

# 画像認識技術を活用した イベント監視プラットフォーム

井上 清司

OKIでは監視カメラの内部で処理する画像認識機能をカメラ自身がWebサービスとして提供し、ネットワークを介してスマートフォンやタブレットで動作するアプリケーションからカメラを直接操作可能とする新たなイベント監視カメラを開発した。認識対象に合わせた画像認識ソフトウェアとその機能呼び出す操作端末上のアプリケーションを入れ替えることによりさまざまなイベント監視に対応可能なプラットフォームとなっている。本稿ではそのプラットフォームの画像処理部に組込用顔認識ミドルウェアFSE<sup>\*1</sup>を活用した実装例を紹介する。

## イベント監視カメラ

現在の日本は少子高齢化の影響で、年々見守られる高齢者が増加し見守る世代が減少していく傾向にある。一方で、社会では常習万引きや窃盗団による被害が多発している上、2020年の東京オリンピックに向けては周辺警備の警備員不足も懸念されている。

こうした中、セキュリティ機器の中でも監視カメラ（特にネットワークカメラ）の出荷台数が伸びている<sup>1)</sup>。単に映像を記録するだけでなく、未然に犯罪を防ぐという観点ではカメラを通して危険を早期に発見する必要があるが、設置カメラの増加に見合った監視人員を用意できない場合には、危険を見逃すリスクの増加も想定される。そこで本稿で紹介するカメラは映像を出力するのではなく、人が見て監視する代わりに、特定の事象を撮像した画像からリアルタイムに認識し通知することを目的としたイベント監視型のネットワークカメラとなっている。

## システム概要

図1に顔認識機能を運用する際のシステムの構成例を示す。図1-(A)は本稿で紹介するイベント監視カメラの最小構成である。カメラ内部で通知するまですべてを処理するので、カメラ以外に必要なものは通

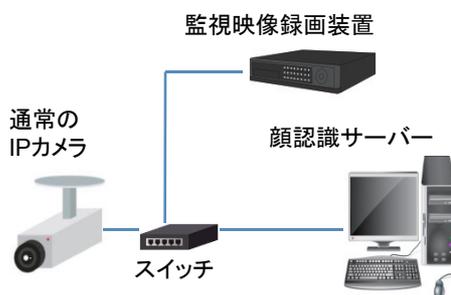
\*1) FSEは沖電気工業の登録商標です。

知を受取るスマートフォン用の操作アプリケーションだけである。図1-(B)に従来型の顔認識システムの構成例を示す。通常のネットワークカメラを利用する場合には、この例の様に監視映像録画装置が必要になる。また、顔認識ソフトウェアを動作させる顔認識サーバーが必要になり、スマートフォンに通知する際にはさらにこの顔認識サーバーからスマートフォンにメールなどで通知が送られることになる。従ってシステム形態としては、従来システムに比べ、導入コストを抑えるだけでなく既存のネットワーク環境と切り離しても動作可能なため、導入・評価が容易なシステムとなっている。

イベント監視カメラ



(A) イベント監視カメラシステム例



(B) 従来型顔認識システム構成例

図1 システム概要

図2にイベント監視カメラ内のソフトウェア概念図を示す。カメラ内では、カメラ独自に常時動作する機能と、外部の操作端末から呼び出されて動作する機能があり、それらはすべてWebサービスとして外部端末から操作される。初期設定で特定の端末からのみの呼

び出しが可能になる。専用アプリケーションがこのカメラの提供するWebAPIを通してサービス呼び出すことにより、認識条件の設定や、認識結果通知の受け取りが可能になる。また、その専用アプリケーションがWebAPIを呼び出すためのライブラリを用意しており、スマートフォンに限らず、ローカルエリアにある複数のイベントカメラを束ねる管理端末や、クラウド上からでもシステムの規模に応じてシームレスにプログラムを移行可能である。

