



マルチセンサとオムニステレオビジョンを用いた人体検出と3次元環境地図の作成

井上, 淳一
滝口, 哲也
有木, 康雄

(Citation)

神戸大学都市安全研究センター研究報告, 12:91-102

(Issue Date)

2008-03

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81001122>

(URL)

<https://hdl.handle.net/20.500.14094/81001122>



マルチセンサとオムニステレオビジョンを用いた 人体検出と3次元環境地図の作成

Human Detection and 3D Environmental Map Estimation

Using Multisensor and Omnistereovision

井上 淳一¹⁾

Junichi Inoue

滝口 哲也²⁾

Tetsuya Takiguchi

有木 康雄³⁾

Yasuo Ariki

概要：地震等による被災者を探索する目的で、駆動機構を持たない小さなセンサを瓦礫の隙間から多数ばら撒く研究が始められている。これは従来の探索ロボットなどと比較し、コスト・探索範囲・探索時間で優位である。しかし、従来の探査ボールは被災者の発見のみを目的とするものであり、被災者がどこに存在するかを救助者が目で確認する必要があった。我々は目的の異なるセンサボールを2種類使い、被災者の検出と瓦礫内の3次元環境地図作成という目的をそれぞれ自動で行うことにより、より安全で迅速な救助を目指す。本稿では、センサボールの小型化・低価格化がなされたとして、低解像度の熱画像データから人体を検出する手法と、狭基線長オムニステレオカメラにおける距離計測法を提案する。

キーワード：センサネットワーク, ステレオビジョン, 位相限定相関法

1. はじめに

阪神・淡路大震災において防災対策の不十分性が露呈し、以後、防災やレスキューに関する研究の必要性が訴えられるようになってきている。文部科学省は先端技術の災害軽減への積極的利活用を進め、平成14年度から「大都市大震災軽減化特別プロジェクト」をスタートさせた。これは、人体検索・情報収集・配信等を支援するロボット・センサ等の開発と実用化を目指すものである¹⁾。

一般に、地震などの災害において、倒壊した家屋の増加に伴い、死亡者数は増加する²⁾。阪神・淡路大震災において、瓦礫内に閉じ込められた被災者が後に起こる火災によって、命を落とす事例が多く見られた。従って、被災者の探索を迅速に行うことによって、被災者の生存率は上昇すると予想される。しかし、救助者が直接、倒壊家屋上部や内部の探索を行うと、2次災害の恐れがあり、危険である。救助者に代わり瓦礫内の被災者を発見、瓦礫内環境を把握できる機器の開発は必要性が極めて高いと言える。

このような被災者の探索や瓦礫内環境の把握の目的で、様々な機器が研究・開発されてきた。電磁波を用いた装置は、呼吸に伴う肺の動きや心臓の動きを捉えることで被災者を発見する³⁾。実際に中越沖地震などで使用されたものもある。瓦礫内の探索を行うロボットも多数開発されている⁴⁾。しかし、電磁波を用いる方法、探索ロボットを用いる方法、どちらも共通して1台にかかるコストは大きく、使用には専門性が伴う。探索ロボットは自分で移動しながら被災者を探索するため、救出までの時間が長期化してしまう。また、移

動機構を持つ必要があるため、機器が大きく瓦礫内の狭い隙間までは探索できない。

最近、駆動機構を持たない小さなセンサを瓦礫内に多数ばら撒き、センサネットワークを用いて被災者を発見する、センサばら撒き型のアプローチに関する研究が始められている⁵⁾。これは従来のアプローチと比較し、1台におけるコストが安く、より広い範囲にわたり、狭い空間を短時間で探索することができる。しかし、従来の探査ボールは、被災者の発見のみを目的とするものであり、瓦礫内の環境がどのようになっているかまでは把握できない。また、被災者がどこに存在するかを救助者が直接目で確認する必要がある。これらの問題を解決すれば、より安全で迅速に救助することが可能になる。

本稿では、目的の異なるセンサボールを2種類使い、被災者の発見と瓦礫内の3次元環境地図作成という目的をそれぞれ自動で行うことによって、より安全で迅速な救助方法を目指す。本稿では、センサボールの小型化・低価格化がなされたとして、熱画像データから人体を検出する手法、狭基線長オムニステレオカメラによる3次元計測法を提案する。

