

02

鉄道と乗降客の安全を 確保する画像認識技術

一踏切、ホーム端からの転落防止一

依田育士

産業技術総合研究所 情報技術研究部門

ステレオカメラを多地点に配置し、ネットワーク利用を前提として、人を中心とする（人の存在、姿勢、顔、意図したジェスチャなどの）リアルタイム認識技術に関する研究を行っている。実時間実環境でのヒューマンセンシングを提供するために、ステレオカメラを利用することで激しい影の変化などを含む屋内外のいかなる場所でも設置可能にすること、距離情報と画像情報を同時に利用して役立てることに研究の中心をおいている。そのために独立した複数ステレオカメラによる複数対象の認識を前提とし、距離情報から対象領域を注目する手法、対象の記述・認識手法を開発している。その典型的な応用事例として、5年以上にわたり安全・安心を実現するための継続的な実験を行っている踏切監視、駅ホーム端監視のためのヒューマンセンシング事例について述べる。

はじめに

近年、社会生活の複雑化に伴い、セイフティ・セキュリティサービスを自動化する技術のニーズは高まっている。現実の交通監視事例など車を中心とした実環境において画像処理が用いられている事例は多数存在する。具体的には混雑状況の監視や、自動車ナンバーの記録などがあげられる。

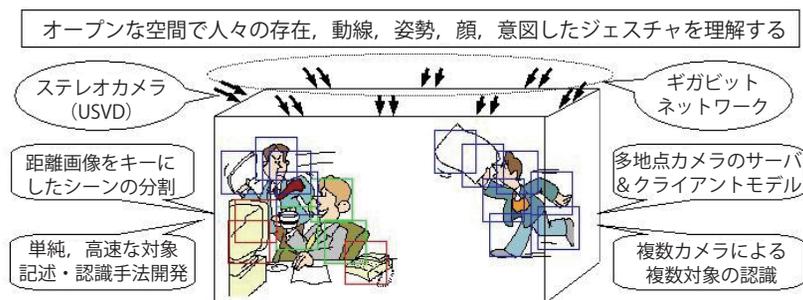
一方、安全の確保を目的とした人を中心とした監視となると現実の適応事例はきわめて限定される。たとえば単眼カメラを用いて人のトラッキングを行う研究^{1), 2)}はいくつか存在する。しかし、論文等が公開されている事例の多くは、正規化されたデータセットもしくは安定した実験室環境から得られたデータを使用している。現実の屋外環境においては、激しい天候変化や影などの外乱要因が複雑に絡み合って存在している。ゆえに、特に高い安定性と認識精度が要求される安全・安心への具体的な適応事例が限定されることとなる。

そもそも、これらの基礎となる画像認識に関連する研究はすでに30年にわたる歴史が存在し、あらゆる手法の開発・適応が試みられてきた。しかしながら、これほど研究開発が進み、適応範囲が広がっているにもかかわらず、一切制約のない実世界を対象とした画像理解、認識といった範疇に踏み込むと、問題の難易度は飛躍的に高まる。よって、共通のフレームワークを作ることがいまだに非常に困難な問題と考えられる。そこで、実世界

では最も困難な認識対象を正確に注目するという問題に対し、距離情報を用いることでセグメンテーション問題を劇的に容易にし、その後の対象認識を実時間実環境で安定的に実現しようとするアプローチが有望な手段となってくる。

具体的には、画像処理を基本にステレオカメラの情報を利用し、かつ、特化された対象から得られる制約をうまく利用することで、実環境での適応を実現した事例がいくつかあげられる。たとえば、辻らの自動車搭載型の歩行者認知システム³⁾の開発があげられる。これは夜間の道路上における人の有無の検出において、赤外線ステレオカメラを用い、車からの距離と人の持つ温度という2つの要素を組み合わせて人検出の信頼性を向上させている。同様に対象の限定された環境をうまく利用したシステムとして、プールのライフガード向け業務補助アプリケーション“Poseidon”⁴⁾があげられる。ここでは、プールの壁面に設置された複数台の単眼カメラから得られる画像に対して、プールという特性を利用して既知であるカメラの位置関係と画像上での位置関係を対応づけることにより、影と溺れている人物の区別を行っている。

