

# 群衆の三次元行動計測に基づく混雑・滞留検知センサ

木村 直哉<sup>†</sup> 市原 淑雄<sup>†</sup> 林 健太郎<sup>†</sup>

<sup>†</sup> 三菱プレジジョン株式会社

〒 247-8505 神奈川県鎌倉市上町屋 345

E-mail: †{nkimura,yichihara,hayashi.kentaro}@mpcnet.co.jp

あらまし 公共スペースの安全・安心を提供する画像センサについて述べる．本センサはステレオカメラにて撮影された映像をリアルタイムに解析し，視野内に発生した混雑や滞留状態を発見する．ステレオカメラにより人の眼と同じ原理で三次元空間情報を知覚し，群衆の密度や歩行速度を計測する．駅構内にて実施した検証実験では危険な混雑状態を正しく検知することができた．本センサはアナウンスによる注意喚起や安全システム制御のためのトリガとして利用される．

キーワード ステレオビジョン，距離画像，オプティカルフロー，群衆行動，混雑検知

## 1. はじめに

イベント会場に向かう歩道橋などでは，群衆が押し寄せることでたびたび事故が起きている [1] ~ [3]．日常においても駅や商業施設の階段，エスカレータなどで危険な状態に遭遇することがある (図 1)．

こうした群衆事故を未然に防止するためには，事前の周到な対策が必要である．ただし，それには限界があるため，加えて，事故発生前の予兆をリアルタイムに捉え，周囲への情報展開による自発的な抑止行動の誘発，分断規制などの強制的な軽減措置を行うことが重要となる．

予兆の捕捉方法としては，現在のところ監視員によるモニタリングが主流である．しかし，昨今では，負荷軽減や人員削減を目的に，監視員に代わり監視カメラの映像から群衆の混雑状態を自動的に検知する手法やセンサとしての実用化が報告されつつある [4]．それらは以下の 2 つに大別される．

- 人を個別に検出し，監視エリアに存在する人数や歩行速度を直接計測する方法
- 画像特徴を解析し，何らかの相関関係に基づいて混雑指標へ換算する方法

前者は，画像から人物そのものや，顔，頭部といった特徴的な部位を検出することを基本としており，さまざま

な方法が提案されている．例えば，予め学習された顔検出器を用いる方法 [5]，人の輪郭特徴に基づいて検出する方法 [6] ~ [8]，局所的な特徴点の移動軌跡情報をクラスタリングし，個々の人物を切り出す方法 [9] ~ [12]，局所的な特徴抽出と大域的な処理の組み合わせで人物を検出する方法 [13] などがある．しかしながら，このようなセグメンテーションをとる方法は，人が密着するようなひどい混雑シーンには弱いという問題がある．

一方後者は，監視カメラにて撮影された画像を解析して，群衆に相当する領域の大域的な特徴量から混雑状態を判断するものである．例えば，領域の総面積 [14] ~ [16]，特徴点の画素数 [4], [14], [15], [17]，エッジ勾配ヒストグラム [18]，フラクタル次元 [19] などの特徴量が計測され，群衆の密度に換算される．また，画像上の各点における見掛けの移動量と方向を表すオプティカルフローを算出し，群衆の歩行速度へ換算している方法もある [4]．これらの方法は，前者の方法に比べて混雑時への対応能力が高く，群衆の状態を大まかに捉える場合には有効であるが，画像上の見かけの特徴量から換算するため正確性には欠ける．また，一般には 1 台のカメラで撮影された画像上の特徴を用いるため，照明変化等に左右されやすいという問題もある．

本稿で述べる画像センサは，監視エリアをステレオカメラで撮影し，人の眼と同じ原理で三次元空間情報を知覚して，直接的に群衆の密度や歩行速度を計測する．そのため，監視エリア内に発生した混雑や，その中でも特にひどい状態である群衆が立ち止まった滞留状態をそれぞれ正確に発見することができる．

