

Le potenzialità dell'“audio in ingresso”

Acquisizione ed elaborazione dell'audio per un quadro migliore della scena e delle prove

Giugno 2021

Sommario

1	Sommario	3
2	Introduzione	4
	2.1 Acquisizione senza registrazione	4
3	Muoversi tra gli ostacoli	4
	3.1 Che cosa dicono le leggi e i regolamenti?	5
	3.2 Indagare sulle possibilità	5
4	Considerazioni sull'installazione	6
5	Preparazione dell'audio	7
6	Topologia delle analitiche	7
7	Casi d'uso ed esempi	8
	7.1 Rilevare eventi con le analitiche audio	9
	7.2 Visualizzazione dell'audio nel video	10
	7.3 Ascolto e interazione	11
	7.4 Registrazione e archiviazione	12
	7.5 Sfruttare al massimo il sistema di sorveglianza	12
8	Monitoraggio e rilevamento	12
	8.1 Caratteristiche sonore	12
	8.2 Elaborazione del segnale	13
	8.3 L'udito	14
9	Dichiarazione di non responsabilità	15
	Appendice Terminologia sulla qualità audio	16

1 Sommario

La funzionalità di acquisizione audio, integrata e pronta per l'uso su una telecamera o disponibile tramite un microfono esterno, può essere utile in molti casi. Un uso responsabile e professionale dell'"audio in ingresso" può offrire vantaggi e un importante valore aggiunto a un sistema di sicurezza. Ad esempio, può fornire la prova mancante in un'indagine forense, o consentire il rilevamento in tempo reale di eventi che richiedono un'attenzione immediata dei vigilanti o del personale ospedaliero. Il semplice fatto che la sorveglianza audio sia in corso può anche avere un effetto deterrente e prevenire i reati.

L'acquisizione audio (spesso combinata a un'analisi) può essere utilizzata come tecnologia standalone, risultando utile in diversi ambiti di prevenzione del crimine, protezione e forensi.

Ma se è abbinata al video, l'acquisizione permette anche di potenziare la videosorveglianza nella maggior parte dei casi. Se il flusso video è integrato da un flusso audio, ad esempio, gli operatori di sicurezza possono avere un quadro nettamente migliore degli eventi di una scena.

Così come è possibile utilizzare diversi tipi di analitiche *video* per il rilevamento automatico degli eventi e gli allarmi a rilevamento visivo, le analitiche *audio* possono monitorare i flussi audio e reagire in caso di fenomeni evidenti.

I software di analisi audio possono essere configurati per attivare allarmi automatici e altre operazioni quando un microfono capta suoni associati a persone che urlano, vetri che si infrangono o colpi d'arma da fuoco. In questo modo, si riceve un avviso tempestivo che consente di reagire e intervenire rapidamente.

Le analitiche audio possono anche rilevare se un suono inatteso provenga da destra o sinistra e puntare automaticamente una telecamera PTZ verso la sorgente sonora. In un ospedale o in una casa di cura, le analitiche audio possono essere utilizzate per captare rumori forti, che implicano che un paziente stia male, e inviare notifiche automatiche a un infermiere. In questo caso, è anche possibile sfruttare le analitiche di visualizzazione audio, che consentono di monitorare facilmente e simultaneamente l'audio in più ambienti.

Tra acquisire i suoni e registrarli c'è una differenza. In molti casi, non è necessario registrare l'audio per raggiungere l'obiettivo: questo può contribuire a rispettare la privacy e i regolamenti sui dati personali. Le applicazioni analitiche, in genere, non registrano l'audio in modo continuo. Normalmente, elaborano solo l'audio in ingresso per cercare frequenze, modelli o livelli specifici. Eseguendo le analitiche in modalità edge (sulla telecamera), nessun dato audio digitale lascia la telecamera: la lasciano solo i risultati delle analitiche, ovvero i metadati o le attivazioni.

Axis non fornisce consulenza legale. Le leggi che regolamentano la sorveglianza variano per regione, stato e Paese. È l'utilizzatore dei prodotti (in genere il cliente finale) ad avere la responsabilità di controllare che la sorveglianza si svolga in maniera conforme. Come nel caso della videosorveglianza, l'installazione della sorveglianza audio deve essere preceduta da uno studio (e dalla comprensione) degli aspetti legali che un sistema del genere può avere.

Una volta intraprese le misure necessarie per rispettare la legge, il sistema deve essere valutato attentamente per quanto riguarda la posizione e la configurazione dei dispositivi. Questo può richiedere una certa preparazione, ma in genere non comporta difficoltà. Con semplici accorgimenti, è possibile avere un'utilizzabilità dell'audio soddisfacente.

2 Introduzione

Le informazioni audio possono essere una risorsa utile per prevenire i reati, per la protezione o per l'uso forense. L'audio acquisito può anche essere processato in tempo reale da un software analitico, consentendo un monitoraggio audio molto efficiente per il rilevamento di attività, comportamenti o eventi.

Questo documento tecnico descrive le potenzialità dell'audio nella sicurezza e alcuni casi d'uso tipici. Vengono presentati vari tipi di analitiche audio con una breve descrizione del loro funzionamento.

Questo documento non fornisce consulenza legale, ma presenta diverse soluzioni tecniche che possono essere utili per configurare un sistema. In base all'implementazione prescelta delle analitiche audio, è possibile consultare le leggi e le normative in vigore a livello locale e utilizzare questo potente strumento ove necessario.

L'ambito di questo documento è limitato all'*acquisizione ed eventuale registrazione* dell'audio, ovvero l'audio *in ingresso*. Un altro uso comune dell'audio nelle soluzioni di sicurezza riguarda la *trasmissione* dell'audio, ovvero l'audio *in uscita*, in genere per riprodurre messaggi vocali/allarmi e dissuadere intrusi o taccheggiatori. Per maggiori informazioni sulla trasmissione dell'audio nella sicurezza, visitare il sito www.axis.com/products/audio.

