ホワイトペーパー

指向性音声検知

12月 2025

目次

1 はじめに 3
2 音声インシデントの検知 3
3 初期設定 3
3.1 設置オプション 3
3.2 指向性音声検知機能の有効化 3
3.3 パン角度の校正 3
3.4 チルト角度のプリセットの調整 4
4 運用 4
4.1 webインターフェース 4
4.2 音声イベントログ 4
4.3 アクションルールのトリガー 4
4.4 オートパイロット機能に対する優先制御 5
5 検知距離 5
5.1 感度の調整 5
5.2 物理的な障害物 5
5.3 気候条件 5
5.3.1 雨音の除去 5
5.3.2 風雑音の除去 5
5.3.3 音の伝搬に対する風の影響 5
6 プライバシーとデータ保全性 6

1　はじめに

従来のセキュリティカメラは、主に映像を用いてイベントを検知します。一方で、音声は重要な状況を提供することができ、特にカメラの視野外で発生する事象において、重要な状況把握の手がかりを提供します。

指向性音声検知機能は、音声の有用性を最大限に活かした機能の代表例です。銃声、ガラスの破損音、悲鳴などの突発的で大きな音を検知します。音を検知すると、この機能は音源の位置を算出します。この情報を利用して、録画を開始したり、オペレーターにアラートを送信したり、対応するPTZカメラを音源に向けて制御したりできます。

このホワイトペーパーでは、この技術の仕組みや動作原理、導入時の検討事項、ならびに制約事項について解説します。

2　音声インシデントの検知

指向性音声検知機能は、カメラに搭載された複数の音声センサーを使用して、周囲の環境を継続的に監視します。本システムは、音量が急激に増加し、設定されたトリガー閾値を上回る音を識別しますシステムは音を検出すると、音源方向のパン角を計算し、可能であればチルト角も算出します。算出された情報はVAPIXイベントとしてPTZカメラに送信され、カメラはそれに応じてパン動作を行い、音源方向を映します。

AXIS Q6300-Eネットワークカメラには指向性音声検知機能が統合されています。AXIS Q6300-Eと互換性のあるすべてのPTZカメラは、指向性音声検知機能に対応しております。

3　初期設定

3.1 設置オプション

AXIS Q6300-Eは、平面の壁、外側の角、ポール、またはマストに取り付けることができます。最適な指向性音声検知機能を得るためには、ポールまたは建物の外側の角への設置が最適です。

インストールガイドに記載されているとおり、AXIS Q6300-E は前面が前方を向くように設置する必要があります。カメラのインストールガイドに示されている通り、正しく設置できるようにカメラ本体の上部には溝が設けられており、外周部には矢印（パンゼロマーク）が表記されています。

一部のPTZカメラは手動での設定が必要です。その場合、設定アシスタントがガイドします。

3.2 指向性音声検知機能の有効化

この機能は初期設定では無効になっています。この機能を使用するには、カメラのWebインターフェースで有効化する必要があります。さらに、この機能でPTZカメラの方向を変換できるようにする場合は、そのオプションも有効化する必要があります。設定は、カメラのWebインターフェースで確認できます：**分析機能** > **AXIS Audio Analytics** > **指向性音声検知機能**

3.3 パン角度の校正

AXIS Q6300-EとPTZカメラでは、パン角の基準が異なる場合があります。一部のPTZモデルは事前に校正されていますが、機種によっては、指向性音声検知機能を使用する前にパン角の校正を行う必要があります。その場合、設定アシスタントが校正プロセスをガイドします。

3.4 チルト角度のプリセットの調整

音が検出されると、指向性音声検知機能は音源までのパン角およびチルト角を算出します。パン角は常に算出されますが、チルト角は条件によっては算出できない場合があります。これは通常、音が下方からではなく、水平方向から聞こえた場合に発生します。チルト角度が算出されない場合、指向性音声検知機能はプリセット値を使用します。この値は、カメラのWebインターフェースのPTZ (パン/チルト/ズーム) タブで調整できます。

4 運用

4.1 webインターフェース

カメラのWebインターフェースでは、指向性音声検知機能の設定を以下で確認できます：**分析機能** > **AXIS Audio Analytics** > **指向性音声検知機能**また、オーバービューカメラから4分割表示なども確認できます。音が検知されると、該当するカメラビューの周囲に黄色い枠が表示され、どのエリアから音が発生したかを確認できます。



図 4.1 *オーバービューカメラの4分割表示と、検知閾値と検知時間を調整するスライダーを備えたWeb インターフェース。*

Webインターフェースでは、閾値と継続時間の設定も制御できます。

- **閾値**周囲の騒音レベルに対して、何dB上回る音を検知対象とするかを定義します。閾値を高く設定するほど、指向性音声検知機能をトリガーするために、より大きな音、またはより急激な音が必要となります。
- **所要時間.**最初の音声イベント発生後、次の音声イベントを無視する時間を指定します。

