

Como tirar proveito do poder dos metadados da cena

Como elevar o nível da consciência situacional, da eficiência e dos insights

Janeiro 2024

Resumo

No contexto do videomonitoramento, os metadados descrevem textualmente o que está no vídeo. Isso pode incluir quais objetos de interesse são visíveis ou uma descrição detalhada da cena em si. Também pode incluir informações relacionadas a atributos associados aos objetos ou à cena, como cores de veículos e roupas, locais exatos ou direção de um trajeto. Os metadados são criados em tempo real. Isso é feito diretamente na câmera ou por outro componente do sistema capaz de executar análises de vídeo.

Além disso, metadados dão contexto a eventos e permitem que grandes quantidades de gravações sejam organizadas e usadas rapidamente em buscas. Assim, esses recursos geram funções que podem ser amplamente categorizadas em três áreas:

- **Investigação forense pós-evento.** Investigação de objetos ou eventos de interesse com base em vários parâmetros que filtram sua busca a um número limitado de candidatos. Os dados de classificação de objetos permitem fazer buscas que envolvem uma ampla gama de detalhes.
- **Utilização em tempo real.** Ajuda os operadores a responderem rapidamente a mudanças situacionais ou fornece informações para ajudar em uma tomada de decisão ou a ativar a reação automatizada.
- **Identificação de tendências, padrões e insights.** As plataformas de IoT e de eficiência operacional que geram relatórios estatísticos podem usar metadados para fazer contagem de visitantes, fazer medição de velocidade, gerar dados de fluxo de tráfego e realizar outros tipos de coleta de dados automatizada.

Algumas câmeras conseguem decodificar áudio e recuperar metadados de áudio. Padrões de som específicos podem ser detectados e classificados de maneira semelhante às classes de objetos detectadas e classificadas em vídeos. Um sistema de reconhecimento de áudio poderia, por exemplo, identificar agressões verbais ou detectar quebras de vidro.

Quando você combina metadados de várias fontes, como fontes visuais, de áudio, relacionadas a atividades e relacionadas a processos, você obtém muito mais insights do que teria obtido de cada fonte isolada. Protocolos abertos e padrões do setor são essenciais para fazer uma integração de metadados de forma tranquila.

Sumário

1	Introdução	4
2	O que são metadados?	4
3	Geração de metadados na borda	4
4	Casos de uso	5
4.1	Uso em tempo real para ações instantâneas	5
4.2	Pesquisa forense	5
4.3	Como identificar tendências e padrões para obter insights	6
5	Onde os metadados são usados?	6
6	Como os metadados são entregues?	7
7	Metadados de áudio	9
8	Combinação de metadados de várias fontes	9

1 Introdução

Os metadados são a base da coleta de dados de inteligência do vídeo. Eles atribuem significado digital ao conteúdo de vídeo, descrevendo os principais detalhes da cena. Usando os metadados, você pode encontrar, avaliar e agir rapidamente, com base em vastas quantidades de vídeo, para lidar com o que é importante. É por isso que os metadados se tornaram uma parte cada vez mais essencial para ter segurança, proteção e operações comerciais eficientes.

Este white paper discute metadados tanto em um contexto de monitoramento quanto em um contexto de eficiência operacional. Ele detalha os benefícios dos metadados e como são usados em sistemas de gerenciamento de vídeo e outros aplicativos.

2 O que são metadados?

Metadados são dados sobre outros dados. No contexto do videomonitoramento, os metadados descrevem textualmente o que está no vídeo, como quais objetos de interesse estão visíveis ou uma descrição detalhada da cena em si. Isso pode incluir informações relacionadas a atributos de objetos ou da cena, como cores de veículos e roupas, locais exatos ou direção de um trajeto. Os metadados são criados em tempo real, seja diretamente na câmera ou por outro componente do sistema capaz de executar análises de vídeo.

