

画像認識プロセッサ Visconti と、その安心・安全への適用事例

中井 宏章, 古川 賢司, 小坂谷 達夫 (株式会社 東芝, 研究開発センター),
宮森 高 (株式会社 東芝, セミコンダクター社, SoC 研究開発センター)

あらまし 安心・安全についてのコンピュータビジョン技術への期待に応えるには、ユーザが求める機能を実現する様々な画像処理手法の確立が必要であるが、一方、この技術の恩恵を誰もが広く日常生活で享受できるようにするには、高度な画像処理の実装に適した高性能・小型で安価かつ信頼性の高い画像処理プラットフォームの実現が大変重要である。本稿では、様々な画像処理を効率的に実行できるアーキテクチャを持つ組み込み用画像認識プロセッサ Visconti[®] と、これを用いた ITS、セキュリティ等の産業分野での開発事例について報告する。

キーワード 画像処理ハードウェア, ITS, 障害物検出, セキュリティ, 顔認識, 呼吸計測

Image Recognition LSI Visconti and its applications to Safety and Security

Hiroaki Nakai, Kenji Furukawa, Tatsuo Kozakaya (Corporate R&D Center, Toshiba Corporation), and
Takashi Miyamori (SoC R&D Center, Semiconductor Company, Toshiba Corporation)

Abstract: In this paper, we present several examples of practical developments in image recognition technologies for use in automotive and security applications. The first example is image processing hardware; where Visconti[®] is perhaps the most efficient platform for visual sensing techniques. The second comprises various image processing algorithms for the sensing of surrounding obstacles to automobiles, human faces and conditions. Combination of such hardware and software form a smart sensor; which are widely expected to become pervasive in many aspects of daily life.

Keywords: Image Processing Hardware, Intelligent Transport Systems, Obstacle Detection, Security, Face Recognition, Respiration Monitoring

1. はじめに

CCD や CMOS カメラの小型化・高信頼性化・低価格化の目覚ましい進展により、多くの産業分野で画像処理機器の利用が検討され始めている。例えば自動車では、ドライバーから見えにくい死角エリアやブラインドコーナーでの視認性向上のために既に多くの自動車にカメラが搭載されているが、不慮の事故も防げるよう危険な状況を事前に察知してドライバーに知らせるセンシング機能の実現が近年期待されるようになってきた。このためには、前方や周囲の走行車両や歩行者、電柱等の周囲の構造物や駐車車両等を正しく検知し距離計測できる手段が必要であり、レーダーを用いた能動的センシングの実用化が進められている。一方、物体の種類や存在する方向、あるいは走行レーンや標識等の周囲情報を詳細に得るにはカメラを用いた画像センシングが優れることから、今後、これら複数センサが組み合わせられて自動車に搭載されるようになると

見込まれ、将来は人間と同じように高度で知的な判断ができる自動車や安全装置が実現すると期待されている。画像処理による物体検知・物体認識は、他にもインフラ系 ITS やバイオメトリクス、自動画像監視等、安心・安全にかかわる産業分野での基盤技術として日に日に期待が高まっている。コンピュータビジョンを自動車やセキュリティ等の産業分野で応用するには、大別して二つの問題が存在する。一つは、ダイナミックに変化する屋外環境でも多種多様な物体や状況を確実に検知できるアルゴリズム(ソフトウェア)、もう一つは、車載や屋外等の厳しい周囲条件でも動作可能な高性能・高信頼性・低価格を全て満たす画像処理ハードウェアである。本稿では、これらの問題に取り組んだ幾つかの開発事例について紹介する。まず、安心・安全のための画像認識を実現する上で基盤となる画像処理ハードウェアについて、続いて、様々な目的に応じた画像認識アルゴリズムの開発事例について述べる。

