mittel zur Verbrechensabschreckung, zur Verringerung von Übergriffen oder zur Anzeige von Personen, die Hilfe benötigen.

Datenschutz ist gegeben, denn die *Erfassung von Hustenanfällen und schreiender Stimme* muss keine Audiodaten speichern, um einwandfrei zu funktionieren. Es wird kein Audio aufgezeichnet, wenn Sie nicht explizit Audioaufzeichnung aktiviert haben.

Sie können verschiedene Einstellungen konfigurieren, damit die Analysefunktionen für Ihren Anwendungsfall optimal arbeiten. Die *Erfassung von Hustenanfällen und schreiender Stimme* führt außerdem ständige Gesundheitskontrollen durch, um den einwandfreien Betrieb zu verifizieren.

Inhalt

1	Einführung	4
2	Melder für Husten und schreiende Stimme	4
3	Optimale Audioerfassung einrichten	5
4	Multisensor-Erkennung	5
5	Overlays	6
6	Ereignistypen und Gesundheitsstatus	6
7	Datenschutz	7

1 Einführung

- audio detection analytics Die Analysefunktionen der Audioerkennung in einer Kamera sind eine leistungsstarke Ergänzung der Videosicherheit. Sie lassen die Früherkennung und die Warnung vor potenziellen Ereignissen zu, womöglich bevor sie im Video entdeckt werden.

Erfassung von Hustenanfällen und schreiender Stimme ist eine Audioanalyseanwendung, die rund um die Uhr die Umgebung überwacht, um Geräusche zu klassifizieren und herauszufiltern. Wenn selbige erkannt werden, generiert die Anwendung einen Alarm.

In diesem Whitepaper stellen wir die *Erfassung von Hustenanfällen und schreiender Stimme* und ihre Konfiguration für optimale Erfassung vor.

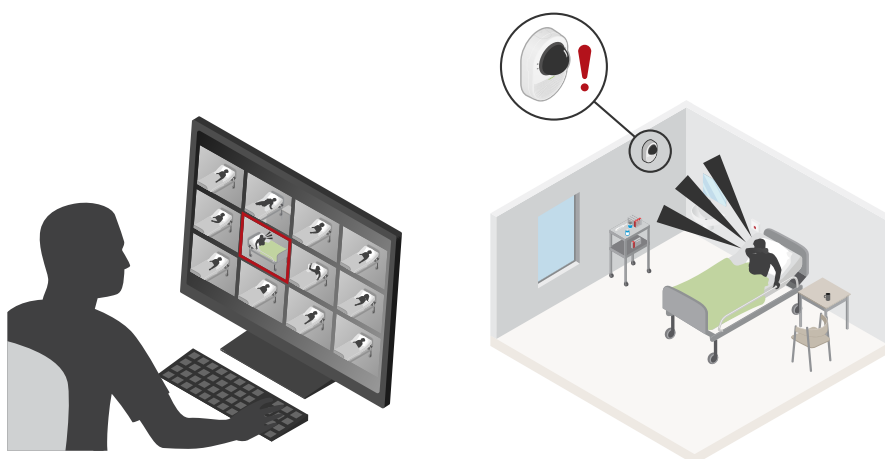
2 Melder für Husten und schreiende Stimme

Erfassung von Hustenanfällen und schreiender Stimme ist in ausgewählten Axis Kameras mit integriertem Mikrofon vorinstalliert. Die Melder erfassen hörbare Indikatoren für Ereignisse in Echtzeit, direkt in der Kamera. Sie können wählen, ob Sie einen der Melder oder beide einsetzen wollen.

Der Hustenmelder erfasst eine bestimmte Anzahl von Hustenereignissen innerhalb eines bestimmten Zeitfensters. Dank des Melders kann das Personal rasch reagieren, wenn jemand hustet oder einen Hustenanfall hat. Er kann Hustenanfälle oder ein einzelnes Husten melden, je nach dem, wie Sie ihn einrichten.

