

平成27年11月

顔画像識別技術と監視カメラが 産み出す「機械の目」の特性

澤田雅之技術士事務所(電気電子部門)

所長 澤 田 雅 之

「機械の目」の優れた識別能力

「機械の目」は、ほぼ正面から緻密かつ鮮明に撮影した顔画像であれば、「人の目」には全く別人の印象を与える場合であっても、例えば、20年～30年の経年変化や整形手術を受けた場合であっても、数十万枚の顔画像の中から本人の顔画像を類似度一位に瞬時検索できる。

顔画像識別エンジンは超高速

- ◎ 識別動作は瞬時 → 数十万の顔画像を、類似度順に瞬時に配列
- ◎ 多数の顔特徴抽出フィルタで、物理的特徴を量的に抽出 → フィルタ枚数に応じた次元数の顔特徴ベクトルを生成 → ベクトル間の距離計算は超高速

識別精度の劣化要因

1 顔画像の鮮明度

ブレ、ボケ、ノイズ、低コントラストが識別精度を劣化

2 顔画像の緻密度

目間画素数が50画素程度であれば十分 → 数百画素を確保しても識別精度の向上には繋がらない。

3 顔の撮影角度

上下方向は約20度まで、左右方向は約30度までOK
→ 角度が大きいく程、他の劣化要因への余裕度が減少

4 顔の経年変化・表情の有無・眼鏡等の有無

カメラ側の工夫では対処不可能 → 顔画像識別エンジン側で対処 → 識別精度劣化の主要因ではなくなっている。

☆「機械の目」の活用方法 ☆

I 被疑者写真検索システム

防犯カメラの録画映像の中から、犯人の遺留顔画像を抽出し、合致する被疑者写真を検索することにより、身元を割るシステム

II ターゲット発見システム

監視カメラのライブ映像の中から、ターゲット(テロリストや指名手配犯)の顔画像と合致する人物を、リアルタイムに発見するシステム

I 被疑者写真検索システム

- ◎ 防犯カメラの録画映像の中から、犯人の遺留顔画像を抽出 → 顔画像識別技術を用いて被疑者写真を検索 → 合致すれば身元が判明
- ◎ 被疑者写真は、緻密かつ鮮明な正面無表情顔 → 検索精度は、犯人の遺留顔画像の品質次第
- ◎ 検索結果の類似度上位約五千枚以内に、犯人の被疑者写真のリストアップが重要

被疑者写真検索時の発見率向上策

- ◎ 録画映像の中から、理想（緻密かつ鮮明な正面無表情顔）に近い遺留顔画像を選択
- ◎ 不鮮明な場合には、画像鮮明化処理を実施 → 照合結果に改善が見られる場合が多い。
- ◎ 防犯カメラを、低照度環境下でもブレやノイズの無い、緻密で鮮明な映像を撮影できる高精細デジタルビデオカメラに更新

防犯カメラの高精細デジタル化

✕ アナログカメラの問題点

52万画素の撮像素子搭載カメラの映像出力端子では、600本程度の水平解像度が実現 → 映像伝送に伴うS/N比の低下に応じて、水平解像度が確実に劣化 → 高精細映像を扱う意味が無い。

◎ デジタルカメラの優位性

映像伝送時にS/N比が低下して0と1の判別誤りが生じてても、デジタル処理による誤り訂正機能により、撮影時の映像品質が保たれる。 → 緻密かつ鮮明な高精細映像を扱うのに最適。4K画質のデジタル監視カメラが既に市販。(拡大しても鮮明)

デジタル防犯カメラの2つのタイプ

◎ ネットワークカメラ

- ・ 撮影した映像をカメラ内部で圧縮し、LANケーブルでTCP/IPネットワークに出力
- ・ カメラサーバ機能を有するため、クライアントであるパソコンや携帯情報端末から自在にアクセス可能

◎ HD-SDIカメラ

- ・ 撮影した映像を非圧縮の1.485Gbpsのデジタル信号として、同軸ケーブルで出力
- ・ アナログカメラ用に既設の同軸ケーブルをそのまま用いて、防犯カメラの高精細デジタル化が可能