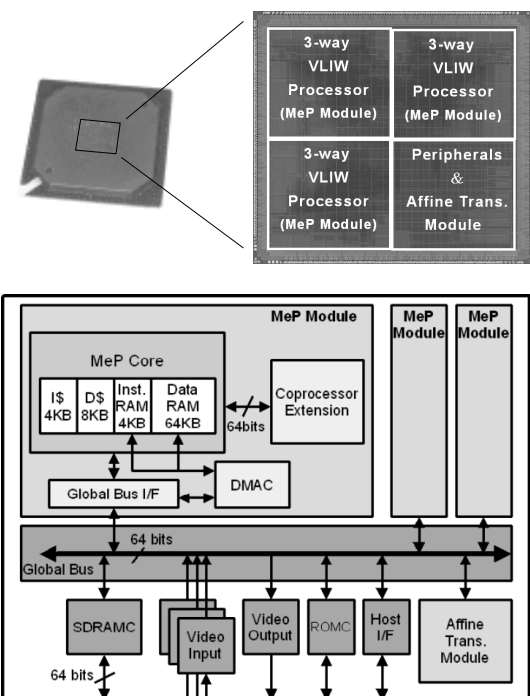


図 1: Visconti[®] チップ写真とアーキテクチャ概略

2. 画像処理ハードウェア

車載カメラを用いた画像認識はレーン逸脱警報や走行補助で既に実用化されており、さらに周辺の障害物検出や車内の乗員検知等への適用も検討が進められている。これらに用いる画像処理ハードウェアは、実時間処理でき(高性能)かつ故障しない(高信頼性)ことが必須であるが、低価格であることも普及には重要である。特定の画像処理を高速実行できる専用回路が組み込まれた LSI は量産時のコスト面では有利であるが、実際のシステム開発では性能向上のための試行錯誤が避けられないため、一部の修正でも全体を作り替えないければならぬ専用 LSI 方式は、かえって開発期間とコストが増大してしまう恐れがある。したがって、ソフトウェアを入れ替えるだけで容易に処理が変更できる汎用プロセッサタイプの画像処理 LSI の実現が望まれる。高度な画像認識には最新鋭パソコンの CPU と同程度の画像メディア処理性能が必要となるが [2]、動作周波数を高くして高性能を実現するパソコン用 CPU は発熱量が多く自動車応用には向かない。よって、高度な並列化処理を実現する方向で車載用の高性能プロセッサの開発が続けられている。その 1 つとして、車載条件で動作し(周囲温度、消費電力、振動)、低価格である(ダイサイズが小さい)という特徴を持つ車載用

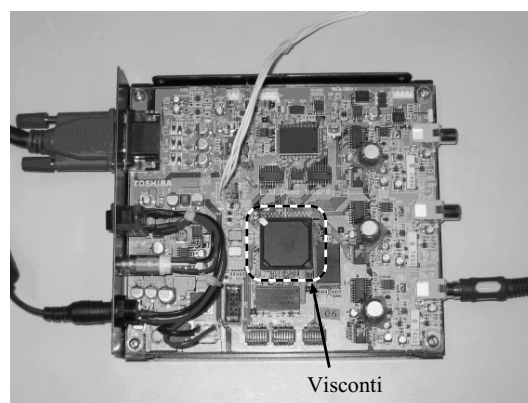


図 2: 評価用画像処理ハードウェア

画像認識プロセッサ Visconti[®] を紹介する [3]。

Visconti[®] は、画像認識に多く共通する演算を最適効率で実行できるよう、プロセッサ設計に様々な工夫が盛り込まれている。概要について図 1 を用いて説明する。1 チップに 3 つのメディア処理モジュール [1] が含まれており、各処理モジュールには SIMD (Single Instruction Multiple Data) と呼ばれる並列演算命令をサポートするコプロセッサ拡張とローカルメモリが含まれ、画像認識で頻りに用いられるフィルタリング、マッチング、ベクトル演算といった処理が最適な効率で実行できる設計となっている。動作周波数 150 MHz での処理性能は 18 GOPS 以上で、1.3 GHz 以上の汎用 CPU と同等のメディア処理性能を発揮するが、平均消費電力は約 1 W と小さく、動作周囲温度は -40 ~ +85 である。また、周辺回路としてメモリコントローラや 3 系統の画像入力、VGA 出力、アフィン変換モジュール等を内蔵しており、画像認識用の様々な機能が 1 チップに集積されている。Visconti[®] は ROM に保存されたファームウェアを実行するため、後述で紹介するように、ファームの変更により様々な画像認

表 1: Visconti[®] 諸仕様

Feature	Specification
Technology	0.13mm CMOS 6-layer metal
Peak performance	18 GOPS (6 GOPS × 3 processors)
Clock frequency	150 MHz
No. of transistors	21 million (17 million for memory)
On-chip memory	260 K bytes
Power consumption	1 W @ 1.5 V
Chip size	6.98 mm × 6.98 mm (48.7 mm ²)
Package	456 pin PBGA