また、歩行者検出を行う研究もいくつか行われている。寺田ら⁵⁾は、人の移動方向が限定される屋内の細い通路の天井に、1台のステレオカメラを真下へ向けて設置し、その下を通過する歩行者をカウントしている。ここでは人の移動方向と垂直に交わる直線上に生成される時空間距離画像を監視することで、人の通過を検出している。



■図-1 ユビキタスステレオビジョンの概念図

Zhao の研究⁶⁾では、自動車にステレオカメラを搭載し、前方に存在する人を検出するシステムが提案されている。ここでは3次元空間情報ではなく視差画像のみを用いてセグメンテーションを行い、複雑なシーンにおける背景と人物の分離を図っている。

また、今回の中心テーマとなる踏切内障害物検知システムの研究として、柴山らの研究⁷⁾があげられる。この研究では、テクスチャ画像と比較して環境光の変動に影響されにくい特徴を持つ距離画像を用いた背景差分を行い、昼夜を問わず、単独のステレオカメラの監視領域内への侵入物体を検知している。

一方我々は、ステレオカメラを固定的な設置で利用すること、最初の認識対象を人とするを前提に研究を進めている。特に、距離情報算出には特定の対象や知識などを用いず、一般的に距離情報を算出しており、実用的、かつ適応範囲が広いことが特長と考えている⁹⁾。このような前提で屋内外に通じるビジョン技術を確認することが目標である。そのために、激しい影の変化などを含む屋内外のどこにでも対応可能なステレオカメラを設置し、そのときに得られる距離情報と画像情報を同時に利用し、いかに役立てるかに注力している。そのためには個々のステレオカメラにおいて、その距離情報から複数の対象領域を見つけ出し、認識のための特徴抽出を行った上で統合する。複数のステレオカメラサーバとそれらを利用するアプリケーションクライアントという枠組みを提供するのがユビキタスステレオビジョンである。その概念図を図-1に示す。

このステレオビジョンによるヒューマンセンシングの典型的な応用事例として、安全・安心のための鉄道乗降客の安全を確保する画像認識技術について述べる。

ユビキタスステレオビジョン

◆ユビキタスステレオビジョンの基本構成

「ユビキタスステレオビジョン」とは、ステレオビジョ

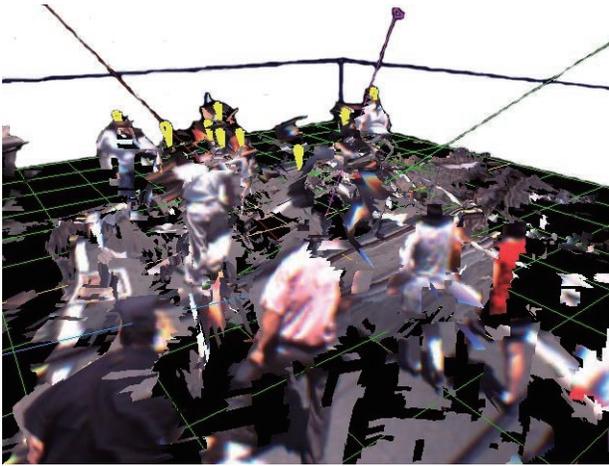
・入力情報は複数のステレオカメラから得られる3次元距離とテクスチャ画像(カラーorモノクロ)
・実時間(リアルタイム)で利用
・実環境で利用(屋内外で利用可能)
・各ステレオカメラはスタンドアロンで稼働
・複数ステレオカメラは静的配置、キャリブレーションは事前に1回
・高速ネットワークの利用を基本とするが、低速でも稼働
・動的制御(パン・チルト・ズームなど)をベースとはしない
・切り出しは距離情報を利用し、対象依存モデルを作らない
・認識手法は対象に依存しないものが基本

