

シーンメタデータのパワーを解放

状況認識、効率、洞察力の向上

1月 2024

概要

映像監視の観点からは、メタデータとはビデオの内容をテキストで説明したものを指します。これには、表示されている対象物体やシーン自体の高レベルの説明が含まれる場合があります。また、車両や衣服の色、正確な位置、進行方向など、対象物体やシーンに関連付けられる属性が含まれる場合もあります。メタデータは、カメラ内で直接、またはビデオ分析を実行できる別のシステムコンポーネントによって、リアルタイムで生成されます。

メタデータにより、イベントの前後状況を把握し、大量の映像を迅速に分類および検索することができます。これにより実現する機能は、大きく以下の3種類に分類されます。

- **イベント発生後のフォレンジック調査。** 検索パラメーターにより、限られた数の候補に検索を絞り込むことができます。こうしたさまざまな検索パラメーターに基づいて、対象の物体やイベントを検索することができます。物体分類データにより、広範囲の詳細を含めた検索が実現します。
- **リアルタイムの使用。** これにより、オペレーターが状況の変化に迅速に対応すること、また意思決定に有益な情報を提供すること、および自動アクションを可能にすることができます。
- **傾向、パターン、洞察の特定。** 統計レポート用のIoTおよび業務効率化プラットフォームは、メタデータを使用することで、訪問者数のカウント、速度の測定、交通量データ、その他の種類の自動データの収集を実現することができます。

カメラによっては、音声をデコードして、音声メタデータを取得することができます。ビデオで対象物体のクラスが検知されラベル付けされるのと同様の方法で、特定の音声パターンを検知してラベル付けすることができます。たとえば、音声認識システムにより、暴言を特定すること、またガラスの割れる音を検知することができます。

ビジュアル、音声、アクティビティ関連、プロセス関連のソースなど、複数の入力からのメタデータを組み合わせることで、個々の入力から個別に得られる情報よりもはるかに多くの洞察を導出することが可能となります。シームレスなメタデータの統合には、オープンプロトコルと業界標準が不可欠です。

目次

1	はじめに	4
2	メタデータとは？	4
3	エッジでのメタデータの生成	4
4	ユースケース	5
	4.1 リアルタイム利用で即座に対応	5
	4.2 フォレンジック検索	5
	4.3 傾向とパターンを特定して洞察を得る	6
5	メタデータの活用方法	6
6	メタデータの配信	7
7	音声メタデータ	9
8	複数のソースからのメタデータを組み合わせる	9

1 はじめに

メタデータは、ビデオからインテリジェンスを収集するための基盤です。シーン内の重要な詳細を記述することで、ビデオコンテンツにデジタル的な意味を付与します。メタデータを使用すると、大量のビデオの中から重要なものを素早く見つけ、評価し、対処することができます。このため、メタデータは効率的なセキュリティ、安全性、ビジネス運営に不可欠な要素となってきています。

このホワイトペーパーでは、監視と運用効率という両方の観点からメタデータについて説明します。メタデータの利点と、それが映像管理システムやその他のアプリケーションでどのように使用されるかについて詳しく説明します。

2 メタデータとは？

メタデータとは、他のデータに関するデータのことです。映像監視の観点でのメタデータは、どんなオブジェクトが表示されているか、シーン自体の高レベルの説明など、映像に何が映っているかをテキストで記述します。これには、車両や衣服の色、正確な位置、進行方向など、オブジェクトやシーンに関連する属性を含めることができます。メタデータは、カメラ内で直接、またはビデオ分析を実行できる別のシステムコンポーネントによって、リアルタイムで生成されます。

