

Analizy audio na rzecz optymalizacji ochrony i bezpieczeństwa

Luty 2024

Streszczenie

Narzędzia do analizy audio na potrzeby ochrony i bezpieczeństwa mogą wykrywać wzorce dźwiękowe i specjalnie oznaczać nieoczekiwane dźwięki w strumieniu dźwięku na żywo. Narzędzie analityczne może na przykład identyfikować agresję werbalną, co pozwala na zapobieganie eskalacji i napaści fizycznej, wykrywać stłuczenie szkła, co pozwala na udaremnianie włamań, oraz wysyłać wczesne ostrzeżenia o osobach potrzebujących pomocy.

Stosowanie różnych typów czujników, na przykład czujników wideo i audio (kamera i mikrofon), zwiększa wiarygodność wyników detekcji i pozwala na uzyskanie bardziej przydatnych informacji.

AXIS Audio Analytics to platforma oprogramowania do analiz brzegowych, która jest zintegrowana z oprogramowaniem urządzeń Axis. To narzędzie analityczne przechwytuje i wykrywa dźwięki bez zapisywania pierwotnego strumienia audio. Jest to sposób na ochronę prywatności – i sprawdza się, ponieważ analizy są wykonywane na brzegu sieci i przesyłane są tylko metadane audio.

Spis treści

1	Wprowadzenie	4
2	Technologia	4
	2.1 Wykrywanie zdarzeń dźwiękowych	4
	2.2 Przechwytywanie i przetwarzanie danych audio	4
	2.3 Analizy na brzegu sieci lub na serwerze	5
	2.4 Metadane	6
3	AXIS Audio Analytics	7
	3.1 Prywatność	7
	3.2 Zastosowania	8
4	Ograniczenia prawne	12
5	Wyłączenie odpowiedzialności	12

1 Wprowadzenie

Narzędzia do analizy audio na potrzeby ochrony i bezpieczeństwa mogą wykrywać wzorce dźwiękowe i specjalnie oznaczać nieoczekiwane dźwięki w strumieniu dźwięku na żywo. Narzędzie analityczne może na przykład identyfikować dźwięki związane z agresją, wykrywać stłuczenie szkła lub generować wczesne ostrzeżenia o osobach potrzebujących pomocy.

Analizy audio w połączeniu z dozorem wizyjnym mogą na bieżąco zapewniać operatorom alerty o ewentualnych incydentach, wskazując widok z odpowiedniej kamery. Pozwala to na wczesną detekcję, szybką interwencję i w wielu przypadkach zapobieżenie eskalacji wydarzeń.

W tym dokumencie wyjaśniono, jak analityka audio może się przyczynić do optymalizacji ochrony i bezpieczeństwa. Omawiamy technologie przechwytywania i przetwarzania dźwięku, ze szczególnym uwzględnieniem analityki wykonywanej w czasie rzeczywistym na brzegu sieci, na przykład klasyfikacji dźwięków opartej na sztucznej inteligencji bezpośrednio w kamerze lub mikrofonie. Pokazujemy również, jak analiza audio na brzegu sieci umożliwia wieloaspektową ochronę prywatności dzięki wykorzystaniu metadanych audio.

Dokument nie zawiera porad prawnych. Przed zainstalowaniem jakiegokolwiek systemu dozoru należy zapoznać się z przepisami prawa i innymi regulacjami obowiązującymi w danym regionie i dotyczącymi danego zastosowania systemu. To właściciel systemu odpowiada za zapewnienie jego zgodności z lokalnymi przepisami, regulacjami i zaleceniami.

2 Technologia

2.1 Wykrywanie zdarzeń dźwiękowych

Zdarzenie dźwiękowe to segment dźwięku, który człowiek zwykle identyfikuje jako charakterystyczny odgłos, na przykład krzyk lub stłuczenie szkła. Tego rodzaju odgłosy mogą być wykrywane i oznaczane w podobny sposób jak klasy obiektów podczas analizy materiału wizyjnego.