本稿は全5章により構成される。第2章ではまず、提案する2種類のセンサボールの概要、センサボールを用いた救助の概略を述べる。第3章では1つ目のセンサボールである人体検出ボールにおける人体検出の方法を述べる。また、比較的安価な2次元放射温度計を用いて、人体を検出する実験を行い、その有効性について述べる。第4章では2つ目の地図作成ボールにおける3次元計測方法を述べる。センサの小型化に伴う狭基線長のオムニステレオカメラにおいて、POC(Phase Only Correlation)を用いた高精度な対応付けのアルゴリズムを述べる。そして、POCと従来手法において精度と速度の比較を行う。最後に第5章で本稿の成果をまとめると共に、今後の課題について述べる。

2. 提案手法の概要

従来のばら撒き型のアプローチである探査ボールを用いた方法は、被災者の検出のみを目的とするものであり、また、被災者がどこに存在するか目で確認する必要がある。本稿では、目的の異なるセンサボールを2種類使い、被災者の検出と瓦礫内の3次元環境地図作成という目的をそれぞれ自動で行うことで、より安全で迅速な救助方法を目指す。次節から、2種類のセンサボールを用いた救助方法、各センサボールの概要を述べる。

(1) 2種類のセンサボールを用いた救助

センサボールは倒壊家屋の瓦礫外のわずかな隙間から投げ入れることを想定している(図1)。



図-1 倒壊家屋における瓦礫

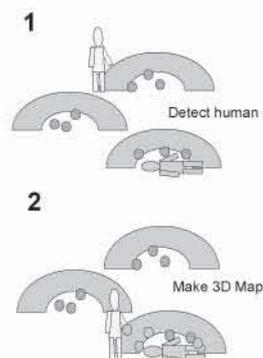


図-2 2種類のセンサボールを用いた救助法

センサボールが瓦礫内の温度、音声、画像といった情報を瓦礫外に送信し、瓦礫外のサーバが生存者の有無の判断及び、3次元距離計測による瓦礫内の地図の作成を行い、救助者に提示を行うことを想定する。よって、センサボール自体には、移動機構、計算機構を積む必要は無く、小型化をはかることが可能である。瓦礫内の全てのセンサボールが直接瓦礫外のサーバと通信を行えるとは限らないので、各センサボールが自律的にネットワークを組み、外部サーバとの通信を行うことを想定する。仮に、1台が故障し、うまく動作しなくとも、他のセンサを介してネットワークを組むことが可能である。

本稿が提案するこの2種類のセンサボールを用いた救助の流れ図2に示す。

1. まず1つ目のセンサボールを瓦礫の隙間から投げ入れ、瓦礫内の人の生存の有無を確認する。
2. 人の生存が確認された瓦礫内のみ2つ目のセンサボールを投げ入れ、瓦礫内の3次元環境地図の作成を行なう。

これらの2つのステップを踏む利点として、人の生存有無の確認を人が目で見て確認しなくても自動で判断できるため、作業効率が上がること、瓦礫内の3次元環境地図を作成することで、より安全な救助活動が可能になることが挙げられる。また、1つ目のセンサボールは瓦礫内の人体の検出を行うので、大量に散布される必要がある。2つ目のセンサボールは1つ目のセンサボールで生存者有りと判定された瓦礫内のみ散布し、瓦礫内の3次元環境地図を作成する。よって少量の散布でよい。

各センサボールの目的と特性を踏まえた上で、次節から各センサボールに積むセンサの選定を行う。

(2) 各センサボールの概要

1つ目のセンサボールを人体検出ボール、2つ目のセンサボールを地図作成ボールと呼ぶことにする。2種類のセンサボールの概要をそれぞれ述べる（表1）。

表-1 センサボールの概要

センサボール名	人体検出ボール	地図作成ボール
目的	生存者の有無の確認	3次元環境地図の作成
搭載センサ	熱画像センサ、マイクロフォンアレイ	オムニステレオカメラ
散布量	大量散布	少量散布
コスト	安価	高価

人体検出ボールは熱画像センサとマイクロフォンアレイを持つ。CO2 センサは生存者の検出には有効であると考えられるが、救助者にかなり近づく必要があるため使用できない。被災者が発する音や、熱といった情報はセンサと被災者までの距離がある程度遠くても検出可能である。よって、熱画像センサとマイクロフォンアレイを採用した。熱と音声の両方を検出することで、音声によって被災者の意識の有無を確認ことができ、また被災者の意識が無い場合でも、熱によって検出を誤ることはない。

地図作成ボールは、オムニステレオカメラを持つ。3次元計測を行なう手法は、電磁波や超音波などを用いたものなどがあるが、測定対象の人体に影響を与えることなく3次元情報を獲得できる、ステレオビジョンを選択した。また、全方位の3次元計測を行う方法として、カメラを回転する方法が考えられるが、そのためには回転機構が必要になり、小型化の妨げになる。よって、全方位ミラーを用いて全周の情報を取得することにした。2つのオムニカメラを垂直に配置することでオムニステレオカメラを構成する。また、複数のセンサの情報を統合することで、1つのセンサで見えなかった瓦礫の隙間も補間し合うことができる。

