

画像認識 LSI “Visconti™2” を搭載した インテリジェントカメラ

Intelligent Camera Incorporating Visconti™2 Image Recognition Processor

小坂谷 達夫

渡辺 友樹

岡田 隆三

■ KOZAKAYA Tatsuo

■ WATANABE Tomoki

■ OKADA Ryuzo

ネットワークカメラの台数は年々増加しており、ネットワークトラフィックやオペレーターによる監視作業など管理・運用コストの増大に対応するため、端末側で画像認識処理が実行可能なカメラに注目が集まってきている。

東芝は、画像認識 LSI “Visconti™2” をネットワークカメラ内に搭載したインテリジェントカメラを開発した。Visconti™2 は、画像認識処理を高速に実行できる様々なハードウェアアクセラレータを内蔵しており、従来はパソコン (PC) で行っていた高度な画像認識処理をカメラ内で撮影と同時に実行でき、オペレーターは、カメラが処理した画像認識結果を参考にすることで、重要な映像を効率よく見ることができる。デモンストレーション用として、人物の顔と上半身、車両、及び猫の顔の4物体を同時に検出するアプリケーションをインテリジェントカメラ上に実装し、リアルタイムで検出処理ができることを確認した。

As the number of network cameras increases year by year, the operation and maintenance costs are growing accompanying the expansion of both the network traffic and the monitoring task load. Demand has therefore been rising for a network camera with an image processing function to improve cost effectiveness and save labor.

Toshiba has developed an intelligent camera incorporating the Visconti™2 image recognition processor, which is equipped with several hardware accelerators and offers high-speed image recognition processing such as object detection with enhanced performance. As the Visconti™2 makes it possible to simultaneously capture video frames and apply advanced image recognition processing to these frames in the intelligent camera, operators can efficiently evaluate the worth of each frame and quickly focus on the important scenes. We have confirmed the effectiveness of the intelligent camera through demonstration experiments on the simultaneous real-time detection of four types of objects including a human face, a human upper body, cars, and a cat.

1 まえがき

近年の安全や防災への関心の高まりとともに、街頭や店舗などに設置されるネットワークカメラの台数は年々増加しており、2015年までは年率10%程度の成長率で推移すると予測されている⁽¹⁾。その一方で、カメラの映像を送信するためのネットワークトラフィックの増大や防犯目的での人間による監視業務など、管理・運用コストの増大が予想されている。

通常の防犯や見守り業務においては、カメラから得られる映像の多くは特段見るべき必要がない定常状態であり、事件や事故などオペレーターが対処すべきイベントが発生することはまれである。そのため、近年では画像認識技術により、映像に動きがあったシーンだけを録画したり知らせたりする補助機能を持つネットワークカメラも発売されている。

しかし、物や車両などの物体を検出し、より意味のあるタグ情報を映像に付加するような高度な画像認識処理は一般的に大きな計算量を必要とする。そのため、従来は外付けのPCなどで画像認識を行うシステムとして構成されるか、あるいは顔検出や差分検知などの特定処理だけをハードウェアで実装し、固定機能として提供されていた。その結果、PCの設置場所や消費電力の増大、及び機能の固定化による汎用性の低下な

ど解決すべき課題が残されている。

東芝は、ネットワークカメラに画像認識 LSI “Visconti™2”⁽²⁾ を搭載したインテリジェントカメラを開発した。Visconti™2 は、種類の異なるマイクロプロセッサで構成されるヘテロジニアス マルチコア アーキテクチャを採用した LSI であり、歩行者の検出や運転者の認証、車両の検出など、複数のアプリケーションを効率的に実行可能である⁽³⁾。

この Visconti™2 をネットワークカメラに搭載することで、従来は PC で行っていた高度な画像処理をカメラ内で実行でき、消費電力や設置性の問題を改善できる。また、Visconti™2 は様々な画像処理を高速に実行できる柔軟な設計になっているため、LSI のソフトウェアを書き換えることで、人物や車両などの物体検出、顔認識を加えたセキュリティ機能、及び車の周囲を表示するアラウンドビューや道路のレーン検出といった視覚サポート機能のように、目的に応じた別の機能を同一のカメラハードウェアで実現できる。