以下では，まず，本センサの概要を紹介する．そして，センサに搭載している混雑・滞留検知アルゴリズムについて述べる．最後に，実映像に対する実験結果を示して本センサの有効性を明らかにする．



図 1 混雑シーン

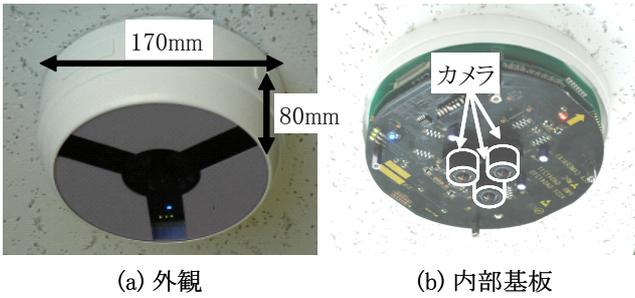


図 2 センサの外観

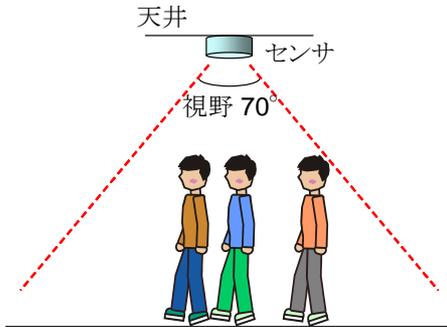


図 3 センサの設置

## 2. センサ概要

図 2 に本センサの外観を示す．本センサは，3つの CCD カメラを搭載した画像撮影部と汎用マイコンによる画像処理部を一体化したコンパクトな構造となっている．3つの CCD カメラは1枚の基板上に離れて配置されており，視点の異なる3枚の画像が同時に撮影できるようになっている．

2台のカメラで構成される一組のステレオ画像対の場合，カメラ同士を結ぶ線分に平行なエッジや模様を有する対象物の距離を得ることは出来ない．より多くのステレオ画像対を組み合わせることで計測精度を向上させることが可能となる [20]．本センサでは3台のカメラを搭載することにより，2台の場合に比べ，対応点の信頼度を向上することが出来る．

さらに，仮に1台のカメラが故障した場合でも，暫定的に残りの2台で三次元情報を得ることにより，最低限の機能を維持しながら修理を待つことが出来るため，システムとしての信頼度も向上することが出来る．

一方，汎用マイコンを搭載していることから，アプリケーションプログラムを変更することにより他の用途にも流用可能である．実際に，同一のハードウェアを用いて，セキュリティ施設への不正侵入を防止するための共連れ検知センサや一方通行ゲートにおける逆流（侵入）を検知するセンサもすでに実用化している [23]．

我々は，本センサを天井やポールなどに取り付け，群衆を見下ろすように撮影することを仮定している．これは，人同士による隠蔽を無くし，より高い検知精度を得るためである（図 3）．

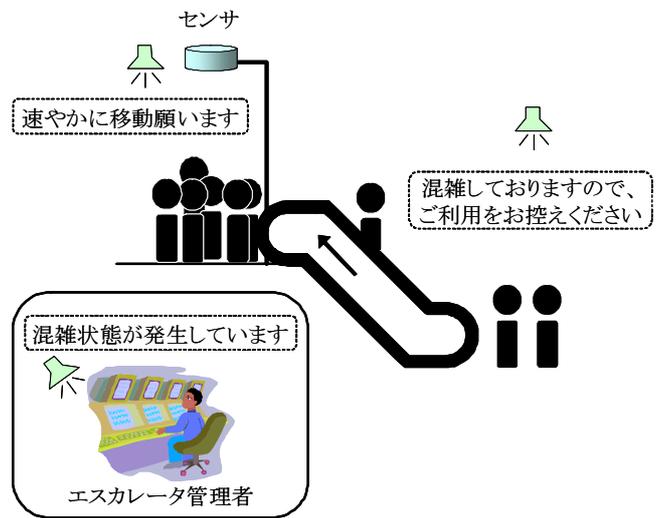


図 4 エスカレータへの適用例

図 4 は，本センサをエスカレータ降り口に設置した適用例である．このシステムでは，降り口に発生した混雑や滞留状態を検知したときに，降り口の群集に対して移動を促すとともに，乗り口の人物に対して利用を控えるようアナウンスを行う．また，エスカレータ管理者に対しても通知を行う．