本稿では、センサボールの小型化、低価格化が達成されたとして、熱画像データから人体を検出する手法、センサの小型化に伴う狭基線長に対応するために、オムニステレオカメラを用いた3次元計測法を提案する。

3. 人体検出ボール

この章では、前章で述べた人体検出ボールの熱画像センサにより、人体を検出する方法について述べる。

人体検出ボールはその用途から大量に散布する必要があるため、できるだけ安価なセンサを積む必要がある。現在売っている熱画像センサは非常に高価である。今回はその中でも解像度が荒いながらも、安価な価格を実現している2次元放射温度計を用いて、人体検出を試みる。

(1) 2次元放射温度計

2次元放射温度計の詳細な仕様を表2に、得られる画像を図3に示す。

枠内の平均の温度が枠の色として表示されている。高温になるにつれて赤く、低温になるにつれて青く表示される。8×8の計64点の温度データがCSV形式で得られ、この64点のデータから人体の検出を行う。

(2) 人体検出手法

人体の検出方法は、学習データの人体部分と背景部分それぞれの平均温度と分散から、人体温度と背景温度の分布を正規分布として算出し、人体温度の最適なしきい値を決定する(図 4)。しきい値以上であれば人体、それ以外は背景と判定する。

表-2 熱画像センサの仕様

名前	2次元放射温度計「アイスクエア」
温度測定部	64点サーモパイル
温度範囲	-50℃～1000℃
分解能	0.1℃
測定エリア	48cm×48cm(距離 1m)
撮影範囲	0.5m～∞
可視撮影部	CMOS カラー
画素数	30万画素
ズーム機能	なし
保存形式	CFカード

(3) 人体検出実験

実験環境は研究室内の実験室を使用した。詳細な実験条件を表 3 に示す。

表-3 人体検出実験条件

気温	27℃
撮影距離	1m, 2m
被験者	1名
人体撮影条件	頭部のみ, 腕のみ, 脚のみ, 全身と分けて撮影
人体撮影枚数	50枚
背景撮影条件	実験室内をランダムに撮影
背景撮影枚数	50枚

人体は被験者1名で、肌の露出部分を頭部、腕、脚、全身、撮影距離を1m、2mと条件を変えて合計50枚撮影した。背景は研究室内実験室をランダムに50枚撮影した。正当率は人体を含む画像を人体であると判定できたかどうか、適合率は人体であるという判定結果が得られた画像に対し、人体が含まれているかどうかの確率である。人体であるという判定は、1点でも人体と一致して判定されていればよいものとした(図 5)。

実験の結果を表 4 に示す。頭部のみが露出している画像を用いて人体の温度分布を学習し、他の部分が露出している人体画像の検出実験を行なった場合、表中に頭部のみと記してある。どのような条件においても正当率は100%であったが、適合率は平均で80.14%であった。これは研究室内実験室にあるコンピュータやコンセントなどの熱源を、人体であると誤って検出してしまった結果である(図 6)。

その中でも、平均体温が高い頭部のみで学習したものは、92%と高い適合率を得ることが出来た。よって、より人体の体温の高い部分のみで人体温度分布を学習すれば、人体と背景の判定はより正確にできる。このことより、64点という非常に粗い熱画像のデータでも、人体の検出が十分可能であることが分かる。

(4) 音声の併用と熱画像センサの全方位化

今後の予定として、音声を用いた人体検出法を併用することで、さらなる人体検出の精度を向上させる予定である。具体的には、マイクロフォンアレイを用いて、瓦礫外の救助作業音のノイズをキャンセルし、被災者の音声を検出するなどの処理を行う。また、熱画像センサ自体の低価格化、高解像度化を期待し、全方位での計測に拡張していく予定である。

表-4 人体検出実験結果

学習条件	適合率(%)	正当率(%)	平均温度(°C)
頭部のみ	92	100	32.41
腕のみ	76	100	30.47
脚のみ	79	100	30.84
全身	82	100	31.03
1mのみ	81	100	31.24
2mのみ	74	100	30.73
全て	77	100	31.05
背景平均温度			27.54