2.1 Acquisizione senza registrazione

È possibile acquisire e utilizzare l'audio senza registrarlo. Acquisire l'audio significa sostanzialmente digitalizzarlo e renderlo utilizzabile nel software. Queste operazioni vengono svolte registrando le vibrazioni sonore nell'aria con un microfono, convertendo i segnali analogici in segnali digitali (tramite un convertitore A/D) e trasferendoli a un processore.

Se l'audio acquisito non viene trasferito su un supporto permanente, come una memoria flash o un disco rigido, non viene registrato. Le registrazioni potrebbero essere superflue in alcuni casi, ad esempio quando un operatore ascolta l'audio in tempo reale. In alcune situazioni, ci sono motivi specifici per *non* registrare l'audio. Le leggi potrebbero prevedere restrizioni diverse a seconda che l'audio sia registrato o solo acquisito.

In generale, le analitiche audio non registrano l'audio in modo continuo. Per funzionare correttamente, eseguono un buffering temporaneo dell'audio. Molti sistemi possono essere configurati per registrare ciò che è stato bufferizzato poco prima e dopo il rilevamento, per consentire alla sicurezza di verificare ed eventualmente preservare l'audio come prova forense.

3 Muoversi tra gli ostacoli

Molte persone si preoccupano dell'uso dei microfoni nelle situazioni di videosorveglianza. In genere si tratta di preoccupazioni legate alla registrazione di normali conversazioni insieme al materiale video.

È possibile superare questo primo ostacolo comprendendo che "l'audio in ingresso" offre molte più possibilità della semplice registrazione. Esistono molti casi in cui non è necessario registrare informazioni sonore.

Le leggi che regolamentano la sorveglianza variano per regione e Paese, quindi occorre sempre sapere quali operazioni sono consentite prima di aggiungere l'audio a un sistema di sorveglianza.

3.1 Che cosa dicono le leggi e i regolamenti?

Come nel caso della videosorveglianza, l'installazione della sorveglianza audio deve essere preceduta da uno studio (e dalla comprensione) degli aspetti legali che un sistema del genere può avere. Se applicabile, è necessario inviare i documenti pertinenti e ottenere i permessi. Ove necessario devono essere utilizzati appositi cartelli o annunci al pubblico.

L'uso e/o la registrazione dell'audio possono essere vietati o richiedere considerazioni particolari per diversi motivi, dalle leggi nazionali o varie regole e norme locali. Mentre in una regione o in un ambiente l'acquisizione dell'audio può essere consentita, la registrazione potrebbe essere vietata. Anche le aziende possono vietare la sorveglianza audio nei loro locali.

3.1.1 Esempi negli Stati Uniti

Negli Stati Uniti, le leggi e i regolamenti variano da uno stato all'altro.

Alcuni stati richiedono il consenso di una sola parte per registrare l'audio. Questo significa che solo un partecipante a una conversazione deve acconsentire alla sorveglianza affinché sia legale.

Altri stati richiedono il consenso di tutte le parti (o di due parti); ovvero, tutte le parti devono acconsentire ad essere registrate prima che possa tenersi una registrazione. Eccezioni al consenso di tutte le parti possono valere negli spazi pubblici, dove una persona non può pensare di trovarsi in un contesto privato.

La valutazione legale può anche portare a un esito diverso in alcuni territori se si utilizza un'applicazione analitica che non registra l'audio. Pertanto, è necessario conoscere le leggi e le regole in vigore nello stato specifico.

3.1.2 Esempi in Europa

La sorveglianza audio è regolamentata dalle leggi di ogni Paese europeo. Pertanto, è necessario conoscere le leggi e le regole in vigore nel Paese specifico.

Le registrazioni audio possono contenere dati personali, soggetti al GDPR (Regolamento generale sulla protezione dei dati). Il GDPR non vieta necessariamente le registrazioni audio, ma l'acquisizione o la registrazione richiedono considerazioni particolari. Quando si aggiunge l'audio a un sistema di videosorveglianza, occorre valutare se sia ancora valida la base giuridica per il trattamento dei dati personali secondo il GDPR.

