

「音声入力」の可能性

撮影シーンの認識と証拠に役立つ、サウンドのキャプチャー
と処理

6月 2021

目次

1	まとめ	3
2	はじめに	4
	2.1 録音を伴わないキャプチャー	4
3	障害をクリアする	4
	3.1 法令の内容は？	5
	3.2 可能性を調査する	5
4	設置の問題	6
5	音声の準備	7
6	分析機能のトポロジー	7
7	ユースケースと例	8
	7.1 音声分析機能によるインシデントの検知	9
	7.2 ビデオでの音声の可視化	10
	7.3 聞き取りと対応	11
	7.4 録音と保存	12
	7.5 監視システムを最大限に活用する	12
8	監視と検知	12
	8.1 音響特性	12
	8.2 信号処理	13
	8.3 人の聴覚	14
9	免責事項	15
付録 1	音質の用語	16

1 まとめ

ビデオカメラに一体化されていてすぐに使用できる、または外部マイクによって提供される音声キャプチャー機能によって、さまざまな重要なユースケースが可能になります。責任とプロフェッショナルな判断に基づいて「音声入力」を使用することで、セキュリティシステムに極めて重要な価値と利点をもたらすことができます。たとえば、現場検証で欠落している証拠を提示したり、警備員や病院スタッフの速やかな対応を必要とするイベントをリアルタイムで検知したりできます。音声監視が行われているという事実だけでも、抑止効果をもたらし、犯罪の防止に役立ちます。

音声キャプチャー(多くの場合、インスタント分析アクションと併用される)は、スタンドアロンテクノロジーとして展開し、犯罪の防止、保護、現場検証における複数のユースケースを可能にすることができます。

ただし、音声キャプチャーは、ビデオと組み合わせることで、既存の監視のユースケースの大部分を強化できる可能性も持っています。たとえば、ビデオストリームに音声ストリームを追加することで、セキュリティオペレーターは、撮影シーンの状況認識力を大幅に高めることができます。

視覚的検知に基づく自動イベント検知とアラーム用として複数のビデオ分析機能を採用するのと同様、音声分析機能は音声ストリームを監視し、何か異常があった場合に反応することができます。

音声分析ソフトウェアは、マイクが人の叫び声、ガラスの割れる音、銃声などに関連する音を拾ったときに、自動アラームやその他のアクションをトリガーするよう設定することができます。これにより、迅速な対応と介入を可能にする早期警告を提供します。

音声分析機能は、予期しない音が左右どちらから聞こえるのかを検知し、自動的にPTZカメラを音源に向けることも可能です。病院や介護施設では、音声分析機能を使用して、患者が苦痛を感じていることを示唆する高音レベルを検知し、看護師に自動通知を送信することができます。このユースケースでは、さまざまな場所からの音の同時監視を容易にする、音の視覚化分析による恩恵も得ることができます。

音声のキャプチャーと録音には違いがあります。多くの用途では、目標を達成するために音声を録音する必要はありません。これは、プライバシーの問題を管理し、個人データに関する規制を遵守するのに役立ちます。一般的に、音声分析アプリケーションは音声を継続的に録音しません。通常は、受信した音声を処理して特定のパターン、音量、または周波数を探します。分析機能を(カメラ内の)エッジで実行する場合、デジタル音声データをカメラから送信する必要はありません。実行された分析の結果、つまりメタデータやトリガーのみがカメラから送信されます。

Axisでは、法的助言は提供していません。監視を規制する法律は地域、州、国によって異なり、監視を法令に準拠した方法で実行する責任は、製品のユーザー(通常はエンドカスタマー)にあります。映像監視の場合と同様、音声監視システムを設置する前に、その設置の法的側面を調査し理解する必要があります。

法的条件を満たすために必要な措置を講じたら、設備の配置と設定について、監視システムを慎重に検討する必要があります。これにはある程度の計画が必要な場合がありますが、一般的には難しくなく、簡単な方法で音声の有用性を確保することができます。

2 はじめに

音声情報は、犯罪の防止、保護、または現場検証向けの貴重な資産になることがあります。キャプチャーされた音声は、アクティビティ、行為、イベントなどを検知するための非常に効率的な音声監視を可能にする分析ソフトウェアによって、リアルタイムで処理することもできます。

このホワイトペーパーでは、セキュリティにおける音声の潜在力について、一般的なユースケースの例とともに説明します。さまざまなタイプの音声分析機能と、その仕組みの概要についても記載されています。

