

# 年齢性別判定のためのカメラ最適制御

Camera Optimum Control for Age and Gender Recognition

外 館 弘 理  
Hiromichi Sotodate

由 雄 宏 明  
Hiroaki Yoshio

矢 野 公 嗣  
Koji Yano

篠 原 利 章  
Toshiaki Shinohara

## 要 旨

ビジネスインテリジェンス拡張キットDG-NVF20は、人数カウント、年齢性別判定機能を搭載した商品である。商品化にあたり、監視システム環境下での「ロバスト性確保」や、システムの低コスト化による「安価な組込み機器への搭載」という課題があった。監視ロバスト性確保については、カメラ内部の信号処理特性を調査し、年齢性別判定に適した映像になるようにパラメータを最適化し、レコーダーからカメラに設定することで実現した。安価な組込み搭載については、並列演算処理に特化したDPP (Data Parallel Processor) ブロックに年齢性別判定処理の一部を並列演算化して搭載することで実現した。

## Abstract

The Business Intelligence Expansion Kit “DG-NVF20,” which can recognize numbers of people and their age and gender, has been launched. There were some problems to overcome before it could be commercialized. The problems related to “ensuring robustness” in a surveillance system environment and “installing it in an inexpensive embedded system” to reduce the price. As for “ensuring robustness,” we investigated the signal processing inside the camera and set the optimum camera parameters to correctly judge numbers of people and their age and gender from the recorder. And with regards to “installing it in an inexpensive embedded system,” we investigated the unused blocks named Data Parallel Processor (DPP) in the embedded chip, which are specialized in parallel arithmetic processing, and we allocated part of the processing of the age, gender and face matching to parallelizing in DPP.

## 1. はじめに

監視業界もIP (Internet Protocol) 化、高画質化が進み、国内・海外ともにメガピクセルIPカメラ、IPレコーダーが主力商品となっている。そのなかで、近年、新興メーカーがコスト面での強力な攻勢をかけており、各商材の低コスト化が進んでいる。

そのため、低コストかつ、他社優位性をもった商品を投入することが必要となっている。筆者らは、そのアプローチの1つが画像認識技術を用いた付加価値機能の搭載であると考えている。その理由は、①技術トレンド、②市場ニーズ、③技術優位性、の3点である。①従来のアナログ監視機器にも、撮像画像の動きを検出しアラーム発報するVMD (Video Motion Detector) などの画像認識機能が搭載されている。しかし、近年のIP化により各機器間の大容量メタデータ通信が容易となり、また、高画質化により鮮明な映像を取得できることから、より高精度かつ多機能な画像認識が可能である。②大量の記録映像から所望のシーンを見つけることは非常に困難である。そのため、特定人物の来場のみをアラーム検知したいなどの、現状人力に依存している作業の自動化・省力化の市場要望が高まっている。③当社グループはカメラとレコーダー両方の機器を開発している数少ないメーカーである。カメラレコーダー間の密な連携を可能とし、それぞれの商品の内部処理特性まで踏み込んだアルゴリズム開発や機能実装することで他社が追いつけない精度を

実現できる。また、カメラ、レコーダー間で機能の分散実装することで安価な組込み機器への搭載も可能である。

このような背景のなか、筆者らは顔照合機能を搭載したIPレコーダーを2010年9月に商品化し、IPレコーダーのアドオンソフトウェアであるビジネスインテリジェンス拡張キットDG-NVF20 (以下NVF20) を2011年5月に商品化した。ビジネスインテリジェンス拡張キットは、顔のしわや輪郭などの顔特徴から年齢性別判定や人数カウントすることで、来店者の客数や客層を分析することができる。

本稿では、ビジネスインテリジェンス拡張キットを商品化するうえで実現した「監視システム環境下でのロバスト性確保手法」と、「安価な組込み機器への搭載手法」について報告する。

なお、本稿記載の精度は、すべて実際のお客様の利用環境における評価結果である。画像認識は設置環境の影響を受けやすく、実験室や恣意的人物の存在する評価では真の実力や課題がわからない。そのため、筆者らは国内や海外で繰り返しフィールド評価を実施し、商品化を実現した。