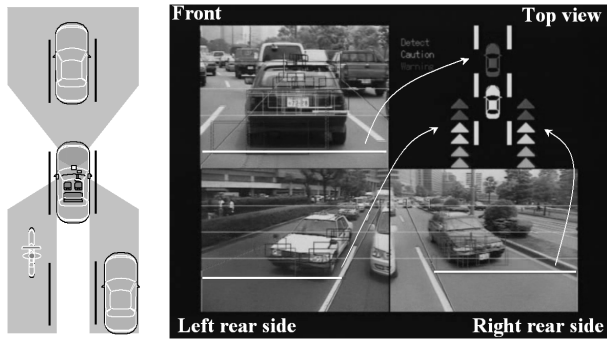


図 3: 車両周辺監視システム

識処理を実行することができる。画像処理専用 LSI のように処理に応じてゲートレベルの設計をする必要がないため、用途に応じて動作を切り替えることも可能である。参考として、Visconti[®] 諸仕様を表 1 に、評価用に試作した画像処理ハードウェアを図 2 に示す。ここで紹介した画像認識 LSI は高性能・高信頼性・小型低価格のいずれの要件も満たし、車載用途だけにとどまらず広い産業応用が可能である。

3. 画像認識システムの開発事例

前述のようなプログラマブルな画像処理ハードウェアとカメラとの組み合わせにより、様々なビジョンシステムが容易に構築できる。本章では、4つのシステム開発事例について概要を説明する。いずれも Visconti[®] 画像処理ハードウェア上に実装されている。ソフトウェア(アルゴリズム)面では、多様な周囲環境やその変化への対応をいかにして可能にするかに開発のポイントが置かれており、周囲環境や対象物についての仮定をできる限り減らしたり、環境変化に影響を受けることの少ない画像特徴や判定方式を用いるアプローチが取られている。これにより、屋内外、天候や昼夜、多様な対象物に対しても調整が少なく、安定動作する画像認識が実現できる。

3.1. 車両周辺監視システム

自動車の前方と左右後側方の3方向に向けて配置された3つのカメラ(図3)を用いた車載周辺監視システムが開発されており[4]、先行車や追い越し車両等、各方向で接近する車両を同時に検出し、運転者に知らせることができる(図右上 印等が車両接近を示している)。警報装置では、ユーザーがスイッチオフしたり、かえって事故を誘発してしまうことのないよう、誤報

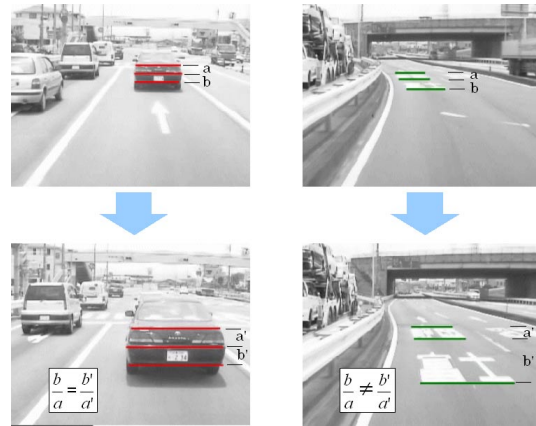


図 4: 立体物判別の簡略な説明

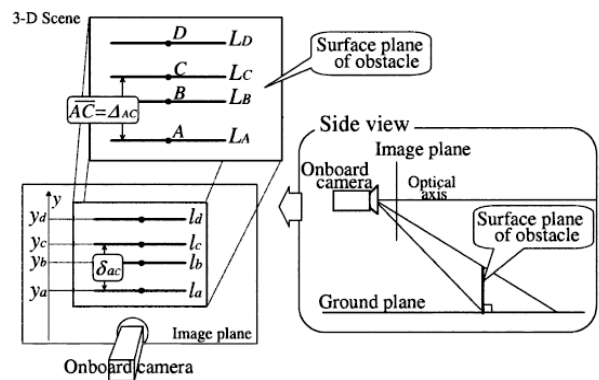


図 5: 直線上の4点 A,B,C,D と画像面への射影

をできる限り減らすことが重要である。このシステムでは複比立体判別法[5]を用い、各方向カメラ1台で立体物である接近車両だけを検知できるようになっている。単眼で立体物有無を判定する方法を端的に説明する。図4中3つの水平線分の間隔 a, b は、その線分が立体物のものであれば、接近する場合でも間隔比が一定に保たれる。これに対し、水平線分が道路面のものであれば、線分の間隔は、遠くの a よりも近くの b の方が大きく広がる。したがって、線分間隔の時間変化から立体物が道路か否かの判定がカメラ1台で行えることになる。具体的には次のように定式化できる。同一直線上の4点から求まる複比は射影変換に関する不変量である。今、図5のように道路面あるいは障害物背面(道路面に垂直と仮定)に水平線分4本が検出されると、それら線分間距離には以下の関係が成り立つ。

$$I_{ABCD} = \frac{\Delta_{AC}}{\Delta_{AD}} \cdot \frac{\Delta_{BD}}{\Delta_{DC}} = \frac{\delta_{ac}}{\delta_{ad}} \cdot \frac{\delta_{bd}}{\delta_{dc}}$$

