

Concrétisez le plein potentiel des métadonnées d'une scène

Meilleure appréciation des situations, gains d'efficacité
et analyses plus précises

Janvier 2024

Avant-propos

Dans le contexte de la vidéosurveillance, les métadonnées décrivent en mots le contenu de la vidéo. Elles peuvent identifier les objets d'intérêt visibles ou donner une description générale de la scène elle-même. Elles peuvent également inclure des attributs associés aux objets ou à la scène, par exemple la couleur des véhicules et des vêtements, les positions exactes ou le sens de déplacement. Les métadonnées sont générées en temps réel, soit directement dans la caméra, soit par un autre composant du système capable d'effectuer des analyses vidéo.

Les métadonnées donnent du contexte aux événements et permettent de trier et de rechercher rapidement dans de gros volumes de données. Elles ouvrent ainsi la voie à des fonctions dans trois domaines généraux :

- **Recherche forensique post-événement.** Recherche d'objets ou d'événements d'intérêt selon une diversité de paramètres de recherche, qui restreignent le nombre de résultats. Les données de classification des objets permettent d'effectuer des recherches en fonction d'un choix étendu de détails.
- **Usage en temps réel.** Permet aux opérateurs de réagir rapidement à l'évolution d'une situation ou d'apporter des informations en appui de la prise de décision ou pour déclencher une action automatique.
- **Identification de tendances, de schémas et de renseignements.** Les plateformes IoT et d'efficacité opérationnelle dédiées aux rapports statistiques peuvent s'appuyer sur les métadonnées pour compter des visiteurs, mesurer des vitesses, recueillir des données sur les flux de circulation et effectuer d'autres types de collecte automatique de données.

Certaines caméras peuvent décoder l'audio pour en extraire des métadonnées. Il est ainsi possible de détecter des motifs sonores spécifiques et de les étiqueter dans la vidéo de la même manière que les classes d'objet. Un système de reconnaissance audio peut par exemple identifier les agressions verbales ou détecter les bris de verre.

En combinant les métadonnées issues de plusieurs sources, par exemple visuelles, audio, liées à l'activité ou liées à un procédé, vous obtenez beaucoup plus de renseignements qu'avec chaque source séparément. Or, les protocoles ouverts et les méthodes normalisées sont fondamentaux pour une intégration transparente des métadonnées.

Table des matières

1	Introduction	4
2	En quoi consistent les métadonnées ?	4
3	Génération de métadonnées en périphérie de réseau	4
4	Scénarios d'utilisation	5
	4.1 Usage en temps réel pour action instantanée	5
	4.2 Recherche forensique	5
	4.3 Identification de tendances et de schémas pour affiner les analyses	6
5	Où sont utilisées les métadonnées ?	6
6	Comment les métadonnées sont-elles fournies ?	7
7	Métadonnées audio	9
8	Combinaison de métadonnées issues de plusieurs sources	9

1 Introduction

Les métadonnées sont à la base de la collecte de renseignements à partir de la vidéo. Elles donnent un sens au contenu vidéo en décrivant numériquement les détails clés d'une scène. Avec les métadonnées, vous pouvez rechercher et évaluer rapidement les événements importants dans de grands volumes de vidéo pour prendre sans délai les mesures nécessaires. Pour cette raison, elles se sont imposées comme une composante essentielle de la sécurité, de la protection et des opérations des entreprises.

Ce livre blanc aborde les métadonnées dans le contexte de la surveillance et de l'efficacité opérationnelle. Il détaille leurs avantages et leurs modalités d'utilisation dans les systèmes de gestion vidéo et d'autres applications.

2 En quoi consistent les métadonnées ?

Les métadonnées sont des données se rapportant à d'autres données. Dans le contexte de la vidéosurveillance, les métadonnées décrivent en mots le contenu de la vidéo, par exemple les objets d'intérêt visibles ou une description générale de la scène elle-même. Elles peuvent inclure des attributs associés aux objets ou à la scène, par exemple la couleur des véhicules et des vêtements, les positions exactes ou le sens de déplacement. Les métadonnées sont générées en temps réel, soit directement dans la caméra, soit par un autre composant du système capable d'effectuer des analyses vidéo.