デジタル防犯カメラ選定上の注意点

◎ 高速シャッター撮影が可能か？

低照度環境下では、シャッター速度が1/30秒まで低下し、動きの速い被写体がブレて不鮮明となる。 → 最低被写体照度が0.1ルクス程度のデイトカメラ又は高感度カラーカメラを使用して、高速シャッター撮影する。

◎ プログレッシブスキャン方式か？

インタレーススキャン方式では、1枚のフレーム画面を2枚のフィールド画面で構成するが、このフィールド画面間には1/60秒の時間差がある。 → 動きの速い被写体にズレが生じて不鮮明となる。 → プログレッシブスキャン方式のカメラを使用して、この問題を回避する。

☆「機械の目」の活用方法 ☆

I 被疑者写真検索システム

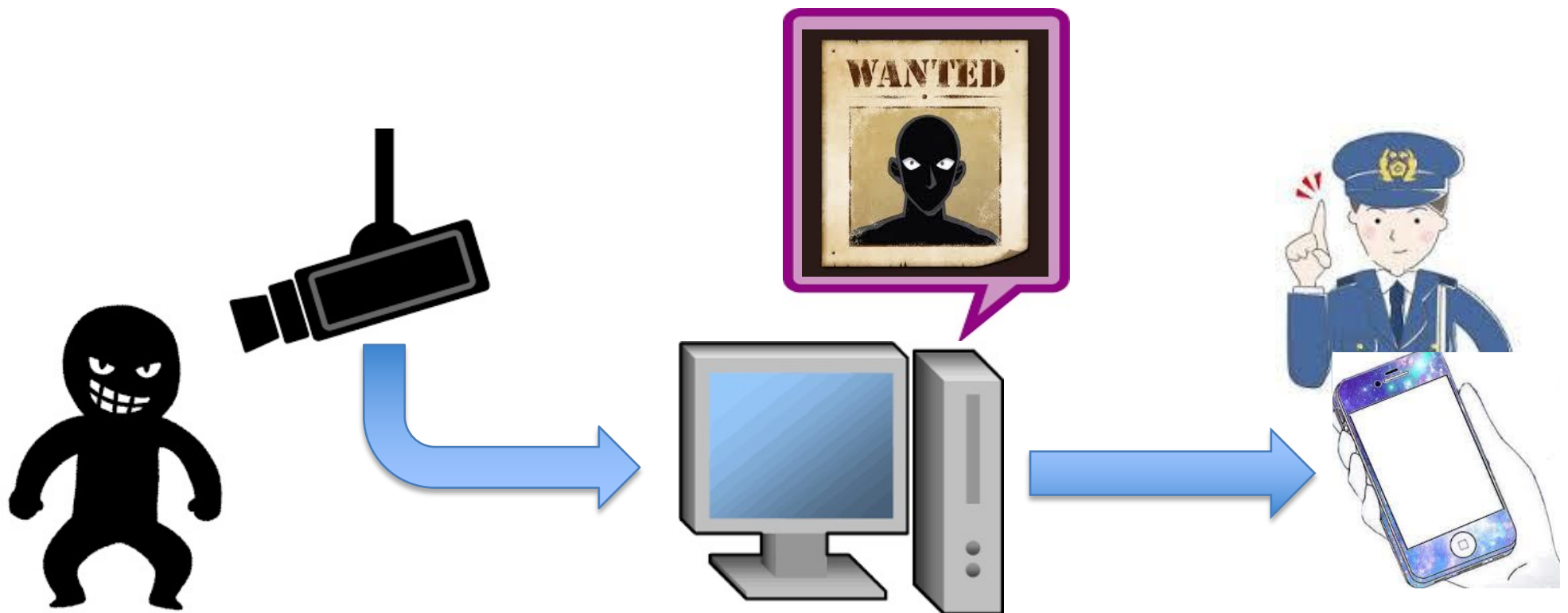
防犯カメラの録画映像の中から、犯人の遺留顔画像を抽出し、合致する被疑者写真を検索することにより、身元を割るシステム