ここで、 Δ_{IJ}, δ_{ij} は三次元空間あるいは画像面での線分 I, J あるいは i, j の距離である。最右辺は線分が同

一平面上にあれば時間に関して不変であること、障害物背面の消失線 $y = \infty$ と道路面の消失線 $y = 0$ を線分 l_d の位置として導入することにより、障害物背面と道路面に関する 3 本の水平線分の運動拘束が各々次のように得られる。

$$\frac{y_a(t_2) - y_c(t_2)}{y_a(t_1) - y_c(t_1)} = \frac{y_b(t_2) - y_c(t_2)}{y_b(t_1) - y_c(t_1)} = M_v$$

$$\frac{y_a(t_2)^{-1} - y_c(t_2)^{-1}}{y_a(t_1)^{-1} - y_c(t_1)^{-1}} = \frac{y_b(t_2)^{-1} - y_c(t_2)^{-1}}{y_b(t_1)^{-1} - y_c(t_1)^{-1}} = M_h$$

ここで $y_i(t)$ は時刻 t での線分 i の y 座標値である。したがって、画像で観測された水平線分 3 本を追跡し、各 y 位置がいずれの拘束を満たすかを判断すれば、それら線分が障害物背面にあるか道路面にあるかが判定できる。

この方法は対象物の種別にはよらず適用できて事前の知識や仮定も必要なく、実装や調整が容易である。処理の流れは 1) 水平線分検出、2) 連続フレーム間での線分追跡、3) 複比計算による判定の 3 段階で、演算量が多い 1) と 2) の画像処理の部分は、前述 Visconti® の SIMD 演算命令を用いた実装で大幅な高速化が可能である [4]。3 方向の接近車両検知を 1 チップの Visconti® で実時間処理できるコンパクトな周辺監視システムが実現している。

3.2. 前方障害物検出システム

高速道路等での追従車間制御や衝突回避を実現するには、遠方の車両や障害物を検出し、距離を正確に計測する必要がある。画像からの距離計測にはステレオ視が適しており、これを基本とした前方障害物検出システムの開発事例を次に紹介する [6]。このシステムでは、障害物検出に平面投影ステレオ法 [7] を使い、天候や照明条件の変化への頑健性を実現している。通常のステレオ視では、カメラの共通視野にある物体の位置を求めるため特徴点抽出とその近傍でのステレオ対応付けが必要となる。ところが、カメラ位置によって抽出される特徴点が変わる、対象物の見え方が変わる、横断歩道等の繰り返し模様がある等により、正確な対応付けが困難となり距離計測を誤る問題がある。平面投影ステレオ法では、この対応付け問題を回避できる。図 6 を用いて処理の流れを説明する。画像中の全ての物が道路平面上にあると仮定し、一方のカメラ画像 a) を他方のカメラ位置から見た画像に平面射影