Um die Anzahl der Ereignisbenachrichtigungen innerhalb eines sehr kurzen Zeitfensters zu reduzieren, ist ab dem Zeitpunkt der Erfassung des ersten Hustenereignisses eine Blockzeit von 5 Sekunden vorgesehen. Wenn die Analysefunktionen beispielsweise auf 3 Hustenereignisse innerhalb von 30 Sekunden eingestellt sind, wird das nächste Hustenereignis erst erfasst, wenn seit dem letzten Hustenereignis mindestens 5 Sekunden vergangen sind. Etwaige Hustenereignisse innerhalb der Blockzeit von 5 Sekunden werden nicht erfasst. Dies bedeutet, dass bei diesen Einstellungen erst dann eine Warnmeldung versendet wird, wenn 3 Hustenereignisse in einem zeitlichen Abstand von mindestens 5 Sekunden erfasst wurden.

Der Melder für schreiende Stimme identifiziert Klangmuster, die mit schreienden Stimmen verbunden sind. Bei einer Erkennung sendet das System in Form eines visuellen Alarms oder durch Auslösung eines Alarms eine automatische Benachrichtigung an das Personal. Mit dieser Frühwarnung kann das Personal rasch reagieren. Es kann einer Person Hilfe anbieten, die sie benötigt, oder eine Eskalation verhindern, die ansonsten zu physischer Aggression führen könnte.