■表-1 USVの定義

ンを用いることで「どんな場所でも頑健に使えるから」、ステレオ利用を前提とした「どんな場所でも使えるソフトウェアとハードウェアの開発を行う」というシンプルな理由で2000年頃に命名した。そもそもの研究の出発点は、RWCプロジェクト後半¹²⁾に関連し、計算機とユーザのより自然な対話を実現するために、基本的にステレオカメラ1台対1人のユーザのための学習型ステレオビジョン⁸⁾を研究していた。そこでプロジェクト最終年度を迎え、研究の総括と今後の展開を考慮し、今までの仕組みを複数ステレオカメラと複数人という枠組みに拡張し、総合的なヒューマンセンシングとして捉え直すことを考えた。

新たな拡張として、多地点配置の高速ネットワーク接続ステレオビジョンの利用を前提とし、人間を中心とする(人の存在、顔や意図したジェスチャの)リアルタイム認識技術や、実際に利用する場に適應できるように学習型の認識技術に関する研究を進めている。

ここで我々の提案するユビキタスステレオビジョン(USV)⁹⁾とは、表-1のような特徴を備えているものと定義した。実時間、実環境、実利用を志向し、あらゆる環境に適應可能、同時に実装を容易にすることを目指している。



■ 図-2 USVによる3次元復元図(踏切内部)

◆ USVのサーバ&クライアントモデル

そこでまず必要となる3次元構造の獲得については、ローカルな処理として各ステレオカメラに1台の計算機において3次元形状の復元計算を行い、取得するカラー画像と併せその計算機内において可能な限り画像特徴の抽出を行う。我々はこれをステレオカメラサーバと呼んでいる。

次にグローバルな処理として、アプリケーションに合わせて必要なデータ、処理特徴を別の計算機1台に統合して全視野内の情報を獲得する。図-2に4台のステレオカメラサーバからの全情報を1台の計算機上に統合した3次元復元図を示した。このように必要に応じ、ステレオカメラサーバからデータを収集、統合、利用する側をアプリケーションクライアントと呼ぶ。収集するデータは、図-2のように3次元情報、画像を収集する場合、また逆に個々のステレオカメラサーバで認識した人位置情報などの結果のみを収集する場合もある。また、両者の中間形態として、必要時のみ画像を転送するという利用方法も多く考えられる。

このとき、全体を統合するアプリケーションクライアントのための基本アルゴリズムは、カメラ台数に依存せず稼働するものでなければならない。後述する人専用踏切の空間監視の事例では4台のカメラを統合して利用している。この際に踏切空間内では完全に情報を同期統合させることを前提とした。1台のカメラが、そのレンズに対する光の差し込みなどで信頼性が低下しても、複数の情報を統合することによりロバストに動作継続することを意図したものである。

一方、駅ホーム端安全管理では実験レベルで5台のカメラを一直線に並べてその空間を監視している。このとき基本的に各カメラは独立して動き、近傍同士のカメラを除き、リアルタイムで統合される情報は画像情報や

i. 強い同期統合型 (高速ネットワークで結合)
距離情報、画像情報ともに統合され、信頼性を向上させる。 ex. 踏切空間監視、室内空間制御
ii. 弱い同期型 (低速ネットワークで結合)
距離情報、画像情報はローカルに処理され、必要情報のみ統合される。 ex. ホーム端安全管理、大空間での動線獲得
iii. 単独型
狭いエリアを1台で稼働するスポットインタフェース。 ex. 重度障害者支援

■ 表-2 USVの統合と独立

3次元情報から得られた非常に少ない記号情報であることを前提としている。

表-2に、複数カメラの統合と協調方法を基にした分類を示す。ここで適応される場面は、以下のいずれかに分類される。いずれの場合も計算機による自動制御を前提としている。自動判定が困難な状況においては、補助的にその画像を単純に管理者などに自動転送する機能を持つことで実用化が可能と考えている。