このホワイトペーパーには、法的助言は含まれていませんが、監視システムの構築に役立つと思われる、さまざまな技術的ソリューションを紹介しています。音声分析機能の実装方法によっては、地域の法律や推奨事項に従って、必要な場所にこの強力なツールを使用できる場合があります。

このホワイトペーパーの内容は、音声のキャプチャーと場合によっては録音、つまり音声入力に限定されています。セキュリティソリューションにおける音声のもう1つの一般的な使用法は、音声のブロードキャスト、つまり音声出力に関係し、通常は侵入者や万引き犯の抑止力として音声メッセージやアラームの再生に使用されます。セキュリティにおける音声のブロードキャストに関する詳細については、www.axis.com/ja-jp/products/audio をご覧ください。

2.1 録音を伴わないキャプチャー

音声を録音せずにキャプチャーして使用することができます。音声のキャプチャーとは、基本的に音声をデジタル化してソフトウェアで使用できるようにすることを意味します。これは、マイクを使用して空気中の音の振動を記録し、このアナログ信号を(A/D変換装置を使用して)デジタル信号に変換し、処理装置に転送することで可能になります。

キャプチャーされた音声は、フラッシュメモリやハードドライブなどの常設メディアに取り込まなければ、録音されません。オペレーターがキャプチャーされた音声をリアルタイムで聞いている場合など、ユースケースによっては録音が不要な場合があります。状況によっては、音声を録音しない特定の理由がある場合もあります。音声を録音するか、キャプチャーのみを行うかによって、法的な制限が異なる場合があります。

一般的に、音声分析機能は音声を継続的に録音しません。正常に機能するには、サウンドを一時的にバッファリングする必要があります。多くのシステムは、保安担当者が検知を検証し、場合によっては現場検証用の証拠として音声を保存できるよう、検知の直前と検知後にバッファリングされた音声を録音するように設定することができます。

3 障害をクリアする

多くの人々は、映像監視下でのマイクの使用について懸念を抱いています。こういった懸念は通常、映像素材とともにプライベートな会話が記録されることに関連しています。

「音声入力」には、単に録音するだけでなく、もっと多くの可能性があることを理解すれば、この最初の障害を乗り越えることができます。音声情報を録音する必要がないユースケースは多数あります。

監視を規制する法律は地域や国によって異なるため、監視システムに音声を追加する前に、何が許可されているかを確認してください。

3.1 法令の内容は？

映像監視の場合と同様、音声監視システムを設置する前に、その設置の法的側面を調査し理解する必要があります。該当する場合は、適切な申請書類を提出し、許可を取得する必要があります。また、必要に応じて、標識や公式声明を使用する必要があります。

音声の使用や録音は、国の立法機関やさまざまな種類の地域の規則や規制によって、複数の理由で禁止されている、または特別な配慮が必要な場合があります。ある地域や環境で音声のキャプチャーが許可されているとしても、録音は禁止されている場合があります。また、企業も、敷地内での音声監視の使用を禁止することができます。

3.1.1 米国の例

米国の法令は、州によって異なります。

一部の州では、音声の録音には1当事者の同意が必要です。これは、会話の一方の当事者のみが同意すれば、監視が合法になることを意味します。

その他の州では全当事者(または2当事者)の同意が必要です。つまり、録音を開始する前に、すべての当事者が録音に同意する必要があります。個人がプライベートを保てることを見込めない公共の場所では、全当事者の同意の例外が適用される場合があります。

音声を録音しない音声分析アプリケーションを使用する場合、一部の地域では法的な評価が別の結果をもたらす可能性もあります。したがって、特定の州で適用される法令を調査する必要があります。

3.1.2 欧州の例

音声監視は、欧州諸国の国内法によって規制されています。したがって、特定の国で適用される法令を調査する必要があります。

録音データには、GDPR (EU一般データ保護規則) の対象となる個人データが含まれている場合があります。GDPRは必ずしも音声録音を禁止してはおりませんが、音声のキャプチャーや録音には特別な配慮が必要です。既存の映像監視に音声を追加する場合、GDPRに基づく個人データの処理に対する法的根拠が引き続き適用されるかどうかを考慮する必要があります。

3.2 可能性を調査する

監視では、音声は決して許可されないという一般的な誤解があります。この誤解が広く流布しているため、多くの場合、音声によって監視システムを強化できる可能性は検討されていません。

しかし、人々に通知した場合や同意を得た場合など、さまざまなタイプの設置が許可される可能性があります。対象地域やユースケースに適用される法令を調査する必要があります。既存のセキュリティシステムで録音と保存のユースケースが許可されない場合でも、*聞き取りと対応*、*聞き取りと証言*、*音声分析機能によるインシデントの検知*など、多くのユースケースはプライバシー権を侵害しないように適応させることができます。

