

ホワイトペーパー

咳検知とストレス音声検知

内蔵マイクを搭載したカメラの音声分析

5月 2024

概要

咳検知とストレス音声検知は、周囲の音声を24時間365日聞いてインシデントを検知する音声分析アプリケーションです。

このアプリケーションは、マイク内蔵の一部のAxisカメラにプリインストールされています。これには2つの個別の検知アルゴリズムが搭載されており、いずれか一方を使用することも、両方を使用することもできます。

咳検知では、1回の咳や咳の発作を検知し、対象の人物に迅速に対応することができます。ストレス音声検知では、強迫、怒り、恐怖に関連する音声パターンを識別するため、犯罪抑止、暴行の抑止または助けを求めている人の特定に効果的なツールとなります。

咳検知とストレス音声検知は、音声データを保存しなくても適切に機能するため、プライバシーを保護することができます。録音を意図的に有効化しない限り、音声は録音されません。

いくつかの設定を構成できるため、ニーズに合わせて最適な分析を実現することが可能です。咳検知とストレス音声検知は適切に動作していることを確認するために継続的なヘルスチェック機能も搭載しています。

目次

1	はじめに	4
2	咳検知とストレス音声検知	4
3	最適な音声検知の実現	5
4	マルチセンサー認識	5
5	オーバーレイ	6
6	イベントの種類とヘルスステータス	6
7	プライバシー	7

1 はじめに

映像監視において、カメラの音声検知分析は強力な補完機能となります。音声検知により、ビデオで発見するよりも前に、潜在的な事故・事態を早期に検知してアラートを発信することができます。

咳検知とストレス音声検知は、周囲の音を24時間365日聞いて音を分類し、フィルタリングする音声分析アプリケーションです。咳やストレス音声を検知されると、アプリケーションでアラートが生成されます。

本ホワイトペーパーでは、咳検知とストレス音声検知の概要および最適な検知を実現するための構成方法についてご説明します。

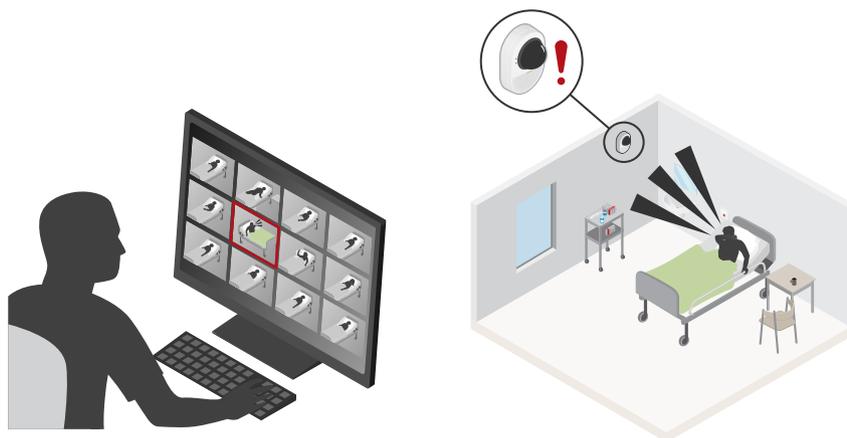
2 咳検知とストレス音声検知

咳検知とストレス音声検知は、マイクを内蔵する一部のAxisカメラにプリインストールされています。この検知では、事故・事態に関連付けられる音声がリアルタイムでカメラに直接キャプチャーされます。検知のいずれか、または両方を使用することができます。

咳検知は、咳を検知し、一定時間内の咳の回数を計測します。対象人物が咳をしている場合や咳の発作を起こした場合に、担当者が迅速に対応できるようになります。この検知機能は、設定方法に応じて、咳の発作または1回の咳を検知することができます。

非常に短い時間内に送信されるイベント通知の数を減らすため、最初の咳がカウントされると同時に5秒のブロック時間が開始します。分析機能が30秒以内に3回の咳をするように設定されている場合、次の咳は前の咳から5秒以上経過した場合にのみカウントされます。この5秒間のブロック時間内の中間の咳はカウントされません。つまり、この設定では、それぞれの咳の間隔が少なくとも5秒以上開いて3回の咳がカウントされた場合にのみアラートが送信されます。