3.2 Indagare sulle possibilità

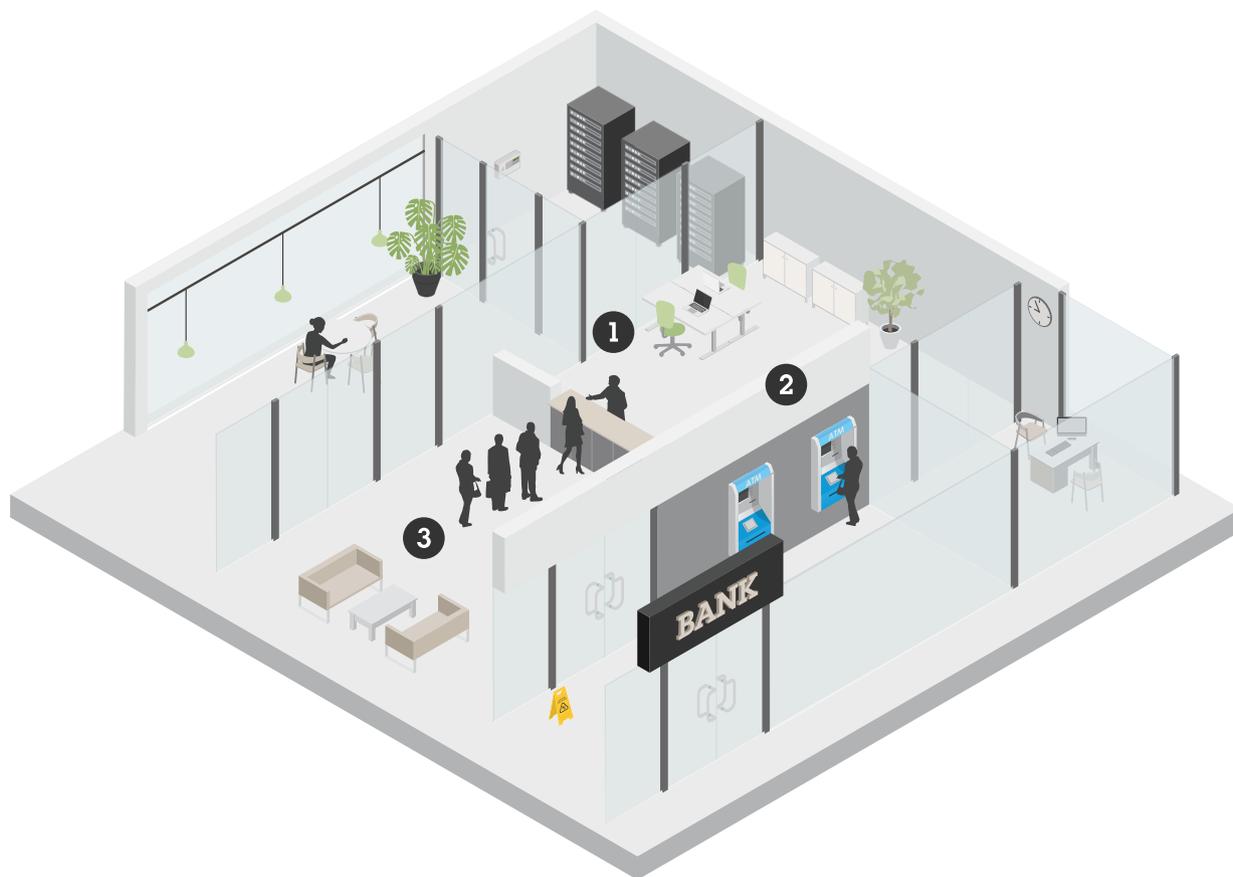
In genere si pensa, e a torto, che l'audio non sia mai consentito nella sorveglianza. Questa idea sbagliata è talmente diffusa che, in molti casi, la possibilità di potenziare un sistema di sorveglianza con l'audio non viene neanche presa in considerazione.

Molti sistemi, però, possono essere consentiti, ad esempio se le persone sono informate, se si ha il loro consenso e così via. Occorre indagare sulle leggi e le regole valide a livello locale e per il caso d'uso specifico. Anche se il caso d'uso *Registrazione e archiviazione* non è consentito in un sistema di sicurezza, molti casi possono essere adattati in modo da non violare la privacy, ad esempio *Ascolto e interazione*, *Ascolto e testimonianza* e *Rilevamento di eventi con le analitiche audio*.

4 Considerazioni sull'installazione

Il posizionamento del microfono in una scena definisce le potenziali applicazioni. Prima di installare dispositivi audio, occorre valutarne attentamente il posizionamento e la configurazione. Questo può richiedere una certa preparazione, ma in genere non comporta difficoltà. Con semplici accorgimenti, l'utilizzabilità dell'audio può aumentare molto.

Valutando il corretto posizionamento di un microfono e scegliendo una buona posizione dal punto di vista acustico, le probabilità di raggiungere gli obiettivi di sorveglianza aumentano. Un microfono, naturalmente, deve essere posizionato in modo da poter acquisire facilmente i suoni rilevanti. Posizioni tipiche sono al centro della stanza, su una telecamera o in prossimità dell'area dove possono svolgersi avvenimenti interessanti. Un microfono non deve essere posizionato vicino a una sorgente sonora, ad esempio un ventilatore o un macchinario, perché possono sovrastare i suoni deboli o lontani.



Posizionamento tipico del microfono

- 1 Dove si svolgono eventi di interesse*
- 2 Su una telecamera*
- 3 Al centro della stanza*

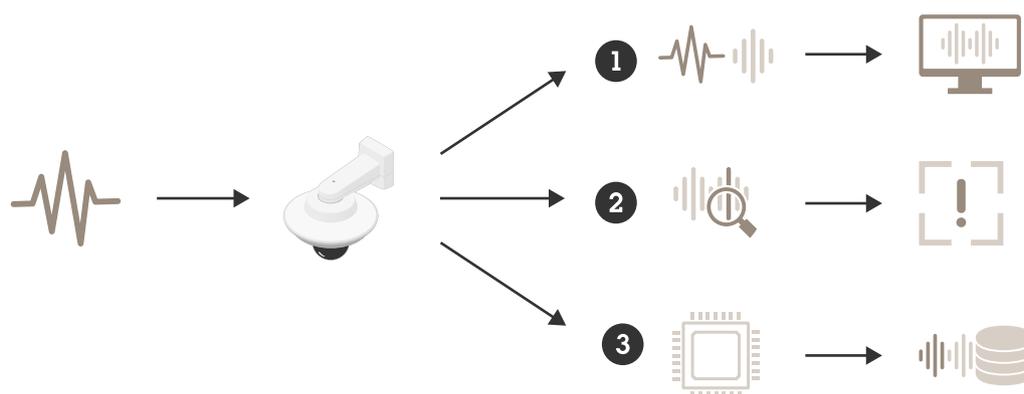
L'ambiente acustico, ad esempio le proprietà di fonoassorbimento delle pareti o del soffitto/pavimento e le complessità dimensionali (es. corridoi molto lunghi), causa echi e riverberi che possono condizionare notevolmente il campo sonoro in alcuni punti. Ad esempio, una voce avrà un suono molto diverso in un'area attenuata (come una sala conferenze trattata acusticamente) rispetto a una chiesa o a un bagno completamente piastrellato. Nelle situazioni difficili dal punto di vista acustico, il posizionamento del microfono può essere fondamentale.