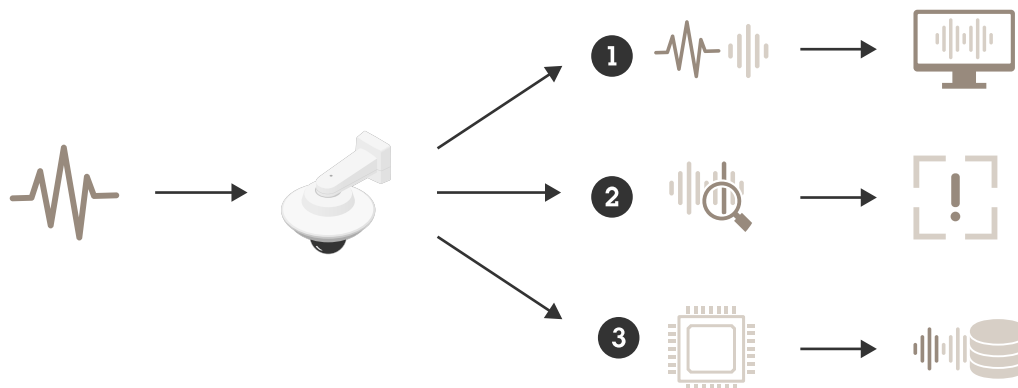
Narzędzia analityczne przeszkolone do rozpoznawania wzorców dźwiękowych zazwyczaj wychwytyją określone połączenia cech, od stopnia natężenia dźwięku po poziom energii na różnych częstotliwościach w funkcji czasu. Po wykryciu określonego wzorca dźwiękowego system może wysłać do pracowników automatyczne powiadomienie w postaci alertu wizualnego lub wyzwolić alarm.

Jeśli wykorzystywane są algorytmy oparte na sztucznej inteligencji, można je szkolić na podstawie dużych ilości danych. Algorytm będzie na przykład potrafił niezawodnie wykrywać ludzkie krzyki po przeszkoleniu z użyciem tysięcy dźwięków tego typu.

2.2 Przechwytywanie i przetwarzanie danych audio

Narzędzia do analiz audio wykorzystują przechwycone dane audio i po analizie ich charakterystyki dźwiękowej generują dane wyjściowe w postaci niedźwiękowej. Zasadniczo przechwytywanie dźwięku oznacza przekształcanie go do postaci cyfrowej i udostępnianie go do użycia w oprogramowaniu. W tym celu mikrofon wychwytyuje drgania dźwięku występujące w powietrzu, a następnie taki sygnał analogowy jest konwertowany na postać cyfrową i przekazywany do modułu przetwarzającego. Jeśli przechwycony materiał dźwiękowy nie zostanie umieszczony na żadnym trwałym nośniku danych, na przykład w pamięci flash lub na dysku twardym, nie zostanie zarejestrowany. W przypadku kamer Axis strumieniowe przesyłanie dźwięku i jego rejestrowanie są domyślnie wyłączone.

Po pierwotnym przechwyceniu dźwięku pozyskane w ten sposób informacje są przygotowywane pod kątem kolejnych etapów przetwarzania. Poszczególne czynności przygotowawcze mogą być wykonywane równoległe lub pojedynczo.



Przygotowanie dźwięku

- 1 Przekształcenie
- 2 Analizy brzegowe w czasie rzeczywistym
- 3 Przetworzenie i zakodowanie w celu przesyłania strumieniowego lub przechowywania – w przypadku korzystania z kamery Axis dźwięk nie jest przesyłany strumieniowo ani przechowywany, chyba że specjalnie włączono strumieniowe przesyłanie dźwięku.

- **Przekształcenie.** Dźwięk zyskuje postać abstrakcyjną i jest konwertowany na informacje wizualne, np. na wykres przedstawiający spektrum dźwięku. Jest to proces nieodwracalny: z wykresu spektrum nie można odtworzyć pierwotnego dźwięku.
- **Analizy brzegowe w czasie rzeczywistym.**

Jeśli dźwięk jest przetwarzany na brzegu sieci, można użyć **klasyfikatora dźwięku**. Wynikiem jego zastosowania są metadane opisujące cechy dźwięku. Z metadanych nie można odtworzyć pierwotnego dźwięku.

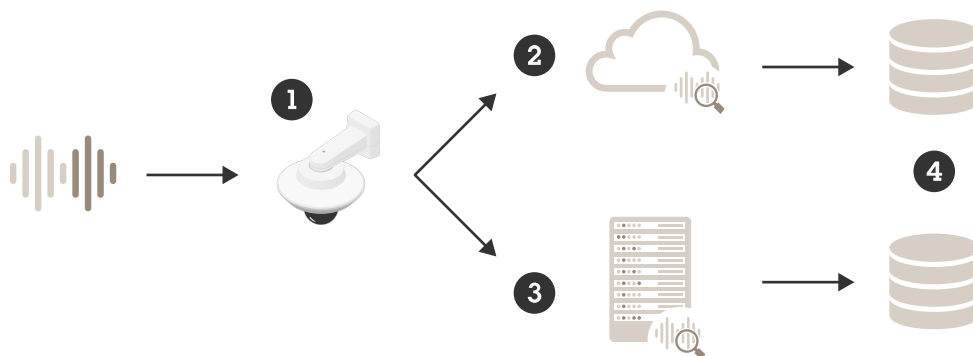
Można także użyć **detektora dźwięku** w celu rozpoznawania wzorców, poziomów lub częstotliwości i uzyskiwania informacji o stanie. Również w tym przypadku nie jest możliwe przywrócenie pierwotnego dźwięku.