4 設置の問題

シーンへのマイクの配置は、潜在的な用途を決定します。音声機器を設置する前に、その配置と設定を慎重に検討する必要があります。これにはある程度の計画が必要な場合がありますが、一般的には難しくなく、簡単な方法で音声の有用性を向上させることができます。

マイクの適切な配置を熟考し、音響的に最適な場所を選択することで、監視目標を達成できる可能性が高まります。言うまでもなく、マイクは、重要な音を簡単にキャプチャーできるように配置する必要があります。一般的には、カメラとともに部屋の中央に、または関連性の高い特定のイベントが発生する可能性のある場所の近くに配置します。マイクは、空調や機械などのノイズ源の近くには配置しないでください。小さな音や遠くから聞こえる音が、かき消されてしまう可能性があります。



一般的なマイクの配置

- 1 関連性の高い行動が発生する場所
- 2 カメラ内
- 3 部屋の中央

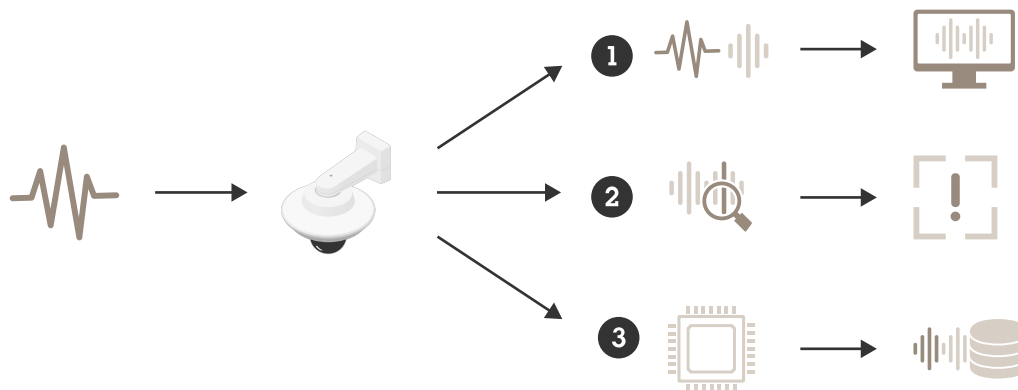
壁、天井、床などの吸音特性や空間の複雑性(非常に長い廊下など)などの音響環境は、特定の場所の音響に深刻な影響を与える可能性のある、さまざまな反響やエコーをもたらします。たとえば、声は、教会や全面タイル張りのバスルームと比較して、防音性の高いエリア(防音処理された会議室など)では、聞こえ方が非常に異なります。音響的に厳しい状況下では、マイクの配置が極めて重要になることがあります。

設備の設置と設置 (音声ゲイン設定など) は共に重要であり、オーディオ機器と監視システムの統合も重要です。システムの設置担当者とインテグレーターは、特定の状況に対する推奨事項を提供することができます。

音声分析の場合、特定の推奨事項が適用されることがありますが、これは一般的な音声録音の推奨事項とは異なる場合があります。かならずユーザードキュメントをよく読み、該当する前提条件を確認してください。

5 音声の準備

最初の音声キャプチャーの後、キャプチャーされた情報は次の処理段階用に準備されま
す。さまざまな準備を並行して行うことも、個々に行うこともできます。



- 1 変換
- 2 リアルタイムでのエッジ分析
- 3 処理およびエンコード

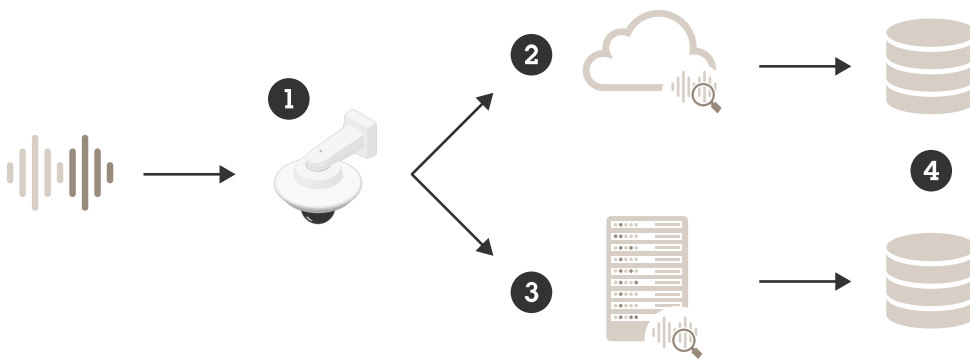
- **変換。** 音が抽象化され、音響スペクトルを示すグラフなどの視覚情報に変換されます。このプロセスを逆にはできません。つまり、スペクトルグラフから元の音声を取り出すことはできません。
- **リアルタイムでのエッジ分析。** 音声をエッジで処理する場合は、**音声分類機能**を使用することができます。これにより、音の特性を説明するメタデータが生成されます。そのメタデータから、元の音声を復元することはできません。**音声検知機能**を使用して、パターン、音量、または周波数を認識し、ステータス情報を提供することができます。この場合も、元の音声を復元することはできません。
- **処理およびエンコード。** 元の音声 (変換または分析されていない) を使用する場合は通常、いくつかの処理とエンコードを実行して目的のユースケースの音声データを準備します。これらのユースケースは、エッジでの音声データの保存、追加処理 (サーバーまたはクラウド上) のための外部クライアントへのストリーミング、または外部ストレージを必要とします。