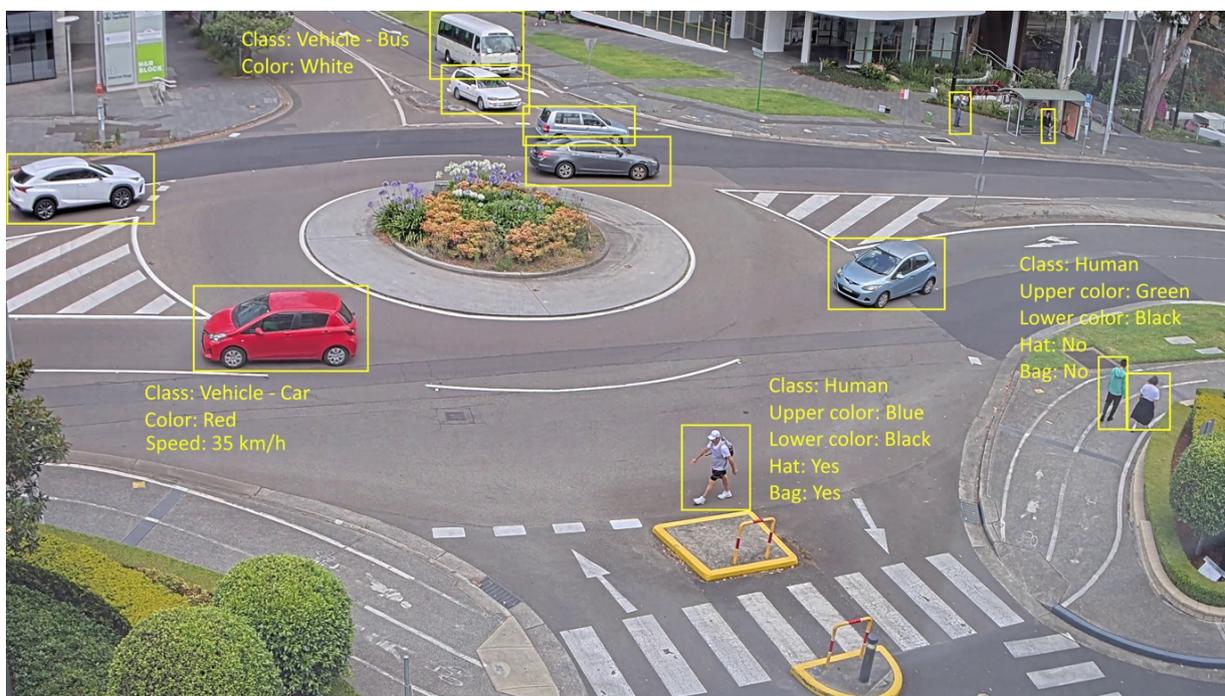


Figure 1. Exemple d'image vidéo où les objets d'intérêt sont détectés et analysés pour produire des métadonnées.

3 Génération de métadonnées en périphérie de réseau

L'analyse vidéo hautes performances était auparavant pratiquée sur serveur, car elle exigeait généralement plus de puissance de traitement que ce que pouvait offrir un dispositif en périphérie de réseau. Mais ces

dernières années, l'amélioration des algorithmes et la hausse des capacités de traitement des dispositifs en périphérie de réseau ont ouvert la voie à l'exécution d'analyses sophistiquées en local. Concrètement, les métadonnées sont produites dans le dispositif lui-même et exploitables directement dans le dispositif par d'autres fonctions d'analyse. Il est également possible de transmettre le flux de données vidéo et le flux de métadonnées au système VMS ou à une autre application pour un traitement plus poussé.