- **Przetwarzanie i kodowanie.** W przypadkach, gdy używane są pierwotne dane audio (bez przekształcenia czy analizy), zazwyczaj wykonywane są określone czynności z zakresu przetwarzania i kodowania, które mają na celu przygotowanie danych dźwiękowych do zamierzonych zastosowań. Zastosowania te mogą obejmować przechowywanie danych audio na brzegu sieci, ich strumieniowe przesyłanie do zewnętrznych urządzeń klienckich na potrzeby dodatkowego przetwarzania (na serwerze lub w chmurze) albo przechowywanie w lokalizacji zewnętrznej. W przypadku kamery Axis trzeba najpierw specjalnie włączyć strumieniowe przesyłanie dźwięku, które jest domyślnie wyłączone ze względu na ochronę prywatności (kontrola prywatności dźwięku).

2.3 Analizy na brzegu sieci lub na serwerze

Umiejscowienie mechanizmu analiz w systemie jest ważne z wielu powodów. Zwłaszcza z perspektywy obaw o ochronę prywatności i przestrzegania przepisów dotyczących danych osobowych duże znaczenie ma to, gdzie algorytm analizuje dane audio. Są sytuacje, gdy danych dźwiękowych nie można przesłać

przez sieć i przechwycone (ale niezachowane) dane audio muszą być analizowane lokalnie. Jeśli stosowane algorytmy wymagają dużej mocy obliczeniowej i nie mogą być wykonywane na brzegu sieci, może być konieczne wysłanie cyfrowych danych audio do chmury lub na serwer.



Możliwe miejsca prowadzenia analizy

- 1 *Brzeg sieci*
- 2 *Chmura*
- 3 *Serwer*
- 4 *Pamięć masowa – pierwotny dźwięk może być zapisywany tylko wtedy, gdy jest włączone strumieniowe przesyłanie dźwięku.*

- **Analizy brzegowe.** Jeśli analizy są wykonywane na brzegu sieci, strumień audio w ogóle nie musi opuszczać kamery. Przesyłane są jedynie dane wyjściowe analizy, czyli metadane lub zdarzenia wyzwalające alarm. W przypadku rozwiązania AXIS Audio Analytics analiza jest wykonywana na brzegu sieci.
- **Analizy serwerowe.** Jeśli analiza jest wykonywana na serwerze, kamera musi przysyłać do niego całe dane audio. W przypadku wstępnego przetwarzania w kamerze transmisja ta może się ograniczać do abstrakcyjnej lub zdepersonalizowanej postaci metadanych. Serwer jest zwykle elementem systemu zamkniętego (tj. właściciel systemu ma nad nim kontrolę), więc można zarządzać aspektami ochrony prywatności związanymi z przesyłaniem danych audio. Niemniej jednak należy zadbać o przestrzeganie obowiązujących przepisów i regulacji.
- **Analizy w chmurze.** Dane audio mogą też być przesyłane do serwera w środowisku chmurowym. Tak jak w przypadku analiz serwerowych informacje audio mogą być wstępnie przetwarzane do postaci metadanych. Chmura ma często charakter zdecentralizowany, więc tym bardziej należy wyjaśnić kwestie dotyczące prywatności i zapewnić przestrzeganie przepisów.

2.4 Metadane

Narzędzia do analizy audio generują stały strumień metadanych na podstawie danych głośności audio. Generowane są również metadane na podstawie zdarzeń wykrytych przez klasyfikatory i adaptacyjną audio detekcję. Narzędzia analityczne działające na brzegu sieci analizują informacje audio wewnątrz kamery. Nie muszą one nigdzie przysyłać strumienia audio – mogą przysyłać tylko metadane, które dostarczają informacji o tym, co się dzieje na dozorowanym obszarze.