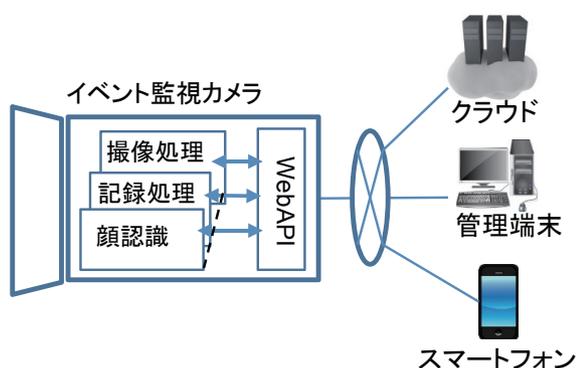


図2 ソフトウェア概念図

スマートフォン側の操作アプリケーションは読み出して画面に表示はするが操作端末内には一切保存しない。

2つ目はイベント通知機能である。カメラは顔画像を通過履歴として保存する以外に常時設定された条件と照合し、合致していればスマートフォンにアラームを通知する機能を備えている。設定可能な条件には以下のようなものがある。

- ①時間範囲指定
- ②顔検出数、顔特徴、年齢、性別
- ③場所

これらの条件を組合せることで、たとえば人がいても良い時間帯（場所）だが年齢・性別で不適切な人が検出された場合の通知や、人がいてはいけない時間帯に顔が検出された場合に通知するなどの使い方ができる。



図3 簡単なイベント監視の例

## Webサービス機能

イベント監視カメラは大きく2つの機能で構成されている。

1つ目は自動顔画像記録機能である。通常の映像監視システムのように、映像の記録をするわけではなく、カメラの前を通過する人物の顔周辺部分だけを切り出して内蔵するSSDに保存する機能である。一旦顔を検出すると顔の追跡ができなくなるまでは、常に同一人物と判断し、顔照合に最適なベストショット画像だけを1枚保存する。また同時に付随する情報として顔画像を保存した時間、推定年齢・性別などの情報をテキストで記録する。通常映像録画装置に記録される映像とは異なり、顔認識をするために必要なバストアップの画像だけを保存する。3~10GB程度あれば、日に1万人程度で約1ヶ月分の顔画像を記録可能であり、サイクリックに上書き保存しながら継続利用が可能である。専用アプリケーションから呼び出されるまでは、これらの記録情報はカメラ内部に保存されるだけで、どこにも出力されない。必要に応じてスマートフォンやタブレットなどの外部端末からこれらの情報を読み出し、検索、削除する機能がWebサービスとして提供される。

顔の識別可能な画像自体は個人情報に当たるため、

## 顔認識処理をカメラに内蔵するメリット

図1に示したように、イベント監視カメラは従来システムを一体化したことによって導入・評価が容易になるメリットに加え、高いフレームレートで処理可能なため被写体が移動する人などの場合に画像認識性能の観点からいくつかのメリットがある。

カメラから取り込んだ画像を顔認識サーバー側で認識・処理する図1-(B)のようなシステムでは、ネットワークの負荷を考慮すると画像を圧縮して送信すべきだが、この画像の送受信過程で行われる圧縮伸長処理による画質の劣化が顔認識精度劣化につながるため、一般的には画像を圧縮せずに送信するケースが多い。この場合カメラ内部で顔認識処理プログラムが画像センサーから受取れる画像のフレームレートに比べ、画像サイズやネットワークの負荷状況に依存して顔認識サーバーが受取るフレームレートは低下する。その結果、目の開いた正面顔のような顔認識処理に適切な画像を見逃す機会が増え、顔照合精度劣化の要因が増加するだけでなく、画像中に検出可能な顔が写っていないことによる見逃し率の増加により顔認識システムの運用全体で見た認識精度の低下を導く原因になる。

さらに図1-(B)の方式では、監視カメラから常に顔認識サーバーに画像を送信するため、画像に顔の写っていない時には、ネットワークに無駄な負荷を与えていることになる。

それに対してカメラ内で顔認識処理をする図1-(A)の場合、ネットワークに与える無駄な負荷が無いことに加え、カメラ内で実行される顔認識処理が取得する画像のフレームレートが高いことからフレーム当たりにより写り込む人の移動量が少なくなり、顔を検出してからその顔がフレームアウトするまで同一人物として追跡が容易になる(図4)。通常顔認識の処理は顔検出と顔照合を繰り返すが、顔追跡が容易になれば画像認識処理の効率化の視点で以下のメリットが得られる。