L'analyse en périphérie de réseau a pour atout de traiter la vidéo non compressée avec une très faible latence. Il est dès lors possible d'exécuter des applications en temps réel rapides, tout en évitant les coûts supplémentaires et la complexité de transférer le traitement vidéo ailleurs dans le système. Les analyses en périphérie de réseau réduisent également les frais de matériels et de déploiement, car les ressources serveur nécessaires dans le système sont moindres.

La production de métadonnées en périphérie de réseau permet d'extraire des données de la vidéo sans pertes d'information dues à la compression ou à la transmission. Il en résulte des métadonnées plus précises et une analyse plus fidèle du contenu vidéo. Sans oublier que plus l'image est de bonne qualité, plus les métadonnées sont pertinentes.

4 Scénarios d'utilisation

En plus de fournir des détails sur les objets d'une scène, les métadonnées donnent du contexte aux événements et permettent de trier et de rechercher rapidement de gros volumes de données. Elles ouvrent ainsi la voie à des fonctions dans les domaines généraux de la recherche forensique post-événement, de l'utilisation en temps réel et de l'identification de tendances, de schémas et de renseignements.

4.1 Usage en temps réel pour action instantanée

Les métadonnées sont utilisables en temps réel pour aider les opérateurs à réagir rapidement face à l'évolution d'une situation. Elles peuvent également contribuer efficacement à la prise de décision ou déclencher des actions automatiques. Les analyses en temps réel en périphérie de réseau qui exploitent des métadonnées de haute qualité peuvent vous aider à protéger les personnes, les sites et les bâtiments contre des préjudices intentionnels ou accidentels. Vous pouvez détecter, vérifier et évaluer rapidement les menaces pour les traiter efficacement.

4.2 Recherche forensique

Les métadonnées permettent de rechercher des objets ou des événements d'intérêt de manière précise et rapide. Elles peuvent faire gagner des heures aux enquêteurs, notamment dans les recherches de très grands volumes de vidéo issus de sources multiples. Vous pouvez rechercher des objets, par exemple des humains ou des véhicules, d'après une variété de paramètres de recherche pour restreindre le nombre de résultats. Les paramètres de recherche peuvent porter sur des caractéristiques de déplacement, des critères temporels et des attributs d'objet.

Table 4.1 Les recherches sont facilitées grâce à plusieurs catégories de métadonnées.

Catégorie de métadonnées	Type de détection	Exemples
Mouvement	Comme progresse un objet	Sens de déplacement, vitesse, autre comportement
Date et heure	Quand apparaît un objet	Jour et heure, durée de séjour

Table 4.1. Les recherches sont facilitées grâce à plusieurs catégories de métadonnées. (Suite)

Catégorie de métadonnées	Type de détection	Exemples
Lieu	Où se trouve l'objet	Position, champ de vision de la caméra
Classification des objets	Quel est le type d'objet	Personne, véhicule (voiture, bus, camion, deux-roues)
Attributs des objets	Quelles sont ses caractéristiques	Tenue, accessoires tels que couvre-chef ou sac, caractéristiques physiques telles que couleur des vêtements

Même si vous n'avez accès qu'à une seule catégorie de métadonnées, par exemple la date et l'heure, ce critère peut se révéler crucial pour trouver les résultats nécessaires.

Les métadonnées à propos du mouvement facilitent les recherches basées sur la vitesse relative d'un objet et son sens de déplacement. Les données de classification des objets limitent les recherches à des détails plus précis. Les caméras équipées de processeurs de deep learning (DLPU) peuvent généralement fournir des métadonnées enrichies grâce à une classification plus granulaire des objets. Par exemple, vous pouvez chercher un camion vert ou une personne portant un manteau bleu.

4.3 Identification de tendances et de schémas pour affiner les analyses

Les plateformes IoT et d'efficacité opérationnelle dédiées aux rapports statistiques peuvent s'appuyer sur les métadonnées pour compter des visiteurs, mesurer des vitesses, recueillir des données sur les flux de circulation et effectuer d'autres types de collecte automatique de données. Ces données sont analysées pour produire des informations exploitables.