Strumień metadanych umożliwia kontrolę spektrum dźwięku i jego reprezentację wizualną, dzięki czemu zdarzenia audio i wideo mogą być zharmonizowane i łącznie obserwowane. Tak właśnie metadane umożliwiają skuteczne i szybkie wyszukiwanie określonych zdarzeń oraz nieoczekiwanych dźwięków. Dzięki

temu śledczy mogą zaoszczędzić wiele godzin, zwłaszcza w przypadku przeszukiwania ogromnych ilości zarejestrowanego materiału.

3 AXIS Audio Analytics

AXIS Audio Analytics to platforma oprogramowania, w ramach której stale będą wprowadzane nowe i ulepszone mechanizmy analiz oraz funkcje. To narzędzie analityczne jest zintegrowane z oprogramowaniem urządzenia AXIS OS i jest bezpłatnie dołączane do kamer oraz innych urządzeń Axis.

Analizy w ramach tej platformy są wykonywane na brzegu sieci – algorytmy działają bezpośrednio w kamerze. Zapewnia to optymalną skalowalność, niski poziom transmisji danych i prywatność. Zapisywane są tylko dane wyjściowe analizy (czyli metadane lub wyzwalacze). Nie jest rejestrowany ani przesyłany strumieniowo z kamery żaden materiał dźwiękowy, a pierwotnego dźwięku nie można zrekonstruować na podstawie metadanych.

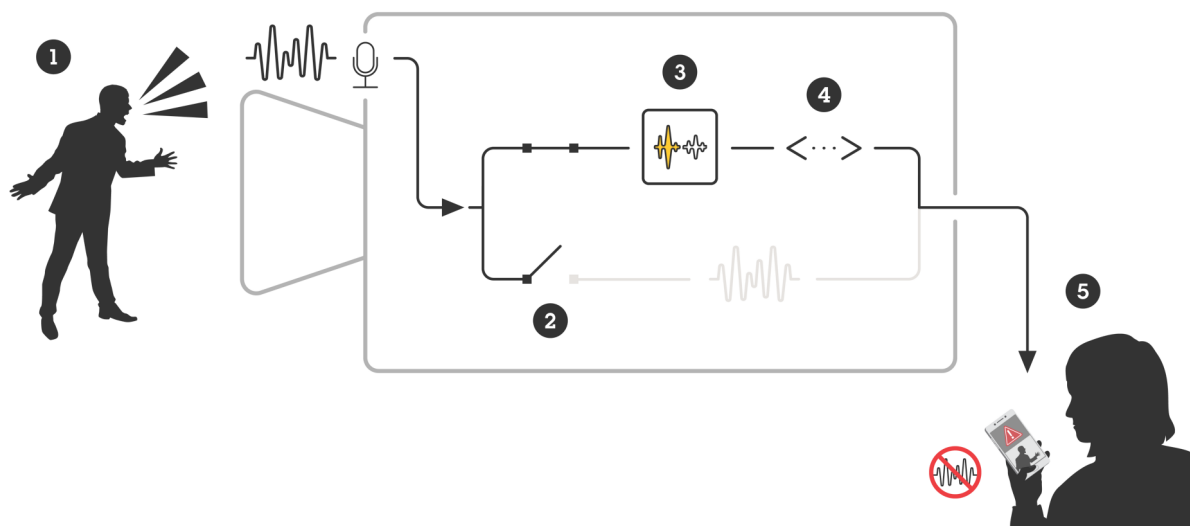
- **Klasyfikacja dźwięku (krzyk).** Jest to klasyfikator dźwięku oparty na sztucznej inteligencji, który wykrywa i analizuje określone dźwięki, na przykład krzyki. Więcej klasyfikacji dźwięku, takich jak stłuczenie szkła, zostanie dodanych w przyszłości. Wynikiem analizy z klasyfikacją są metadane opisujące charakterystykę dźwięku.
- **Adaptacyjna audio detekcja.** Jest to detektor dźwięku, który generuje zdarzenie w przypadku nagłej zmiany poziomu dźwięku. Wykrywa skoki natężenia dźwięku dowolnego rodzaju, a ma przy tym tę zaletę, że dostosowuje się do poziomu hałasu otoczenia, nawet jeśli ten poziom jest zmienny.

3.1 Prywatność

Narzędzia do analiz audio generalnie nie rejestrują dźwięku przychodzącego ani nie wysyłają go z kamery. One po prostu przetwarzają dźwięki, aby umożliwić wyszukiwanie określonych zdarzeń, wzorców lub poziomów dźwięku w systemie odbiorczym, na przykład na pulpicie w ramach prac wyjaśniających lub w oprogramowaniu do zarządzania materiałem wizyjnym w celu generowania alertów dla operatorów. Nie można zrekonstruować żadnych pierwotnych danych audio ani zarejestrować żadnej prywatnej rozmowy. Jest tak ze względu na to, że analizy są wykonywane na brzegu sieci i przesyłane są tylko metadane audio.