Sono importanti sia l'installazione che la configurazione delle apparecchiature (ad esempio l'impostazione del guadagno), nonché l'integrazione delle apparecchiature audio con il sistema di sorveglianza. In situazioni specifiche, gli installatori e gli integratori di sistemi possono dare consigli utili.

Per le analitiche audio, a volte valgono raccomandazioni specifiche che possono essere molto diverse da quelle adottate per la registrazione audio generica. Occorre sempre studiare la documentazione per l'utente per conoscere i prerequisiti validi.

5 Preparazione dell'audio

Dopo l'acquisizione iniziale, i dati vengono preparati per le fasi di elaborazione successive. I vari preparativi possono essere eseguiti in modo parallelo o esclusivo.



- 1 *Trasformazione*
- 2 *Analitiche edge in tempo reale*
- 3 *Elaborazione e codifica*

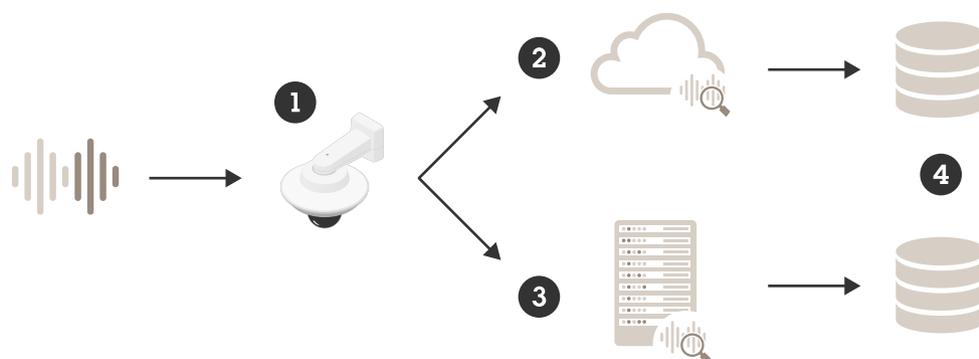
- **Trasformazione.** L'audio viene reso astratto e convertito, ad esempio in informazioni visive come un grafico che mostra lo spettro sonoro. Questo processo non può essere invertito: non è possibile recuperare l'audio originale dal grafico dello spettro.
- **Analitiche edge in tempo reale.**
Se l'audio viene processato in modalità edge, è possibile utilizzare un **classificatore audio**. Il risultato è una serie di metadati che descrive le caratteristiche sonore. L'audio originale non può essere ricreato a partire dai metadati.
È possibile utilizzare un **rilevatore audio** per riconoscere modelli, livelli o frequenze e fornire informazioni di stato. Anche in questo caso, l'audio originale non è ripristinabile.
- **Elaborazione e codifica.** Nei casi in cui si utilizza l'audio originale (non trasformato o analizzato), in genere si eseguono alcune operazioni di elaborazione e codifica, in modo da preparare i dati audio per i casi d'uso previsti. Questi casi d'uso possono prevedere l'archiviazione dei dati audio in modalità edge, lo streaming a client esterni per un'ulteriore elaborazione (su server o cloud) o l'archiviazione esterna.

6 Topologia delle analitiche

La posizione del motore analitico all'interno del sistema è importante per molti motivi. Soprattutto per rispettare la privacy e i regolamenti sui dati personali, è importante sapere *dove* vengono analizzati i dati audio dall'algorithm software. Esistono casi in cui i dati audio non possono essere inviati in rete; dunque,

è fondamentale che i dati audio acquisiti (ma non archiviati) possano essere analizzati localmente. Se gli algoritmi sono molto complessi e non possono essere eseguiti in modalità edge, può essere necessario inviare i dati audio digitali in cloud o a un server.

- **Analitiche edge.** Quando le analitiche vengono eseguite in modalità edge, i dati audio digitali non devono lasciare la telecamera. Nel caso dell'acquisizione audio senza registrazione, vengono inviati solo i risultati delle analitiche, ovvero i metadati o le attivazioni.
- **Analitiche su server.** Se le analitiche vengono eseguite su un server, i dati audio devono lasciare la telecamera. Se pre-processati sulla telecamera (edge), questi dati possono essere resi astratti o depersonalizzati. Normalmente, un server fa parte di un sistema chiuso (controllato dal proprietario), quindi è possibile gestire la privacy dell'audio trasferito. Tuttavia, occorre sempre seguire le regole e le normative applicabili.
- **Analitiche su cloud.** L'audio digitale può anche essere trasferito a un server operativo in un contesto cloud. Come nel caso delle analitiche su server, i dati audio possono essere pre-processati sotto forma di metadati. L'uso del cloud è spesso decentralizzato, quindi è ancora più importante curare la privacy e garantire il rispetto delle regole.

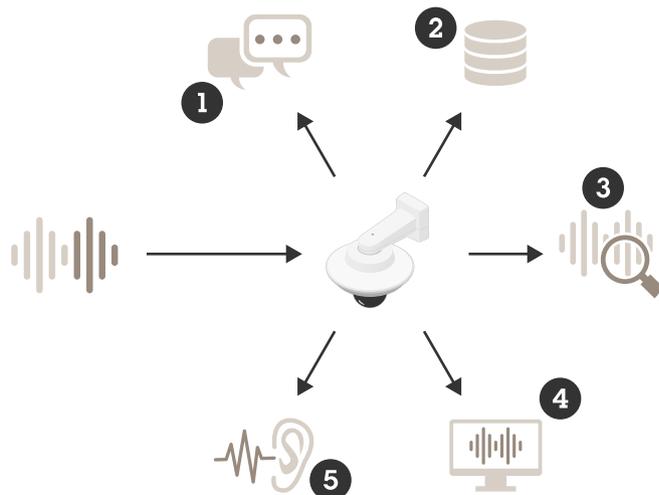


- 1 Edge
- 2 Cloud
- 3 Server
- 4 Archiviazione

7 Casi d'uso ed esempi

Spesso, le funzionalità audio sono integrate e pronte per l'uso sulle telecamere. Esistono vari casi in cui l'uso responsabile e professionale dell'"audio in ingresso" può essere di grande valore e offrire vantaggi. Ad esempio, può essere utilizzato per fornire la prova mancante in un'indagine forense, o consentire il rilevamento in tempo reale di eventi che richiedono un'attenzione immediata dei vigilanti o del personale

ospedaliero. Il fatto che la sorveglianza audio sia in corso può anche avere un effetto deterrente e prevenire i reati.