ストレス音声検知は、強迫、怒り、恐怖に関連する音声のパターンを識別します。音声を検知されると、システムは視覚的なアラートまたはアラームトリガーを通じてスタッフに自動的に通知します。早期警告により、担当者は迅速に対応できます。困っている人に助けを提供したり、身体的な攻撃に繋がる可能性のある行為を未然に防ぐことができます。



医療機関における咳検知とストレス音声検知の利用

3 最適な音声検知の実現

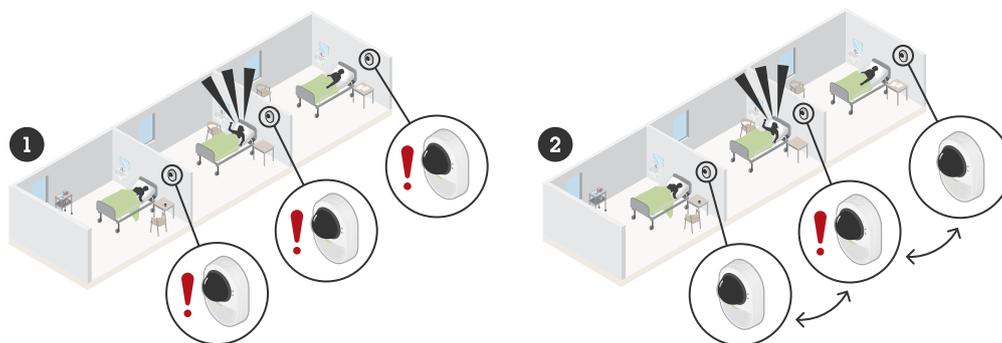
- **カメラの設置位置**：分析機能を備えたカメラは、HVACシステム、PAシステム、スピーカー、ボタンと閉まるドアなどの干渉ノイズ源から少なくとも1.5メートル離れた場所に設置する必要があります。また、カメラは音声を検知したいエリアがまっすぐ見える場所に設置することが好ましいです。この設置方法は厳密な要件ではありませんが、これにより正確な検知が可能となります。音は角や障害物で曲がったときに影響が出る可能性があるためです。例えば、すべての周波数が同じ角度で屈折するわけではありません。
- **感度**：感度設定により、検知システムを微調整することができます。感度が高いほど、より多くの音声を検知することが可能となります。望ましくない検知（誤報）のリスクが高まりますが、検知を見逃さないことが非常に重要な場合には、感度を上げる必要があります。感度を低くすると、音声は正しく分類されたことが確実な場合にのみ検知が通知されます。これにより、潜在的なインシデントを見逃すリスクが高まりますが、感度を低くしないと誤報が多発する場合には、感度を下げる必要があります。
- **データ収集モード**：インストール後の一定期間、データ収集モードを使用して、検知された音声の種類を把握することができます。結果と分析から得た情報により、特定の設置状況下で最適と考えられる感度レベルを把握することが可能となります。
- **咳検知の閾値**：咳の回数の閾値を設定することができます。設定された時間内に検知された咳の回数が閾値に達すると、アラームがトリガーされます。
- **高度な設定**：高度な設定は、エキスパートユーザーのみを対象とした機能です。変更により、誤った検知が発生したり、まったく検知されなくなる可能性があります。ただし、特定のシナリオではこれらの設定の変更が必要になる場合があります。これは、システムの専門家からアドバイスを受けた場合、またはシステムの専門家と相談した場合にのみ実行してください。

4 マルチセンサー認識

隣接する部屋など、カメラが互いに近くに配置されている場合、同じ音声インシデントが複数のカメラで検知される場合があります。この場合、インシデントが発生している場所を特定することが困難になることがあります。

ストレス音声の検知には、マルチセンサー認識機能が役立ちます。この機能が有効の状態、複数のカメラが同じストレス音声を検知した場合、最初に検知したカメラのみが

通知をトリガーします。このように、近くのカメラが連携して誤イベントを減らし、同じイベントに対する重複通知を減らすことができます。



- 1 マルチセンサー認識機能なし：隣接する部屋のカメラが同じストレス音声インシデントを検知し、複数のアラームを生成します。
- 2 マルチセンサー認識機能あり：最も近いカメラのみが検知を報告します。