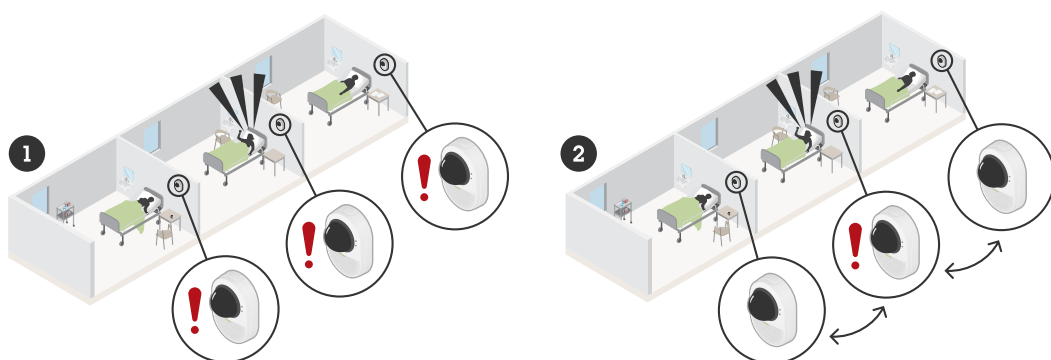
3 Optimale Audioerfassung einrichten

- **Kamerapositionierung.** Die Kamera mit den Analysefunktionen sollte mindestens 1,5 Meter (5 feet) von Störgeräuschquellen wie beispielsweise HLK-Anlagen, Beschallungsanlagen oder Lautsprechern und knallenden Türen entfernt angeordnet werden. Außerdem sollte die Kamera vorzugsweise in der Sichtlinie des Bereichs platziert werden, in dem Sie Audio erfassen wollen. Die Sichtlinie ist zwar keine zwingende Anforderung, kann jedoch eine genauere Erfassung zulassen. Das liegt daran, dass Geräusche beeinträchtigt werden können, wenn sie Ecken oder Hindernisse überwinden müssen. Zum Beispiel lassen sich nicht alle Frequenzgänge im gleichen Umfang biegen.
- **Empfindlichkeit.** Das Erfassungssystem kann mit den Empfindlichkeitseinstellungen feinabgestimmt werden. Höhere Empfindlichkeit führt zu mehr Erfassungen. Damit wächst das Risiko unerwünschter Erfassungen (Fehlalarme); sie kann jedoch erforderlich sein, wenn es darauf ankommt, keine Erfassung zu versäumen. Mit geringerer Empfindlichkeit werden Erfassungen nur gemeldet, wenn es sehr sicher ist, dass das Geräusch richtig klassifiziert wurde. Damit wächst das Risiko, potenzielle Ereignisse zu verpassen, aber eine niedrige Empfindlichkeit kann erforderlich sein, wenn es anderenfalls viele Fehlalarme gäbe.
- **Datenerfassungsmodus.** Über einen bestimmten Zeitraum nach der Installation können Sie den Datenerfassungsmodus nutzen, um Erkenntnisse dazu zu gewinnen, welche Audiotypen erfasst werden. Die Ergebnisse und Analysen können Informationen dazu liefern, welcher Empfindlichkeitsgrad für die jeweilige Installation optimal ist.
- **Grenzwert der Hustenerfassung.** Sie können den Grenzwert dafür festsetzen, wie viele Hustenereignisse erforderlich sein sollten. Ein Alarm wird erst ausgelöst, wenn die Anzahl der Hustenereignisse innerhalb des zugewiesenen Zeitraums den Grenzwert erreicht.
- **Erweiterte Einstellungen.** Erweiterte Einstellungen sind nur für erfahrene Benutzer bestimmt. Änderungen können zu falschen Erfassungen oder gar keinen Erfassungen führen. Für bestimmte Szenarien müssen Sie allerdings diese Einstellungen möglicherweise ändern. Das sollte nur erfolgen, wenn es von einem Systemexperten geraten oder mit ihm abgesprochen wird.

4 Multisensor-Erkennung

Wenn Kameras unweit voneinander angeordnet sind, beispielsweise in benachbarten Räumen, kann dasselbe Audioereignis von mehreren Kameras erfasst werden. Das kann die Lokalisierung des Ereignisses erschweren.

Bei der Erfassung schreiender Stimme kann in solchen Fällen die Multisensor-Erkennungsfunktion hilfreich sein. Wenn aktiviert und mehrere Kameras dieselbe schreiende Stimme erfassen, löst nur diejenige Kamera, die sie zuerst erfasst hat, eine Benachrichtigung aus. Auf diese Weise arbeiten benachbarte Kameras zusammen und reduzieren falsche Ereignisse und Doppelbenachrichtigungen über dasselbe Ereignis.



- 1 Ohne Multisensor-Erkennung: Kameras in benachbarten Räumen erfassen dasselbe Ereignis schreiender Stimme und lösen mehrfachen Alarm aus.
- 2 Bei Multisensor-Erkennung meldet nur die nächstliegende Kamera eine Erfassung.

Mit Multisensor-Erkennung richten Sie Peergroups ein, um Kameras in der Nähe zusammenzustellen, die in einem gemeinsamen Audioempfangsbereich positioniert sind. Es gelten einige Einschränkungen:

- Alle Peers sollten für die Nutzung von NTP-Zeitsynchronisation konfiguriert sein.
- Alle Peers sollten mit der gleichen Version der *Erfassung von Hustenanfällen und schreiender Stimme* laufen.
- Alle Peers sollten sich gegenseitig über das Netzwerk erreichen können.

Bei Nichterfüllung einer der vorstehenden Bedingungen fällt der Peer in den Standalone-Modus zurück und wird als verschlechtert ausgewiesen.

5 Overlays

Der Video-Feed kann mit einem Live-Spektrogramm und Anwendungsbenachrichtigungen überlagert werden. Sie können die Größe der Overlays individuell anpassen und sie in die gewünschte Position ziehen. Mit einem Schieber lässt sich die Opazität des Overlays anpassen.

Anwendungsbenachrichtigungen zeigen von der Kamera erfasste Ereignisse und den Anwendungsstatus.

Das Spektrogramm bietet eine visuelle Darstellung des Audios. Wenn man den Ton hört und gleichzeitig sieht, wodurch er ausgelöst wurde, lässt sich rasch bestimmen, wie schwerwiegend ein Ereignis ist.