## 3. 混雑・滞留検知アルゴリズム

### 3.1 三次元空間情報の抽出

前述の通り，本センサでは視点の異なる3枚の画像が同時に撮影される．この3枚の画像から視野内における三次元空間情報を抽出する．

複数台のカメラ画像から三次元空間情報を抽出する手法としては，古くより「ステレオ視」が知られている．これは三角測量の原理に基づいている．以下では説明を簡単にするために，焦点距離の等しいカメラ2台が，光軸が平行で，かつ水平に設置されている場合について述べる．ここで，図 5 に示すように，空間中の点  $P(X, Y, Z)$  が各画像においてそれぞれ  $p_l(x_l, y_l), p_r(x_r, y_r)$  として観測されたとすると，以下の関係式が成り立つ．

$$X = \frac{b(x_l + x_r)}{2d} \quad (1)$$

$$Y = \frac{b(y_l + y_r)}{2d} \quad (2)$$

$$Z = \frac{bf}{d} \quad (3)$$

ただし， $f$  は焦点距離， $b$  は基線長， $d$  は視差 ( $x_l - x_r$ ) である．したがって， $f, b$  を既知とすると，点  $P(X, Y, Z)$  の画像間での対応点  $p_l(x_l, y_l), p_r(x_r, y_r)$  から，点  $P(X, Y, Z)$  の三次元位置が算出される．なお，画像間の対応点の探索にはブロックマッチング手法を用いる．この方法では，一方の画像中の各画素ごとに，その周辺の局所的なパターンを用いて，他方の画像から類似するパターンの位置を求めることで対応点を検出する．

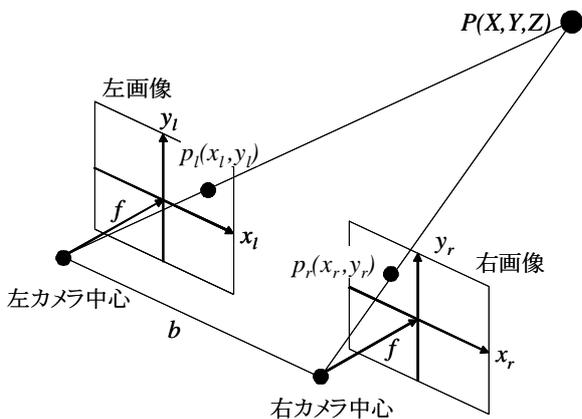
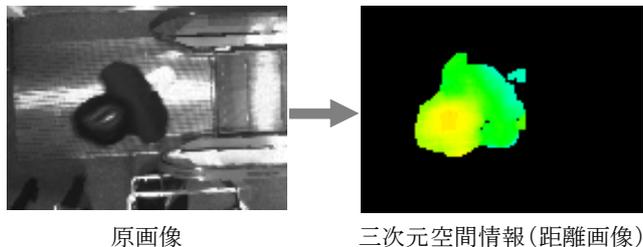


図 5 ステレオ視の原理



原画像 三次元空間情報 (距離画像)

