

## ユビキタスステレオビジョンによる 安全・安心のための実時間実環境ヒューマンセンシング

依田 育士 坂上 勝彦  
産業技術総合研究所 情報技術研究部門

あらまし：ステレオカメラを多地点に配置し、高速ネットワーク利用を前提として、人間を中心とする（人の存在、姿勢、その顔、意図したジェスチャの）リアルタイム認識技術に関する研究を行っている。ここで提唱しているユビキタスステレオビジョンとは、複数のステレオカメラとそれらを利用するアプリケーションという枠組みで構成され、実時間実環境でのヒューマンセンシングを提供するものである。研究のフォーカスは、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どこに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である。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着目する手法、対象の記述・認識手法を開発している。本論文ではこのユビキタスステレオビジョンと、安全・安心のための利用例として、踏切監視、駅ホーム端監視、広域空間監視などの実時間実環境ヒューマンセンシング事例について述べる。

### Real-time Human Sensing for Safety and Security in Real Environment by Ubiquitous Stereo Vision

Ikushi YODA and Katsuhiko SAKAUE  
Institute of Information Research Technology  
National Institute of Advanced Industrial Science and Technology (AIST)

**Abstract:** We carried out experiments on real-time recognition technology mainly for humans (of human existence, the face and intentional gestures) in the real-world environment. We placed stereo cameras ubiquitously and used a large-capacity network. The “Ubiquitous stereo vision” consists from multi stereo cameras and its applications, and provides real-time human sensing in real environment. We focused on the realizing human sensing in real environment including strong changes of weather and shadows. The objectives of this research were to use 3D and color information, and to develop a real-time new human sensing method. This paper describes the Ubiquitous stereo vision and the safety applications (railroad crossing, platform of railroad station, and huge event space).

#### 1 はじめに

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、その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応できるような学習型の認識技術に関する研究を2000年頃から行っている。当時のRWCプロジェクト後半[5]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[1]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そうと考えた。当時は「ユビキタス」というキーワードが、一昔前の「マルチメディア」というように使われることを予想しておらず、ステレオビジョンを用いることで「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命名した。ここでは、我々の提案するユビキタスステレオビジョンとその安全・安心のためのアプリケーション事例について述べる。

本研究の究極の目的は実時間実環境におけるシーン理解であるが、これに関連する画像認識の研究は既に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、ど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り、共通のフレームワークすら未だに見えてこないのが現状である[6-7]と考えている。そこで、認識対象を正確に着目するという現実的には最も困難な問題に対し、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することで、その後の認識問題に注力するためのフレームワークを作る。

現実の交通監視事例など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し、特に混雑状況の監視や、ナンバーの記録などは多数存在する。一方でリアルタイムでの安全のための人を中心とした監視となると現実の適応事例は極めて限定される。具体的な適応例は、自動車搭載型の歩行者認知への適応事例[8]などである。具体的な適応事例が極端に限定されるのは、特に屋外での激しい天候変化や影への対応が困難なため、シビアナ安全・安心への応用が限られるためである。

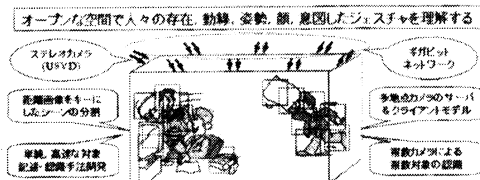


図1 ユビキタスステレオビジョンの概念図

そこで我々は、ステレオカメラを固定的な設置で利用すること、認識対象は人々を中心とすること、距離情報算出には特定の対象知識などを用いずに一般的に距離情報を算出することを前提に研究を行っている[2]。この前提でまずは屋内外に通じるビジョン技術を確立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フォーカスしている。そのためには独立した個々のステレオカメラにおいて、その距離情報からローカルに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。

## 2 ユビキタスステレオビジョン

### 2.1 ユビキタスステレオビジョンの基本構成

ここで我々の提案するユビキタスステレオビジョン (USV) [2]とは、表1のような特徴を備えているものとした。実時間、実環境、実利用を志向し、あらゆる環境に適応可能、同時に実装を容易にすることを目指している。