◆ステレオカメラ設置形態と距離情報の処理方法

当然のことながらステレオカメラの設置方法と、その距離情報の処理方法は、アプリケーションの目的に依存する。しかし、可能な限りアプリケーションに依存しない共通の距離情報処理方法も考え得るはずである。そこで、その共通の距離情報処理手法として、引き出し射影法(Draw Out Method)を提案している¹⁰⁾。これは、カメラから得られる距離情報をグローバル座標に変換後、空間中に含まれる距離情報を垂直方向に複数段階に分けて射影し、高さ別の空間情報を持った画像群(ベクトル情報)として利用する手法である。このときカメラの設置形態に応じ、階層的に距離情報を処理する手法と、加算的に処理する手法を使い分ける。表-3にカメラの設置形態とその対象、手法についてまとめた。斜め見下ろし交差型は個人の個別特徴を保持し、認識することに用いている。一方、垂直見下ろし型は大勢の人々に対して個別特徴を排除し、集団を分節している。これらの手法を組み合わせることで、より広い目的にも対応可能である。

◆現実の利用シーン

具体的な利用シーンとして、表-4のような屋内(ダイナミックレンジが小さく、照明変動もある程度一定)と屋外(ダイナミックレンジが大きく、照明変動も不安定)

A. 斜め見下ろし交差型	
対象：テキスト情報利用も含む少人数の精細、高信頼性な認識 手法：段違い引き出し法（情報を重複し加算することで形状を保存） 視野：基本的にカメラ間の視野が重複することを想定	
B. 垂直見下ろし型	
対象：不特定多数内の個々のトラッキング、人数計測 手法：加算引き出し法（全体形状を安定的に抽出） 視野：カメラ間の視野が一部重複、または独立	

■表-3 ステレオカメラの設置形態と処理手法

a. 室内での人の認識とインタラクション（個人識別も可能）	
・人行動ログの取得	・インタフェースとしての利用
・特定作業への支援	・アミューズメントへの応用
b. 屋外での人の認識（トラッキングが中心）	
・安全監視	・大空間での動線獲得

■表-4 屋内外でのUSV利用シーン

での人の認識を想定している。ただし、示した事例以外にも、適応シーンを替えて同様な枠組みでより広いシーンに適応可能と考えている。

以下、次章において広域な空間での不特定多数の認識を対象とする安全管理での利用シーンを中心に事例を説明する。

オープンな屋外空間での人の(安全)監視

安全・安心を目的にすると、屋外など設置する場所を一切選ばないこと、より広い制約のない空間に適応させることが必要になってくる。我々はこのような応用例として、踏切内監視、駅ホーム端監視などについて継続的に実験を行っている。これらは、真の実環境において、どの程度の解像度画像によって何が可能なか。そのときに必要となるアルゴリズムは何なのか。実環境に耐えられるハードウェアスペックはどの程度なのか。これらを実地に掴むために実験を行っている。表-5に、今までに行ってきた主要な実験例を一覧にした。以下個々の

	日時	場所	目的	カメラ台数
踏切	2003/1-2	東急東横線 元住吉駅踏切	踏切空間の人の監視	6台 斜め見下ろし交差型
	2004/6	東急東横線 自由が丘駅踏切	踏切空間の人の監視	4台 斜め見下ろし交差型
駅ホーム	2001/11	JR 牛久駅 3番線ホーム	ホーム端転落監視実験 転落状況のデータ取得	4台 垂直見下ろし型
	2001/11	東急東横線 横浜駅 1番線ホーム	ホーム端監視実験 乗客のデータ取得	5台 垂直見下ろし型
	2003/1-2	東急東横線 元住吉駅 3番線ホーム	ホーム上の人の監視 車いすの認識 白杖の認識	1台 斜め見下ろし型