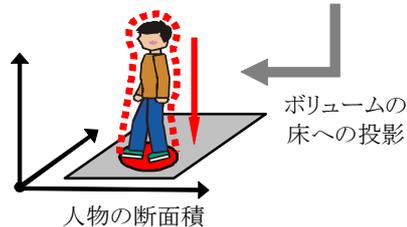
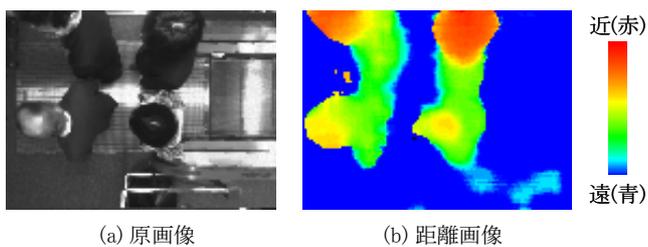
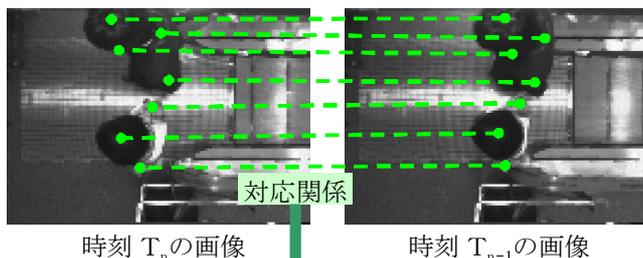


図 7 密度の計測



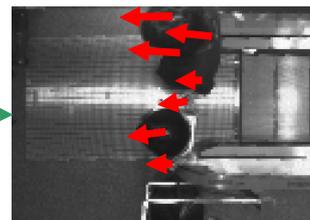
(a) 原画像 (b) 距離画像

図 6 三次元空間情報の抽出



時刻  $T_n$  の画像 時刻  $T_{n-1}$  の画像

対応関係



オプティカルフロー

図 8 オプティカルフローの算出

対応点探索による三次元位置の算出をすべての画素に対して行うことで三次元空間情報をあらゆる距離画像が生成される。図 6 に距離画像の例を示す。図 6(a) は原画像のひとつであり、同図 (b) はそれに対する距離画像である。距離画像は、カメラから物体までの距離に応じた色で表示されている。

一般に、画像上のすべての画素で対応点を正しく求めることは非常に難しい。シーンによっては似たようなパターンが複数存在したり、正しい位置であっても視点の違いによる見え方やランダムノイズの影響により画像間で完全に一致することがないためである。

本センサでは、カメラを 3 台に増やすことで、対応点の信頼度を向上させている。この 3 眼ステレオ視では、ひとつの基準カメラに対して、2 組のステレオ画像対を形成し、それぞれのステレオ画像対から得られる結果を統合することで、対応誤りを軽減させることができる。

### 3.2 群集の密度と歩行速度の計測

#### (1) 密度計測

得られた三次元空間情報から群集の密度を計測する。これは次の手順により行われる (図 7)。

- ① 視野内の所定の空間 (検知エリアとして直方体の空間を設定) に存在する立体物を群集として抽出する
- ② 立体物の総断面積を求める
- ③ 検知エリアの断面積 (底面の面積) に対する立体物の総断面積の密度を求める

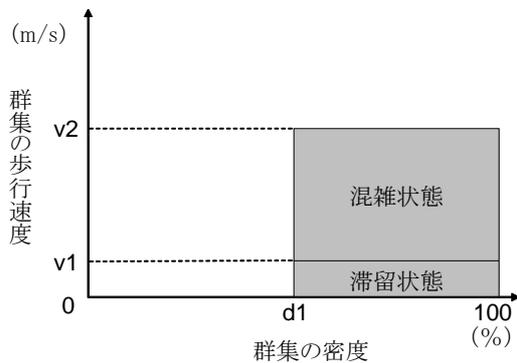
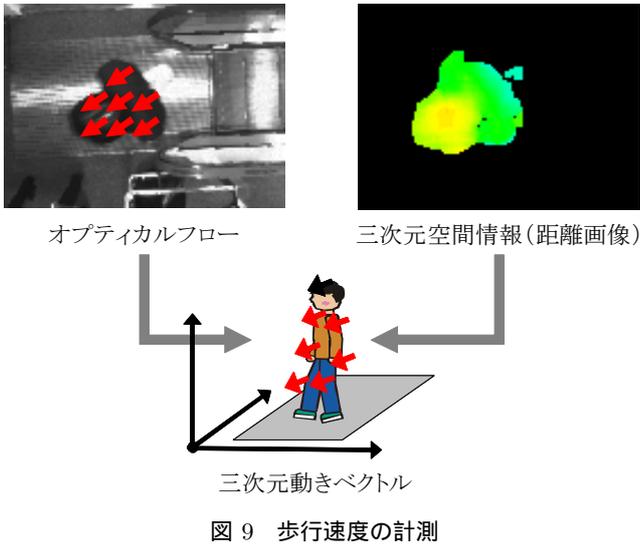
立体物の抽出方法としては、背景差分を行った画像同士でステレオマッチングを行う方法 [21], [22] などがある。