Finalità tipiche dell'acquisizione audio:

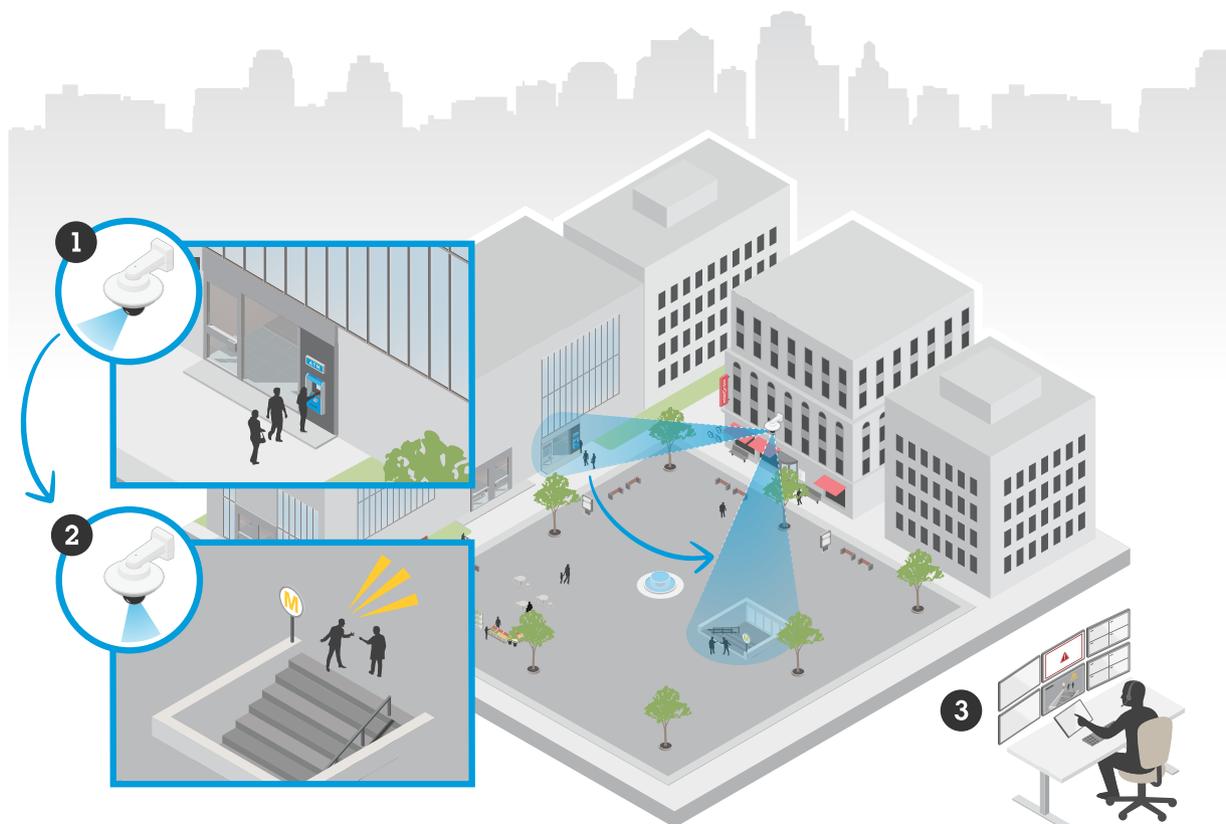
- 1 *Comunicazione*
- 2 *Registrazione*
- 3 *Analisi*
- 4 *Visualizzazione*
- 5 *Ascolto*

7.1 Rilevare eventi con le analitiche audio

Le analitiche audio sono programmi che elaborano l'audio per trovare ed estrarre informazioni specifiche. Vengono utilizzate per rilevare eventi come colpi d'arma da fuoco, vetri che si infrangono o aggressioni. Ad esempio, possono elaborare l'audio in ingresso per rispondere a una domanda come "si è rotta una finestra?" Dopo il rilevamento, in genere il sistema invia una notifica automatica al personale mediante un alert visivo o un allarme. In questo modo, si riceve un avviso tempestivo che consente di reagire e intervenire rapidamente.

7.1.1 Riposizionamento di una telecamera

Un altro esempio di analitica audio è un'applicazione per il riposizionamento di una telecamera PTZ. L'applicazione combina le funzionalità audio e video: dopo aver rilevato da dove proviene l'audio, orienta automaticamente la telecamera verso la sorgente sonora.



- 1 Una telecamera PTZ monitora un bancomat.
- 2 Il microfono della telecamera capta un rumore forte e improvviso e la telecamera viene riposizionata istantaneamente verso l'evento.
- 3 L'operatore riceve un allarme e verifica l'evento.

7.2 Visualizzazione dell'audio nel video

L'audio acquisito in un video può essere visualizzato come spettro sonoro su un monitor. Se si supera una soglia prestabilita, il grafico indica un allarme.