## 2. ビジネスインテリジェンス概要

### 2.1 商品の位置づけ

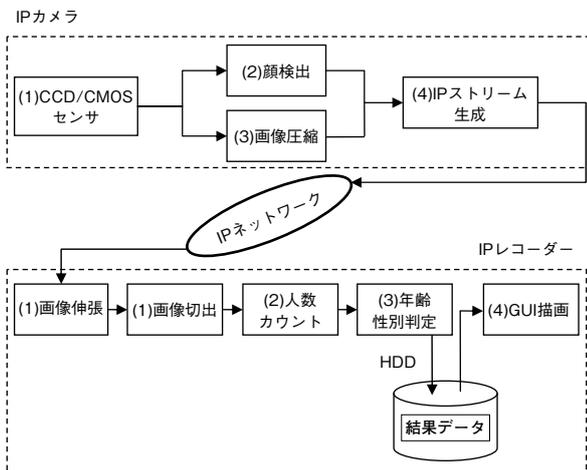
現在の監視システムの運用手法は、カメラの映像を24

時間レコーダーに録画し、盗難・事件などのイベントが発生したとき映像再生するといった事後検証のみに利用される場合が大半である。しかし、これらイベントは頻繁に発生するものではないため、監視システムというものは大半の間利用されていないのが現実である。

そこで、監視システムが利用されていないときに来店のお客様数や客層分析といったマーケティング用途を提供することで、従来の万引き抑止といった監視用途に加え、効果的な販売戦略づくりにも利用することができる戦略的な位置づけをもった商品である。

## 2.2 システム構成

第1図に示すとおり、IPカメラで顔検出処理、IPレコーダーで人数カウント・年齢・性別判定処理をさせることで、実現機能の処理分散を図っている。



第1図 システム構成  
Fig. 1 System structure of NVF20

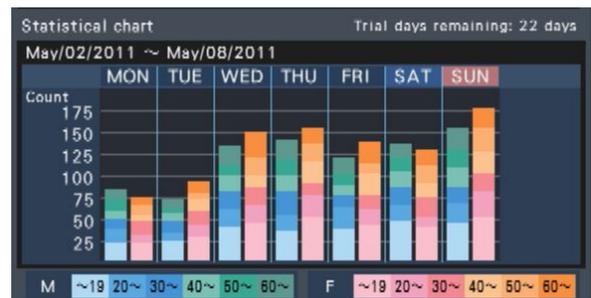
### 【IPカメラ】

- (1) メガピクセル対応のCCD (Charge Coupled Device) もしくは CMOS (Complementary Metal Oxide Semiconductor) センサから得られた画像は、顔検出部と画像圧縮部に送られる。
- (2) 顔検出部では、Ada-boost学習を拡張した顔向きにロバストな顔検出器[1]を用いて、画像中から最大8つまでの顔を検出する。
- (3) 画像圧縮部では、H.264もしくはMotion JPEGに圧縮する。
- (4) IPストリーム生成部にて、圧縮画像をRTP (リアルタイムデータ転送プロトコル) で配信する。検出した顔の矩形 (くけい) 座標はRTPヘッダの拡張領域に埋め込む。

### 【IPレコーダー】

レコーダーでは、カメラから配信された顔矩形座標を利用して、人数カウント、年齢性別判定を行う。

- (1) 年齢性別判定に利用する顔画像を切り出すため、画像伸張部と画像切出部で、顔画像の切り出しを行う。
- (2) 人数カウント部は、各顔の座標や移動方向と検出間隔を利用する。
- (3) 年齢性別判定部は、輝度値や画像周波数成分から顔特徴量を生成し、個人識別するための特徴空間へ正規化し、あらかじめ用意したモデルに対するスコアから年齢性別を推定する。その際、顔向き角度の大きい顔画像は年齢性別精度に悪影響を与えることから、顔向きなどから算出される照合信頼度の高い顔画像のみを利用する。
- (4) HDDに蓄積された推定結果データを読み出し、ユーザー指定された期間の人数カウント、年齢性別結果をGUI (Graphic User Interface) 表示する (第2図)。