■表-5 継続的な広域空間監視実験状況



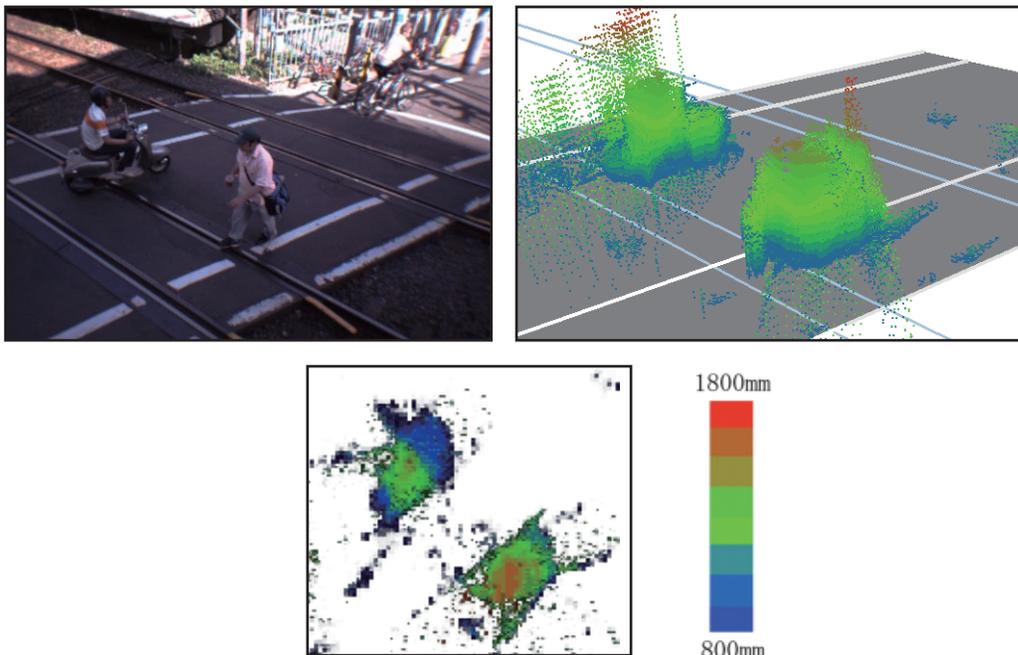
■図-3 USVによる踏切空間監視

目的別に実環境実験について紹介する。

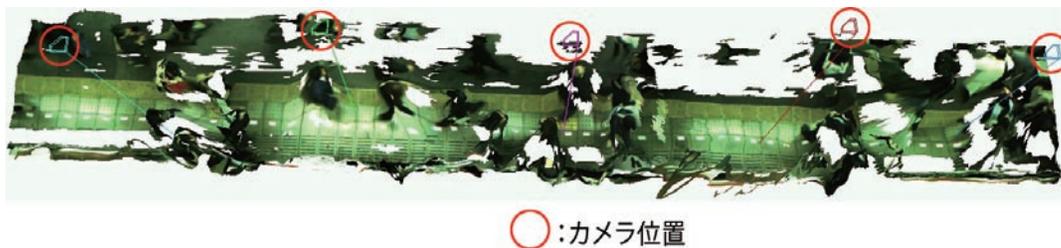
◆踏切空間監視

現在の踏切内安全監視装置は、赤外線センサの送受信の組を複数設置する構成が主にとられている。しかしながら、赤外線センサによって構成されるライン上に障害物があった場合のみ反応するので、自動車は確実に検知できても、人や自転車などがそのラインの間にいた場合は、無力なセンサとなっている。そこで、ステレオカメラからの距離情報を利用して、踏切内を認識する研究も始まっている⁷⁾。しかし、ここでは1台のステレオカメラにより精度を安定させることを行っており、複数台による精度向上、安定性確保は行われていない。

我々が提案するユビキタスステレオビジョンによる監視方法では、踏切の四隅に踏切空間を取り囲むようにステレオカメラを設置し(斜め見下ろし交差型)、4台のステレオカメラで完全に空間を監視する。図-3の画像群が、4台のステレオカメラから得られたカラー画像であり、図-2が統合された3D情報を示している。この