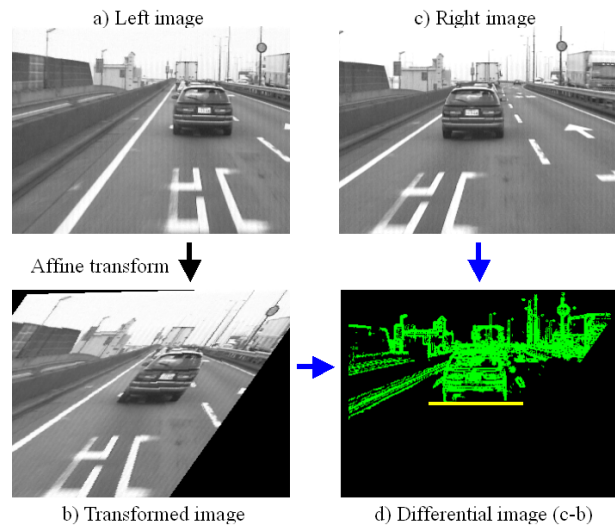


図 6: 平面投影ステレオ法の流れ

変換する b)。変換画像と他方のカメラの画像 c) とを比較すると d)、道路面にあるという仮定が成り立たない立体物の部分だけに画像の違いが生じ、差分等の単純な処理だけで障害物位置が求められる。これを発展させ、2 つのカメラが前方遠方の同方向を向いている場合には上記平面射影変換が単純なアフィン変換に近似できる性質を利用すれば [8](図 7)、道路面位置から視差を予測できるため水平方向の対応探索が不要となる (図 8 上)、雨天や夜間等の悪条件の元でも対象物位置決定が正確に行える (図 8 下、R-R' は対応し Q-Q' は対応しない)、さらに、左右カメラ画像と平面投影画像との 3 画像間でのブロックマッチングにより立体物検出のため近似的な統計判別が行える等、様々な天候や道路条件でも検出性能が劣化しない障害物検出方式が実現できる (図 9)。この方式は、画像の幾何変換の他は差分処理やブロックマッチング等の並列処理に適した演算から構成されるため、前述の Visconti® 画像処理ハードウェアを用いた実装では、アフィン変換モジュールの活用等により毎秒 30 フレームの処理が実

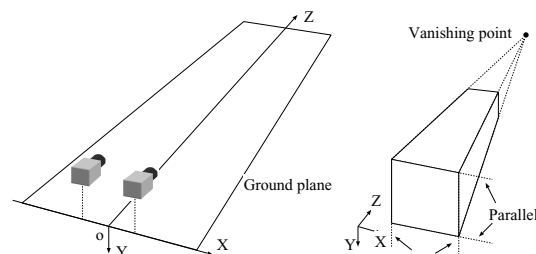


図 7: カメラモデル

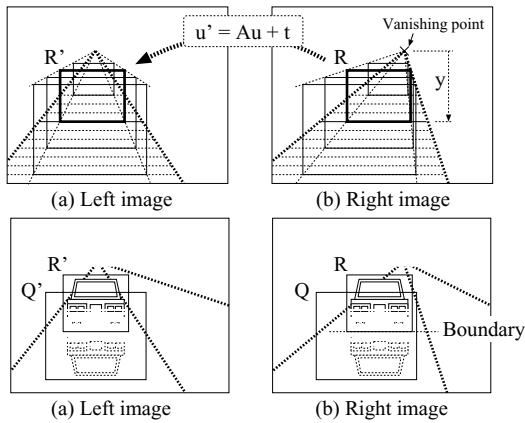


図 8: 対応探索領域と接地位置決定



図 9: 悪条件下での先行車両検出例 (雨天, 夜間)

現している。

3.3. 顔認証システム

近年、セキュリティ分野では生体情報を用いた個人認証技術が注目を集めているが、中でも顔画像による個人認証は、非接触に認証でき、同時に顔画像が記録できる等、利便性の面で優れた方式である。個人認証を正確に行うには、顔の向きや表情の変化、周囲環境の照明変化に影響を受けにくい高精度で頑健なパターン認識処理が不可欠である。これまでは、計算量の多さから高性能のパソコン用 CPU を用いた装置が開発されてきたが [9]、この高精度なパターン認識を Visconti[®] を用いて実現した例を述べる [10]。

顔認証処理の流れは顔検出と顔認識の 2 段階からなる。顔位置の検出精度は認証性能に大きな影響を与えるため、顔部品形状情報 (エッジ) と顔部品の濃淡パターン情報を組み合わせて判定する高精度な顔検出方法を用いる [11]。まず、円形形状を持つ画像特徴のみを安定に検出できる分離度フィルタにより特徴点候補 (瞳と鼻孔の 4 点) を検出する。図 11 に円形分離度フィルタの模式図と顔特徴点候補の検出結果を示す。図中丸印が検出された特徴点候補である。次に正しい候補