第2図 年齢性別のGUI表示例  
Fig. 2 Example of displaying output

## 3. 監視システム環境下でのロバスト性確保

顔検出・顔照合は、デジタルカメラにも搭載されているが正面撮像を基本としている。監視カメラは上部から撮像されるため下向きに撮像される傾向があり、かつ人物挙動に制約をもたすことができない。また、24時間365日常時運用であるため、低照度時のオートゲインコントロール (AGC) によって生じるノイズや日照変動の影響なども考慮しなければならない。

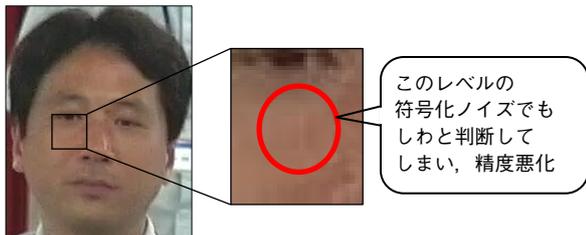
そのため、監視システム環境下のロバスト性を確保しなければ、実運用に耐えることができない。

### 3.1 課題1：カメラ信号処理の影響

現状の監視カメラは、“人が見やすい”を基準とした証跡用途の画像補正処理が実装されている。例えば、当社カメラのスーパーダイナミック機能は、明暗が激しい場

面で暗所のみダイナミックレンジを自動調整することにより、人が見やすい画像にしている。また、照度が低い環境下では、AGCなどでダイナミックレンジを上げて見やすくしている。

しかし、このようなカメラ内部の画像補正は、後段レコーダーの画像認識精度には悪影響を与えてしまう。例えば、第3図のようなAGCや符号化によって生じるノイズや人物の歩行速度によって生じる被写体ボケなど、従来の目視ベースの監視では問題とならなかったことが、年齢性別判定精度に大きな影響を与えてしまう。



第3図 AGCや符号化によって生じるノイズ  
Fig. 3 Cause of accuracy degradation (e.g. Encoding noise)

3.2 課題2：日照変動

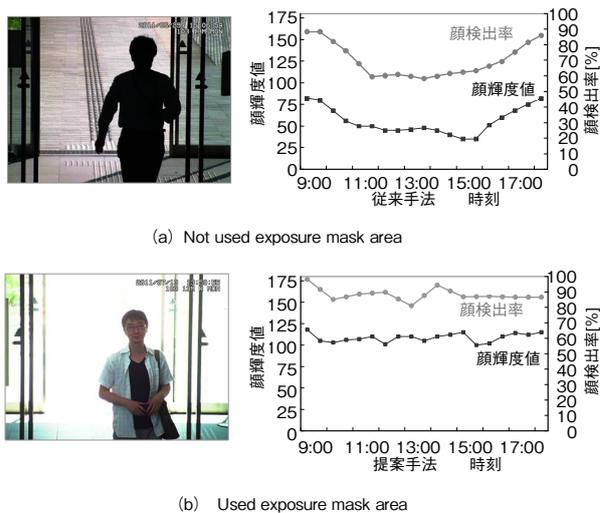
監視システムは、24時間365日運用が必須であるため、日照変動の影響を考慮する必要がある。

来店客数をカウントするため、カメラは室内からドア（屋外）側に向けて設置される。蛍光灯下の室内照度は500 lx程度に対し、屋外の照度は夏晴天時50 000 lx、曇りでも10 000 lxあるため、室内から屋外を撮像するとハレーションが起きてしまい画面が真っ白になってしまう。当社のIPカメラはハレーション対策として、入射光量に応じてアイリスやシャッター速度などを自動調整することで露光制御しているが、昼間時の屋外照度が大きい場合は、屋外照度に引っ張られ第4図 (a) のように顔が真っ黒になってしまう。

3.3 アプローチ1：パラメータ設定の最適化

カメラ信号処理の影響を回避するため、当社のカメラ信号特性を分析し、年齢性別判定をより生かす観点からカメラ信号処理パラメータの最適値（第1表）を決定、レコーダーからカメラに自動設定した。

年齢性別判定は、高画質/高解像度であるほうが精度確保しやすいため、動画配信形式をH.264、解像度を4VGA（Video Graphics Array）（最大）、ビットレートを4096 kbit/sとした。スーパーダイナミックなどの内部補正処理