■ 図-4 加算引き出し法による距離情報利用



■ 図-5 USVによるホーム端3次元復元図

ような構成で、単体ではその必要解像度、太陽光の影響、設置俯角の影響、また、複数台化の効果などあらゆることを評価した。

通行人を含む踏切内のシーンにおける距離情報に対し、加算引出法を適用している事例を図-4に示した。上段左図がカラーテクスチャ画像、上段右図が3次元空間において多段階射影画像を1台のカメラと同じ角度から見た鳥瞰図、下段の図がそれを2次元に射影したものである。この図より歩行者とスクーター搭乗者の形状が、射影された地平面で明確に出ていることが分かる。ここでは、加算引き出し法を利用して不安定な3次元情報を射影することで、人などの存在を安定的に検知、追跡している。

◆ 駅ホーム端安全管理

駅のホーム端からの転落事故による死亡者だけでも、年間数十件に達し¹³⁾、特に都市部での鉄道、地下鉄においては緊急に解決すべき課題となっている。そのためにはホームドアの設置が最も優れた解決法である。しか

し、既存電車がホームドアに対応する車両にすべて置き換えられるだけでなく、さらに既存のホームもドアの荷重に耐えられるように改良されなければならない。現実的にこれらインフラの整備を考えると、ホームドアの完全配備には数十年を要することが予想されている。そこで、既存のセンシング技術としては、落下検知マットが利用されている。しかし、これは接触センサが線路上の枕木横からホーム端の下部までに配置されるのみであり、構造上すべての場所を覆えないという欠点を持つ。また、超音波センサ、赤外線センサといった非接触、かつアクティブなセンサの利用も検討されているが、列車が頻繁に出入りする環境であることや鳥やゴミなどの遮蔽への対応が困難なため実用に至っていない。そこで、ステレオカメラを駅ホーム端の天井から下を見るように直線に並べ(垂直見下ろし型)、ホーム端の人々の状況を監視する¹¹⁾。実際に東急東横線横浜駅の旧地上ホームで5台のカメラから取得した3次元画像を図-5に示した。これは5台のステレオカメラから得られた3次元情報からポリゴンを作りテクスチャマッピングをした図である。