II ターゲット発見システム

監視カメラのライブ映像の中から、ターゲット(テロリストや指名手配犯)の顔画像と合致する人物を、リアルタイムに発見するシステム

II ターゲット発見システム

監視カメラのライブ映像の中から、ターゲット（テロリストや指名手配犯）の顔画像と合致する人物を、リアルタイムに発見するシステム



これ迄のターゲット発見システム

✕ 2001年に、フロリダ州タンパ警察が世界で初めて導入したが、全く発見できないため2年後に撤去

✕ その後、ロンドンやバーミンガムなどでも導入したが、これまで発見事例は無い。

→ 効果的なターゲット発見システム
の実現事例は皆無

ターゲット発見システム失敗の原因

★ 顔画像識別エンジンと監視カメラ双方の性能が低かった。

→ 高い本人発見率と極めて低い他人誤認率を、両立できなかった。

◎ 近年、顔画像識別エンジンと監視カメラの性能が飛躍的に向上

特に、我が国の顔画像識別技術及び監視カメラ技術は、いずれも世界のトップランナー

→ 効果的なターゲット発見システムを、世界で初めて実現できる環境

ターゲットを発見するためのポイント

- 1 監視カメラで顔を捉えてからターゲット発見の警報出力までに要する時間は、数十秒の単位ではなく、数秒の単位であること。
- 2 高い本人発見率(90%以上)と、極めて低い他人誤認率(0.01%以下)を両立させること。

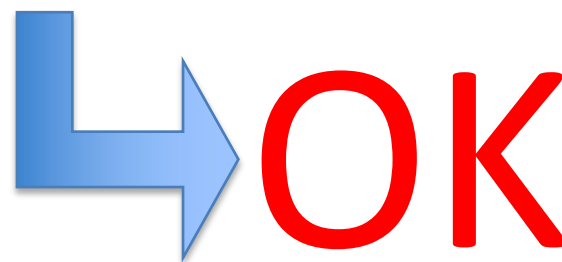
顔画像識別エンジンは超高速

◎ コンピュータの中では、数十万の顔画像を、類似度順に瞬時に配列

→ 監視カメラで顔を捉えた時点から、ターゲット発見の警報出力までに要する時間は、数秒が実現可能

ターゲットを発見するためのポイント

- 1 監視カメラで顔を捉えてからターゲット発見の警報出力までに要する時間は、数十秒の単位ではなく、数秒の単位であること。



OK

残る問題は

- 2 高い本人発見率(90%以上)と、極めて低い他人誤認率(0.01%以下)を両立させること。

ターゲット発見システム実現の要諦

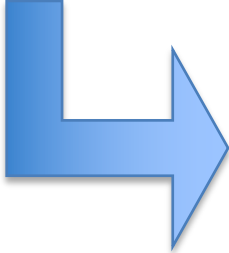
★ ターゲット発見システムの識別精度は、ターゲットとする顔画像及び監視カメラで捉える顔画像の品質次第

→ 90%以上の本人発見率と0.01%以下の他人誤認率を両立させる鍵は、次の2つ

- (1) ターゲットとする顔画像は、運転免許証写真、パスポート写真、被疑者写真等の緻密かつ鮮明な正面無表情顔であること。
- (2) 監視カメラで捉える顔画像は、ほぼ正面から撮影した緻密かつ鮮明な顔画像であること。

ターゲット発見システム実現の鍵(1/2)

- (1) ターゲットとする顔画像は、運転免許証写真、パスポート写真、被疑者写真等の緻密かつ鮮明な正面無表情顔であること。

 OK

残る問題は 

- (2) 監視カメラで捉える顔画像は、ほぼ正面から撮影した緻密かつ鮮明な顔画像であること。

ターゲット発見システム実現の鍵(2/2)

(2) 監視カメラで捉える顔画像は、ほぼ正面から撮影した緻密かつ鮮明な顔画像であること。



では、どのような顔画像であれば良いのか？
その裏付けとなるのは、次の2つの実験

- 1 米国立標準技術研究所の「顔認識技術に係るベンダーテスト」
- 2 法務省の「日本人出帰国審査における顔認証技術に係る実証実験」

1 米国立標準技術研究所の「顔認識技術に係るベンダーテスト」

(NIST Interagency Report 8009より引用)