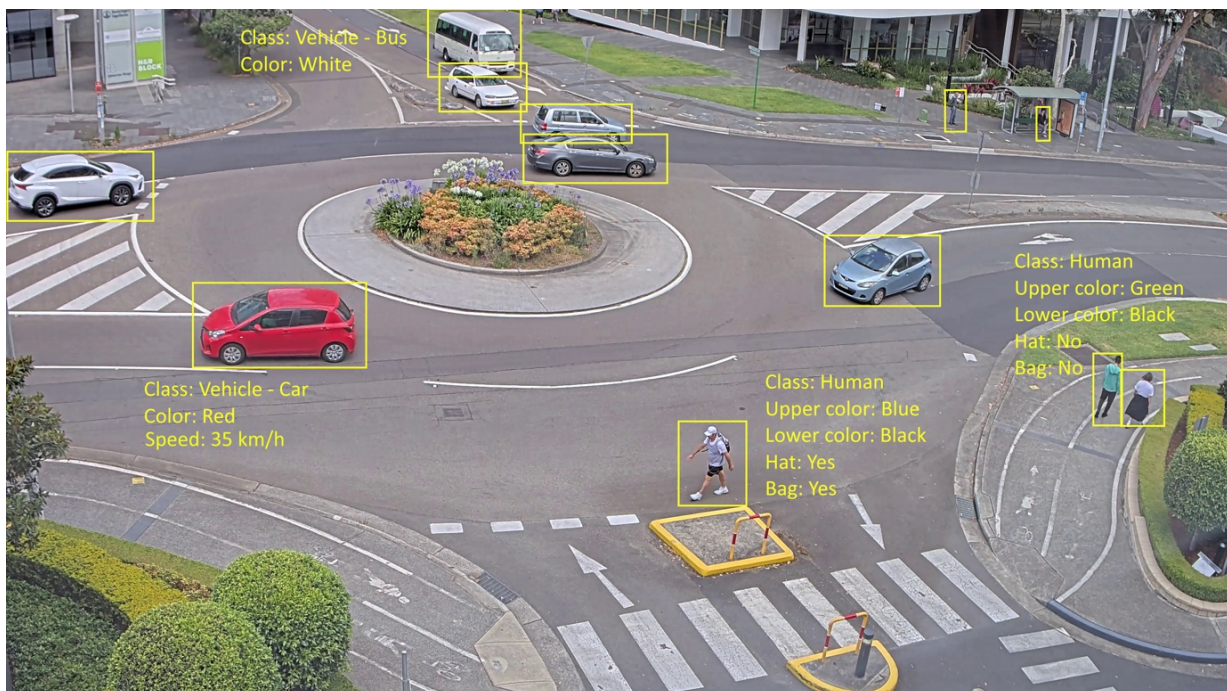


Figure 1. Exemplo de um quadro de um vídeo no qual objetos de interesse são detectados e analisados para criar metadados.

3 Geração de metadados na borda

A análise de vídeos de alto desempenho costumava ser feita no servidor, pois normalmente exigia mais processamento do que um dispositivo na borda era capaz de oferecer. Nos últimos anos, o desenvolvimento de algoritmos e o aumento do poder de processamento dos dispositivos na borda tornaram possível executar análises avançadas na borda. Isso significa que os metadados são gerados no dispositivo e podem

ser usados diretamente nele por outras ferramentas de análises. O fluxo de vídeo e o fluxo de metadados também podem ser transmitidos ao VMS ou outro aplicativo para processamento adicional.

As ferramentas de análise na borda têm acesso a material em formato de vídeo não compactado e com latência muito baixa. Isso possibilita a utilização rápida e em tempo real, além de evitar os custos e complexidades adicionais envolvidos em mover todo o material em vídeo para processamento em outras partes do sistema. Além disso, as ferramentas de análise na borda têm custos mais baixos no que tange hardware e implantação, já que diminuem a necessidade de usar recursos de servidores no sistema.

Gerar metadados na borda significa extrair dados do vídeo sem perder nenhuma informação na compressão ou transmissão. Isso permite obter metadados mais precisos e dados para análise mais exatos sobre o conteúdo do vídeo. Quanto melhor for a qualidade da imagem, melhores são os metadados.

4 Casos de uso

Os metadados não apenas fornecem detalhes sobre objetos em uma cena. Esses dados também dão contexto a eventos e permitem que grandes quantidades de gravações sejam organizadas e usadas em buscas rapidamente. Isso possibilita a existência de funções que podem ser amplamente categorizadas como diferentes áreas de investigação forense pós-evento e para uso em tempo real, identificação de tendências, padrões e insights.

4.1 Uso em tempo real para ações instantâneas

Os metadados podem ser usados em tempo real para ajudar os operadores a responderem rapidamente a mudanças situacionais. Esses dados também podem fornecer informações valiosas para embasar tomadas de decisão ou permitir que haja uma reação automatizada. A análise da borda de tempo real, que trabalha com metadados de alta qualidade, pode ajudar você a proteger pessoas, locais e edifícios de danos intencionais ou acidentais. Detecte, verifique e avalie ameaças rapidamente para que elas possam ser resolvidas com eficiência.