4.2 音声イベントログ

音声イベントは記録され、CSVファイルとしてエクスポートすることができます。

4.3 アクションルールのトリガー

指向性音声検知機能は、VMSにおいて録画の開始などの自動アクションをトリガーするために使用できます。いくつかのトリガーが用意されています。

- **音声の検知**このトリガーは、すべての音声イベントに対して作動します。これは、たとえば録画を開始したり、ブックマークを作成したり、オペレーターが容易に検索や確認できるように、VMSに検知結果をログとして記録したりするために使用できます。

- **カメラ「X」の表示内で音声が検知されました。**検知された音のパン角およびチルト角が、固定カメラのいずれかと一致する場合、このトリガーが作動します。このトリガーは、音が検知された際に当該のカメラビューにオーバーレイを表示するために使用でき、音源を特定しやすくします。
- **カメラの視野外で音が検知されました。**このトリガーは、検知された音が固定カメラの表示範囲内にない場合に作動します。

4.4 オートパイロット機能に対する優先制御

オートパイロットは、全指向性カメラで対象範囲を監視し、PTZカメラを用いて人物、車両、その他の移動物体を追跡・分類する機能です。

指向性音声検知機能は、オートパイロットよりも優先されます。オートパイロットが物体を追跡中に指向性音声検知機能がトリガーされた場合、その追跡は中断され、PTZカメラは音源方向へ向けられます。一時停止後、オートパイロットは新たな位置から追跡を再開し、音源となった物体が視野内にあれば、それを追尾します。

5 検知距離

検知性能は、音源の大きさや距離といった要因はもちろん、周囲の騒音レベルやその他の環境要因にも左右されます。

5.1 感度の調整

閾値を変更することで、検知感度を用途に合わせて調整できます。周囲の騒音レベルに対して、どの程度大きな音であれば指向性音声検知機能をトリガーするかを定義します。

5.2 物理的な障害物

建物や大きな物体は、音を反射したり遮ったりすることがあります。その結果、特定の方向からの音については、検知範囲が想定より短くなる場合があります。

5.3 気候条件

雨風は、音の伝搬および知覚に影響を与える可能性があります。

5.3.1 雨音の除去

激しい雨は大きな音を発生させることがあります。ただし、指向性音声検知機能がトリガーされるには、音の振幅が急激に増加し、背景ノイズレベルに設定された閾値を加えた値を超える必要があります。雨音は背景ノイズの一部として扱われるため、フィルタリングされ、通常は指向性音声検知機能をトリガーすることはありません。

5.3.2 風雑音の除去

風雑音は、風が建物、ポール、電線などの固定物に当たることで発生します。これにより、うなり音やその他の持続的な音が発生することがあります。また、強風や突風にさらされると、カメラ自体が風雑音の原因となる場合もあります。突風は通常、数秒程度と短時間ですが、平均風速の2〜3倍の強さになることがあり、一時的な雑音を引き起こすことがあります。ただし、風雑音はカメラの音声センサーによって異なる特性として検知されるため、指向性音声検知機能では除外されます。

5.3.3 音の伝搬に対する風の影響

強風は音の伝搬に影響を与えるため、検知範囲に影響を及ぼす可能性があります。

風のある環境を伝わる音は、風の影響の受け方によって、実際より大きく聞こえたり小さく聞こえたりすることがあります。一般的に、地表付近の風は、上空の風よりも速度が遅くなります。

この速度差によって風速勾配が生じ、音波の進行方向が実質的に変更されます。その結果、風下に向かって伝わる音は下方に曲がりやすく、風上に向かって伝わる音は上方に曲がる傾向があります。すなわち、音源から見て風下（かつ地上）に立っている人のほうが、風上に立っている人よりも、音は大きく聞こえます。この影響は、距離が長くなるほど、また風速が高いほど大きくなります。

6 プライバシーとデータ保全性

指向性音声検知機能は、プライバシーおよびデータ保護を基本原則として設計されています。

指向性音声検知機能は、音響センサーを使用して、音量測定および基本的な音源位置の三角測量を行います。本システムは音量レベルを監視しますが、会話や音声を録音またはストリーミングしたり、単語・言語・話者を認識したり、音声コンテンツを外部システムへストリーミングしたりすることは、技術的にできません。

すべての音声処理はデバイス内部で完全に行われ、音声データがカメラの外部に送信されることは一切ありません。設定された音量閾値を超えた場合、システムはメタデータイベントを生成し、VMSに送信します。これらのイベントには、検知タイムスタンプ、音圧レベル、PTZ (パン/チルト/ズーム) の位置決め用に算出された座標が含まれています。音声コンテンツ、録音データ、または識別可能な音のパターンが送信されることはありません。