5 Où sont utilisées les métadonnées ?

L'exploitation des métadonnées présente de nombreux avantages pour comprendre les caractéristiques et le contenu d'une scène. Les principaux consommateurs de métadonnées appartiennent aux catégories ci-dessous.

Applications en périphérie de réseau. Les fonctions d'analyse exécutées sur la caméra peuvent appliquer des règles et des filtres logiques aux informations relatives à un objet de la scène. Par suite, elles peuvent déclencher des actions d'après des seuils prédéfinis ou des comportements spécifiques, par exemple pour commander une caméra PTZ sur détection et déplacement d'un individu dans la scène.

Systèmes de gestion vidéo (VMS). Dans le contexte de la vidéosurveillance, les métadonnées sont couramment employées dans un système VMS pour afficher des incrustations autour d'objets d'intérêt potentiels dans la scène. Avec le développement d'algorithmes plus sophistiqués de détection et de classification des objets, les opérateurs sont désormais en mesure de localiser des objets d'intérêt d'après des caractéristiques particulières, comme la couleur des vêtements. La possibilité d'exécuter des requêtes de recherche avec ces points de données évite dans une large mesure la revue manuelle de grands volumes de vidéo.

Plateformes IoT. Les métadonnées peuvent être agrégées et présentées visuellement dans des plateformes de business intelligence pour produire des informations concrètes par l'analyse des tendances en temps réel

et sur le long terme. Les analyses statistiques basées sur les flux de clients ou l'expérience client facilitent les prises de décision pilotées par les données pour améliorer les opérations.

Analyse de deuxième niveau. Certaines applications nécessitent une combinaison de traitement local et de traitement serveur pour aboutir à des analyses plus fines. Le traitement préalable est réalisable sur la caméra et le traitement ultérieur sur un serveur. Un tel système hybride peut favoriser l'extension des analyses à moindre coût par la seule transmission des séquences vidéo et métadonnées pertinentes au serveur.

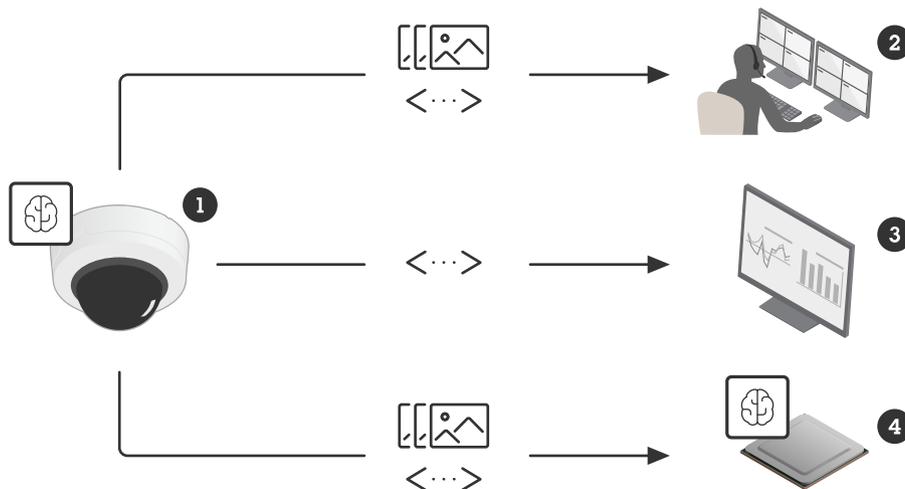


Figure 2. Consommateurs de métadonnées

- 1 Applications en périphérie de réseau
- 2 VMS
- 3 Plateformes IoT
- 4 Analyse de deuxième niveau

6 Comment les métadonnées sont-elles fournies ?

Les métadonnées produites peuvent être transmises selon plusieurs approches, en fonction de leur usage prévu. Dans les applications en temps réel, les métadonnées sont constamment diffusées au consommateur sur demande, cette condition étant essentielle pour garantir une réponse appropriée et mieux appréhender une situation. Dans d'autres applications moins critiques qui n'exigent pas d'action en temps réel, les métadonnées peuvent faire l'objet d'une consolidation complémentaire, basée par exemple sur la trajectoire

de chaque objet particulier de la scène, avant leur transfert au consommateur. Cette méthode réduit le volume total de données à stocker et à traiter.