4.2 Pesquisa forense

Os metadados possibilitam que haja buscas eficientes e rápidas de objetos ou eventos de interesse. Isso pode gerar economia de muitas horas para os investigadores, especialmente em buscas envolvendo vastas quantidades de vídeo provenientes de várias fontes. Procure objetos, como seres humanos e veículos, usando vários parâmetros de busca para filtrá-la e gerar um número limitado de candidatos. Os parâmetros da busca podem incluir, por exemplo, características relacionadas a movimento, tempo e objeto.

Tabela 4.1 As buscas podem ser facilitadas se você usar várias categorias de metadados.

Categoria de metadados	O que é detectado?	Exemplos
Movimento	Como um objeto se move	Direção, velocidade, outros comportamentos
Tempo	Quando um objeto aparece	Dia da semana, hora do dia, tempo de permanência
Localização	Onde o objeto está	Local, campo de visão da câmera

Tabela 4.1. As buscas podem ser facilitadas se você usar várias categorias de metadados. (Continuação)

Categoria de metadados	O que é detectado?	Exemplos
Classificação de objetos	Qual é o tipo de objeto	Pessoa, veículo (carro, ônibus, caminhão, bicicleta/motocicleta)
Atributos do objeto	Quais são suas características	Roupas, acessórios como chapéus ou sacos, características físicas, como cor de roupas

Mesmo se você tiver acesso a apenas uma categoria de metadados, como tempo, ela pode ser crucial para você encontrar os resultados necessários.

Os metadados sobre movimento permitem buscas baseadas na velocidade relativa do objeto e na direção do movimento. Os dados de classificação de objetos permitem fazer buscas que envolvem uma gama mais ampla de detalhes. Câmeras com unidades de processamento de deep learning (DLPU) geralmente são capazes fornecer metadados enriquecidos, como uma classificação de objetos mais granular, permitindo que você procure, por exemplo, um caminhão verde ou uma pessoa com um casaco azul.

4.3 Como identificar tendências e padrões para obter insights

As plataformas de IoT e de eficiência operacional que geram relatórios estatísticos podem usar metadados para fazer contagem de visitantes, fazer medição de velocidade, gerar dados de fluxo de tráfego e realizar outros tipos de coleta de dados automatizada. Os dados são analisados para gerar insights úteis.

5 Onde os metadados são usados?

Os metadados geram muitos benefícios, possibilitando entender as características e o conteúdo de uma cena. Os principais consumidores de metadados podem ser categorizados como descrito a seguir.

Aplicativos na borda. As ferramentas de análise sendo executadas na câmera podem aplicar filtros e regras lógicas às informações relacionadas ao objeto na cena. Assim, as ferramentas de análise podem ativar ações baseadas em limites definidos ou comportamentos específicos, como o controle de uma câmera PTZ baseado na detecção e movimento de uma pessoa na cena.

Sistemas de gerenciamento de vídeo (VMS). No contexto do videomonitoramento, os metadados têm sido comumente usados nos VMS para gerar sobreposições visuais em torno de possíveis objetos de interesse na cena. Com o desenvolvimento de algoritmos mais avançados de detecção e classificação de objetos, os operadores agora também são capazes de localizar objetos de interesse com base em características específicas, como a cor das roupas. Ter a capacidade de executar consultas usando esses pontos de dados reduz bastante a necessidade de revisar manualmente grandes quantidades de gravações.

Plataformas de IoT. Os metadados podem ser coletados e exibidos visualmente em plataformas de inteligência de negócios para a geração de insights úteis por meio da análise de tendências históricas e em tempo real. As análises estatísticas baseadas no fluxo de clientes ou na experiência dos clientes possibilitam uma tomada de decisão orientada por dados, o que aprimora as operações.

Segunda camada de ferramentas de análise. Alguns aplicativos precisam usar processamento na borda e no servidor combinados para executar análises mais avançadas. O pré-processamento pode ser executado na câmera, e o processamento adicional, em um servidor. Esse sistema híbrido pode facilitar a ampliação

do uso de ferramentas de análise de forma econômica por meio da transmissão apenas dos vídeos e metadados relevantes para o servidor.