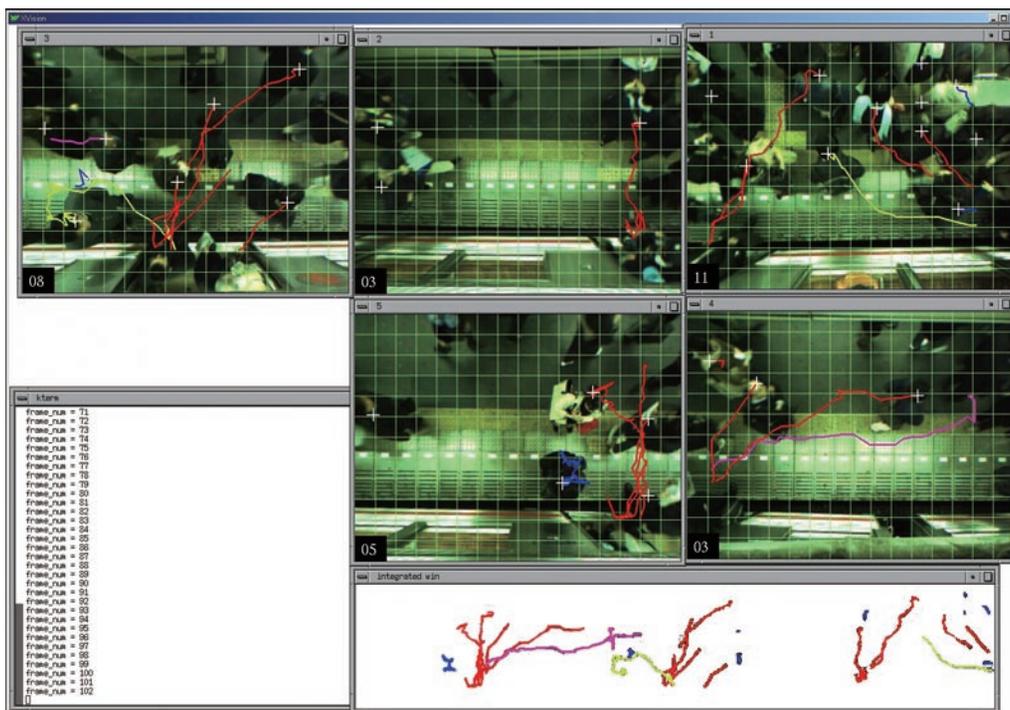


図-6 USVによる全体の動線管理

下側がホーム端(電車)側で、中央の黄色部分が点字ブロックである。

このような複雑な状況下での人を観測する基礎技術は、駅だけでなくあらゆるオープンな空間でも使える高い汎用性を持っている。そこで、現実の複雑な状態の人々を対象に実験を行った。JR牛久駅ではホームからの転落実験を、東急東横線横浜駅ではホーム端の軌跡獲得を、東急東横線元住吉駅ではホーム上の人に着目した実験を行い、それぞれデータを取得した。

JR牛久駅ではホーム上と線路面上両方で、人の落下、座り込み、寝込み姿の時系列3次元情報から、全時間帯における(夕日の差し込み時、夜間を含む)人の存在を確実に検知できることを確認した。

東急東横線横浜駅においては、始発から終電までの全時間帯において、実際の乗客の出入りを撮影した。図-6には、全乗客の動線管理を示した。各カメラの画像は、カメラ単位に接続されたコンピュータで個々に処理され、その3次元情報から乗客の位置とその移動方向を計算している(各カメラから得られたカラー画像上に乗客の動線を色付きで図示。各色は乗客の動線を示す)。そこから得られたわずかな結果情報より全体の動線を作成する弱い同期統合型で実行されている(図-6の右下の枠が5

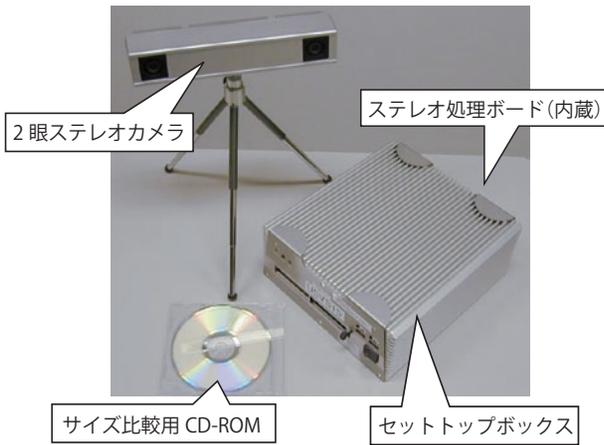
台のカメラから得られた全視野が統合された空間を示している。動線情報のみが転送、統合され、全乗客の動線を表示している)。

この処理戦略により、今後数十台単位でのカメラの増加にも対応可能となっている。もし危険状況の自動判定が困難な場合は、単純に必要なカメラ画像を管理者などに転送するといった利用方法を前提として研究を行っている。

まとめと今後の展開

ユビキタスステレオビジョンの基本構成、ならびに安全・安心のための鉄道関連のアプリケーション事例について述べた。本研究では、ステレオカメラを静的に適宜配置することを前提としている。その前提の上で、各ステレオカメラから得られるシーン内の必要な情報を掌握するために、3次元情報と画像情報を同時に利用する有用な手法の研究開発が目的である。

そして、屋外での安全性向上支援の事例として、踏切監視、ホーム端監視の事例を示した。それぞれの目的に合わせて複数のステレオカメラが強い同期で動作する場合、弱い同期で動作する場合、あるいは単に単数で動



■図-7 ユビキタスステレオビジョンデバイス(USVD)