図-3 熱画像例

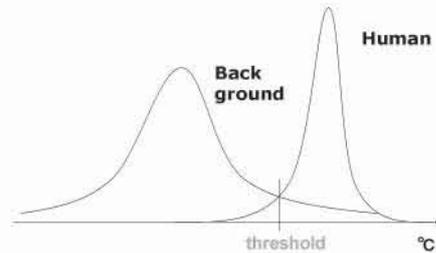


図-4 人体検出方法

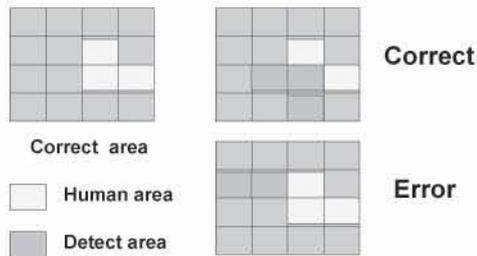


図-5 正解判定法



図-6 誤検出例

4. 地図作成ボール

地図作成ボールは、オムニステレオカメラを用いて、全方位の3次元計測を行う。また、この地図作成ボールを複数台用いることによって、瓦礫内の障害物で隠れていた箇所も1つの地図として取得できる。

(1) オムニステレオビジョン

今回用いた全方位ミラーである HyperOmniVision はその光学的特性から、他の全方位ミラーである球面ミラーや円錐ミラーなどと異なり、ミラーの焦点から見た画像(一般のカメラで撮像した画像)や、カメラを鉛直軸周りに回転して得られるパノラマ画像(円筒状の全方位画像)に、簡単に変換可能である。よって、多彩な画像処理を施すことができる。HyperOmniVision の詳細な仕様を表5に、得られる全方位画像を図7に示す。

まず、HyperOmniVision の視覚系について述べる。

図8に示すようにZ軸を鉛直軸とする3次元座標系0-XYZを考える。この時、双曲面は次式で表わせる。

$$\frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (1)$$

表-5 全方位ミラーの仕様

名前	HyperOmniVision
ミラー形状	双曲面ミラー
高さ	118.75mm
直径	90.2mm
仰角	30°
俯角	50°

a, b は双曲線の形状を定義するパラメータで, a は漸近線からミラー頂点までの直線距離を示している。また, 双曲面ミラーは $(0, 0, +c)$ と $(0, 0, -c)$, ($c = a^2 + b^2$) の 2 つの焦点を持つ。ミラーの焦点 O_M 及びカメラのレンズ O_C はこの 2 焦点に位置し, 画像面はカメラのレンズ中心 O_C からカメラの焦点距離 f だけ離れた平面になる。

この HyperOmniVision をつけた視覚センサを垂直に 2 つ配置することにより, 全方位のステレオ画像を得る。HyperOmniVision を用いて全方位ステレオ画像を取得するには, 全方位視覚センサを水平方向に配置する方法と, 垂直に配置する方法が考えられるが, 水平方向に設置した場合, 全方位ステレオ画像間で対応関係が求まる領域がかなり制限され, 探索ラインが複雑になる。一方, 垂直に設置した場合は, 全方位画像の放射線上にエビポーラ拘束が成り立つ。

垂直に 2 つ配置したときの位置関係を, 図 9 を用いて述べる。

3 次元座標系の原点は相対的に低い位置の HyperOmniVision の視覚系の原点とする。今, 垂直上方から得られる全方位画像 I_u , 下方から得られる全方位画像を I_l とし, 3 次元空間中の点 $P(X, Y, Z)$ に対する I_l, I_u 上での写像点を $p(r_1, \theta), p(r_2, \theta)$ とする。図 9 より, 3 次元環境中の同一対象が全方位画像 I_u 上に写像される位置は, 全方位画像 I_l 上に写像される位置よりも画像中心に近い内側の位置にある。つまり, 3 次元環境中の同一対象物体が写像される際の画像中心からの半径は, 全方位画像 I_u 上の r_2 の方が, 全方位画像 I_l 上の r_1 に比べ必ず小さいという関係がある。よって, $p(r_1, \theta)$ を $p(r_2, \theta)$ と対応付けるには, 図 10 の放射線上の太く示した探索範囲に限定することができる。今回は, 全方位画像をパノラマに展開した後に対応付けを行う。よって展開後の探索の方向は上下となり, 計算も容易に可能である。