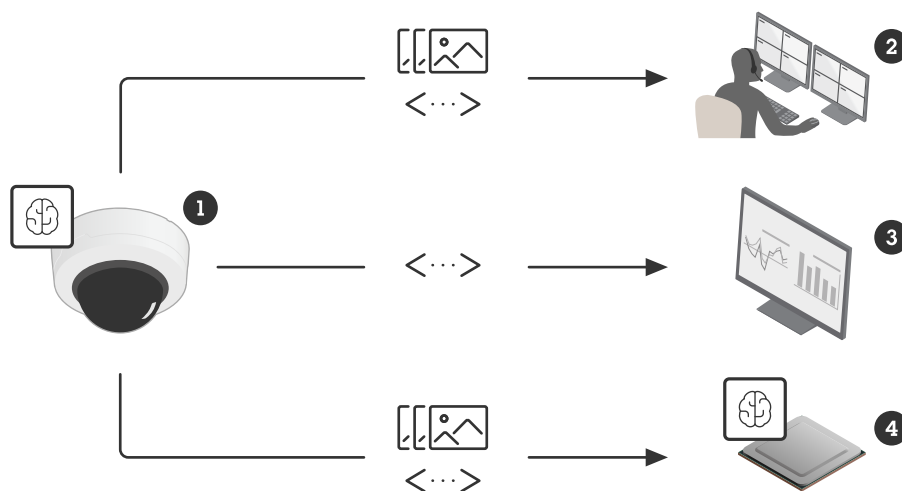


Figure 2. Consumidores de metadados

- 1 Aplicativos na borda
- 2 VMS
- 3 Plataformas de IoT
- 4 Segunda camada de ferramentas de análise

6 Como os metadados são entregues?

Os metadados gerados podem ser entregues por meio de abordagens diferentes, com base no uso pretendido. Em contextos com utilização em tempo real, os metadados precisam ser constantemente transmitidos sob demanda para o consumidor. Isso é essencial para garantir que haja resposta e consciência situacional adequadas. Em outras situações menos críticas, em que não é necessário realizar ações em tempo real, os metadados podem passar por uma consolidação ainda maior, como a aglutinação do trajeto

de cada objeto específico na cena , antes de serem entregues ao consumidor. Isso reduz a quantidade total de dados que precisam ser armazenados e processados.

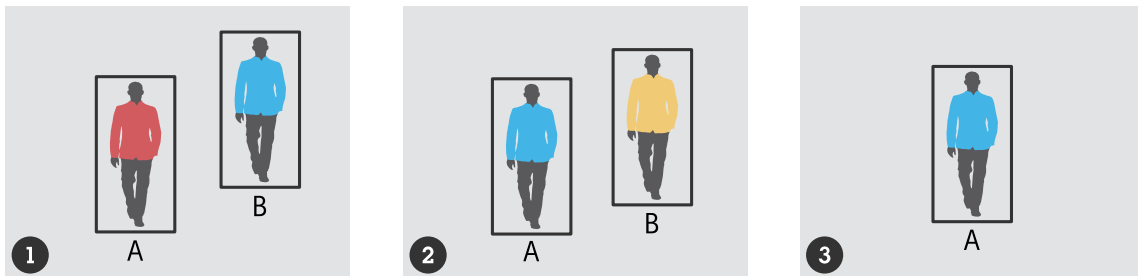


Figure 3. Esta figura ilustra a transmissão de metadados. Aqui, os quadros contínuos de metadados da câmera fornecem informações da cena em tempo real. Cada quadro captura a cena em um momento específico, independente dos eventos anteriores.

- 1 O quadro 1 detecta os objetos A e B. O objeto A é classificado como humano em roupas vermelhas, e o B, como humano em roupas azuis.
- 2 No quadro 2, a câmera atualiza a classificação, determinando que o objeto A realmente usa roupas azuis e que o objeto B usa roupas amarelas. Embora os objetos sejam os mesmos do quadro 1, seus atributos de cor mudam e isso se reflete nos metadados.
- 3 O quadro 3 mostra a ausência do objeto B, pois o rastreamento da câmera está apenas no objeto A, ainda classificado como um humano em roupas azuis.

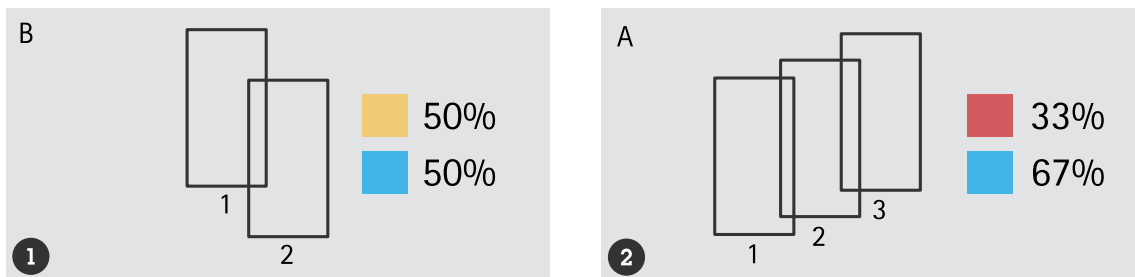


Figure 4. Esta figura demonstra uma entrega de metadados consolidados. Neste processo, a câmera fornece informações em um formato unificado e baseado no trajeto detectado dos objetos na cena. Os quadros referentes a cada objeto abrangem todos os detalhes conhecidos ao longo da vida útil do trajeto do objeto.