6 Ereignistypen und Gesundheitsstatus

Vom Melder für Hustenanfälle und schreiender Stimme generierte Ereignisse sind *zustandslos*. Sie treten kurzzeitig auf und werden durch eine Erfassung ausgelöst. Nach Ablauf der Blockzeit eines Ereignisses (fünf Sekunden, konfigurierbar) generiert eine Erfassung ein neues Ereignis.

Der Gesundheitsstatus des *Melders für Hustenanfälle und schreiender Stimme* zeigt sich in der Anwendung *zustandsabhängiger Ereignisse*. Ein zustandsabhängiges Ereignis bleibt so lange aktiv, wie die Bedingung anhält, und ändert sich erst bei Auflösung der Bedingung.

Gesundheitskontrollen sind integriert, um die einwandfreie Funktion zu überprüfen und Alarm zu geben, wenn etwas außer Betrieb ist. Bei der *Erfassung von Hustenanfällen und schreiender Stimme* lassen sich drei Gesundheitszustände unterscheiden:

- Gesunder Zustand: Normalbetrieb. Erfassungen sind möglich.
- Beeinträchtigter Zustand: Betrieb im beeinträchtigten Modus. Dies wird in der Regel durch vorübergehende Faktoren wie beispielsweise der Verlust einer Peer-Kamera, Audio-Übersteuerung aufgrund sehr lauter Geräusche oder Audiopuffer-Überlauf verursacht. Im beeinträchtigten Zustand sind Erfassungen möglich, aber es können mehr Fehlerfassungen oder versäumte Erfassungen auftreten. Der beeinträchtigte Zustand löst sich in der Regel selbst.
- Störungszustand: kein Betrieb. Es sind keine Erfassungen möglich. Dies wird in der Regel durch Faktoren verursacht, die sich nicht von selbst beheben lassen, z. B. wenn die Audiounterstützung in den Geräteeinstellungen deaktiviert oder die Audioeingangsverstärkung stummgeschaltet ist.

Der Zustand "beeinträchtigt" und "Störung" wird im Info-Panel und auch im Text-Overlay (falls aktiviert) angezeigt, so dass der Bediener weiß, dass die Anwendung mit beeinträchtigtem Zustand oder erkannter Störung läuft.

Alle 60 Sekunden wird ein Heartbeat-Ereignis ausgelöst (sofern aktiviert, konfigurierbare Zeiteinstellung). Das kann auf der Empfängerseite genutzt werden, um zu überprüfen, ob die Analysefunktionen intakt sind und laufen, und um zu warnen, wenn keine Heartbeats empfangen werden. Solange der Störungszustand aktiv ist, werden keine Heartbeat-Ereignisse gesendet.

7 Datenschutz

Audiodaten werden in der Kamera verarbeitet und analysiert, und es müssen keine Audiodaten gespeichert werden, damit die *Erfassung von Hustenanfällen und schreiender Stimme* einwandfrei funktioniert. Nur bei expliziter Aktivierung wird Audio bei Ereignissen aufgezeichnet. Das kann für forensische Analysen bei der Untersuchung von Ereignissen nützlich sein, zur Fehlersuche bei Falschmeldungen oder zum Nachhören von Ereignissen, wenn dies vom Video Management System nicht unterstützt wird.

Über Axis Communications

Axis ermöglicht eine smartere und sichere Welt durch die Verbesserung von Sicherheit, Schutz, betrieblicher Effizienz und Geschäftsanalytik. Als Technologieführer im Bereich Netzwerk-Video bietet Axis Videosicherheits-, Zutrittskontroll-, Intercom- und Audiolösungen. Die branchenweit anerkannten Schulungen der Axis Communications Academy vermitteln fundiertes Expertenwissen zu den neuesten Technologien.

Das 1984 gegründete schwedische Unternehmen beschäftigt etwa 5.000 engagierte MitarbeiterInnen in über 50 Ländern und bietet mit Technologie- und Systemintegrationspartnern auf der ganzen Welt kundenspezifische Lösungen an. Der Hauptsitz ist in Lund, Schweden.