AXIS Audio Analytics nie rejestruje dźwięku ani nie przesyła go strumieniowo – przesyła tylko wspomniane metadane. Ze względu na ochronę prywatności strumieniowe przesyłanie dźwięku jest domyślnie wyłączone w urządzeniach Axis (kontrola prywatności dźwięku), co oznacza, że dźwięk nie jest przesyłany strumieniowo ani rejestrowany i nie można go zrekonstruować. W razie potrzeby można włączyć strumieniowe przesyłanie dźwięku, ale jeśli wydarzy się coś istotnego, powiadomienia będą generowane nawet wtedy, gdy strumieniowe przesyłanie dźwięku jest wyłączone.

Ilustracja przedstawia ogólnie, jak AXIS Audio Analytics współpracuje z kontrolą prywatności dźwięku w kontekście wychwytywania dźwięków i wykorzystywania metadanych do generowania alertów.



Kontrola prywatności dźwięku i AXIS Audio Analytics w kamerach Axis – od krzyku (1) do alertu (5).

- 1 Mikrofon wychwytuje dźwięki.*
- 2 Strumieniowe przesyłanie dźwięku jest domyślnie wyłączone.*
- 3 Klasyfikator audio rozwiązania AXIS Audio Analytics wykrywa krzyki w dźwięku przychodzącym.*
- 4 Klasyfikator audio generuje metadane, w tym powiadomienie o zdarzeniu.*
- 5 Odpowiednie podmioty otrzymują alert na podstawie powiadomienia o zdarzeniu i metadanych. Mogą go zweryfikować, sprawdzając strumień wideo. Strumień audio nie jest dostępny.*

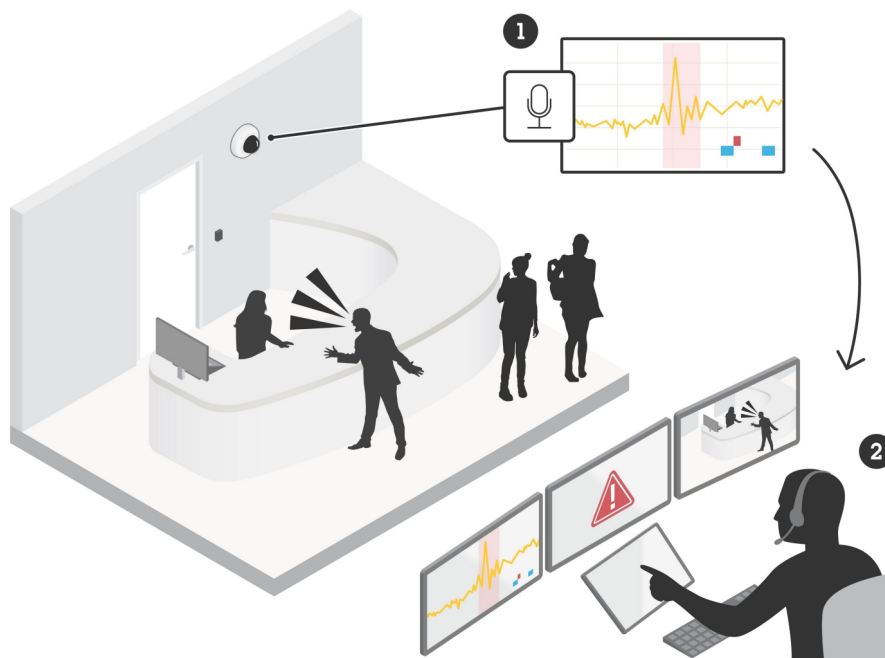
3.2 Zastosowania

Analizy oparte na sztucznej inteligencji wprawdzie mają spory potencjał w zakresie odfiltrowywania nieistotnego szumu, jednak mogą generować fałszywe alerty w przypadku wysokiego poziomu hałasu tła. Deszcz uderzający o szyby, grzmoty, syreny, muzyka czy intensywny ruch rozmawiających osób mogą wyzwać fałszywe alerty. Dlatego typowe zastosowania dotyczą raczej cichych obszarów, takich jak banki czy recepcje, ale też wiele innych przestrzeni zamkniętych po godzinach pracy, takich jak sklepy, restauracje czy biura, a nawet klatki schodowe.