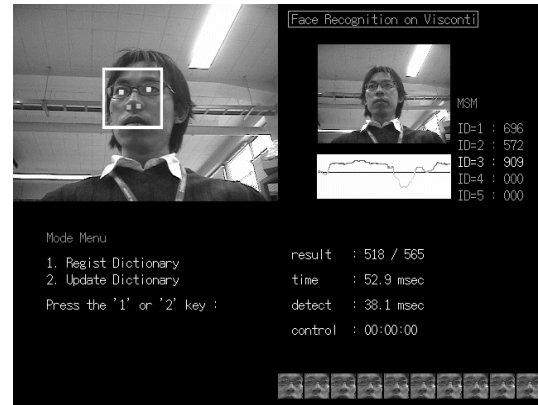


図 10: 顔認証システム

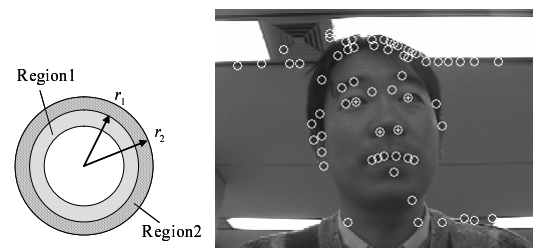


図 11: 円形分離度フィルタと顔特徴候補

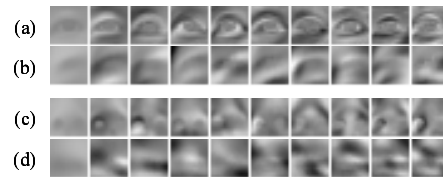


図 12: 部分特徴辞書

点を選択するため、目鼻の各部品パターンから辞書を作成しておき (図 12 a,c)、部分空間法により濃淡パターンの検証を行う。また、目じりや眉毛など似たパターンが誤検出される場合があるので、誤りやすいパターンからも辞書を作成し (図 12 b,d)、正しい辞書と誤り辞書との両者の類似度を用いることで、誤検出を減らしロバストに特徴点を検出する。図 11 中 (+) で示された目鼻の 4 点が選択された特徴点を示す。この特徴点を基準として、アフィン変形により向きと大きさを正規化して顔パターンを切り出す。

高精度な顔認識の実現には、顔パターンの変動による影響が少ない方式を用いることが重要である。このシステムでは、制約相互部分空間法 (CMSM: Constrained Mutual Subspace Method) による動画像系列を用いた顔認識を行う [12]。まず CMSM の元となる相互部分空間法 (MSM: Mutual Subspace Method) の概要を図 13

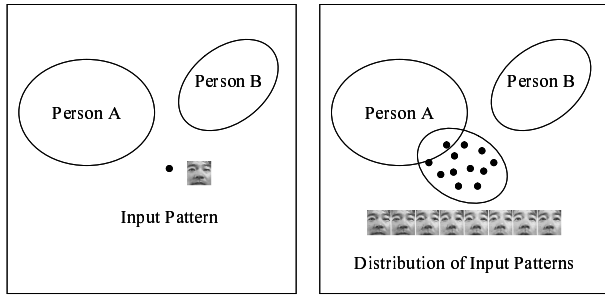


図 13: 相互部分空間法の概念図

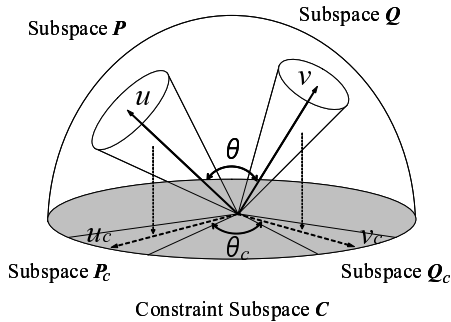


図 14: 制約相互部分空間法

を用いて簡単に説明する。従来では、1枚の静止画像と登録人物の顔パターン分布(部分空間)との類似度を計算するため、左図のように顔向きや表情変化により入力と登録データの距離が大きく変化する場合があって安定な認識が難しい。これに対して右図のように動画像系列を用いた手法では、入力にも分布を用いるため顔向きや表情変化によるパターン変動に対しロバストな認識が可能となる。さらに CMSM では、登録時と認証時で照明条件が異なるといった照明条件変動にも対応できるよう、登録パターンと入力パターンの部分空間の類似度を計算する前に、照明条件を抑制できるような制約部分空間(と名付けた空間)へ射影する処理を行う。具体的には、2つの部分空間 P, Q を考え、予め様々な環境で撮影したデータをもとに差分部分空間と呼ばれる部分空間を計算し(詳しくは [12])、射影後の部分空間 P_C, Q_C との類似度を計算することで認証処理を行う。図 14 に制約相互部分空間の概念図を示す。類似度には、2つの部分空間のなす角度として定義される正準角 θ

$$\cos^2 \theta = \sup_{\|u_C\| \neq 0, \|v_C\| \neq 0} \frac{|(u_C, v_C)|}{\|u_C\|^2 \|v_C\|^2}$$