6 分析機能のトポロジー

システム内の分析エンジンの場所は、多くの理由で重要です。特にプライバシーの問題を管理し、個人データに関する規制を遵守するためには、ソフトウェアアルゴリズムが

どこで音声データを分析するかが重要です。音声データをネットワーク経由で送信できない状況があり、この場合、キャプチャーした(保存されていない)音声データをローカルで分析する必要があります。エッジで実行できないような、計算負荷が非常に高いアルゴリズムが必要な場合は、デジタル音声データをクラウドまたはサーバーに送信する必要があります。

- **エッジ分析。** 分析機能をエッジで実行する場合は、デジタル音声データをカメラから送信する必要はありません。録音を伴わない音声キャプチャーの場合、実行された分析の結果、つまりメタデータやトリガーのみがカメラから送信されます。
- **サーバー分析。** 分析機能をサーバーで実行する場合は、デジタル音声データをカメラから送信する必要があります。カメラ(エッジ)で前処理された場合、このデータは抽象化または匿名化されたメタデータである可能性があります。サーバーは通常、クローズドシステム(システム所有者が管理)の一部であるため、転送される音声のプライバシーに関する問題を管理することができます。この場合でも、適用される規則や規制に準拠していることを確認する必要があります。
- **クラウド分析。** デジタル音声は、クラウド分析でもサーバーに転送することができます。サーバー分析の場合と同様に、音声情報を前処理してメタデータにすることが可能です。クラウドの使用は分散化されていることが多いため、プライバシーの問題に対処し、規制に準拠していることを確認することがさらに重要です。

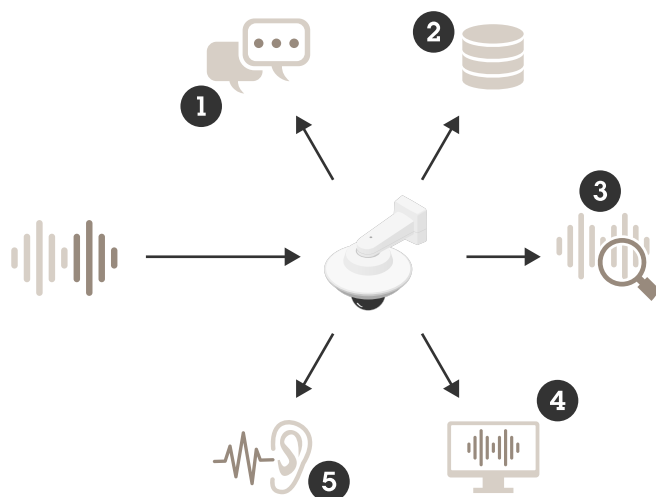


- 1 Edge
- 2 クラウド
- 3 サーバー
- 4 ストレージ

7 ユースケースと例

多くの場合、音声機能はビデオカメラに一体化されており、すぐに使用できます。責任とプロフェッショナルな判断に基づいて「音声入力」を使用することで、重要な価値と複数の潜在的な利益を提供できる、さまざまなユースケースがあります。たとえば、現場検証で欠落している証拠を提示したり、警備員や病院スタッフの速やかな対応を必要とするイ

イベントをリアルタイムで検知したりするために使用できます。音声監視が行われているという事実も、抑止効果をもたらし、犯罪の防止に役立ちます。



音声キャプチャーの一般的な目的:

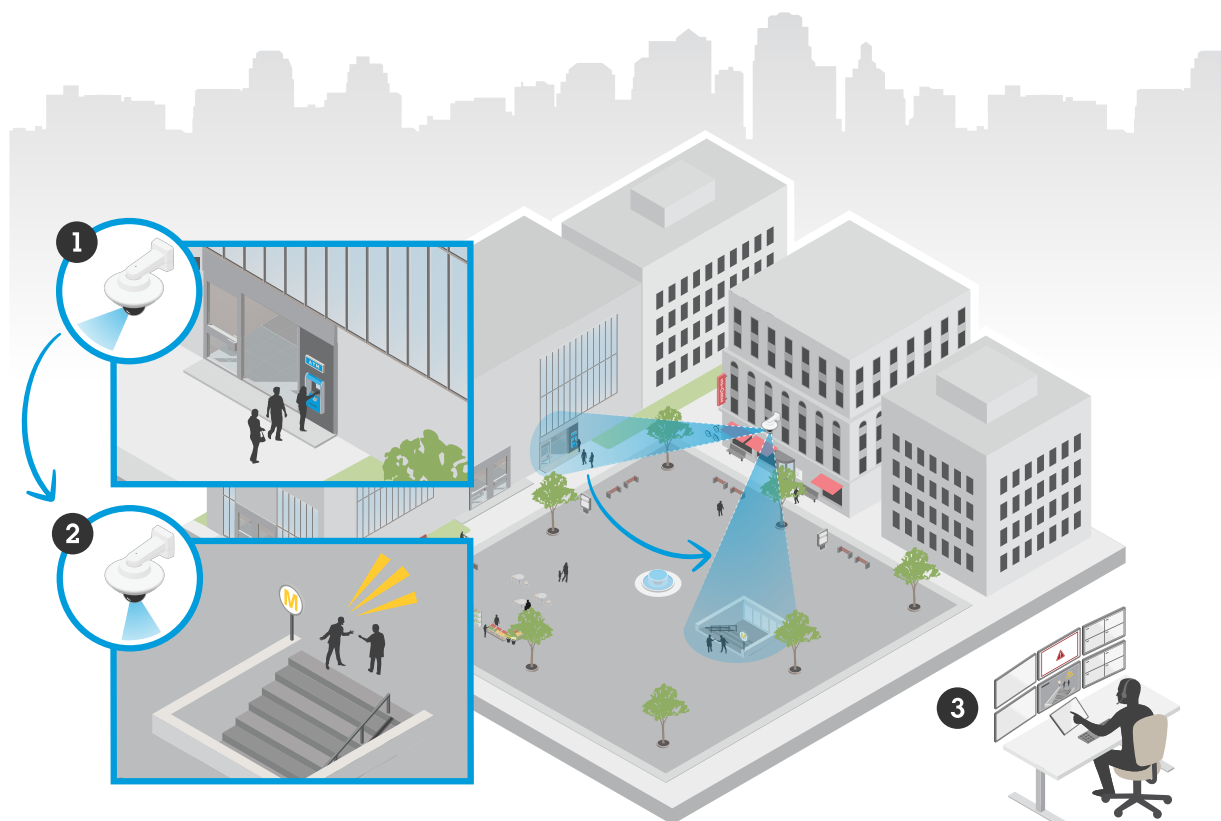
- 1 コミュニケーション
- 2 録音
- 3 分析
- 4 可視化
- 5 聞き取り

7.1 音声分析機能によるインシデントの検知

音声分析アプリケーションは、キャプチャーした音声进行处理して特定の情報を検索・抽出するソフトウェアプログラムです。銃声、ガラスの破損、攻撃などのイベント検知に使用されます。たとえば、入力音声を処理することで、「窓が割れたか？」という質問に「はい」または「いいえ」で答えることができます。イベントを検知すると、システムは通常、視覚的なアラートで、またはアラームを作動させて、スタッフに自動通知を送信します。これにより、迅速な対応と介入を可能にする早期警告を提供します。

7.1.1 カメラの向きの変更

音声分析機能のもう1つの例が、PTZカメラの向きを変更するアプリケーションです。これは、音声はどこから来ているかを検知し、カメラの向きを音源の方に自動的に変更する、音声とビデオの機能を組み合わせたものです。



- 1 PTZカメラがATMを監視しています。
- 2 カメラのマイクが突然発生した大きなノイズを拾い、カメラが即座にインシデントの方向に向きを変えます。
- 3 オペレーターがアラームを受信し、インシデントを検証します。

7.2 ビデオでの音声の可視化

ビデオでキャプチャーした音声は、視覚化して、モニターにスペクトルグラフとして表示することができます。設定しきい値を超えると、グラフはアラームの表示を開始します。

このような音の視覚化は、病院内の病室など、一度に複数の音源からの音を監視する必要がある状況で役立ちます。多数の音源(多数の部屋からの音)を同時に聞くことは困難ですが、ナースステーションのモニターで多数の視覚化情報を見るほうがはるかに簡単