本センサはハードウェアで三次元空間情報を取得しているため、人の存在しない背景のみの三次元空間情報を用意しておき、背景の三次元空間情報と現在の三次元空間情報を比較することにより行う。なお、背景の三次元空間情報は、人が存在しない間に徐々に更新していく。

#### (2) 歩行速度計測

群集の歩行速度は、物体各点における三次元の動きベクトル<sup>(注1)</sup>を求め、その移動量の平均値として算出する。そのために、まず、基準カメラ画像上でオプティカルフローを求める。オプティカルフローは、画像上の各点における見掛けの移動量と方向を表した移動ベクトルであり、時間的に連続した画像間の対応点を探索することで求められる。例えば、図 8 に示すように、時刻  $T_n$  および  $T_{n-1}$  の画像において対応関係を求めると、その始点・終点座標値から移動ベクトルが求められる。

(注1): 厳密には鉛直方向成分を無視している。人物が水平な床面を歩行している場合、鉛直方向成分の移動量は小さいと考えられるためである。



フローを計算する点は、あらかじめ画像中に離散的に配置され、立体物としてみなされた領域内に含まれる点を選択する。これは、計算時間を短縮するとともに、人物以外の背景物体の影響を軽減させるためでもある。

求められたオプティカルフローは、三次元の動きベクトルへ変換される(図9)。すでに画像上の各点に対応する空間の三次元情報が得られているため、その情報をもとに二次元から三次元への変換を行う。最後に、三次元の動きベクトルの平均移動量を求め、群衆の歩行速度とする。

### 3.3 混雑・滞留状態の判定

計測された群衆の密度と歩行速度から混雑や滞留状態を推定する。ここでは、図10に示すように、密度と歩行速度の2次元マップにおいて、混雑状態と滞留状態を定義している。具体的には、密度  $d_1$  %以上、かつ、歩行速度が  $v_1 \sim v_2$  m/sの状態を混雑状態とし、密度  $d_1$  %以上、かつ、歩行速度が  $v_1$  m/s以下の状態を滞留状態としている。いずれも、人が多く、動きが遅いことを想定している。

## 4. 実験

本センサに実装した混雑・滞留検知アルゴリズムの有

表 1 意図的に発生させた混雑状態による実験

No.	写真	状況
1	図 11(a)	群衆の密度しきい値 $d_1$ % (図 10) に達しないため、混雑および滞留状態とはならない。
2	図 11(b)	群衆の密度しきい値 $d_1$ % (図 10) との関係により判定結果が異なる。 $d_1$ % を超えた場合で、群衆の歩行速度が $v_1 \sim v_2$ m/s の状態なら混雑状態、群衆の歩行速度が $v_1$ m/s 以下ならにより混雑または滞留状態となる。
3	図 11(c)	群衆の密度しきい値 $d_1$ % (図 10) を超えているため、群衆の歩行速度が $v_1 \sim v_2$ m/s の状態なら混雑状態、群衆の歩行速度が $v_1$ m/s 以下ならにより混雑または滞留状態となる。