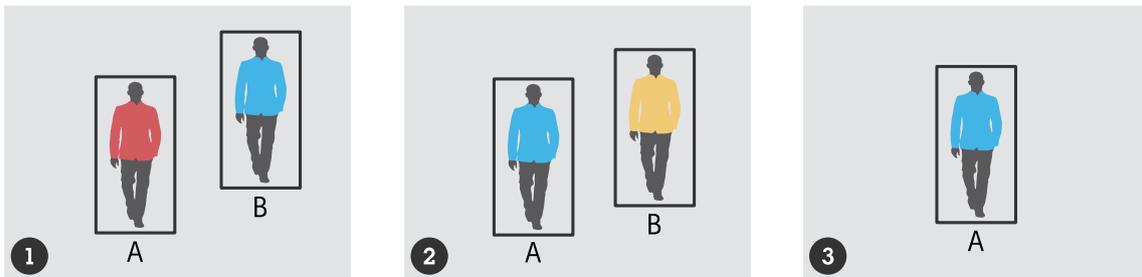


Figure 3. Cette figure illustre la diffusion des métadonnées, dans laquelle les cadres de métadonnées de la caméra fournissent des informations en temps réel à propos de la scène. Chaque image capture la scène à un instant particulier, indépendamment des événements passés.

- 1 L'image 1 détecte les objets A et B, en classant A en tant qu'humain en tenue rouge et B en tant qu'humain en tenue bleue.
- 2 Dans l'image 2, la caméra actualise la classification et détermine que l'objet A porte en réalité une tenue bleue et l'objet B une tenue jaune. Bien que les objets restent identiques à ceux de l'image 1, le changement de leurs attributs de couleur est transposé dans les métadonnées.
- 3 L'image 3 montre l'absence de l'objet B et la caméra qui suit seulement l'objet A, toujours classé en tant qu'humain en tenue bleue.

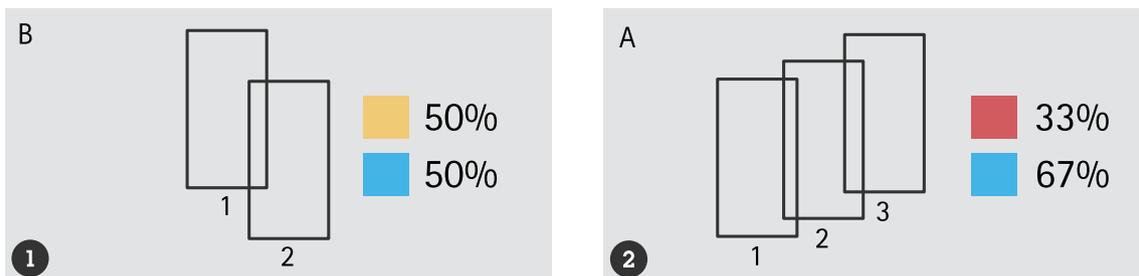


Figure 4. Cette figure illustre la fourniture de métadonnées consolidées, où la caméra transmet des informations dans un format unifié basé sur la trajectoire détectée des objets dans la scène. Les images de chaque objet englobent tous les détails connus sur toute la durée du parcours de l'objet.

- 1 Dans la première image, les détails sur l'objet B sont présentés, notamment sa première et sa dernière détection, le récapitulatif de sa trajectoire et les attributs détectés au cours du suivi. L'objet B porte une tenue jaune avec une probabilité de 50 % et une tenue bleue avec une probabilité de 50 %.
- 2 La deuxième image transpose ce format pour l'objet A, avec une probabilité de 33 % pour une tenue rouge et une probabilité de 67 % pour une tenue bleue.