なはずです。部屋からのビデオ画像を利用できる場合、視覚化情報をビデオ画像にオーバーレイとして追加できます。



病院のビデオ画像にオーバーレイとして追加した音の視覚化情報。

7.3 聞き取りと対応

おそらく最も基本的で直感的なユースケースは、オペレーターの直接的な対応によってシーンの状況認識力を高める音声監視です。一般的な例としては、疑わしい会話をとらえることで、警備員を派遣して調査を進めることができます。また、病院や介護施設で、患者が苦痛を感じていないかどうかを聞き取り、看護師を派遣することができます。「奇妙な」音が聞こえる方向を検知し、PTZカメラを音源に向けることも可能です。

これらのユースケースでは、1人または複数のオペレーターが制御室から、またはモバイルデバイスのセキュリティアプリケーションを介して音声環境にアクセスします。人の耳で音をとらえ、脳でシーンや状況に関連する要素を抽出します。映像監視と組み合わせて使用した場合、音声は意思決定のための情報に新たな次元をもたらします。音源がカメラの視野外にある場合や光条件が厳しい場合など、場合によっては、音声だけが唯一の次元になります。

7.3.1 聞き取りと証言

音声監視は、目撃した(聞いた)イベントに基づく直接的な証言を目的として使用することもできます。このユースケースは、意思決定が目的ではないため、*聞き取りと対応*の使用方法とは異なりますが、多くの場合、これらのユースケースは共存します。たとえば、有罪を示す激しい口論を聞いた場合、オペレーターは警備員を派遣するだけでなく、後で聞いたことを証言することもできます。

7.4 録音と保存

音声データをキャプチャーして録音するユースケース (適切な場合) は、非常に有益な追加の証拠を提供することができます。これには、有罪を示す会話や銃声などがあります。録音された音声は、誰が何を言ったか、何発の銃弾が発砲されたかなど、現場検証が必要なイベントの証拠を提供することができます。

現場検証目的で音声を録音する場合、元のデータを保存し、処理しないように注意する必要があります (他の用途では、必要または有益な場合があります)。現場検証用の録音では、どんな処理でも証拠の改ざんが見なされる可能性があります。音声強調アルゴリズムを使用して、録音された会話の可聴性を高めることができます。これにより、証拠能力が高まる場合があります。ただし、このようなアルゴリズムは、後で録音素材のコピーに適用する必要があります。録音データを可能な限り未処理の状態に維持することで、その後の素材の使用方法に対する選択肢を広げておくことができます。

7.5 監視システムを最大限に活用する

監視システムには、多くの場合、複数のタイプのセンサーが組み込まれています。カメラのイメージセンサーは、シーンの視覚的側面を記録するものです。レーダー技術に基づく動体検知器や赤外線放射などの非視覚センサーもよく使用されます。映像監視が適さず、非視覚センサーがスタンドアロンデバイスとして使用されることもあります。しかし、多くの場合、非視覚センサーは、他の種類の情報を追加することによって、カメラによる監視システムを補完するために使用されます。

また、監視システムにオーディオセンサー (マイク) を採用することで、ユースケースの大部分が強化されます。非音声システムに音声機能を追加すると、分析機能またはオペレーターの対話操作によるマルチセンサーの相互作用が可能になります。

「聞き取りと対応」のユースケースはシンプルな例で、ビデオに加えて音声ストリームを受信することで、オペレーターの撮影シーンの状況認識力が大幅に高まります。人を見ているだけでは攻撃的な行動を見つけるのは難しいかもしれませんが、人の声も聞こえればはるかにそれが簡単になります。

もう1つの典型的な例は、ビデオ動体検知などのビデオ分析機能の使用です。低光量など、分析アプリケーションにとって困難な条件がある場合、音声分析機能の存在によって、検知の信頼性が向上することがあります。

8 監視と検知

音声には、監視と音声分析の両方に使用できる複数の種類の情報が含まれています。さまざまな処理と特性評価が、この情報の抽出と絞り込みを支援し、周囲のシステムとの使用と相互作用を促進します。

8.1 音響特性

監視用途において、ラウドネス (大きさ) やピッチ (高さ) などの特性は、重要な情報を含む場合があります。聞こえる時間、移動の有無、音源までの距離などはすべて、聞こえる音について結論を出すときのパズルのピースとなります。音声の監視と検知用のハードウェアとソフトウェアは、同じタイプの情報を処理するように設計されており、音量や時間の経過に伴うさまざまな周波数のエネルギーなどの特性の複雑な組み合わせを「聞き取り」ます。