米国立標準技術研究所では、2000年、2002年、2006年、2010年、2013年に顔認識技術に係るベンダーテストを実施。

最新の2013年のテストでは、次の2種類の顔画像を用いて、1:Nの顔画像照合精度を多角的に評価。

- 1 Mugshot 画像 : 米国の警察の実際の現場で、デジタルスチルカメラにより撮影。我が国の被疑者写真に相当。
- 2 Webカメラ画像 : 米国の国境警備隊が勾留した中南米からの不法入国者を、Webカメラにより撮影。国境警備隊職員の指示に基づき、Webカメラに顔を向けた瞬間を捉えている。

Mugshot画像の品質

○ オリジナル画像サイズ

480 × 640、240 × 240、768 × 960

○ JPEG圧縮後の画像情報量

平均48kB

○ 目間画素数

平均107画素で、標準偏差は40画素

○ ポーズ

顔を真正面から捉えた画像が多いが、中には、5度から10度ほど左右を向いている画像がある。また、視線がカメラに向けられていない画像が少なくない。

○ 照明

適切な照明のもとで撮影されている画像が多い。

Webカメラ画像の品質

- オリジナル画像サイズ

240 × 240

- JPEG圧縮後の画像情報量

平均5.7kB

- 目間画素数

平均45画素で、標準偏差は12画素

- ポーズ

顔を真正面から捉えた画像は少なく、5度から10度ほど左右を向いている画像や、下から見上げるように撮影した画像が少なくない。

- 照明

殆どが室内照明下で撮影されているため、照明が不十分で顔に影ができている画像が大半である。

Mugshot画像とwebカメラ画像の検索精度の比較

1 Mugshot画像で、1万枚の高品質顔画像を1万回検索

- (1) 閾値を設けず、他人誤認率は無制限とした場合
→ 本人発見率(類似度順位の1位検索率)は、97%
- (2) 閾値を設けて、他人誤認率を0.01%とした場合
→ 本人発見率(類似度順位の1位検索率)は、90%強

2 Webカメラ画像で、1万枚の高品質顔画像を1万回検索

- (1) 閾値を設けず、他人誤認率は無制限とした場合
→ 本人発見率(類似度順位の1位検索率)は、89%
- (2) 閾値を設けて、他人誤認率を0.01%とした場合
→ 本人発見率(類似度順位の1位検索率)は、70%弱