3.2.1 Alarmowanie na podstawie detekcji dźwięku

W banku lub w obszarze recepcji narzędzie analityczne oparte na dźwięku może wspomagać dozór i wykrywać dźwięki związane z agresją lub rozbiciem szkła. W przypadku detekcji informacja na temat

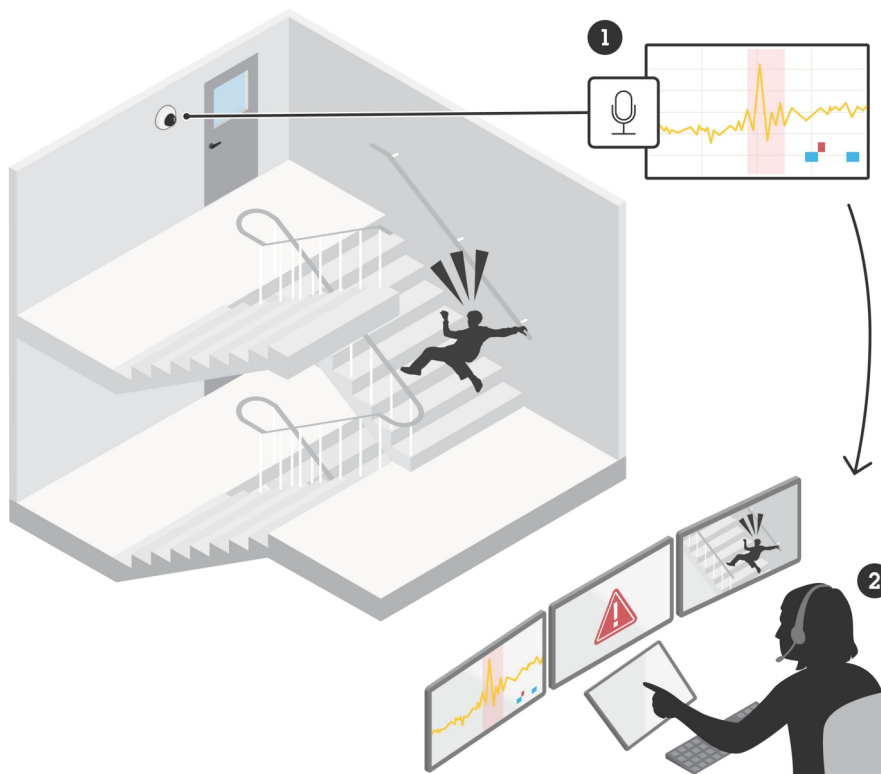
zdarzenia oraz metadane audio są wysyłane do obsługi w postaci alertu wizualnego lub mogą wyzwolić alarm. Jest to wczesne ostrzeżenie, które umożliwia szybką reakcję i interwencję.



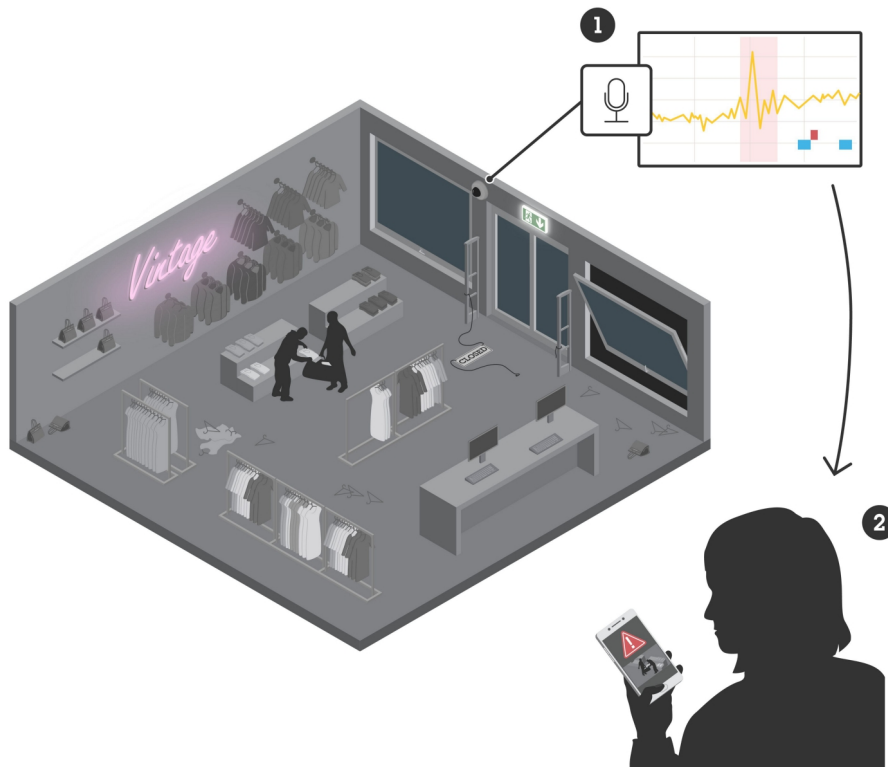
- 1 Kamera z narzędziem analitycznym obsługującym klasyfikację dźwięku wykrywa krzyk w recepcji.
- 2 Operator otrzymuje powiadomienie i może je zweryfikować, sprawdzając strumień wideo, zanim podejmie dalsze kroki.