- 1 No primeiro quadro, são apresentados detalhes sobre o objeto B, incluindo sua primeira e última detecção, o resumo da sua trajetória e os atributos detectados durante o trajeto. O objeto B tinha uma probabilidade de 50% de estar usando roupas amarelas e uma probabilidade de 50% de estar usando roupas azuis.
- 2 O segundo quadro reflete esse formato para o objeto A, revelando uma probabilidade de 33% de estar usando roupas vermelhas e uma probabilidade de 67% de estar usando roupas azuis.

O benefício do método consolidado é que a câmera reduz significativamente a quantidade de dados enviados ao consumidor, fornecendo metadados somente quando há objetos presentes na cena. Neste caso, os dados estão resumidos (consolidados) para facilitar a interpretação. O método transmitido fornece uma descrição completa da cena englobando todos os quadros, mesmo quando não há atividade ou objetos presentes. O consumidor precisa analisar esses dados de acordo com sua necessidade específica. Como mencionado, o método transmitido é benéfico para utilização em tempo real, enquanto o consolidado é ideal para pós-processamento, quando o consumidor não precisa realizar ações imediatas.

Entender os pontos fortes e as limitações de cada método é essencial na hora de projetar a arquitetura do sistema. Por exemplo, uma plataforma de IoT que gera insights com base nos metadados se beneficiaria do recebimento de um resumo pós-incidente dos objetos na cena, pois esses serviços normalmente são restritos em termos de largura de banda e têm limitações de armazenamento.

Além disso, os metadados podem ser entregues por meio de vários protocolos de comunicação e formatos de arquivo diferentes a depender das necessidades e preferências específicas do consumidor pretendido.

7 Metadados de áudio

Algumas câmeras conseguem decodificar áudio e recuperar metadados de áudio. As ferramentas de análise de reconhecimento de áudio podem detectar padrões de som e destacar sons de interesse em áudio ao vivo e gravado. Assim, os sistemas de reconhecimento de áudio, juntamente com os dispositivos de videomonitoramento, podem alertar os operadores sobre os incidentes em andamento, mostrando a eles as imagens relevantes da câmera. O sistema é capaz de, por exemplo, identificar uma agressão verbal para evitar uma escalada no conflito e agressões, detectar quebras de vidros para evitar arrombamentos ou fornecer avisos precoces de pacientes em sofrimento. Como possibilitam que operadores não apenas vejam, mas também ouçam o que está acontecendo em uma cena, os sistemas de reconhecimento de som podem permitir a detecção precoce, a intervenção rápida e, em muitos casos, a prevenção de uma escalada de um conflito. O reconhecimento de som pode servir também como uma forma secundária de verificação.

As ferramentas de análise treinadas para reconhecer padrões de som costumam captar uma combinação de características, que pode englobar do nível de decibéis à energia em diferentes frequências ao longo do tempo. Padrões de som específicos podem ser detectados e classificados de maneira semelhante às classes de objetos detectadas e classificadas em vídeos.

8 Combinação de metadados de várias fontes

O verdadeiro potencial dos metadados se concretiza quando eles são aplicados a várias fontes, como de dados visuais, de áudio, de atividades e de processos. Fontes de dados como rastreamento de RFID, coordenadas de GPS, alertas de adulteração, leituras de medidores (como temperatura ou níveis químicos), detecção de ruído e dados transacionais de ponto de venda são valiosos no gerenciamento de qualquer local. Os dados de todas as fontes podem ser alinhados de acordo com seus registros de data e hora.

Combinar metadados de fontes diferentes resulta na obtenção de muito mais insights do que se pode obter de cada fonte isolada. Protocolos abertos e padrões do setor são essenciais para fazer uma integração de metadados de forma tranquila.

Sobre a Axis Communications

A Axis torna possível um mundo mais inteligente e seguro criando soluções para melhorar a segurança e o desempenho dos negócios. Como empresa de tecnologia de rede e líder do setor, a Axis oferece soluções em videomonitoramento, controle de acesso, intercomunicação e áudio. Nossas soluções são aprimoradas por aplicativos de análise inteligentes e apoiados por treinamento de alta qualidade.

A Axis tem cerca de 4.000 funcionários dedicados em mais de 50 países e colabora com parceiros de tecnologia e integração de sistemas em todo o mundo para fornecer soluções aos clientes. A Axis foi fundada em 1984 e tem sede em Lund, Suécia