Z 軸からの対象物体までの奥行き D を求める。この奥行き D は, 上側, 下側ミラー焦点との俯角を α' および, α とすると, カメラ間の距離である基線長 B を用いて以下のように表せる。

$$D = \frac{B}{|\tan \alpha - \tan \alpha'|} \quad (2)$$

α および, α' は, 視差 d と各全方位画像の中心から注目点までの半径 r_1, r_2 を用いて, 以下のようになる。

$$\alpha = \tan^{-1} \frac{(b^2 + c^2) \sin(\tan^{-1}(\frac{f}{r_1})) - 2bc}{(b^2 - c^2) \cos(\tan^{-1}(\frac{f}{r_1}))} \quad (3)$$

$$\alpha' = \tan^{-1} \frac{(b^2 + c^2) \sin(\tan^{-1}(\frac{f}{r_2})) - 2bc}{(b^2 - c^2) \cos(\tan^{-1}(\frac{f}{r_2}))}$$

$$= \tan^{-1} \frac{(b^2 + c^2) \sin(\tan^{-1}(\frac{f}{r_1 - d})) - 2bc}{(b^2 - c^2) \cos(\tan^{-1}(\frac{f}{r_1 - d}))} \quad (4)$$

上式から、下側の全方位画像 I_l 上の位置 $p(r_1, \theta)$ に対する奥行き D を求めるには、 b, c は双曲面ミラーの形状パラメータであるので、カメラの焦点距離 f 、基線長 B 、ステレオ画像から得られる視差 d が求まれば、計測可能であることが分かる。カメラキャリブレーションを行い、焦点距離、基線長、ステレオ画像間のピクセルの対応付けを行い、視差を求める。カメラキャリブレーションは、今回 Zhang の方法を用いた⁶⁾。

山澤は垂直にオムニカメラを2台設置し、対応付けに SSD を用いて環境のセンシングを行っている⁷⁾。しかし、この手法をそのまま瓦礫環境内に適用する場合、対応付けにおいて、以下のような問題点が挙げられる。

1. センサの小型化に伴い、基線長が短くなり、通常の対応付けでは十分な精度の視差を得るのが難しい。
2. オクルージョンや特徴が少ないテクスチャが多く、誤対応を起こしやすい。

1 はカメラ間の距離が短くなることで、視差がピクセル分解能以下になることで生じる。基線長 B が小さくなれば、同じ距離分解能を得ようとすると、必然的に俯角の差も小さくなる。よって得られる視差も小さくなり、ピクセル精度の対応付けでは、精度の良い距離計測が困難になるのが分かる。

2 は瓦礫環境下で木材などが無秩序に散乱するため、予想される。

本稿では、1 に対してはサブピクセル精度の高精度な対応付けが可能な位相限定相関法、POC (Phase Only Correlation) により解決する。未実装ではあるが、2 に対しては、グラフカットにより視差画像の補間を行うことで解決することが可能である⁸⁾。

本稿では、POC を用いて 1 の問題を解決する。



図-7 全方位画像例

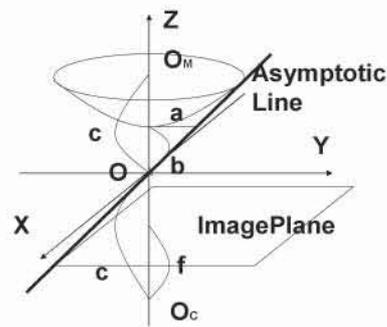


図-8 HyperOmniVision の視覚系

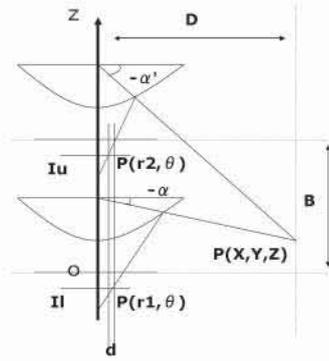


図-9 オムニステレオビジョン

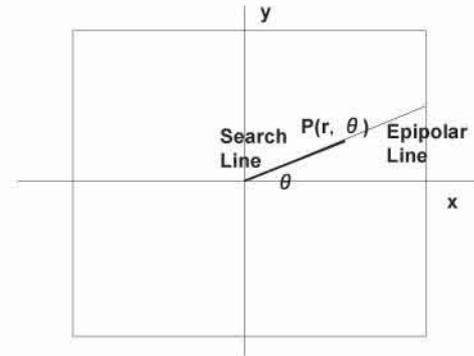


図-10 探索範囲の限定

(2) POC による対応付け

対応付けのアプローチは大きく 2 つに大別される。1 つはエッジやコーナーなど特徴的な点のみ対応付けする特徴ベースの手法、もう 1 つは領域ベースの手法である。領域ベースの手法は特徴ベースの手法と比較して、計算が大きくなるというデメリットがあるが、密な視差画像が得られるというメリットがある。瓦礫内の環境を取得するという目的において、できるだけ多くの情報を持っているほうが良い。よって、今回は領域ベースの手法を選択する。