La méthode consolidée a pour avantage de réduire considérablement le volume de données transmises par la caméra au consommateur. De fait, la caméra fournit des métadonnées uniquement lorsque des objets sont présents dans la scène, auquel cas elles sont résumées (consolidées) pour faciliter l'interprétation. Cette méthode en flux fournit une description complète de la scène dans chaque image, même en l'absence d'activité ou d'objet. Le consommateur doit interpréter ces données d'après sa fonction prévue. Comme indiqué, la méthode en flux est avantageuse pour les scénarios d'utilisation en temps réel, tandis que la méthode consolidée est à privilégier pour le post-traitement, lorsque le consommateur n'a pas besoin de prendre des mesures immédiates.

Il est essentiel de bien comprendre les atouts et les limitations de chaque approche pour la conception de l'architecture du système. Par exemple, une plateforme IoT produisant des analyses basées sur les métadonnées bénéficiera d'un récapitulatif post-incident des objets de la scène, car ces services sont généralement limités en termes de bande passante et de stockage.

Par ailleurs, les métadonnées peuvent être transmises par une diversité de protocoles de communication et de formats de fichier en fonction des besoins et préférences du consommateur destinataire.

7 Métadonnées audio

Certaines caméras peuvent décoder l'audio pour en extraire des métadonnées. Des fonctions de reconnaissance peuvent détecter des motifs sonores et signaler les bruits d'intérêt dans l'audio en direct et enregistrée. De cette manière, les systèmes de reconnaissance audio associés aux caméras de vidéosurveillance peuvent avertir les opérateurs d'incidents potentiels en cours et les orienter vers les vues des caméras correspondantes. Le système peut par exemple identifier les agressions verbales pour éviter toute escalade vers des violences physiques, détecter les bris de verre pour prévenir les effractions ou fournir des alertes précoces de patients en détresse. En donnant les moyens aux opérateurs de voir et d'entendre ce qui se passe dans une scène, les systèmes de reconnaissance sonore peuvent contribuer à une détection précoce, une intervention rapide et, dans de nombreux cas, à désamorcer les situations tendues. La reconnaissance sonore peut également offrir un deuxième moyen de vérification.

Les fonctions d'analyse entraînées pour reconnaître certains motifs sonores écoutent généralement une série de caractéristiques, du niveau sonore à l'évolution de l'énergie des différentes fréquences en fonction du temps. Il est ainsi possible de détecter des motifs sonores spécifiques et de les étiqueter dans la vidéo de la même manière que les classes d'objet.

8 Combinaison de métadonnées issues de plusieurs sources

Le véritable potentiel de métadonnées se concrétise lorsqu'il associe plusieurs sources, par exemple visuelles, audio, liées à l'activité ou liées à un procédé. Suivi RFID, coordonnées GPS, alertes de sabotage, mesures de procédé (température, composition chimique, etc.), détection de bruit, données de transaction des points de vente... Combinées, ces sources de données sont des alliés précieux dans la gestion d'un site. Les données de toutes ces sources peuvent être synchronisées d'après leur horodatage.

La combinaison des métadonnées de diverses sources se traduit par des analyses beaucoup plus poussées qu'avec chaque source séparée. Or, les protocoles ouverts et les méthodes normalisées sont fondamentaux pour une intégration transparente des métadonnées.

À propos d'Axis Communications

En concevant des solutions qui améliorent la sécurité et les performances de l'entreprise, Axis crée un monde plus clairvoyant et plus sûr. En tant qu'entreprise de technologie de réseau et leader de l'industrie, Axis propose des solutions de vidéosurveillance, de contrôle d'accès, d'interphonie et de systèmes audio. Les performances de ces solutions sont améliorées grâce à des applications d'analyse intelligentes et une formation de haute qualité.

Axis emploie près de 4 000 personnes dans plus de 50 pays et collabore avec des partenaires technologiques et d'intégration de systèmes dans le monde entier pour fournir des solutions clients adaptées. Axis a été fondée en 1984 et le siège social se trouve à Lund, en Suède.