インテリジェントカメラは、カメラで撮影した映像をカメラ内部でそのまま画像認識処理できるので、映像と同時に画像認識結果を送信できる。情報を受け取ったシステムやオペレーターは、例えば人数や顔認識結果などの画像認識結果に従って見るべき映像の重要度を判断できる。デモンストレーション

用として、人の顔、人の上半身、車両、及び猫の顔の4物体を同時に検出するアプリケーションをインテリジェントカメラに実装し、リアルタイムで検出処理が実行可能であることを確認した。

2 インテリジェントカメラの概要

インテリジェントカメラは図1に示すように、Visconti™2, CMOS (相補型金属酸化膜半導体) センサ, NOR (Negative OR Circuit) フラッシュ, DDR2 (Double Data Rate 2) メモリ, 及びネットワークコントローラで構成されている。

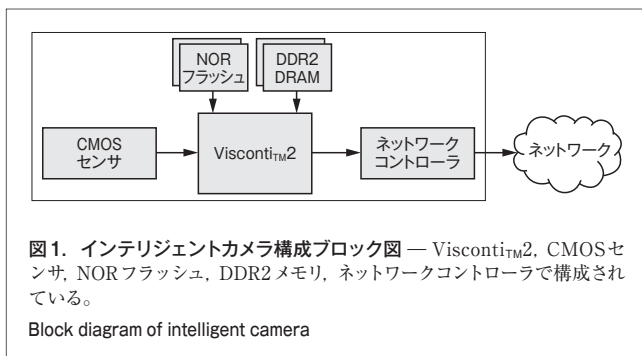
Visconti™2は、五つのMeP (Media Embedded Processor) を装備したマルチコアプロセッサ搭載のLSIとして設計されている(図2)。各コアは独立してそれぞれ別の処理を実行可能であり、更に、そのうちの4コアにはVLIW (Very Long Instruction Word) 型のSIMD (Single Instruction, Multiple Data) 演算コプロセッサIVC2が装備されている。また、多様な画像認識処理を高速に実行できるように、特定の処理に特化した画像処理アクセラレータ (IPA) を搭載している。

IPAは、HOG (Histograms of Oriented Gradients) 特徴量⁽⁴⁾を計算するHOGアクセラレータ、ヒストグラムを生成するヒストグラム アクセラレータ、各画素とその周辺への各種演算を行うフィルタ アクセラレータ、SAD (Sum of Absolute Difference) による画像比較を行うマッチング アクセラレータ、及び画像の拡大、縮小、回転、平行移動などを行うアフィン変換を実行するアフィン アクセラレータの計5種類を装備している。