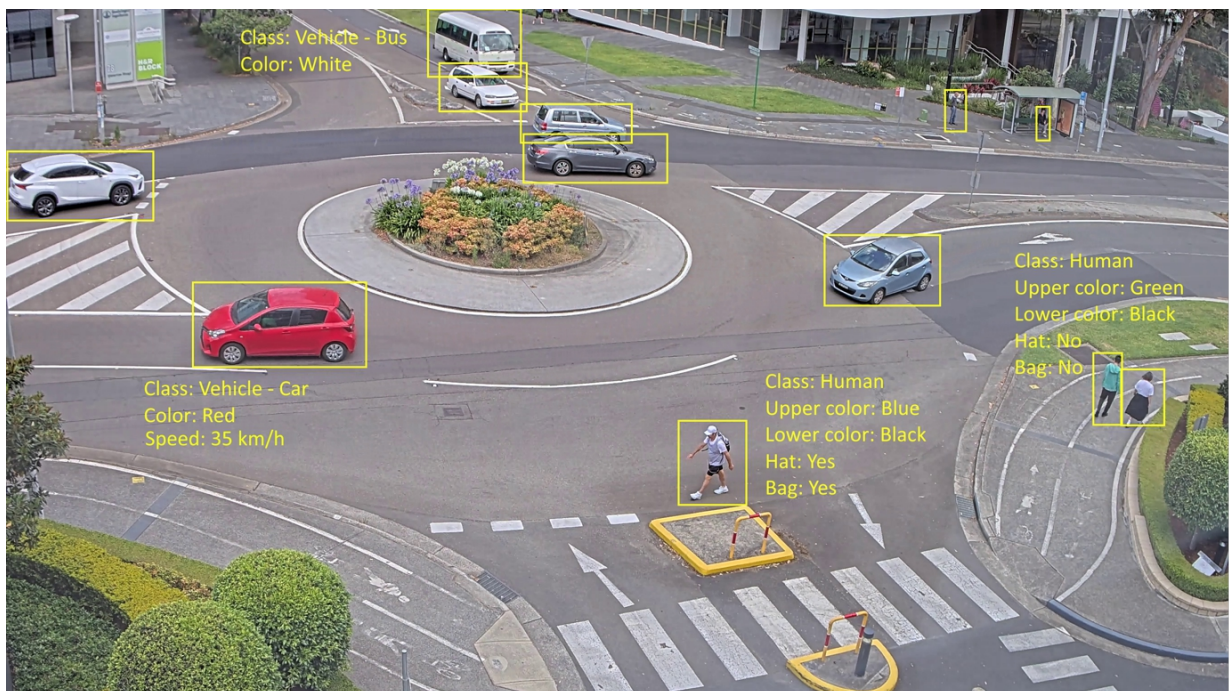


Figure 1. 関心のあるオブジェクトが検知および分析されてメタデータが生成されるビデオフレームの例

3 エッジでのメタデータの生成

高性能なビデオ分析は、一般的にエッジデバイスが提供できる以上の処理能力を必要とするため、サーバーベースで行われていました。しかし、近年では、アルゴリズムの開発とエッジデバイスの処理能力の向上により、エッジで高度な分析を実行できるようになりま

した。これは、メタデータがデバイス内で生成され、他のアナリティクスによってデバイス内で直接使用できることを意味します。ビデオストリームとメタデータストリームは、さらに処理するためにVMSや別のアプリケーションに送信することもできます。

エッジベースの分析では、非常に低い遅延で非圧縮ビデオ素材にアクセスできます。これにより、高速なリアルタイム利用が可能になると同時に、すべてのビデオをシステムの他の場所で処理するための追加コストや複雑さを回避することができます。また、エッジベースの分析は、システム内に必要なサーバーリソースが少なく済むため、ハードウェアと導入のコストも低く抑えることができます。

エッジでメタデータを生成することで、圧縮や送信時に発生し得る情報の損失を心配せずに、ビデオからデータを抽出することができます。これにより、一段と正確なメタデータを生成でき、より正確なビデオコンテンツの分析が実現します。画質が良くなればなるほど、メタデータの質も上がります。

4 ユースケース

メタデータは、シーン内のオブジェクトに関する詳細を提供するだけではありません。イベントのコンテキストを提供し、大量の映像を素早くソートして検索することもできます。これにより、イベント発生後のフォレンジック検索、リアルタイムの使用、傾向・パターン・洞察の識別といったことができるようになります。

4.1 リアルタイム利用で即座に対応

メタデータをリアルタイムで使用すると、オペレーターが状況の変化に迅速に対応できるようになります。また、意思決定をサポートしたり、自動化された行動を有効にするために有益な情報を提供することもできます。高品質のメタデータと連動するリアルタイムのエッジ分析は、人、サイト、建物の安全を確保し、故意または偶発的な危害からそれらを保護するのに役立ちます。脅威を迅速に検知、検証、評価できるため、脅威に効率的に対処できます。