領域ベースの手法としてオーソドックスなものに SAD(Sum of Absolute Difference)、SSD(Sum of Squared Difference)、NCC(Normalized Cross-Correlation)がある。NCC は位相情報を用い、SAD、SSD などと比較して精度良く対応付けできるが、1 点の対応付けの度に二次元離散フーリエ変換を伴うため計算量が大きい。しかし、近年、位相限定相関法(POC)と呼ばれる NCC と比べて大幅に計算量が小さく、より高精度な位置合わせの手法が登場している⁹⁾。この手法はピクセル分解能以下のサブピクセル精度で高精度な位置合わせを行うことができ、指紋認証、顔認証、超解像など、高精度な位置合わせが必要な分野へ応用されている。この手法を用いてピクセル分解能以下の視差を得る。

POC における位相限定相関関数の説明を行う。

今、 $N_1 \times N_2$ ピクセルの 2 つの画像 $f(n_1, n_2)$ 及び $g(n_1, n_2)$ が与えられたとする。画像 $f(n_1, n_2)$ 及び $g(n_1, n_2)$ の二次元離散フーリエ変換(DFT)を次式で定義する。

$$\begin{aligned}
 F(k_1, k_2) &= \sum_{n_1, n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\
 &= A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)}
 \end{aligned} \tag{5}$$

$$\begin{aligned}
 G(k_1, k_2) &= \sum_{n_1, n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\
 &= A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)}
 \end{aligned} \tag{6}$$

ここで $k_1 = -M_1, \dots, M_1$, $k_2 = -M_2, \dots, M_2$ は離散周波数インデックス(整数)であり、回転因子を $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$ 及び $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$ と定義する。 $A_F(k_1, k_2)$ 及び $A_G(k_1, k_2)$ は振幅スペクトルであり、 $\theta_F(k_1, k_2)$ 及び $\theta_G(k_1, k_2)$ は位相スペクトルを表す。このとき、 $F(k_1, k_2)$ と $G(k_1, k_2)$ の正規化相互パワースペクトルを次式で与える。

$$\begin{aligned}
R(k_1, k_2) &= \frac{F(k_1, k_2) \overline{G(k_1, k_2)}}{|F(k_1, k_2) G(k_1, k_2)|} \\
&= e^{j(\theta_F(k_1, k_2) - \theta_G(k_1, k_2))}
\end{aligned} \tag{7}$$

ここで、 $\overline{G(k_1, k_2)}$ は $G(k_1, k_2)$ の複素共役を表す。また、 $\theta_F(k_1, k_2) - \theta_G(k_1, k_2)$ は、二つの画像の位相差スペクトルである。この位相差スペクトルは画像マッチングにおいて、重要な性質を示している。しかし、これを直接利用するより、次のような関数を定義するほうが便利である。

位相限定相関関数 (POC 関数) $r(n_1, n_2)$ を、正規化相互パワースペクトルの二次元逆離散フーリエ変換 (IDFT) として定義する。

$$r(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1, k_2} R(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \tag{8}$$

2つの画像が類似している場合、POC 関数は、デルタ関数に近い極めて鋭いピークを有する。この相関ピークの高さは2つの画像の類似性を示し、相関ピークの座標は2つの画像の相対的な位置ずれに対応している。

以下では、画像 $g(n_1, n_2)$ が画像 $f(n_1, n_2)$ をサブピクセル移動量 (δ_1, δ_2) だけ微小に平行移動させた画像である場合を考える。このとき、 $f(n_1, n_2)$ と $g(n_1, n_2)$ の POC 関数は、次式で与えられる。

$$r(n_1, n_2) = \frac{\alpha}{N_1 N_2} \frac{\sin(\pi(n_1 + \delta_1))}{\sin(\frac{\pi}{N_1}(n_1 + \delta_1))} \frac{\sin(\pi(n_2 + \delta_2))}{\sin(\frac{\pi}{N_2}(n_2 + \delta_2))} \tag{9}$$

α は、相関ピーク値の高さを示すものである。画像に対し、ノイズが加わると α の値が減少する。この相関ピークのモデルに基づく関数のフィッティングにより、パラメータ α , δ_1 , δ_2 を推定することで、画像のサブピクセル精度の移動量を求めることができる。

POC 関数を用いて対応付けを行なう際に、重要な高精度化手法について述べる。

DFT において、信号が周期的に循環することが仮定されているため、画像端における信号の不連続性は問題になる。この不連続性の影響を減らすため、窓関数である以下のハンギング窓を適用する。

$$\omega(n_1, n_2) = \frac{\left\{1 + \cos\left(\frac{\pi n_1}{M_1}\right)\right\}}{2} \frac{\left\{1 + \cos\left(\frac{\pi n_2}{M_2}\right)\right\}}{2} \tag{10}$$