Adaptacyjna detekcja audio może służyć do wykrywania nieoczekiwanych dźwięków poza godzinami pracy. Narzędzie analityczne analizuje dźwięki w otoczeniu i inicjuje reakcje w przypadku wykrycia

głosów, stłuczenia szyby lub innego nagłego, krótko trwającego odgłosu. Po wykryciu zdarzenia narzędzie analityczne przekazuje metadane w celu odpowiedniego powiadomienia operatorów.



- 1 Kamera z narzędziem analitycznym obsługującym klasyfikację dźwięku wykrywa niespodziewane dźwięki po godzinach pracy.
- 2 Operator otrzymuje powiadomienie i może je zweryfikować, sprawdzając strumień wideo, zanim podejmie dalsze kroki.



- 1 Kamera z narzędziem analitycznym obsługującym klasyfikację dźwięku wykrywa w sklepie dźwięki poza godzinami pracy.
- 2 Właściciel sklepu otrzymuje powiadomienie i może je zweryfikować, sprawdzając strumień wideo, zanim podejmie dalsze kroki.

3.2.2 Stosowanie różnych czujników w celu lepszego wykorzystania systemu dozoru

Systemy dozoru często obejmują kilka rodzajów czujników. Jednym z nich jest przetwornik obrazu w kamerze, który rejestruje dane wizualne na temat obserwowanej sceny. Często stosowane są też czujniki niewizualne, na przykład detektory ruchu oparte na technologii radarowej lub emisji promieniowania podczerwonego. Czujniki niewizualne stanowią uzupełnienie instalacji kamery, zapewniające innego rodzaju dane wejściowe.

Przez poszerzenie systemu dozoru o czujniki dźwiękowe (mikrofony) zyskuje się nowe możliwości w ogromnej większości zastosowań. Dodanie rozwiązań audio i analityki audio do pozbawionego ich wcześniej systemu otwiera drogę do interakcji wieloczujnikowej. W przypadku korzystania z analityki wideo dodanie analityki audio przekłada się na większą wiarygodność detekcji. Przydaje się to zwłaszcza wtedy, gdy analiza wideo jest utrudniona, na przykład w warunkach słabego oświetlenia.

System można skonfigurować – na przykład w oprogramowaniu do zarządzania materiałem wizyjnym – w taki sposób, aby wyzwalał działania dopiero wtedy, gdy reakcja wystąpi zarówno w obszarze analizy wideo, jak i analizy audio. Na przykład w ramach analizy audio zostaje wykryty krzyk, a w ramach analizy

wideo – osoba w polu widzenia kamery. W niektórych środowiskach taka kombinacja stanowi właściwy poziom bezpieczeństwa.

3.2.3 Dane wejściowe dla wykresów

Metadane audio mogą stanowić dane wejściowe dla wykresów analitycznych takich platform jak VSaaS (Video Surveillance as a Service – dozór wizyjny jako usługa), IoT (Internet of Things – Internet rzeczy) czy business intelligence, które gromadzą metadane i prezentują je w formie wizualnej. Prowadzą one analizy w czasie rzeczywistym i analizy historycznych trendów w celu błyskawicznego generowania ogólnego obrazu sytuacji i przydatnych w praktyce informacji. Analizy statystyczne oparte na przepływie klientów lub ich wrażeniach umożliwiają podejmowanie decyzji opartych na danych w celu usprawniania działalności operacyjnej.

Dzięki reprezentacji graficznej można poznać wyniki bez odsłuchiwania faktycznego dźwięku, a więc nie jest potrzebny dostęp do pierwotnego dźwięku. Zamiast tego uzyskuje się przydatne informacje, na przykład wynikające z policzenia zdarzeń. Nie powinno więc być żadnych wątpliwości co do prywatności danych. Należy pamiętać, że mogą obowiązywać różne ograniczenia prawne, w zależności od tego, czy materiał dźwiękowy jest rejestrowany czy tylko przechwytywany (monitorowany).