- (1) フレーム間で同一人物の顔領域(画像中にすでに顔があるとわかっている領域)の顔検出処理を省ける。
- (2) 顔追跡中の顔画像を比較し、認識に適した顔画像だけを選択し保存可能である。
- (3) 顔追跡中に顔照合をする場合、一度誰かの顔に一致すれば、それ以降のフレームでその人物に対する照合処理を省ける。

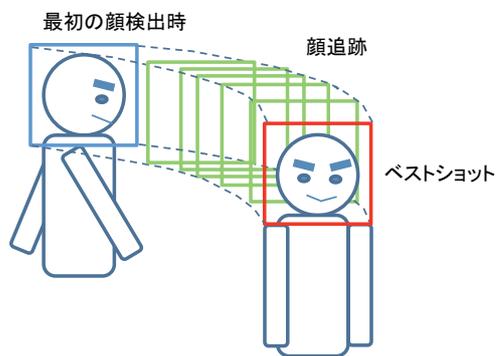


図4 カメラ内の顔追跡処理概念図

## 万引き防止システムへの応用

国内における小売店舗における万引被害は4615億円(2010年万引防止官民合同会議推定)であり、万引き対策が小売店舗の重要な課題の1つとなっている。万引きは、従来青少年の犯罪であったが、2012年以降は65歳以上の高齢者の万引きによる検挙件数の方が多く、時代を反映した社会問題の1つとなっている。2016年2月に東京万引き防止合同会議が発行した「万引き被疑者などに関する実態調査分析報告書」<sup>2)</sup>によると、万引きを諦める原因として初犯、再犯を含め約7割から8割が「店員の声掛け」と回答しており、接客の重要性が防犯の視点でも重要な役割であることが確認されている

(図5は「万引き被疑者などに関する実態調査分析報告書」<sup>2)</sup>より引用)。

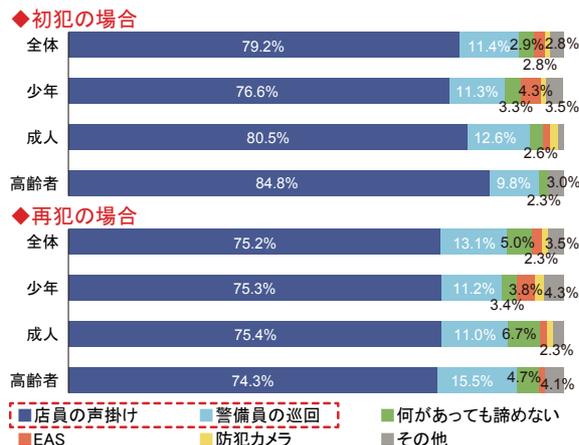


図5 被疑者が万引きを諦める原因

しかし、店内利用客が従業員の数よりも遥かに多く、すべての顧客に接客することが困難な業態では、効率的に声かけするためにも挙動不審人物や万引き常習者の早期発見が重要度を増している。

そのような中、近年総合スーパーや大手書店ではこうした監視業務の一部に顔認証システム(特定の人物を識別するシステム)の導入が始まっている。

そこでイベント監視カメラを万引き犯早期発見に活用するに当たって、カメラの顔認識機能を呼び出すアプリケーションをカスタマイズした例を以下に示す。

### (1) システムの外側での人物の特定

一般の小売店舗で利用する場合は特に誤認識による弊害を避けたい。そのため照合結果を通知する際には類似度数と共に、登録画像と照合画像の両方を表示し、あくまでも最後は「人」が見て同一人物かどうか判断した上で対処して運用する。

### (2) 外部入力画像による登録と照合

運用開始時の利便性向上のため、既に利用者が発見したい人物の画像を保有している場合には、その画像を操作端末上のアプリケーションからイベント監視カメラに送り込み、監視画像との照合を可能にしている。質の悪い画像の登録による認識精度の低下が予想されるが、(1)のように最終的な判定を人に委ねることで安全に運用しつつシステムの利便性向上を図っている。