第4図 逆光による顔輝度値と顔検出精度の関係  
Fig. 4 Example of a backlight image and relation between brightness and face detection accuracy

第1表 最適パラメータ

Table 1 Optimum camera parameters

設定パラメータ	設定値
動画配信形式	H.264
解像度	1280×960 (4VGA)
ビットレート	4096 kbit/s
スーパーダイナミック	Off
暗部補正	Off
逆光補正	Off
AGC	High

はすべて無効になるように設定した。ただし、夜間撮像対応のためAGCはHighとした。AGCや符号化によって生じるノイズ対策として、レコーダー内部にメディアンフィルタを実装した。メディアンフィルタの有効性については、照度250 lxの低輝度環境下で顔照合評価を行い、正報率38.10%から52.38%の精度向上を確認した。

すなわち、安定した画像認識の精度を確保するために、複雑な制御パラメータが変化する画像補正を極力抑え、信号レベル確保とノイズを抑える単純で安定した処理を行っている。

3.4 アプローチ2：部分的露光制御

当社カメラの露光制御機能は、画面輝度平均値がユーザー設定した目標値に近づくように、アイリス/シャッター速度/AGC/逆光補正などを複雑に制御している。年齢性別判定を安定して処理させるためには、これらの露光制御の変更量を少なくする必要がある。そこで、画面領域内を、輝度変化が激しい領域と安定した領域に分類し、安定した領域のみで露光制御する部分的露

光制御手法を開発した。

第4図に、(a)露光マスク設定なし、(b)露光マスク設定ありの顔輝度値および顔検出率を示す。(a)は昼間になるにつれ顔輝度が小さくなり顔検出率が60%程度まで低下しているが、(b)は昼間でも安定した顔輝度が確保でき、顔検出率も90%程度確保できていることがわかる。

## 4. 安価な組込みチップへの搭載

システムの低コスト化のため画像認識用チップ追加は不可能であり、既存搭載チップの空資源内に機能実装する必要がある。

### 4.1 課題：CPU負荷

第2表(a)が、従来の最大負荷時(16chデコード/顔照合)のCPU処理負荷である。CPUを安定動作させるための定常負荷は70%以下である必要があり空資源は全くなかった。しかしながら、新たに年齢性別処理用チップを追加することは低コスト化の観点から到底できない。

そのため、既存のアーキテクチャ内で新たな資源を生み出し、かつ年齢性別判定の高速化が必要であった。

第2表 CPU負荷  
Table 2 CPU utilization

処理ブロック	(a)従来負荷 [%]	(b)提案負荷 [%]
画像伸張	33.8	33.8
顔照合	37.1	8.9
顔特徴生成	24.5	0.0 (DPP処理)
正規化	3.7	
Medianフィルタ	4.9	4.9
ほか(顔切り出し/顔照合など)	4.0	4.0
NVF20(新規追加機能)		1.9
人数カウント		0.6
顔特徴生成		0.0 (DPP処理)
正規化		
年齢性別モデル化処理		1.3
増加分 (DPP用Endian変換など)		4.4
TOTAL	70.9	49.0

### 4.2 アプローチ：DPPブロック利用

本システムのCPUはメディアプロセッサであり、一般的なソフトウェア演算ブロック以外にもハードウェア演算ブロックが複数存在する。全演算ブロックの使用状況を調査し、ハードウェア資源のDPP(Data Parallel Processor)ブロック(画像データの並列演算処理に特化したブロック)が利用可能であることを見つけ出した。

年齢性別判定の高速化を図るために、顔照合と年齢性別の特性を分析したところ、性別判定は骨格形状といった全体的な顔特徴量の違いで識別できるのに対して、顔照合と年齢判定は局所的な違いで識別できることを見つけた。そこで、顔照合と年齢判定の類似処理ブロックの一部を共通化したアルゴリズムを開発し、DPPに並列演算化して実装することで高速化を実現した。

## 5. 実験結果

### 5.1 ロバスト性に関する実験結果

ロバスト性確保に関する精度評価および他社比較を実施した。カメラ設置環境差を吸収するため同一のカメラを利用して、従来手法、提案手法のそれぞれを評価した。評価指標として、人数カウントは、24時間累計のドア通過数誤差とした。年齢・性別は、3名の主観年齢・性別の平均値を正解年齢としたときの24時間累計の各年代の分布比率誤差の最大値とした。