表1 USVの定義

・入力情報は複数のステレオカメラから得られる3次元距離とカラー画像
・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

ここでまず必要となる3次元構造の獲得については、各ステレオカメラに1台の計算機においてローカルに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ローカルな計算機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サーバと呼んでいる。



図2 USVによる3次元復元図 (踏切内部)

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1台の計算機上に統合されたユビキタス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。統合するデータは、図2のように全てのデータを集積する場合、また逆に個々のステレオカメラサーバで認識した結果のみを利用する場合もある。また、両者の中間形態として、必要時にのみ画像を転送するといった利用方法も多く考えられる。

## 2.2 USVのサーバ&クライアントモデル

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に稼働するものでなければならない。あ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した。複数の情報を統合することで、1台のカメラが角度、光の差し込みなどで信頼性が低下しても、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、多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。しかしながら、自動判定が難しい場合は、その画像を単純に管理者などに自動転送する機能を併せ持つことを想定している。

表2 USVの統合と独立

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

## 2.3 ステレオカメラ設置形態と距離情報の処理方法

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、当然のことながら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[3]。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、加算的に処理する方法を使い分ける。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。ただ、目的に応じてこれらの手法は混在もする。

表3 ステレオカメラの設置形態と処理手法

A. 斜め見下ろし交差型
対象: テクスチャ情報利用も含む少人数の精細、高信頼性な認識 手法: 段階引き出し法 (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

## 2.4 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 (照明がある程度安定) と屋外 (ダイナミックレンジ大) 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、3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。

表4 屋内外でのUSV利用シーン

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 ・ 特定作業への支援	・ インタフェースとしての利用 ・ 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

## 3 屋外大規模空間での人の状態 (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、また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。このような目的の応用事例として、駅ホーム端監視、踏切内監視、イベント空間監視事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のか、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。これらを実地に掘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した。以下個々の目的別に実環境実験について紹介する。

表5 継続的な広域空間監視実験状況

	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	ホーム端端落監視実験 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
広域	2005.3-9	愛知万博 グローブハウス オレンジホール	会場内部の人数計測 行動軌跡の解析 RF-ID との連携実験	7台 垂直見下ろし型

### 3.1 踏切空間視

現在の踏切内安全監視装置は、赤外線センサの組を複数設置する構成が主に取られている。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、全く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[9]。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、安定性確保は全く行われていない。



図3 USVIによる踏切空間監視

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、設置角度の影響、複数台の効果などあらゆることを評価している。

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2次元に射影したものである。歩行者とスクーター搭乗者の頂上部が、各プレーン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。

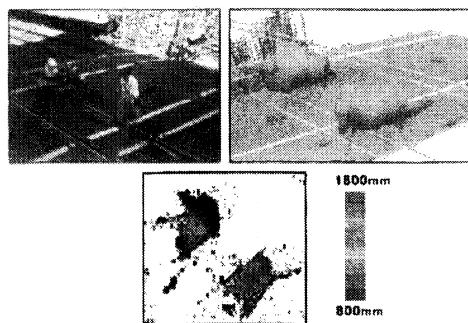


図4 加算引き出し法による距離情報利用

### 3.2 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[10]、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。ホームドアの設置が最も優れた解決法であるが、そのためには既存電車がホームドアに対応する車両に全て置き換えられる必要性があり、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、これは接触センサを線路上の枕木横からホーム端の下部までに配置され、構造上全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り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[4]。実際に東急東横線横浜駅で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。そのために、現実の複雑な状態の人々を対象に実験を行い、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。



図5 USVIによるホーム端3次元復元図

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。

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、ローカルに処理され、その3次元情報から乗客の位置とその方向を計算し、そこから得られた僅かな結果情報から全体の動線を作成する弱い同期統合型の実行されている。