1 米国立標準技術研究所の「顔認識技術に係るベンダーテスト」に基づく考察

ターゲット発見システムにおいて、0.01%以下の他人誤認率と90%以上の本人発見率を両立させるには

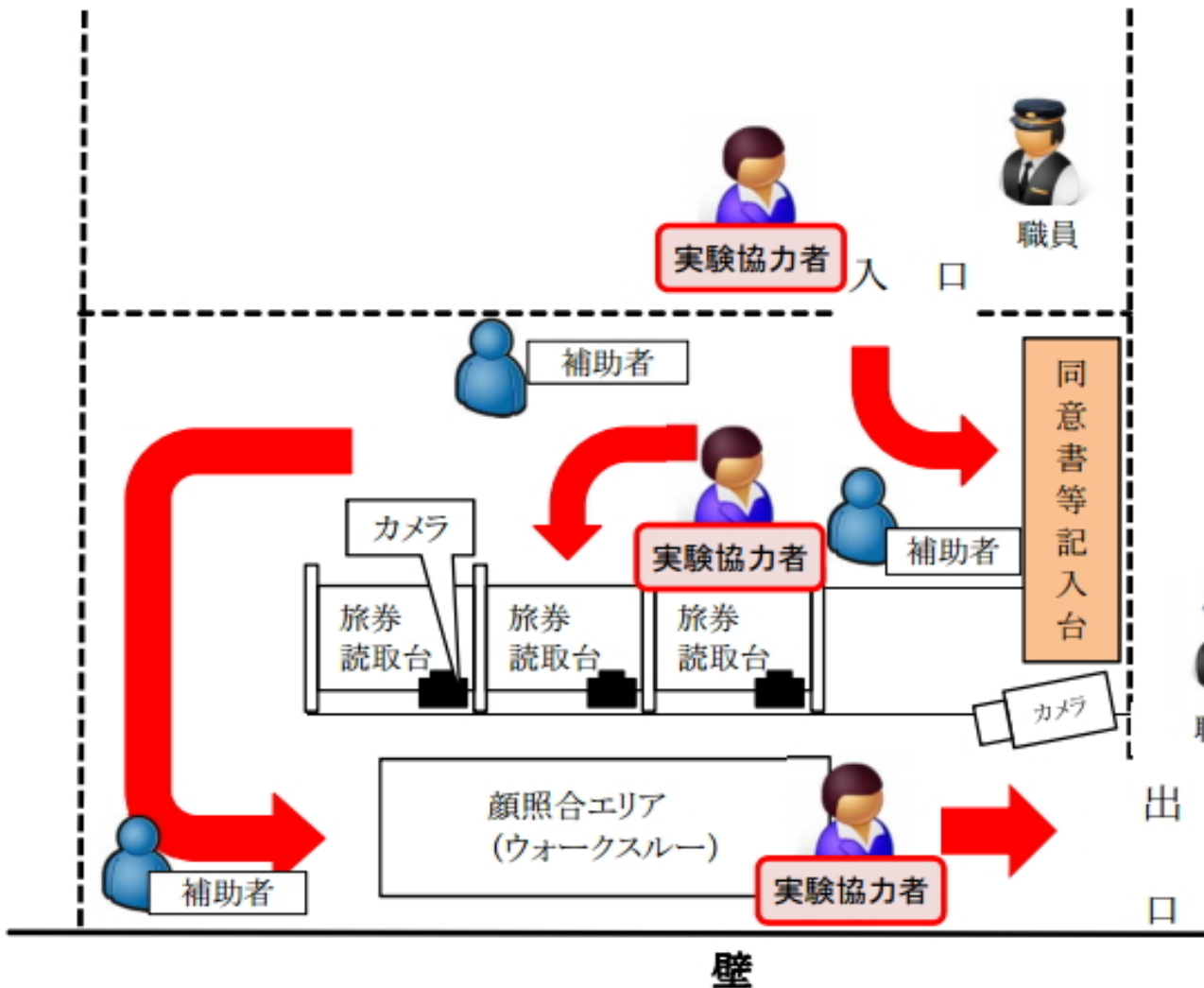
- (1) Mugshot画像品質に相当する緻密さ、鮮明さ及び撮影角度(最大でも10度程度)であれば両立できる。
- (2) Webカメラ画像に相当する品質では不十分であり、緻密さ、鮮明さ及び撮影角度について、以下の改善が必要
 - ⇒ 高精細デジタル監視カメラを用いて、緻密さを改善する。
 - ⇒ 照明の工夫により、又はワイドダイナミックレンジ機能付きデジタル監視カメラを用いて、鮮明さを改善する。
 - ⇒ 同一の人物を捉えた一連のライブ映像の中から、サンプリングの手法により、照合に適した顔画像を抽出してターゲット顔画像との照合に供する。

2 法務省の「日本人出帰国審査における顔認証技術に係る実証実験」

法務省では、我が国のICパスポートに記録されている顔画像を利用した邦人の出帰国審査の自動化を目指し、平成24年及び平成26年に、成田空港及び羽田空港で実証実験を実施。

平成26年の実験では、下記の、静止中撮影時と歩行中撮影時における認証精度の実証が中心。

- 1 **静止中撮影実験**：旅券読取台で読取操作中の実験協力者の顔を、旅券読取台に設置したビデオカメラでほぼ正面から撮影。
- 2 **歩行中撮影実験**：顔照合エリアである通路を歩行中の実験協力者の顔を、通路前方に設置したビデオカメラで撮影。実験協力者には、ビデオカメラを注視して歩行するように依頼 → ビデオカメラにほぼ正対した顔を撮影