4.2 フォレンジック検索

メタデータを使用すると、関心のあるオブジェクトやイベントを効率的かつ迅速に検索できます。特に複数のビデオソースから膨大な量の映像を検索する場合、メタデータを利用することで、調査にかかる時間を大幅に短縮することができます。さまざまな検索パラメーターに基づいて、人物や車両などのオブジェクトを検索し、限られた数の候補に絞り込むことができます。検索パラメーターには、たとえば、動き、時間、オブジェクトの特徴などがあります。

表 4.1 さまざまなメタデータのカテゴリを使用することで、容易な検索が実現

メタデータのカテゴリ	何を検知するか？	例
動き	オブジェクトの 動き	方向、速度、その他の動作
時間	オブジェクトの 出現時間	曜日、時間帯、滞留時間
場所	オブジェクトの 位置	場所、カメラの視野

表 4.1. さまざまなメタデータのカテゴリを使用することで、容易な検索が実現 (継続中)

メタデータのカテゴリ	何を検知するか？	例
オブジェクトの分類	どのようなオブジェクトか	人物、車両（自動車、バス、トラック、自転車/バイク）
物体の属性	どのような特徴があるか	衣服、帽子やバッグなどのアクセサリ、衣服の色などの身体的特徴

時間など、メタデータの1つのカテゴリにアクセスするだけでも、必要な結果を見つけるために極めて重要であることがわかります。

動きに関するメタデータは、オブジェクトの相対速度と移動方向に基づく検索を可能にします。オブジェクト分類データにより、より広範囲の詳細を含む検索が実現します。ディープラーニングプロセッシングユニット（DLPU）を搭載した搭載カメラは、通常、より詳細なオブジェクト分類を含む充実したメタデータを提供できるため、たとえば緑色のトラックや青いコートを着た人などを検索することができます。

4.3 傾向とパターンを特定して洞察を得る

統計レポート用のIoTおよび業務効率化プラットフォームは、メタデータを使用することで、訪問者数のカウント、速度の測定、交通量データ、その他の種類の自動データの収集を実現することができます。データを分析することで、実用的な洞察を得ることが可能になります。

5 メタデータの活用方法

メタデータを活用して、シーンの特徴や内容を把握することで、多くのメリットがもたらされます。メタデータの主な使用先として、以下が挙げられます。

エッジアプリケーション。カメラで分析を実行することで、シーンの対象物体に関する情報に論理フィルタとルールを適用することができます。そのため、分析により、定義されている閾値や特定の動作に基づいて、アクションをトリガーすることが可能となります。たとえば、シーンの人物の検知と動きに基づいて、PTZカメラを制御する例が挙げられます。

ビデオ管理ソフトウェア（VMS）。一般的に、映像監視においては、シーンに存在する対象物体の周囲に視覚的なオーバーレイを表示するために、メタデータがVMSで使用されています。より高度な物体検知と分類アルゴリズムが開発されたことで、衣服の色といった特定の特徴に基づいて、対象物体の位置を特定できるようになりました。こうしたデータ点を使用して検索クエリを実行できるようになったため、大量の映像を手動で確認する必要性が大幅に削減されました。

IoTプラットフォーム。メタデータを収集してビジネスインテリジェンスプラットフォームに視覚的に表示し、リアルタイム分析を実行する、および過去の傾向を分析することで、実用的な洞察を導出することができます。顧客の流れや顧客体験に基づいた統計分析を取得すれば、データに基づいた意思決定と業務の改善が実現します。

分析の第2層。アプリケーションによっては、より高度な分析を実行する場合に、エッジベースの処理とサーバーベースの処理を組み合わせなければならないことがあります。この場合は、前処理をカメラで実行し、その後の処理をサーバーで実行することができます。