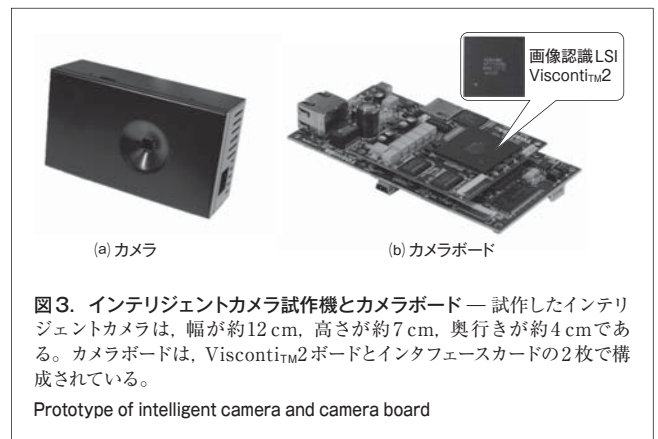
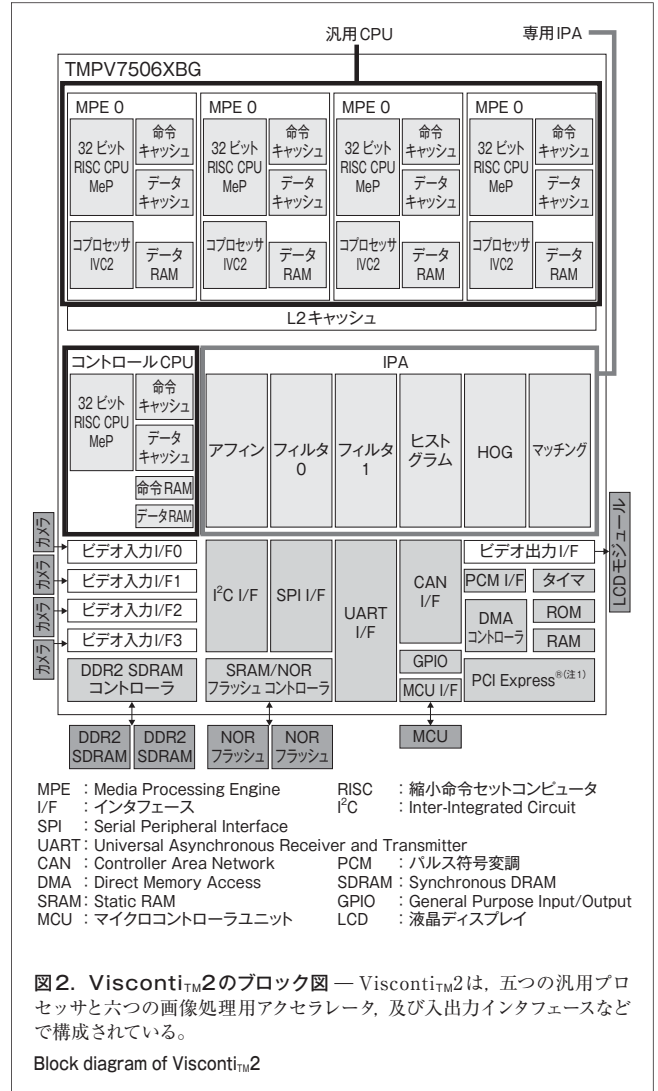
更に、MePやメモリ間をつなぐ経路のクロスバースイッチによる動的選択や、自由度の高いメモリアクセス最適化により、SIMD演算によるデータレベルから、VLIWによる命令レベル、IPAによるモジュールレベル、及びマルチコアによるアプリケーションレベルまで様々なレベルで柔軟な並列化が可能になり、全体として高い性能が発揮できるように設計されている。

インテリジェントカメラの試作機を図3(a)に示す。サイズは、幅が約12 cm、高さが約7 cm、奥行きが約4 cmで、電源はPoE (Power over Ethernet) で供給されるため、1本のケーブルでデータ通信と電源供給を同時に賄うことができる。

図3(b)はカメラボードである。カメラボードは、Visconti™2



ボードとインタフェースカードの2枚から構成されている。このインタフェースカードを交換することで、ネットワークカメラ以外にUSB (Universal Serial Bus) で接続するWebカメラとしても利用できるなど、拡張性が考慮されている。また、インタ



(注1) PCI Expressは、PCI-SIGの商標又は登録商標。

フェースカードには30万画素のCMOSセンサ、映像をMotion JPEG (Joint Photographic Experts Group) 形式に圧縮して送信するネットワークコントローラなども実装されている。

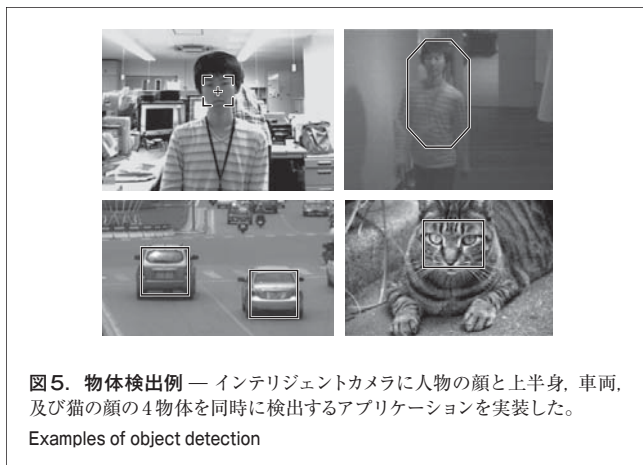
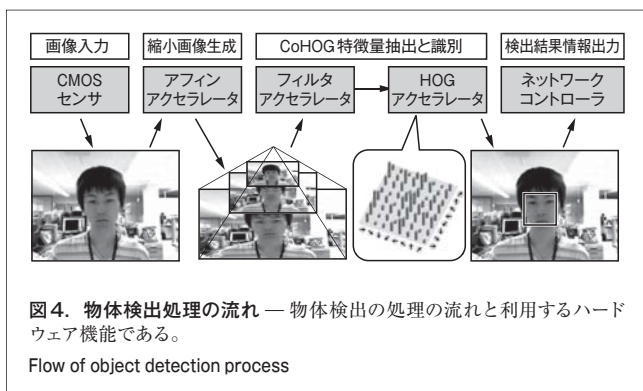
Visconti_{TM}2の制御プログラムはフラッシュメモリに格納され、更にワーキングメモリとしてDDR2メモリが利用できる。ネットワークコントローラはVisconti_{TM}2のビデオ出力とシリアルインタフェースで接続されており、映像だけでなく画像処理結果も送信することが可能である。