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。

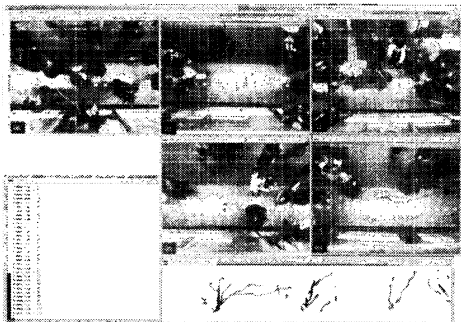


図6 USVIによる全体の動線管理

### 3.3 広域イベント空間監視

上述の監視技術は、ホーム端や踏切といった安全監視だけの目的以外にも、より大きな空間での人流の計測、監視に適応可能である。そこで2005年に半年間にわたり開催された愛知万博(愛・地球博)のグローバルハウス・オレンジホールにおいて常設実験を行った。具体的には会期中における観客滞在時間のカメラ視野内の全軌跡をRF-IDから得られる情報と同期を取りながら追跡した。1台のカメラあたりのフレーム数は総計73Mを超えた。実験の目的は以下の3つである。

1. 会場の領域単位でのリアルタイム人数計測
2. 視野内の人の行動分析
3. RF-IDタグ(Aimulet)との連携実験

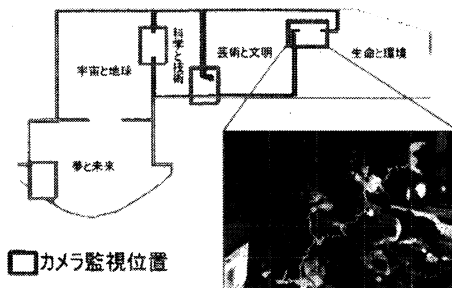


図8 愛知万博グローバルハウス・オレンジホール

あらゆるセンシングには向き、不向きがあり万能センサというのは現実問題困難である。そこで、他のセンサとビジュアルセンサをどのように組み合わせるのかなどへの研究の発展、展開は無限の可能性があると考えている。そこで、万博での実験結果を受け、同様のフレームワークを用いた実証実験が、秋葉原にあるダイビル10F産総研フロアにおいて、発展的に研究が開始されたところである。

### 4 まとめと今後の研究

ユビキタスステレオビジョンの基本構成、ならびに現実のアプリケーション事例について述べた。本研究の目的は対象に依り、ステレオカメラを静的に適宜配置することにより、そのシーン内の必要な情報を掌握するために、3次元情報とカラー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。

このときの具体的なアプリケーションとしては、大規模空間を対象とする主に屋外での安全性向上支

援の事例を示した。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、あるいは単純にシングルで動作する場合などが考えられる。

さらに最重要事項として、現在までの全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) が完成している。これはステレオ距離計算部分を全てハードウェアで実装した PCI ハーフボードと、専用ステレオカメラから構成される。これにより、ボードを挿すマザーボードの CPU を全て認識処理に使うことを可能にしている。これにより最終形態はランチボックスサイズのセットトップボックス化を可能にし、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。今後は、この USVD をベースとして実装の展開を図る。同時にその中で必要となる人を中心とした対象の自動的な学習方法、対象の記述方法、現実シーン内での認識方法の開発を進める。

## 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] I. Yoda and K. Sakaue: "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).
- [2] I. Yoda and K. Sakaue: "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).
- [3] I. Yoda, Y. Yamamoto, D. Hosotani, and K. Sakaue "Human Body Sensing Using Multi-Point Stereo Cameras," *Proceedings of International Conference on Pattern Recognition 2004 (ICPR 2004)*, Vol.4, pp.1010-1015 (2004).
- [4] 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 2004*, pp.805-811 (2004).
- [5] 産総研表世界知能研究班: "産業技術総合研究所における RWC 研究研究発表資料集" 2000 (2001), 2001~2002 (2002).
- [6] A. Pentland: "Looking at People: Sensing for Ubiquitous and Wearable Computing" *IEEE Trans. PAMI Vol.22, No.1*, pp.107-119 (2000).
- [7] 松山隆司: "分散協調視覚による動的 3 次元理解" 日本学術振興会未来開拓学術研究推進事業研究成果報告書 (2001).
- [8] 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).
- [9] 丸山, 高荷: "Space-Sweep 法を用いた屋外障害物検出法" 計測自動制御学会 第 59 回パターン計測部会研究資料 pp.15-22 (2002)
- [10] 運輸省: "事業者別運輸事故件数及び死傷者数" 平成 6-10 年度鉄道統計年報 (1994-2000)