こうしたハイブリッドシステムを用いて、関連性の高いビデオとメタデータのみをサーバーにストリーミングすることで、コスト効率が高く拡張性ある分析が可能となります。

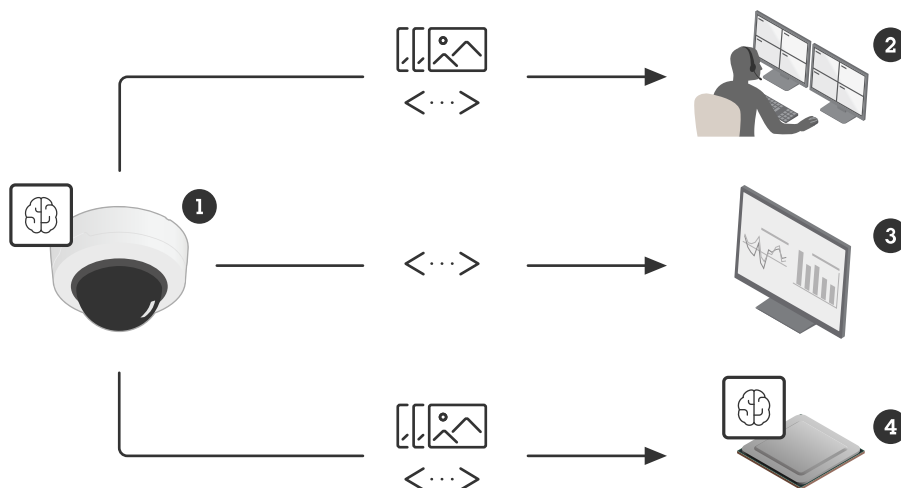


Figure 2. メタデータの使用先

- 1 エッジアプリケーション
- 2 VMS
- 3 IoTプラットフォーム
- 4 分析の第2層

6 メタデータの配信

生成されたメタデータは、使用目的に応じてさまざまなアプローチで配信することができます。リアルタイムアプリケーションの場合は、適切な応答と状況認識を確保するため、オンデマンドでメタデータを受信者/使用者に継続的に配信する必要があります。リアルタイムアクションが必要ない場合で、アプリケーションがそれほど重要でない場合は、受信者/使用者に配信する前にメタデータを統合することができます。たとえば、シーンにおけ

る特定の対象物体の追跡に基づいて、データを統合することができます。これにより、保存および処理する必要のあるデータの総量が削減されます。

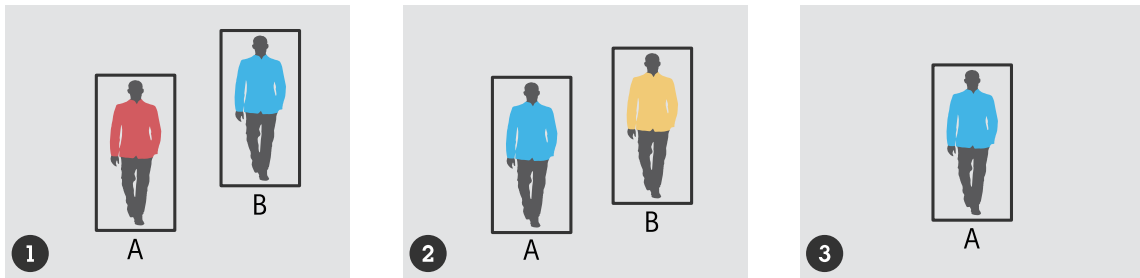


Figure 3. メタデータのストリーミングの一種。上図の場合は、カメラからの連続的なメタデータフレームにより、リアルタイムのシーン情報が提供されます。過去のイベントとは関係なく、各フレームで特定時間のシーンがキャプチャーされています。

- 1 フレーム1では、対象物体Aと対象物体Bが検知され、Aが赤い服を着た人物、Bが青い服を着た人物として分類されています。
- 2 フレーム2では、カメラで分類が更新され、対象物体Aが青い服、対象物体Bが黄色の服を着た人物として判断されています。対象物体はフレーム1と同じままですが、色の属性が変更されています。これがメタデータに反映されます。
- 3 フレーム3では、対象物体Bが不在です。カメラで対象物体Aのみが追跡され、依然として青い服を着た人物として分類されています。