1 静止中撮影実験
 旅券読取台で読
 取作業中の顔を撮
 影して、ICパスポ
 ート内の顔画像と照
 合

2 歩行中撮影実験
 顔照合エリアを
 歩行中の顔を撮影
 して、ICパスポ
 ート内の顔画像と照合

日本人出帰国審査における顔認証技術に係る実証実験結果報告
 (法務省、平成26年11月18日)より引用

「日本人出帰国審査における顔認証技術に係る実証実験」の結果（1/2）

実証実験に参加したA社～E社の中で、静止中撮影実験における認証精度が最も高かったのはA社であり、歩行中撮影実験における認証精度が最も高かったのはD社

他人誤認率が0.001%、0.01%、0.1%となる閾値を設定した場合における、A社及びD社の本人見逃し率（本人発見率の補数）は、次のとおり

（A社の本人見逃し率）

静止中撮影 (0.001%)0.26%、(0.01%)0.17%、(0.1%)0.16%

歩行中撮影 (0.001%)1.24%、(0.01%)0.55%、(0.1%)0.33%

（D社の本人見逃し率）

静止中撮影 (0.001%)0.54%、(0.01%)0.27%、(0.1%)0.03%

歩行中撮影 (0.001%)0.91%、(0.01%)0.32%、(0.1%)0.11%

「日本人出帰国審査における顔認証技術に係る実証実験」の結果（2/2）

(A社の本人見逃し率)

静止中撮影 (0.001%)0.26%、(0.01%)0.17%、(0.1%)0.16%

歩行中撮影 (0.001%)1.24%、(0.01%)0.55%、(0.1%)0.33%

(D社の本人見逃し率)

静止中撮影 (0.001%)0.54%、(0.01%)0.27%、(0.1%)0.03%

歩行中撮影 (0.001%)0.91%、(0.01%)0.32%、(0.1%)0.11%

- ◎ A社の歩行中撮影時の本人見逃し率は、静止中撮影時と比べて数倍悪化。一方、D社の歩行中撮影時の本人見逃し率は、静止中撮影時と比べて数割の悪化に留まり、静止中撮影時の本人見逃し率に優れるA社を凌駕

→ カメラの機能・性能・設置方法、照明方法を総合した歩行中撮影システムの構築にD社は成功

2 法務省の「日本人出帰国審査における顔認証技術に係る実証実験」に基づく考察

ターゲット発見システムにおいて、0.01%以下の他人誤認率と90%以上の本人発見率を両立させるには

- (1) 歩行中撮影システムの映像品質を、静止中撮影システムの映像品質相当に改善する。
 - ⇒ 高感度で高精細なデジタル監視カメラを用いて、緻密さを改善する。
 - ⇒ 照明の工夫により、又はワイドダイナミックレンジ機能付きデジタル監視カメラを用いて、鮮明さを改善する。
- (2) ターゲット発見システムでは、監視カメラを注視して歩行するように依頼はできない。
 - ⇒ 同一の人物を捉えた一連のライブ映像の中から、サンプリングの手法により、照合に適した顔画像を取り出してターゲット顔画像との照合に供する。

効果的なターゲット発見システムの実現には

◎ 顔画像識別エンジンの照合精度と照合速度は、十分に実用水準

◎ 監視カメラ側の工夫により、効果的なシステムを実現できる。

→ 高感度で高精細なデジタル監視カメラを用いて、緻密さを改善

→ 補助照明の設置、又はワイドダイナミックレンジ機能付デジタル監視カメラを用いて、鮮明さを改善

→ 同一の人物を捉えた一連のライブ映像の中から、サンプリングの手法により、照合に適した顔画像を抽出

☆「機械の目」の活用方法 ☆

I 被疑者写真検索システム

防犯カメラの録画映像の中から、犯人の遺留顔画像を抽出し、合致する被疑者写真を検索することにより、身元を割るシステム

II ターゲット発見システム

監視カメラのライブ映像の中から、ターゲット(テロリストや指名手配犯)の顔画像と合致する人物を、リアルタイムに発見するシステム

平成27年11月

顔画像識別技術と監視カメラが
産み出す「機械の目」の特性

終

澤田雅之技術士事務所(電気電子部門)

所長 澤田雅之