Una visualizzazione dell'audio simile può essere utile quando occorre monitorare suoni provenienti da più sorgenti contemporaneamente, ad esempio da più stanze d'ospedale. Mentre ascoltare tante sorgenti audio (stanze) simultaneamente è difficile, è molto più facile guardare più visualizzazioni su un solo monitor in

sala infermieri. Se sono disponibili i video delle stanze, è possibile aggiungere le visualizzazioni audio in sovrapposizione.



Visualizzazione audio in sovrapposizione nei video di un ospedale.

7.3 Ascolto e interazione

Forse, il caso d'uso più basilare e intuitivo è la sorveglianza audio con interazione diretta con l'operatore per avere un quadro più completo della scena. Un esempio si ha quando si percepisce una conversazione sospetta e si invia un vigilante sul posto per indagare. Oppure, in ospedale o in una casa di cura, quando si sente un paziente che si lamenta e si chiama un infermiere. In altri casi è possibile rilevare se un suono "strano" provenga da destra o sinistra e puntare una telecamera PTZ verso la sorgente sonora.

Questi casi d'uso coinvolgono uno o più operatori, che hanno accesso all'ambiente audio da una sala di controllo o da un'applicazione di sicurezza su dispositivo mobile. L'orecchio capta i suoni, mentre il cervello estrae ciò che è rilevante nella scena o nella situazione. Se utilizzato insieme alla videosorveglianza, l'audio rappresenta una fonte di informazioni in più per il processo decisionale. In alcuni casi, l'audio è l'unica fonte possibile, ad esempio se la sorgente audio è esterna al campo visivo della telecamera o se le condizioni di illuminazione sono difficili.

7.3.1 Ascolto e testimonianza

La sorveglianza audio può anche essere utilizzata per testimonianze dirette degli eventi (visti o ascoltati). Questo caso d'uso è diverso da *Ascolto e interazione* perché lo scopo non è prendere una decisione; spesso, però, i casi d'uso coesistono. Ad esempio, se una lite degenera in parole che possono incriminare, un operatore può non solo inviare una guardia sul posto, ma anche testimoniare successivamente ciò che ha ascoltato.

7.4 Registrazione e archiviazione

Se opportuno, l'acquisizione e la registrazione dei dati audio possono fornire utili prove supplementari, ad esempio se qualcuno ha pronunciato parole che possono incriminare o ha sparato un colpo d'arma da fuoco. L'audio registrato può provare chi ha detto cosa, quanti colpi di pistola sono stati sparati e simili eventi di interesse forense.

Quando l'audio viene registrato in contesto forense, occorre sempre conservare i dati originali ed evitarne l'elaborazione (che in altri casi potrebbe essere necessaria o vantaggiosa). Per le registrazioni forensi, qualsiasi tipo di elaborazione potrebbe essere considerato manomissione delle prove. È possibile utilizzare algoritmi di ottimizzazione vocale per aumentare l'intelligibilità del parlato, incrementandone il valore forense. Questi algoritmi, però, devono essere applicati successivamente, su una copia del materiale registrato. Mantenendo la registrazione il meno elaborata possibile, si avranno più alternative per utilizzare il materiale in seguito.

7.5 Sfruttare al massimo il sistema di sorveglianza

Spesso, i sistemi di sorveglianza integrano vari tipi di sensori. Uno di questi, ovviamente, è il sensore d'immagine della telecamera, che registra l'aspetto visivo di una scena. Inoltre si utilizzano spesso sensori non ottici, come quelli di movimento basati sulla tecnologia radar o sull'emissione di radiazioni a infrarossi. A volte la videosorveglianza non è opportuna, quindi i sensori non ottici vengono utilizzati come dispositivi standalone. In molti casi, però, i sensori non ottici vengono utilizzati per completare il sistema di telecamere aggiungendo altri tipi di informazioni.

Aggiungendo sensori audio (microfoni) a un sistema di sorveglianza, si dispone di un'arma in più in gran parte dei casi d'uso. Aggiungere funzionalità audio a un sistema che non ne è dotato consente l'interazione audio tramite le analitiche o l'operatore.

Il caso d'uso *Ascolto e interazione* è un esempio semplice: ricevendo anche un flusso audio, l'operatore ha un quadro nettamente migliore di quanto accade nella scena. Rilevare comportamenti aggressivi può essere difficile solo osservando le persone, ma è molto più facile se è possibile ascoltarle.

Un altro caso tipico di utilizzo sono le analitiche video, ad esempio per rilevare il movimento. Se l'applicazione analitica si trova a funzionare in condizioni difficili, come una scarsa illuminazione, la presenza di analitiche audio può aumentare l'attendibilità del rilevamento.

8 Monitoraggio e rilevamento

L'audio contiene vari tipi di informazioni che possono essere utilizzate sia per il monitoraggio che per le analitiche. Elaborazioni e caratterizzazioni di vario tipo aiutano a estrarre e rifinire queste informazioni, per un'interazione e un uso più semplici con il sistema circostante.

8.1 Caratteristiche sonore

Caratteristiche quali l'intensità e l'altezza possono contenere informazioni importanti in un contesto di sorveglianza. La durata, il movimento e la lontananza di un suono sono tessere che compongono il mosaico quando traiamo conclusioni su ciò che sentiamo. L'hardware e i software di monitoraggio e rilevamento audio sono progettati per funzionare con le stesse informazioni e "ascoltano" combinazioni complesse di caratteristiche, dal livello di decibel all'energia delle varie frequenze nel tempo.