また、比較した他社製品は実環境に設置済みの商品であり、赤外線カメラを上部から真下方向に撮像し、カメラ下を通過した人数をカウントする人数カウントに特化した商品で、NVF20より高価格である。

評価結果を第3表に示す。人数カウント誤差は、(従来手法) > (他社製品) > (提案手法)となり、従来手法は他社製品より誤差が大きいのに対し、提案手法を用いることで人数カウントに特化した他社製品よりも8.34%誤差が少ないことを確認した。

第3表 評価結果  
Table 3 Evaluation result

	誤差[%]		
	従来手法	提案手法	他社製品
人数カウント	27.4	3.96	12.3
年齢	10.9	8.4	機能なし
性別	7.5	2.4	機能なし

年齢性別については、従来手法に比べ年齢2.4%、性別5.0%の精度が向上した。従来手法でもある程度の精度が確保できているのは、年齢性別判定は、顔向きなどから算出される照合信頼度の高い顔画像のみを利用しているためと思われる。

### 5.2 組込み搭載に関する実験結果

最大負荷時(16chデコード/顔照合/NVF20)のCPU処理負荷を第2表(b)に示す。

顔照合処理の顔特徴量生成処理(24.5%) / 正規化処

理 (3.7%) をDPP処理することで、28.2%のCPU資源を確保した。次に、年齢性別判定用の顔特徴量生成処理と正規化処理をDPPブロックで顔照合処理と共通動作させることで、NVF20のCPU処理負荷を1.9%にすることができた。

結果、顔照合のみしか動作しない従来手法の負荷70.9%に対し、顔照合およびNVF20の両方を動作させても負荷49.0%となり、21.9%の負荷削減を実現した。

## 6. まとめ

本稿では、監視ロボバスト性確保のため、年齢性別に適した映像になるような最適パラメータの自動設定および部分的露光制御手法を提案した。従来手法では他社製品より誤差が大きかったのに対し、提案手法を用いることで他社製品を上回る結果を得ることができた。また、並列演算処理に特化したDPPブロックに顔照合処理と年齢性別判定処理の一部を並列演算化することで、安価な商品への機能搭載を実現した。

このことは、本商品が①精度面、②コスト面で非常に強い競争力をもっていることを意味している。そして、カメラ・レコーダーの両方を開発している当社グループの優位性を利用したことで成し得た。①カメラとレコーダー間でそれぞれの内部特性まで踏み込んだ連携をすることで、人数カウントに特化した他社製品よりも高精度を実現し、加えて、年齢性別判定機能も提供している。②カメラとレコーダーで機能分散させ安価機器に搭載実現することで、低コストで市場提供可能である。

また、本稿の結果から画像認識技術は環境ロボバスト性が非常に重要であることが理解できる。さまざまな設置環境に対するロボバスト性を確保することが、市場での優位性をもつ商品の鍵となる。筆者らは現在も、アルゴリズム開発部門、カメラ開発部門とレコーダー開発部門が参画したフィールド評価を継続実施しており、さまざまな市場環境に対応できるようさらなる精度改善を図っている。

## 参考文献

- [1] 中野渡祥裕 他, “向き・照明にロボバストな顔認証,” パナソニック技報, vol. 54, no. 4, pp. 24-29, 2009.

## 執筆者紹介



外館 弘理 Hiromichi Sotodate  
パナソニック システムネットワークス (株)  
セキュリティシステム事業部  
Security Systems Business Div.,  
Panasonic System Networks Co., Ltd.



由雄 宏明 Hiroaki Yoshio  
パナソニック システムネットワークス (株)  
先行技術開発センター  
Advanced Technology Development Center,  
Panasonic System Networks Co., Ltd.



矢野 公嗣 Koji Yano  
パナソニック システムネットワークス (株)  
セキュリティシステム事業部  
Security Systems Business Div.,  
Panasonic System Networks Co., Ltd.



篠原 利章 Toshiaki Shinohara  
パナソニック システムネットワークス (株)  
セキュリティシステム事業部  
Security Systems Business Div.,  
Panasonic System Networks Co., Ltd.