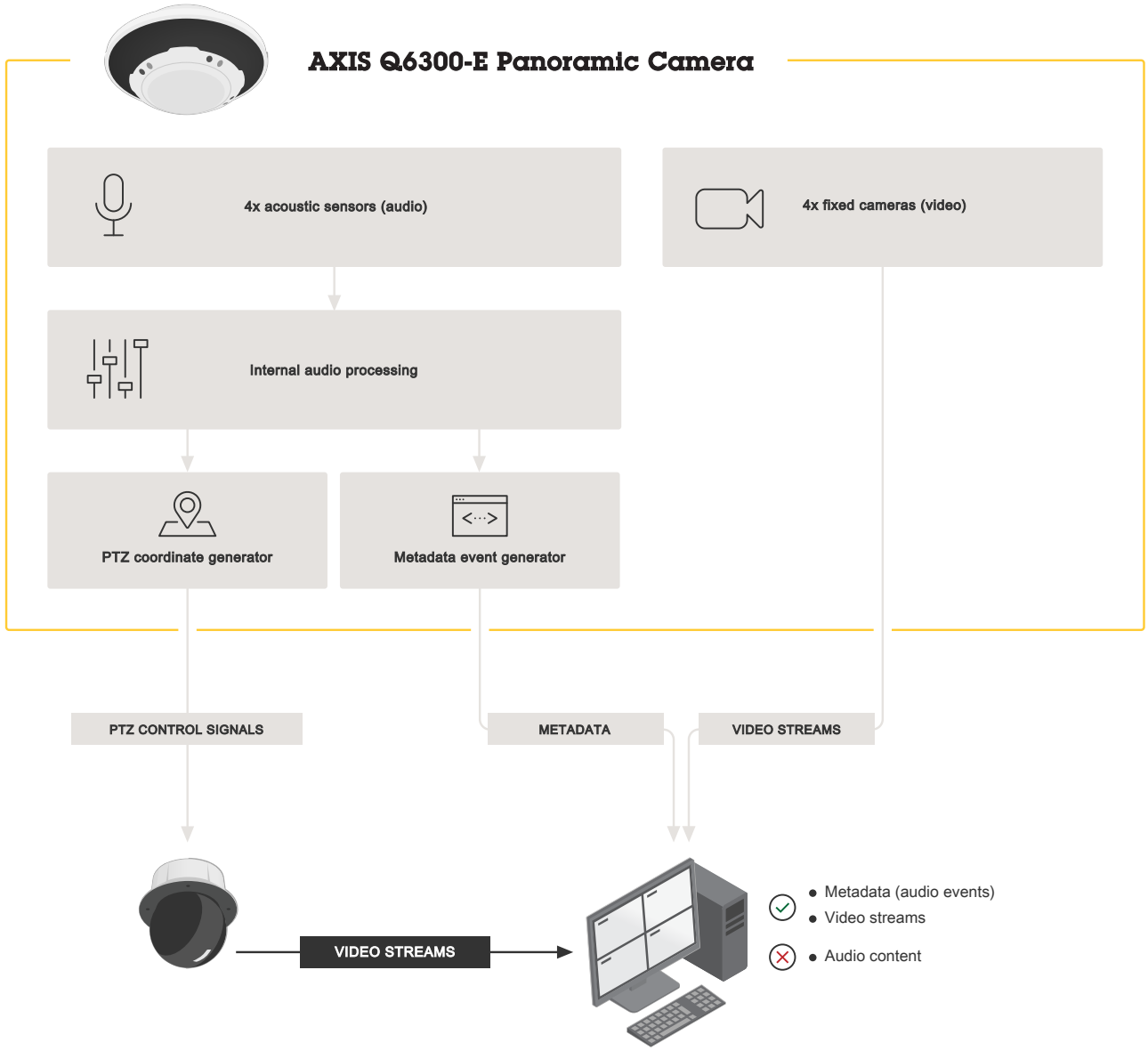


図 6.1 *指向性音声検知機能におけるシステムアーキテクチャおよびデータフロー音声コンテンツがカメラの外部に送信されることはありません。*

T10234321_JA/JA/M4.2/2512

Axis Communicationsについて

Axisは、セキュリティ、安全性、運用効率、ビジネスインテリジェンスを向上させることで、よりスマートでより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、また業界をけん引するリーダーとして、Axisは映像監視、アクセスコントロール、インターコム、音声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、インテリジェントアプリケーションによって強化され、質の高いトレーニングによってサポートされています。

Axisは50ヶ国以上に5,000人を超える熱意にあふれた従業員を擁し、世界中のテクノロジーパートナーや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年に創業し、本社はスウェーデン・ルンドにあります。

© 2025 — Axis Communications AB. AXIS COMMUNICATIONS、AXIS、ARTPEC、VAPIXは、スウェーデンおよびその他の国で登録されたAxis ABの登録商標です。その他のすべての商標は、それぞれの所有者に帰属します。記載内容は予告なく変更されることがあります。