- **空間情報。**これは、場所、方向、距離などの概念を含む、私たちの周りの物理的な世界に関係しています。空間情報を使用することで、音声キャプチャーをさまざまな方向にフォーカスまたはズームして、より高品質な録音が可能になります。また、分析機能でを使用して、音源の方向や音源までの距離を判断することもできます。
- **時間情報。**時間情報は、動的な意味(時間の経過に伴う変化)と絶対的な意味(発生日時)の両方で重要であり、ビデオなどの他のセンサーからの情報に関連して多く見られます。時間情報は、何がいつ、どのくらいの期間起こったかを把握するための行動分析において重要な役割を果たします。
- **スペクトル情報。**これは、音の高さや、複雑な音の高さの組み合わせなど、周波数に関係します。音声監視で使用されるマイクは、フラットな周波数特性を持つように設計されています。つまり、可聴帯域(20 Hz~20 kHz)内のすべての周波数を均等にキャプチャーしようとしています。これは、人の聴覚系の動作とは異なります。人間は、人の会話で通常発生する周波数を、他の周波数よりも簡単に検知できるためです。
- **振幅情報。**これは、音の強度や大きさについてです。振幅情報は、スペクトル情報を補完し、一緒に使用することで、聞こえてくる音声の構造をイメージすることができます。

8.2 信号処理

音声監視では、信号処理は通常、送信、ストレージ効率、主観的な品質などの改善、または対象となる音の要素の強調や検知に関係します。これは、さまざまな方法で音声を変更または分析するソフトウェアアルゴリズムを介して行われます。

8.2.1 信号の修正

アルゴリズムを使用して、以下のような特定の目的で信号を変更することができます。

- 自動ゲインコントロールで可聴性を高めるなど、信号を改善するため。
- イコライザーで相対周波数成分を変更することなどにより、信号を変更するため。
- 特定の周波数や振幅を削除して信号を制限するため。これには、圧縮によるデータ量の削減や、音声のスクランブルによるプライバシーの確保などが挙げられます。

8.2.2 信号の分析

音声分析機能は、キャプチャーされた(ただし通常は録音されない)音声データを使用し、関連する音響特性を分析して、音声以外の結果を生成します。アプリケーションは基本的に、音声データを別の形式の、より実用的なアセットに変換します。攻撃、銃声、ガラスの破損、車の警報などを検知するために特別に開発された分析アプリケーションがあります。

機械学習アルゴリズムが使用されている場合、大量のデータから学習し、明確にプログラムされていなくても予測できるようになります。音声における1つの例としては、アルゴリズムが何千回もドアが閉まる音を学習すると、その音を確実に検知できるようになります。

8.3 人の聴覚

人の耳は、音声の検知と分析に使用できる最善のツールの1つです。ほとんどのアルゴリズムでは、対応が困難な騒音の激しい環境でも、人の耳と脳は音声を検知して解釈することができます。

耳を使うことで、音がどこから来ているのか、音源が動いているのかなど、シーンから空間情報を引き出すことができます。耳は2つあるため、音が左、右、またはその中間から聞こえてくるのかを聞き取ることができます。耳だけでなく脳も、音が上下前後、どこから来ているのかを聞き取れるよう作られています。脳内の複数の「フィルタリングステップ」は、左右の耳の時間差を処理し、マイクロ秒程度の小さな差異を即座に検知して、特定の種類のイベントを認識できるようにします。私たちは、特に人の声に対して、また、過去に経験した危険に関連する音に関しても、音声信号の分析が可能な非常に発達した能力を備えています。

適切な状況(高音質、ステレオサウンド、過度の遅延がないなど)では、オペレーター(人)が強力な「分析ツール」となり、検知用のハードウェアやソフトウェアを補完することができます。オペレーターは、2つのマイクが装備された音声監視製品を使用して、シーンから音源の場所や移動などの空間情報を引き出すことができます。

9 免責事項

本書とそのコンテンツはAxisの厚意により提供されており、本書に対するすべての権利またはそれに関連する知的財産権(商標、商品名、ロゴタイプ、および同様のマークを含むがこれらに限定されない)は法律によって保護され、本書に対するすべての権利、権原、および/または利益、またはそれに関連する知的財産権は、Axis Communications ABに帰属し、今後も継続して帰属するものとします。