を用いる。ここで u_C, v_C は $u_C \in P_C, v_C \in Q_C$ を満

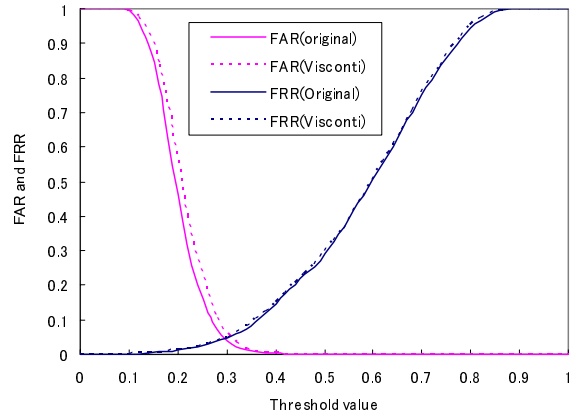


図 15: 認識性能比較結果

たすベクトルである。

以上の顔認証処理では、顔検出のための全画面探索、顔認識のための高精度なベクトル行列演算が処理の中心となる。Visconti[®] を用いて実時間処理を実現するため、画像処理の並列化だけではなく、演算の固定小数点化、3つの処理モジュールへのタスクの動的割り当て等、様々な最適化を行ったところ、秒 20 フレームの準実時間で顔検出から認識までを行えるシステムが実現した。また、認証性能においても高性能 PC を用いて実装した場合と遜色のない結果が得られている(図 15)。このシステムには冷却ファンといった機械構成要素が一切ないため、盗難防止や乗員検知等の車載応用に使えるだけでなく、小型で低価格、設置場所を選ばない顔認識システムとして、家庭用の高機能インターホン、情報家電やロボット等、従来では適用が難しいと考えられていた幅広い応用分野での顔認識技術の展開が期待できる。

3.4. 就寝モニタリングシステム

高齢者の増加や睡眠時無呼吸症候群、乳幼児突然死症候群といった睡眠障害や病疾への関心の高まりから、睡眠時の生理状態を自動監視できる装置の実現が望まれており、病院や老人ホーム等でのフェイルセーフ用途も含め、一般に広く在宅で使用できるような操作が容易かつ安価な装置の実現が期待されている。このような睡眠時の呼吸計測では、非接触計測により被験者が真に自然な状態で睡眠できること、オペレータの操作なしに長時間の連続計測が可能であることが重要であり、これを実現する画像計測による就寝モニタリ

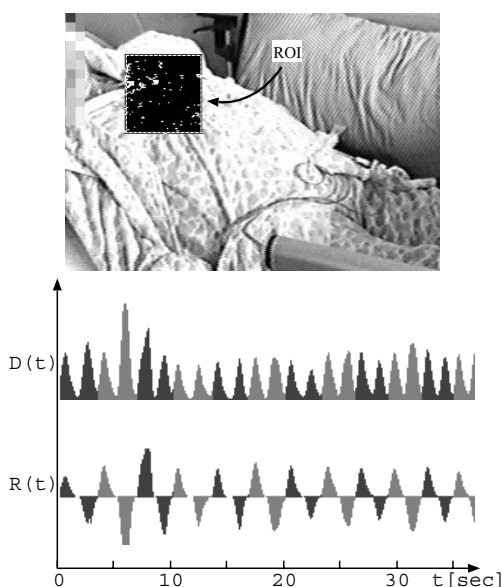


図 16: 呼吸計測のための ROI 設定例と呼吸計測例

グシステムが報告されている [13]。このシステムでは、暗環境でも S/N の高い呼吸信号が獲得できるような画像計測手法が採用されており、睡眠中の呼吸停止や非呼吸体動等の自動判定を実装することで長時間の自動計測を可能としている。

処理の流れは、1) 呼吸計測のための ROI を被験者胸部に自動設定する (図 16 上図)、2) ROI 内の差分変化量から被験者の呼吸波形を獲得し、同時に呼吸周期や画像変化量を算出する、3) 被験者の就寝状態を判定する、の 3 段階からなる。呼吸波形はフレーム間絶対値差分の ROI 内和から求めるが (図 16 $D(t)$)、被験者の吸気呼気の動作切り替わり時に極小となるため、この極小のタイミングで $D(t)$ 波形を正負反転させることにより呼吸波形 $R(t)$ を求める (図 16 $R(t)$)。この呼吸波形 R のゼロ交差間隔から呼吸周期 T は容易に計測できる。暗環境での呼吸計測を容易にするため差分フレーム間隔を呼吸周期に合わせて長く設定する。また、呼吸計測においては睡眠中の寝返り体動やトイレのための離床・着床といった所在変化は避けられないことと、暗環境にて S/N の高い呼吸波形を獲得するために、最適な位置に ROI を自動設定できる機能が不可欠となる。ROI 設定は、フレーム間差分値を一定時間積算して最大の積算値となった領域に設定する方式としている。さらに、呼吸間隔、呼吸波形パワースペクトラム、明度分布ヒストグラムの時間変化等を利用し、正常呼吸、無呼吸、非呼吸体動、所在変化等の被