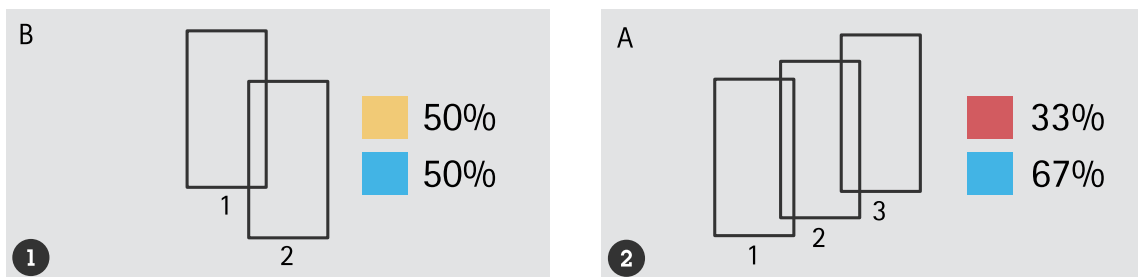


Figure 4. 統合されたメタデータ配信。上図では、シーンで検知された対象物体の追跡に基づいて、カメラで統一された形式で情報が提供されています。各対象物体が映っているフレームには、人物の追跡期間全体にわたる既知の詳細がすべて含まれています。

- 1 最初のフレームでは、最初と最後の検知、軌道の概要、追跡中に検知された属性など、対象物体Bに関する詳細が提示されています。対象物体Bが黄色の服を着ている可能性は50%、青色の服を着ている可能性は50%となっています。
- 2 2番目のフレームには、対象物体Aのこの形式が反映されています。これにより、赤い服を着ている可能性が33%、青い服を着ている可能性が67%であることが分かります。

統合することのメリットとして、シーンに対象物体が存在する場合にのみメタデータが配信されるため、カメラから受信者/使用者に配信されるデータ量が大幅に削減され、要約（統合）の結果として解釈が容易になることが挙げられます。ストリーミング方式の場合は、活動や対象物体が存在しなくても、すべてのフレームにシーンの完全な説明が含まれます。そのため、受信者/使用者は特定のニーズに基づいて、このデータを判断する必要があります。前述のように、ストリーミング方式はリアルタイムのユースケースに有益で、統合方式は受信者/使用者が直ちにアクションを起こす必要のない場合の後処理に最適です。

システムアーキテクチャを設計する上で、それぞれのアプローチの長所と短所を理解することが不可欠となります。たとえば、メタデータに基づいて洞察を導出できるIoTプラットフォームの場合は、シーンの対象物体に関するインシデント後の概要を受信することでメリットが得られます。これは、こうしたサービスは通常、帯域幅とストレージの容量によって制限されるためです。

また、メタデータは、受信者/使用者の特定のニーズや優先事項に基づいて、さまざまな通信プロトコルやファイル形式で配信することができます。

7 音声メタデータ

カメラによっては、音声をデコードして、音声メタデータを取得することができます。音声認識分析では、音声パターンを検出し、ライブ音声や録音音声の中から関心のある音をハイライトすることができます。音声認識システムと映像監視デバイスを組み合わせることで、進行中の潜在的なインシデントをオペレーターに警告し、関連するカメラビューに誘導することができます。このシステムは、たとえば、言葉による攻撃を検知してエスカレーションや暴行を防止したり、ガラスの割れる音を検知して侵入を防止したり、苦しんでいる患者を早期に発見したりすることができます。音声認識システムを活用すれば、オペレーターが現場で発生している事態を目だけでなく耳で確認できるため、早期発見や迅速な介入が可能となり、多くの場合、事態が悪化するのを防止することができます。音声認識システムは、二次的な検証手段としても利用することが可能です。

音声パターンを認識できるように訓練されている分析機能は、一般的に音量レベルからさまざまな周波数のエネルギーに至るまで、その特性の組み合わせを経時的に判断できるように構成されています。ビデオで対象物体のクラスが検知されラベル付けされるのと同様の方法で、特定の音声パターンを検知してラベル付けすることができます。

8 複数のソースからのメタデータを組み合わせる

メタデータの真の可能性は、視覚、音声、アクティビティ関連、プロセス関連の入力などの複数の入力に適応されたときに発揮されます。RFID追跡、GPS座標、いたずら警告、メーター読み取り値（温度や化学薬品レベルなど）、騒音検知、POSトランザクションデータといったデータソースは、あらゆるサイトの管理において貴重です。すべてのソースから得たデータは、タイムスタンプに基づいて整列させることができます。

さまざまなソースからのメタデータを組み合わせると、それぞれのソースから得るよりもはるかに多くの洞察が得られることとなります。シームレスなメタデータの統合には、オープンプロトコルと業界標準が不可欠です。

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。