以上のようにイベント監視カメラが提供するWebAPIを利用するアプリケーションをカスタマイズすることでさまざまな分野に適合したシステムの構築が可能となっている。

## 万引き対策におけるソリューション

従来の監視映像から犯人を特定する場合、①犯人の顔が判っていて記憶をたどり監視映像を見つけて特定するケースと、②逆に映像から犯行状況を確認して犯人の顔を特定するケースがある。どちらも人が映像を見つけて犯人の顔を特定する必要がある。イベント監視カメラの提供する第1のソリューションは①の探索業務時間の短縮である。自動顔画像記録機能により通過した人物の顔画像だけを時間帯、見た目の年齢層・性別などの条件による絞り込み検索が可能になり、従来の映像による探索業務時間を大幅に短縮することが可能である。

第2のソリューションは常習万引き犯などの早期発見である。現状は店舗の従業員が接客の合間に探索するケースや、私服警備を雇って探索させるケースがある。どちらのケースにおいても人が目で見て探すという監視業務を軽減し、本来の業務への集中を支援する。たとえば接客業務への集中や、私服警備ならば犯人追跡と確保に集中するなどがあげられる。

このシステムは直接万引き犯を減らすわけではなく、従業員による接客や私服警備の行動が最終的には防犯に貢献する。イベント監視カメラを万引き対策に活用する目的は、探索したい人物を早期に発見し、従来人が監視していた時間の短縮や、見逃しによるリスクを軽減することである。

## 個人情報保護法と顔認識技術

昨年度の個人情報保護法改正により顔特徴データも個人情報の範疇として扱われることになり、イベント監視カメラもカメラ内に一部の画像と顔特徴データを保存するため、法律の遵守が必要である。カメラ内に保存した顔画像自体は、通常の防犯画像同様、特定の個人を識別可能な場合には個人情報に当たる。もちろん、そこから抽出した顔特徴データも個人情報となる。

したがって、イベント監視カメラを利用されるお客様には、防犯カメラの録画映像と同様、顔認識に関するデータも企業内のルールに基づいて正しく運用管理をしていただく必要がある<sup>3)</sup>。

## まとめ

本稿では顔認識技術を活用したイベント監視型ネットワークカメラの実装例を紹介した。従来の顔認識システムでは監視映像を全て録画装置に記録していたのに対し、イベント監視カメラでは、人の顔画像だけを

記録し、時間、場所、年齢・性別、顔特徴などの条件設定による検索を可能にし、カメラから直接タブレットやスマートフォンに認識結果を通知するようにした。イベント監視カメラが外部に公開するWebAPIを利用することにより操作端末側でさまざまな業務アプリケーションを構築可能である。またこのカメラが提供するWebサービスをマッシュアップすることで更なるサービス機能も構築可能である。

イベント監視カメラは、小売だけでなく金融、鉄道、病院、宿泊施設、建設現場、老人ホームなどさまざまな分野での応用が可能であり、今後はお客様と新たなデジタルトランスフォーメーションを実現していくIoTプラットフォームの構成要素として整備していく予定である。◆◆

## 参考文献

- 1) 矢野経済研究所：IPカメラ国内市場に関する調査を実施（2016年）、2016年10月19日発行
- 2) 警視庁生活安全総務課：万引き被疑者などに関する実体調査分析報告書（平成26年度調査）、2016年2月
- 3) 全国万引犯罪防止機構：万防時報 第18号、p6、2015年2月12日

## ● 筆者紹介

井上清司：Seiji Inoue. 情報通信事業本部 新規事業開発室

## TIP 【基本用語解説】

### フレームレート

動画において、単位時間あたりに処理させるフレーム数（静止画像数、コマ数）のこと。

### 顔検出

背景の中から顔の位置を見つける処理

### 顔追跡

一度検出した顔が次のフレームでどこにあるかを探す処理

### 顔照合

予め、顔の特徴データとして登録されている複数の人の顔の情報とカメラに写った人の顔の特徴がどのくらい似ているかを判定する処理