4 Ograniczenia prawne

Wiele osób ma zastrzeżenia dotyczące używania mikrofonów w kontekście dozoru wizyjnego. Zastrzeżenia te zwykle dotyczą rejestrowania głosu razem z materiałem wizyjnym. Jednak w przypadku narzędzi do analizy audio zwykle żaden dźwięk nie jest rejestrowany ani przesyłany strumieniowo. Przepisy regulujące kwestie związane z dozorem różnią się w zależności od kraju i regionu, więc przed zastosowaniem funkcji audio w systemie dozoru należy koniecznie się zapoznać z obowiązującymi przepisami.

Przechwytywanie i rejestrowanie materiału dźwiękowego może być zabronione lub objęte szczególnymi obostrzeniami wynikającymi z ustawodawstwa krajowego bądź różnego rodzaju przepisów i regulacji lokalnych. W pewnych regionach lub środowiskach może być dopuszczalne przechwytywanie dźwięku przy jednoczesnym zakazie jego rejestrowania. Firmy również mogą zabraniać prowadzenia dozoru dźwiękowego na swoim terenie.

5 Wyłączenie odpowiedzialności

Ten dokument i jego treść są udostępniane dzięki uprzejmości firmy Axis, a wszelkie prawa do dokumentu oraz związane z nim prawa własności intelektualnej (w tym, między innymi, znaki towarowe, nazwy handlowe, logotypy i podobne zawarte w nim znaki) są chronione na mocy prawa oraz wszelkie prawa, tytuły własności i/lub udziały w dokumencie i wszelkie związane z nim prawa własności intelektualnej należą i będą należeć do Axis Communications AB.

Zwracamy uwagę, że ten dokument jest udostępniany w swojej aktualnej postaci bez jakichkolwiek gwarancji wyłącznie do celów informacyjnych. Informacje zawarte w tym dokumencie nie stanowią i nie mają stanowić porady prawnej. Ten dokument nie ma tworzyć i nie tworzy żadnego zobowiązania prawnego po stronie firmy Axis Communications AB i/lub jakichkolwiek podmiotów z nią powiązanych. Zobowiązania firmy Axis Communications AB i/lub jakichkolwiek podmiotów z nią powiązanych w odniesieniu do dowolnych produktów Axis są w całości uregulowane warunkami umowy zawartej między firmą Axis i podmiotem, który kupił takie produkty bezpośrednio od firmy Axis.

DLA UNIKNIĘCIA WĄTPLIWOŚCI WYJAŚNIA SIĘ, ŻE CAŁE RYZYKO ZWIĄZANE Z UŻYCIEM TEGO DOKUMENTU I JEGO REZULTATAMI PONOSI UŻYTKOWNIK DOKUMENTU, A AXIS W MAKSYMALNYM

ZAKRESIE DOZWOLONYM PRZEZ PRAWO WYKLUCZA I WYŁĄCZA WSZELKIE USTAWOWE, WYRAŻNE LUB DOROZUMIANE GWARANCJE, W TYM, MIĘDZY INNYMI, WSZELKIE DOROZUMIANE GWARANCJE WARTOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU, TYTUŁU I NIENARUSZENIA PRAW ORAZ ODPOWIEDZIALNOŚCI ZA PRODUKT, A TAKŻE WSZELKIE GWARANCJE WYNIKAJĄCE Z JAKIEJKOLWIEK OFERTY, SPECYFIKACJI LUB PRÓBKİ, KTÓRE MOŻNA BY ODNIEŚĆ DO TEGO DOKUMENTU.

O firmie Axis Communications

Axis umożliwia tworzenie mądrzejszego i bezpieczniejszego świata, tworząc rozwiązania zwiększające bezpieczeństwo i wydajność biznesową. Jako firma z branży technologicznej będąca liderem na rynku, Axis oferuje systemy dozoru wizyjnego, kontroli dostępu, domofonowe i rozwiązania audio. Rozwiązania te są wzbogacone o inteligentne aplikacje analityczne i wysokiej jakości szkolenia

Firma Axis zatrudnia około 4000 zaangażowanych pracowników w ponad 50 krajach i współpracuje z partnerami z sektora technologii oraz integracji systemów na całym świecie, aby dostarczać rozwiązania dla klientów. Firma Axis powstała w 1984 roku, a jej siedziba znajduje się w Lund w Szwecji