3 画像認識処理の流れ

インテリジェントカメラでは、CoHOG (Co-occurrence HOG) 特徴量⁽⁵⁾と線形SVM (Support Vector Machine) 識別器⁽⁶⁾を用いたパターン認識による物体検出が可能である。CoHOG特徴量は、画像から抽出した輝度の勾配方向を共起ヒストグラムにより表現する特徴量であり、線形SVMと組み合わせることで高精度に物体を検出できる。Visconti_{TM}2に搭載されているハードウェアアクセラレータを用いることで、高速な処理を実現している。

物体検出処理の流れを図4に示す。

まず、画像の変形を高速に行うアフィン アクセラレータを用いて、画像を複数の大きさに縮小する。次に、縮小画像から



CoHOG特徴量を抽出し、線形SVMによる識別処理を画像全体について繰り返し行う。フィルタ アクセラレータとHOG アクセラレータを用いることで特徴量抽出と識別を高速で処理する。識別に用いる辞書データは動的に変更可能である。

デモンストレーション用として、人物の顔と上半身、車両、及び猫の顔の4物体を同時かつリアルタイムに検出するアプリケーションを実装した(図5)。ソフトウェアで識別対象を瞬時に切り替えながら画面内の物体を検出し、検出結果の座標値などをネットワークコントローラにより外部に出力する。映像からの物体検出処理は、反復処理が多いためにPCでも負荷の大きい処理である。インテリジェントカメラは、その検出処理を4種類同時に実行させても10フレーム/s程度でリアルタイム動作しており、その有効性を確認できた。

4 あとがき

画像認識LSI Visconti_{TM}2を搭載したインテリジェントカメラを開発した。物体検出などの高度な画像処理をカメラ内で実行でき、辞書を切り替えることで人物や車両など、用途や目的に応じて処理内容を決定することが可能である。

文献

- (1) 矢野経済研究所. "世界の監視カメラ市場に関する調査結果2011". Yano ICT. <http://www.yanoict.com/yzreport/185>. (参照2012-12-20).
- (2) Tanabe, Y. et al. "A heterogeneous multi-core SoC for image-recognition applications". ISSCC 2012. San Francisco, 2012-02, IEEE. p.222 - 223.
- (3) 岡田隆三 他. "次世代車載画像認識LSIによる安全運転支援". Vision Engineering Workshop (ViEW)2011, 横浜, 2011-12, 精密工学会. p.268 - 274.
- (4) Dalal, N.; Triggs, B. "Histogram of oriented gradients for human detection". Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, 2005-06, IEEE. p.886 - 893.
- (5) Watanabe, T. et al. "Co-occurrence Histogram of Oriented Gradients for Human Detection". The 3rd Pacific-Rim Symposium on Image and Video Technology, Lecture Notes in Computer Science, 5414. Tokyo, 2009-01, IEEE. p.37 - 47.
- (6) Vapnik, V. The Nature of Statistical Learning Theory. Berlin, Springer-Verlag GmbH, 1995, 188p.



小坂谷 達夫 KOZAKAYA Tatsuo

研究開発センター インタラクティブメディアラボラトリー研究主務。画像処理及びパターン認識の研究・開発に従事。情報処理学会会員。
Interactive Media Lab.



渡辺 友樹 WATANABE Tomoki

研究開発センター インタラクティブメディアラボラトリー。車載及び監視カメラ向けの人物検出技術の研究・開発に従事。
Interactive Media Lab.



岡田 隆三 OKADA Ryuzo, Ph.D.

研究開発センター インタラクティブメディアラボラトリー主任研究員、博士(工学)。画像認識技術の研究・開発に従事。電子情報通信学会会員。
Interactive Media Lab.