また、正規化相互パワースペクトル $R(k_1, k_2)$ の計算の際に、低域型のスペクトル重み付け関数 $H(k_1, k_2)$ を適用する。これは一般的に、自然画像のエネルギーは低周波領域に集中しており、エイリアシング、ぼけ、雑音、ひずみなどが加わると、高周波成分の SN 比が大幅に劣化するためである。スペクトル重み付け関数として、以下のガウス関数を適用する。

$$H(k_1, k_2) = e^{-((n_1 + \delta_1)^2 + (n_2 + \delta_2)^2) / 2\sigma^2} \tag{11}$$

ここで、式(9)に対応する相関ピークモデルは、次式のように変化する。

$$\hat{r}(n_1, n_2) = \frac{\alpha}{2\pi\sigma^2} e^{-((n_1 + \delta_1)^2 + (n_2 + \delta_2)^2) / 2\sigma^2} \tag{12}$$

一般に、移動量(δ_1, δ_2)は実数値であり、POC 関数のピーク座標がサンプリング格子点の間に存在するため、正確に移動量を推定するのは困難である。そこで、実際に計算された POC 関数の数値データに対して、本モデルをフィッティングすることで、ピクセル間に存在するサブピクセル精度のピークの位置を推定する。これら各種の高精度化手法を施したものが POC と呼ばれる。

また、オクルージョンされた点や視差が計算できない点に関しては、誤って計算される可能性がある。POC 関数のピーク相関値を用いることにより、これらの点を除去する。

(3) 評価実験

POC の性能を評価する。以下の実験においては、事前に入力画像に対して、画素の値を平均 0、分散 1 に正規化して処理を行っている。また、POC において、スペクトルの重み付け関数の分散値を 2.8、窓関数のサイズを 41×41 、フィッティングは相関ピーク値周辺の 5×5 のポイントを使用した。

POC のサブピクセル単位の位置合わせの精度評価と、計算速度の評価を行なった。精度評価は 1010×510 の Lena の画像を平行移動させ、Lanczos の方法で 10 分の 1 にリサイズすることで、擬似的にサブピクセル単位の移動をさせた画像を使用した(図 11)。計算速度の評価は 288×384 の tsukuba の画像に対し、水平方向 1line の対応付けの計算速度を評価した(図 12)。また、比較手法として SSD, NCC を実装し、さらにパラボラフィッティングを適用してサブピクセル精度の対応付けを行なった¹⁰⁾。

実験結果を表 6 に示す。

表-6 POC 評価実験結果

	平均誤差(pixel)	計算速度(cputime)
POC	0.015	4.844
SSD	0.069	1.344
NCC	0.062	11.28

従来手法である SSD と比較して、速度は劣るものの、4 倍以上の精度で位置合わせができた。また NCC と比較し、4 倍以上の精度、2.3 倍以上の速度で位置合わせができた。POC の SSD や NCC に対する優位性を示せた。

精度評価実験の平行移動量に対する POC の相関値と誤差値の関係を図 13 に示す。ピクセル分解能の半値、つまり ± 0.5 pixel になるほど誤差が大きくなるのが分かる。誤差は最大 0.05 pixel まで大きくなってしまっている。また、半値に近づくにつれて相関値が小さくなっていることも分かる。よって、相関値は対応付けにおける信頼度を表す指標であるといえ、これを用いて誤対応点を除去することが可能である。



図-11 Lena

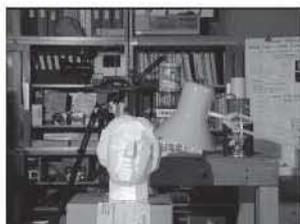


図-12 tsukuba

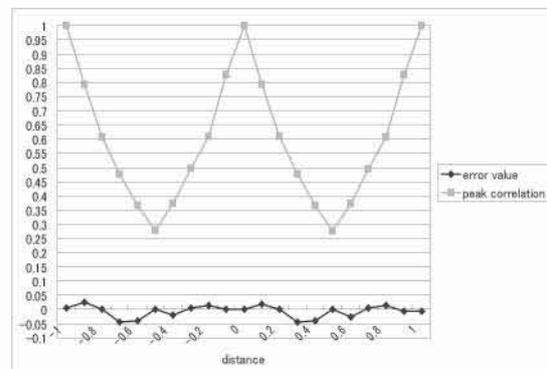


図-13 相関値と誤差との関係

5. まとめ

本稿は、目的がそれぞれ異なるセンサボールを2種類使い、被災者の発見と瓦礫内の3次元環境地図作成をそれぞれ自動で行うことにより、より安全で迅速な救助を目指した。

人体検出ボールでは、2次元放射温度計を用いて得られた64点のデータから、人体と背景の平均温度と分散を学習し、正規分布でしきい値を求め、人体の検出を行った。実験結果より、人体の温度の高い部分を人体の熱分布として学習することにより、正当率100%、適合率92%を得ることができた。64点という非常に粗い熱画像センサにおいても、人体の検出は可能であることを示した。