- **Dati spaziali.** Riguardano il mondo fisico che ci circonda e comprendono parametri come la posizione, la direzione e la distanza. I dati spaziali possono essere utilizzati per concentrare o intensificare l'acquisizione audio in varie direzioni, consentendo registrazioni migliori. Possono anche essere utilizzati dalle analitiche per determinare la direzione di provenienza di un suono o la lontananza della sorgente sonora.
- **Dati temporali.** I dati temporali sono importanti sia in senso dinamico (variazione nel tempo) che in senso assoluto (quando si è verificato un evento?), spesso in relazione a informazioni raccolte da altri sensori (ad esempio video). I dati temporali hanno un ruolo importante nell'analisi comportamentale, per sapere che cosa è accaduto, quando e per quanto tempo.
- **Dati spettrali.** Riguardano le frequenze, ad esempio l'altezza di un suono o, nei suoni più complessi, una combinazione di altezze. I microfoni utilizzati nella sorveglianza audio sono progettati per avere una risposta in frequenza piatta, ovvero cercano di acquisire allo stesso modo tutte le frequenze nello spettro dell'udibile (20 Hz – 20 kHz). Funzionano diversamente dall'udito, perché l'uomo riesce a captare le frequenze del parlato più facilmente rispetto alle altre.
- **Dati sull'ampiezza.** Riguardano la sensazione sonora o l'intensità di un suono. I dati sull'ampiezza possono essere utilizzati insieme ai dati spettrali per disegnare un quadro strutturale dell'audio in ingresso.

8.2 Elaborazione del segnale

Nella sorveglianza audio, l'elaborazione del segnale ha in genere lo scopo di migliorare la trasmissione, l'efficienza di archiviazione o la qualità soggettiva, oppure enfatizzare o rilevare componenti di interesse. Queste operazioni vengono svolte con algoritmi software che modificano o analizzano l'audio in diversi modi.

8.2.1 Modifica dei segnali

Gli algoritmi possono essere utilizzati per modificare il segnale per uno scopo specifico, come:

- migliorare il segnale, ad esempio aumentandone l'intelligibilità tramite il controllo automatico del guadagno;
- alterare il segnale, cambiandone ad esempio la frequenza relativa con un equalizzatore.
- limitare il segnale eliminando frequenze o ampiezze specifiche, per ridurre il volume di dati tramite la compressione o garantire la privacy alterando le voci.

8.2.2 Analisi dei segnali

Le analitiche audio utilizzano i dati acquisiti (ma in genere non registrati) e ne analizzano le caratteristiche acustiche per generare risultati non audio. Sostanzialmente, le applicazioni convertono i dati audio in una risorsa più fruibile e disponibile in un altro formato. Ad esempio, esistono applicazioni analitiche specifiche per rilevare aggressioni, colpi d'arma da fuoco, vetri che si infrangono o antifurto per auto.

Se si utilizzano algoritmi di machine learning, questi possono essere addestrati con grandi quantità di dati e imparare a fare previsioni senza essere programmati esplicitamente. Un esempio in un contesto audio può essere un algoritmo in grado di rilevare con affidabilità il suono di una porta che si chiude dopo essere stato addestrato con migliaia di suoni simili.

8.3 L'udito

L'orecchio è uno dei migliori strumenti che abbiamo a disposizione per captare e analizzare l'audio. Negli ambienti molto rumorosi, l'orecchio e il cervello riescono a captare e interpretare il parlato, mentre la maggior parte degli algoritmi non ne è in grado.

Con l'udito, possiamo estrapolare informazioni spaziali da una scena, ad esempio la direzione di provenienza di un suono, o sapere se la sorgente sonora si muove. Poiché abbiamo due orecchie, possiamo sapere se un suono arriva da sinistra, da destra o da una posizione intermedia. Le orecchie e la testa sono anche strutturate in modo da consentirci di sapere se un suono proviene dall'alto o dal basso, da davanti o dietro. Il cervello applica diversi "stadi di filtraggio" alle differenze temporali che riscontra tra le due orecchie, rilevando istantaneamente le discrepanze (anche di microsecondi) per renderci consapevoli di eventi specifici. Abbiamo una buona capacità di analizzare i segnali sonori; soprattutto le voci, ma anche i suoni storicamente associati a pericoli.

Nelle circostanze giuste (buona qualità audio, suono stereofonico, ritardo non troppo elevato), un operatore può essere un potente "strumento di analisi" che completa l'hardware o il software di rilevamento. Utilizzando un prodotto di sorveglianza audio con due soli microfoni, un operatore può ricavare informazioni spaziali da una scena, capendo ad esempio da dove arriva un suono e il suo movimento.

9 Dichiarazione di non responsabilità

Il presente documento e i suoi contenuti sono forniti per gentile concessione di Axis. Tutti i diritti relativi al documento o i diritti di proprietà intellettuale correlati (compresi a titolo esemplificativo marchi commerciali, denominazioni commerciali, logotipi e marchi simili) sono protetti per legge. Tutti i diritti, i titoli e/o gli interessi in relazione al documento o i diritti di proprietà intellettuale sono e rimangono assegnati ad Axis Communications AB.

Il documento viene fornito nello stato in cui si trova, senza garanzie di alcun tipo e a solo scopo informativo. Le informazioni fornite nel presente documento non costituiscono e non intendono costituire consulenza legale. Il documento non ha lo scopo di creare né crea obblighi legali per Axis Communications AB e/o le sue affiliate. Gli obblighi di Axis Communications AB e/o delle sue affiliate in relazione ai prodotti Axis sono soggetti esclusivamente ai termini e alle condizioni contrattuali tra Axis e l'entità che ha acquistato i suddetti prodotti direttamente da Axis.