本書は情報提供のみを目的として「現状有姿」で提供されており、いかなる保証も伴いません。本書に記載されている情報は、法的助言を意図したものではありません。本書は、Axis Communications ABおよび/またはその関連会社に、いかなる法的義務も発生させるものではありません。Axis製品に関連するAxis Communications ABおよび/またはその関連会社の義務は、Axisとかかる製品をAxisから直接購入した事業者との間の契約条件のみに従うものとします。

疑義を避けるために記すと、本書の使用、結果、および実行に関するすべてのリスクは、本書のユーザーが負うものとし、Axisは、法律で認められる最大限の限度で、明示的、黙示的、または制定法上のいずれであるかにかかわらず、商品性、特定目的の適合性、権原、非権利侵害性、および製造物責任の黙示的な保証、または本書に関連する提案、仕様、サンプルから生じる保証を含むがこれらに限定されず、あらゆる保証を否認し、除外します。

付録 1 音質の用語

デジタル音声:

デジタル音声とは、デジタル形式で録音されたアナログ音声(多くの場合、マイクでキャプチャーされた音響信号)を指します。デジタル音声では、音声信号の音波は通常、連続したシーケンスの数値サンプルとしてエンコードされます。精度は、エンコーダが記録する有効桁数に依存します。たとえば、CD音声では、1秒間に44,100回サンプリングされ、各サンプルのビット深度は16ビットです。

ノイズ:

ノイズは、ラウドネスレンジのサイレントエンドを定義または制限する不要な(場合によっては避けられない)音です。これは、録音されたソース(部屋のファンなど)から、マイク(例セルフノイズ、振動、風)、配線(例干渉、クロストーク)、キャプチャーデバイス(例セルフノイズ、デジタルサンプリングノイズ)まで、オーディオチェーンのすべての部分によって生成され、すべての合計から一般的にノイズフロアと呼ばれるものが作成されます。

ノイズは通常、定義されたレベル(場合によってはシステムが処理できる最大の音)からノイズフロアまでの全範囲、SNR(信号対雑音比)によって定義されます。

映像で相当するのは、ランダムな(通常は)静的ピクセルパターン、「スノー」として現れる映像ノイズで、暗い画像で見られるものを制限します(ノイズがサイレント信号で聞こえるものを制限するのと同様)。

歪み:

元の「真の」音声から差し引かれる信号の不要な変更すべてを歪みと呼びます(上述のように、ノイズは通常、歪みの仕様から除外されます)。歪みは主観的な品質を低下させ(通常、「良い音」として聞こえる歪みがあります)、客観的な情報コンテンツを曖昧にして特にコンテンツ分析のための信号を聞き取りにくくし、分析機能を低下させます。

THD(全高調波歪み)とIMD(相互変調歪み)は、一般的に歪みを定量化するために使用される2つの値です。

映像における歪みは、色収差、口径食、ブレなどのアーティファクトと関連し、画像の見栄えを悪くするとともに確認できる部分の詳細が制限されます。

サンプル周波数と周波数特性

デジタルシステムでは、音声は1秒間に設定された回数、標本化されます。これをサンプル周波数(通常は1秒あたり8000~48,000回またはHz)と呼びます。信号理論(ナイキスト-シャノンの標本化定理)では、音を適切にキャプチャーするには、サンプル周波数がアナログ信号の望ましいまたは必要な最大周波数の少なくとも2倍である必要があることが示されています。

通常、人の耳は、年齢やその他の要因に応じて、20Hzから約15~20kHzの周波数を認識することができます。大まかに言えば、およそ数百Hzの低周波数帯域は、特定の音の基本周波数(声の基本周波数など)を定義する場合が少なくありませんが、数千Hzを超える高周波数帯域には、より多くの「詳細」が含まれます。

音声の周波数特性は、ビデオの解像度とフレームレートと関連し、設定を低くすると詳細部分が少なくなります。

ビット深度:

音声はサンプリングされるたびに、アナログ値がキャプチャーされ、デジタル値に変換されます。デジタルドメインには無限大がないため、詳細部分の量は定義されたビット深度に制限されます。各ビットは2つの可能値(0または1、低または高など)を表し、定義された振幅範囲(例:選択された電圧または音圧レベル)と組み合わせられて、この範囲の断片を作成します。2ビットは4つの断片を、3ビットは8つの断片をもたらします。簡単に言うと、3ビットでサンプリングされた1ボルトの信号は、分割されて1/8ボルトのステップで表されます。

十分な音質を得るには、少なくとも人の耳には、通常16ビットで十分です(65,536ステップ)。CD音声にはこれが使用されています。分析や、より要求の厳しい使用には、24ビットのほうが適しています。

ビット深度は、ビデオのコントラスト、各ピクセルが再現できる輝度またはクロミナンスの範囲と関連します。

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。