マルチセンサー認識を使用すると、ピアグループを作成して、相互の音声ピックアップ範囲内にある近くのカメラをグループ化することができます。一部制限あり：

- すべてのピアが、NTP時刻同期を使用するように設定されていること。
- すべてのピアで同じバージョンの咳検知とストレス音声検知を実行していること。
- すべてのピアがネットワーク経由で相互に通信できる状態になっていること。

上記のいずれかの条件が満たされていない場合、ピアがスタンドアロンモードに戻り、デグレードとマークします。

5 オーバーレイ

ライブスペクトログラムとアプリケーション通知をビデオフィードの上にオーバーレイすることができます。オーバーレイのサイズをカスタマイズし、希望の位置にドラッグできます。スライダーを使用して、オーバーレイの不透明度を調整できます。

アプリケーション通知には、カメラで検知されたイベントとアプリケーションのステータスが表示されます。

スペクトログラムは音声を視覚的に表現します。音声を耳で聞きながら、その視覚的表現を確認すると、インシデントの重大度を迅速に判断するのに役立ちます。

6 イベントの種類とヘルスステータス

咳検知とストレス音検知で生成されるイベントはステートレスです。つまり、検知に基づくトリガーによって瞬間的に出力が決定されるということです。イベントブロック時間（5秒、設定可能）が経過すると、新しい検知イベントが生成されます。

咳検知とストレス音声検知のヘルスステータスは、ステートフルイベントを使用することで反映されます。ステートフルイベントでは、状況が発生している限りイベント状態はアクティブのままとなり、状況が解決された場合にのみトグルします。

ヘルスチェック機能が組み込まれており、適切な動作を確認し、何かがオフの場合にアラートします。咳検知とストレス音声検知では、3つの状態を区別できます。

- 健全な状態：通常の動作。検知可能な状態です。
- デグレード状態：デグレードモードで動作しています。これは通常、ピアカメラの喪失、非常に大きな音の音声クリッピング、音声バッファオーバーランなどの一時的な要因によって発生します。デグレード状態でも検知は可能ですが、誤検知や検知漏れが増える可能性があります。デグレード状態は通常、自動的に回復します。
- 機能不良状態：動作しません。検知はできません。これは通常、デバイス設定で音声サポートが無効になっている、あるいは音声入力ゲインがミュートにされているなど、自然には解決されない要因によって発生します。

デグレード状態と機能不良状態は、情報パネルとテキストオーバーレイ（有効な場合）に表示されるため、オペレーターはアプリケーションがデグレード状態で実行されているか、または機能不良状態かを把握することができます。

ハートビートイベントは60秒ごとにトリガーされます（有効な場合。時間設定は設定可）。これを受信側で使用して、分析機能が機能しているかどうかを確認し、ハートビートが受信されない場合はアラートを発信することが可能となります。機能不良状態の間は、ハートビートイベントは送信されません。

7 プライバシー

咳検知とストレス音声検知では、音声データがカメラ内で処理・分析されるため、音声データを保存する必要がありません。意図的に有効化した場合にのみ、イベント発生時に音声の録音が行われます。録音すれば、インシデントを調査する際のフォレンジック調査、誤検知が通知された場合のトラブルシューティング、ビデオ管理システムでサポートされていない場合に事故・事態の音声を聴く場合に役立ちます。

Axis Communicationsについて

Axisはセキュリティとビジネスパフォーマンスを向上させるソリューションを生み出すことで、よりスマートで安全な世界の実現を目指しています。ネットワークテクノロジー企業として、また業界のリーダーとして、Axisはビデオ監視、アクセスコントロール、インターコム、音声システムなどのソリューションを提供しています。これらのソリューションはインテリジェントな分析アプリケーションによって強化され、高品質のトレーニングに支えられています。

Axisは50ヶ国以上に約4,000人の熱意にあふれた従業員を擁し、世界中のテクノロジーおよびシステムインテグレーションパートナーと連携することで、カスタマーソリューションをお届けしています。Axisは1984年に設立され、本社はスウェーデンのルンドにあります。