作する場合などが考えられる。

また、現在までのすべての実験結果を反映させたハードウェアであるユビキタスステレオビジョンデバイス(USVD)が完成した(図-7)。これはステレオ距離計算部分をすべてハードウェアで実装したPCIハーフボード、それを搭載するセットトップボックス、専用ステレオカメラから構成される。これにより、対象に依存しない距離計算部分がハード化され、セットトップボックスのCPUをすべて認識処理に使うことを可能にする。同時に、セットトップボックスは完全ファンレス化を実現した。ゆえに、文字通りにどんなところにも設置可能なハードウェアである。

そして、2007年1月から通行者の安全確保を目的とした常設踏切監視実験が開始される。ここでは、このUSVDを踏切内に実装し、現状の赤外線監視装置の検知情報、ならびに遮断機開閉情報とリアルタイムで連動

しながら年間を通して稼働し、ハードとソフト両方の性能評価と実証実験を行い、実環境でのさらなる性能向上が図られる予定である。

謝辞 本研究の一部は都市エリア産学官連携促進事業(筑波研究学園都市エリア)によって行われた。

参考文献

- 1) Zhao, T. and Nevatia, R. : Tracking Multiple Humans in Complex Situation, IEEE Transaction on PAMI, Vol.26, No.9, pp.1208-1221 (2004).
- 2) Mittal, A. and Davis, L. S. : M2Traker: A Multi-view Approach to Segment and Tracking People in a Cluttered Scene, Int. Journal of Computer Vision, Vol.51, No.3, pp.189-203 (2003).
- 3) 辻, 長岡: 夜間の歩行者認知支援システムの開発, 第11回画像センシングシンポジウム講演論文集, pp.387-390 (2005).
- 4) Poseidon System Form Vision-IQ Works : <http://www.poseidon-tech.com/us/index.html>
- 5) Terada, K., Yoshida, D., Oe, S. and Yamaguchi, J. : A Method of Counting the Passing People by Using the Stereo Images, Proceedings of ICIP, Vol.2, pp.338-342 (1999).
- 6) Zhao, L. : Stereo- and Neural Network-based Pedestrian Detection, IEEE Transaction on ITS, Vol.1, No.3, pp.148-154 (2000).
- 7) 柴山憲文, 堀 圭二, 太田 勝: 踏切障害物検知用のステレオ画像装置の開発, 第9回画像センシングシンポジウム講演論文集, pp.1-6 (2003).
- 8) Yoda, I. and Sakaue, K. : Utilization of Stereo Disparity and Optical Flow Information for the Computer Analysis of Human Interaction Machine Vision and Applications, Vol.13, No.4, pp.185-193 (2003).
- 9) Yoda, I. and Sakaue, K. : Concept of Ubiquitous Stereo Vision and Applications for Human Sensing, Proceedings IEEE Inter. Symposium on Computational Intelligence in Robotics and Automation, pp.1251-1257 (2003).
- 10) Yoda, I., Yamamoto, Y., Hosotani, D. and Sakaue, K. : Human Body Sensing Using Multi-Point Stereo Cameras, Proceedings of International Conference on Pattern Recognition 2004 (ICPR 2004), Vol.4, pp.1010-1015 (2004).
- 11) 依田育士, 細谷大輔, 坂上勝彦: ユビキタスステレオビジョンによる駅ホーム端安全管理, 電学論 C, Vol.124, No.3, pp.805-811 (2004).
- 12) 産総研実世界知能研究班: 産業技術総合研究所における RWC 研究発表資料集, 2000 (2001), 2001 ~ 2002 (2002).
- 13) 運輸省: 事業者別運転事故件数及び死傷者数, 平成 6-10 年度鉄道統計年報 (1994-2000).

(平成 18 年 12 月 1 日受付)

依田育士(正会員)

yoda@ieee.org

1992年東京都立科学技術大学大学院工学研究科修士課程修了。同年電子技術総合研究所入所。コンピュータビジョンの研究に従事。ユビキタスビジョン研究班長。筑波大学連携大学院助教授(併任)。工学博士。