PER EVITARE OGNI DUBBIO, TUTTI I RISCHI INERENTI L'USO, I RISULTATI E L'ESECUZIONE DEL PRESENTE DOCUMENTO SONO ASSUNTI DALL'UTILIZZATORE DEL DOCUMENTO. AXIS DECLINA ED ESCLUDE, NELLA MASSIMA MISURA CONSENTITA PER LEGGE, TUTTE LE GARANZIE STATUTARIE, ESPLICITE O IMPLICITE, COMPRESA A TITOLO ESEMPLIFICATIVO GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ, IDONEITÀ PER UNO SCOPO PARTICOLARE, TITOLARITÀ E NON VIOLAZIONE E RESPONSABILITÀ SUI PRODOTTI, O GARANZIE DERIVANTI DA QUALSIASI PROPOSTA, SPECIFICA O ESEMPIO IN RELAZIONE AL PRESENTE DOCUMENTO.

Appendice Terminologia sulla qualità audio

Audio digitale:

L'audio digitale è una rappresentazione dell'audio analogico (spesso un segnale acustico acquisito con microfono) in formato digitale. Negli audio digitale, l'onda sonora del segnale viene in genere codificata con una sequenza continua di campioni numerici. La precisione dipende dal numero di cifre significative registrate dal codificatore. Ad esempio, nell'audio in qualità CD, i campioni vengono acquisiti 44.100 volte al secondo con una profondità di 16 bit.

Rumore:

Il rumore è composto dai suoni indesiderati (e a volte inevitabili) che definiscono o limitano l'estremo della gamma di intensità definito come silenzio. È generato da tutte le parti di una catena sonora: sorgente registrata (ad esempio un ventilatore nella stanza), microfono (es. rumore proprio, vibrazioni, vento), cavi (es. interferenza, diafonia) e dispositivo di acquisizione (es. rumore proprio, rumore di campionamento digitale). Tutti questi elementi si combinano e creano quello che in genere viene detto "rumore di fondo".

In genere, il rumore viene definito con il parametro SNR (rapporto segnale/rumore), ovvero l'intero intervallo da un livello prestabilito (a volte il suono più intenso che il sistema è in grado di gestire) al rumore di fondo.

Quando si parla di immagine, l'equivalente è il rumore video, una disposizione (normalmente) casuale e statica di pixel che causa un "effetto neve". Questo fenomeno limita ciò che si riesce a vedere nelle immagini scure (così come il rumore audio limita ciò che si riesce a sentire per i segnali silenziosi).

Distorsione:

Tutte le alterazioni indesiderate di un segnale sottraggono qualcosa alla "fedeltà" all'originale e vengono dette distorsioni (il rumore, come spiegato sopra, è normalmente escluso). La distorsione riduce la qualità soggettiva (in genere esistono distorsioni che hanno un suono "gradevole") e oscura le informazioni oggettive; rende il segnale più difficile da ascoltare, specialmente per l'analisi dei contenuti, e riduce le funzionalità analitiche.

Due proprietà che vengono normalmente utilizzate per quantificare la distorsione sono la distorsione armonica totale (THD) e la distorsione di intermodulazione (IMD).

Nel video, la distorsione si manifesta con artefatti come aberrazioni cromatiche, vignettature, sfocature, ecc. Fa sembrare un'immagine "di cattiva qualità" e limita i dettagli visibili.

Frequenza di campionamento e risposta in frequenza:

In un sistema digitale, l'audio viene campionato un numero prestabilito di volte al secondo: si tratta della frequenza di campionamento (in genere da 8000 a 48.000 volte al secondo, o Hz). Per acquisire adeguatamente un suono, la teoria dei segnali (e nello specifico il teorema del campionamento di Nyquist Shannon) afferma che la frequenza di campionamento deve essere almeno pari al doppio della frequenza più alta desiderata o richiesta del segnale analogico.

Una persona normale è in grado di percepire le frequenze da 20 Hz a circa 15-20 kHz, a seconda dell'età e di altri fattori. In linea generale, la gamma delle basse frequenze (da un centinaio di Hz in su) definisce la componente di fondo di suoni specifici (come la frequenza fondamentale della voce), mentre la gamma di frequenza più alta, sopra le migliaia di Hz, contiene più "dettagli".

La gamma di frequenza utilizzata nell'audio è paragonabile alla risoluzione e alla velocità in fotogrammi nel video; più è bassa, meno si hanno dettagli.

Profondità di bit:

Ogni volta che l'audio viene campionato, un valore analogico viene acquisito e tradotto in una rappresentazione digitale. Nel dominio digitale non esistono infinitesimi, quindi la quantità di dettagli si limita a una profondità di bit prestabilita. Ogni bit rappresenta un valore binario (0 o 1, alto o basso, ecc.): combinandosi con un intervallo di ampiezza predefinito (es. una tensione o un livello di pressione sonora prestabiliti) crea frazioni di questo intervallo. Due bit generano quattro frazioni, tre bit ne generano otto, e così via. Semplificando, un segnale da 1 V campionato a 3 bit viene suddiviso e rappresentato con incrementi di $1/8$ di V.

Per avere una qualità audio sufficiente, almeno per l'uomo, in genere sono sufficienti 16 bit (che rappresentano 65.536 incrementi). Questo è il valore utilizzato nell'audio in qualità CD. Per le analitiche o utilizzi più complessi, sono più indicati 24 bit.

La profondità di bit è paragonabile al contrasto nel video, ovvero la gamma di luminanza o crominanza che ogni pixel è in grado di riprodurre.

Informazioni su Axis Communications

Axis permette di creare un mondo più intelligente e sicuro grazie a soluzioni di rete che migliorano la sicurezza e forniscono nuove opportunità di business. In qualità di leader nel settore dei video di rete, Axis offre prodotti e servizi per videosorveglianza e analisi dei video, controllo degli accessi, sistemi audio e di interfono. Axis ha oltre 3.800 dipendenti in più di 50 paesi e collabora con partner in tutto il mondo per fornire soluzioni ai clienti. Fondata nel 1984, Axis è una società con sede a Lund, in Svezia.

Per ulteriori informazioni su Axis, si prega di visitare il nostro sito Web axis.com.