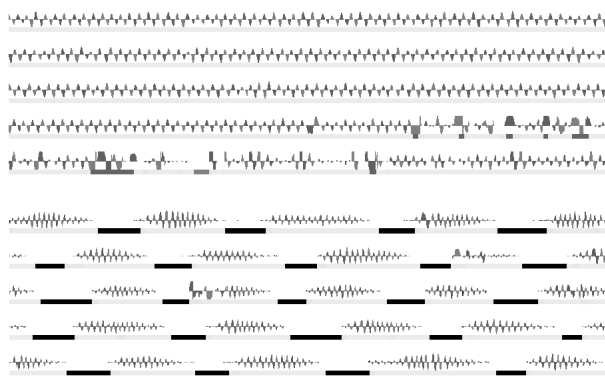


図 17: 呼吸計測例 (上段: 健康者 安静睡眠時、下段: 睡眠時無呼吸症候群患者)

験者の就寝状態を自動判別する機能を実装し、危険な無呼吸状態が発生したときには警報を発することができるシステムとなっている。実際に就寝時の呼吸計測を行った例を図 17 に示す。このシステムでの主要な画像処理はフレーム間差分という単純な処理であるため、Visconti[®] 画像処理ハードウェアで高速処理することは容易であり、毎秒 30 フレームの呼吸計測システムが実現している。

4. おわりに

本稿では、画像処理ハードウェアと画像認識アルゴリズムの具体的な開発事例を通じ、コンピュータビジョン分野での安心・安全への取り組みの一端について紹介した。安心・安全に関する産業分野ではコンピュータビジョンへの期待がますます高まる一方で、普及のための頑健性と低コスト化がさらに強く求められるようになっている。この期待に応えるには、様々な応用で共通して採用されるアルゴリズムや演算を明確にしてハードウェア設計に取り入れ、利用可能なリソースを無駄なく最大限に利用できるシステムを開発する等、ハード・ソフトの両面での同時進行的な研究開発が今後ますます重要になっていくと考えられる。

参考文献

- [1] Media embedded Processor,
<http://www.MePcore.com>
- [2] Y. Kondo, et al., A 4GOPS 3Way-VLIW Image Recognition Processor Based on a Configurable

Media-processor, ISSCC, digest of technical papers, pp.148-149, 2001.

- [3] J. Tanabe, et al, Visconti: Multi-VLIW Image Recognition Processor based on Configurable Processor, CICC, pp.185-188, 2003.
- [4] 古川 他、車載用画像処理 LSI を用いた車両周辺監視システム, SSII, pp.227-232, 2003
- [5] 岡田 他、複比と消失線に基づく車載単眼障害物検出, 信学論 D-II, Vol. J87-D-II, No. 12, pp. 2165-2175, 2004.
- [6] H Nakai, et al., A Practical Stereo Scheme for Obstacle Detection in Automotive Use, ICPR, Vol. 3, pp.346-350, 2004.
- [7] 小野口 他、平面投影ステレオ法を用いた道路領域抽出, 情処研報 CVIM-93-7, pp.61-68, 1995.
- [8] H. Hattori et al., Stereo without Depth Search and Metric Calibration, CVPR, vol.1, pp.177-184, 2000.
- [9] 顔照合セキュリティシステム FacePass[®], <http://www3.toshiba.co.jp/snis/scd/fps/>
- [10] 小坂谷 他、Visconti[®] を用いた顔認識システムの開発, SSII, pp.483-488, 2004.
- [11] 福井 他、形状情報とパターン照合の組み合わせによる顔特徴点抽出, 信学論 D-II, Vol.J80-D-II, No.8, pp.2170-2177, 1997.
- [12] 福井 他、制約相互部分空間法を用いた環境変化にロバストな顔画像認識 – 照明変動を抑える制約部分空間の学習 –, 信学論 D-II, Vol.J82-D-II, No.4, pp.613-620, 1999.
- [13] 中井 他、動画画像処理による呼吸モニタリングシステム, 信学論 D-II, Vol.J83-D-II, No.1, pp.1-9, 1999.