地図作成ボールでは、垂直に2台のオムニカメラを配置し、オムニステレオカメラを構成した。垂直の全方位画像の対応付けには、サブピクセルレベルでの高精度な位置付けが可能なPOCを用い、センサの小型化に伴うサブピクセル化に対応した。実験結果より、POCを用い、従来手法であるNCCと比較し、4倍以上の平均誤差0.015pixelの精度、2.3倍以上の速度で対応付けを行うことが可能であり、POCの優位性を示すことができる。

今後の課題として、人体検出ボールにおいては、熱画像センサの全方位化、音声の併用が挙げられる。地図作成ボールにおいては、POCで対応付けが不可能なテクスチャが少ない部分に対して、グラフカットを用いて補間を行い、対応する予定である。

参考文献

- 1) <http://www.rescuesystem.org/ddt/>
- 2) 河田恵昭：大規模地震災害による人的被害の予測，自然災害科学，Vol.16-1，pp.3-13，1997.
- 3) 森茂雄，後藤眞二，荒井郁男：電波による埋没生存者の存在および位置計測の一方方法，SANE2000，Vol.180，pp.33-38，2000.
- 4) 日笠博史，石田 宏，永谷 圭司，田中 豊：距離センサと視覚センサの統合による倒壊環境の三次元地図の構築．ロボティクス・メカトロニクス講演会'04 予稿集，2P2-H-23，2004.
- 5) Inoue, K., Yamamoto, M., Mae, Y., Takubo, T., and Arai, T.: Design of Search Balls with Wide Field of View for Searching Inside of Rubble, International Workshop on Safety, Security and Rescue Robotics, pp.170-175, 2005.
- 6) Zhang, Z.: A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pp.1330-1334, 2000.
- 7) 茶園篤，山澤一誠，横矢直和，竹村治雄：HyperOmni Visionによる全方位ステレオ視，IE96-122，pp.96-103，February 1997.
- 8) Tappen, F.M, Freeman, T.W.: Comparison of GraphCuts with Belief Propagation for Stereo, using Identical MRF Parameters, ICCV03, pp.900-907, 2003.
- 9) TAKITA, K., Aoki, T., Sasaki, Y., Higuchi, T., Kobayashi K.: High-Accuracy Subpixel Image Registration Based on Phase-Only Correlation, IEICE Vol.E86-A, No.8, pp.1925-1934, 2003.
- 10) 西口仁視，野村由司彦，坂本良太，杉浦徳弘：誤差特性を考慮した2次元サブピクセルマッチングの高精度化，画像の認識・シンポジウム(MIRU2007)，pp.863-868，2007.

筆者： 1) 井上淳一，工学部情報知能工学科，学生；2) 滝口哲也，都市安全研究センター，講師；3) 有木康雄，都市安全研究センター，教授

HUMAN DETECTION AND 3D ENVIRONMENTAL MAP ESTIMATION USING MULTISENSOR AND OMNI-STEREO VISION

Junichi Inoue
Tetsuya Takiguchi
Yasuo Arika

Abstract

Utilization of Information Technology is aggressively promoted for disaster prevention after the Great Hanshin Awaji Earthquake. Devices with electromagnetic radiation or search robots have been developed for detection of victims and estimation of three-dimensional environmental map in the house collapsed by the earthquake. However, they have problems in terms of the cost and the search time.

Recently, researches arise that scatter a large amount of small sensors. This approach has advantages of the lower device cost and the faster detection of human bodies in the narrow space over the larger area than the conventional ones. However, it aimed only at the victim's detection, and the rescuers had to confirm by their eyes where the victims were.

We propose, in this paper, two kinds of sensor balls with different purposes. One automatically detects human bodies and the other estimates three-dimension environmental map in collapsed house. Using these two kinds of sensor balls, we aim at a safer and prompter rescue. Under the assumption that the sensor ball is miniaturized and the price is reduced, we propose human detection by thermal image with low resolution as well as three-dimensional coordinate estimation by omni stereo camera with narrow base length using POC (Phase Only Correlation).