(a) 群衆の密度 小



(b) 群衆の密度 中



(c) 群衆の密度 大

図 11 意図的な混雑状態

効性を確認するため、検証実験を行った。

### 4.1 意図的に発生させた混雑状態による実験

12名の人物により群衆の密度および速度を変化させて、混雑状態および滞留状態を意図的に作りだし、センサの検知性能を確認した。意図的な混雑状態による実験の状況を表1、および図11に示す。

すれ違い、一部立ち止まり、およびランダム歩行など様々なバリエーションを含む各20回の試行に対し、結

表 2 検知された混雑状態

No.	状況	備考
1	ベビーカーを押した人が降り口で方向転換したために、後続利用者が詰まり、混雑した。	図 12(a)
2	杖をついた足の不自由な方が降りる際に時間を要し、後続利用者が詰まり、混雑した。	図 12(b)
3,4,5	電車から下りた乗客が殺到し、混雑した。	図 12(c)



(a) ベビーカーの方向転換による混雑



(b) 杖をついた方の低速歩行による混雑



(c) 電車から降りた乗客の殺到による混雑

図 12 検知された混雑状態

果として、混雑状態・滞留状態ともに 100 % 検知することができた。

#### 4.2 実環境下における実験

実環境下における実験は何力所かで実施した。一例として、某駅の上りエスカレータの降り口にセンサを設置し、土・日曜をまたいで 4 日間（7時から 23時まで）センサを稼働させた。その間、延べ 131,918 人の通行があり、センサは 5 回の混雑状態を検知した。ほとんど移動のない危険な滞留状態は検出されなかった。

検知した状況を表 2、および図 12 に示す。これを見ると、混雑状態として妥当な状況が検知されていることが分かる。

## 5. む す び

本稿では、公共スペースの安全・安心を提供する画像センサとして、群集の三次元行動計測に基づく混雑・滞留検知センサについて述べた。本センサは 3 台のカメラにて撮影された映像をリアルタイムに解析し、視野内に発生した混雑や滞留状態を発見する。某駅にて実施した実験では、延べ 131,918 人が通行する間に 5 回の混雑状態を正しく検知することができ、有効性が確認できた。今後は、さらなる性能向上や、応用アプリケーションの開発について検討する予定である。

## 文 献

- [1] "第 32 回明石市民夏まつりにおける花火大会事故調査報告書", 2002.
- [2] "東京ビッグサイトにおけるエスカレーターでの事故発生について", <http://www.bigsight.jp/topics/press080804.pdf>, 2008.
- [3] "名古屋市交通局久屋大通駅におけるエスカレーター事故とその対応について", [http://www.mlit.go.jp/report/press/tetsudo09\\_hh\\_000006.html](http://www.mlit.go.jp/report/press/tetsudo09_hh_000006.html), 2008.
- [4] 渡邊 清高, 羽下 哲司, "映像監視による群集監視システム", 画像センシングシンポジウム SSII, No.IN2-18, 2008.
- [5] S.-F. Lin, J.-Y. Chen, H.-X. Chao, "Estimation of number of people in crowded scenes using perspective transformation", IEEE Trans. on Systems, Man, and Cybernetics, Vol.31, No.6, pp.645-654, 2001.
- [6] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors", ICCV, Vol.1, pp.90-97, 2005.
- [7] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations", CVPR, Vol.2, pp.459-466, 2003.
- [8] M. R. Sullivan and M. Shah, "Detecting and Segmenting Humans in Crowded Scenes", ACM Multimedia Conference, 2007.
- [9] D. Biliotti, G. Antonini, J. P. Thiran, "Multi-layer hierarchical clustering of pedestrian trajectories for automatic counting of people in video sequences", Proc. IEEE Workshop on Motion and Video Computing, 2005.
- [10] V. Rabaud and S. Belongie, "Counting Crowded Moving Objects", IEEE CVPR, Vol.1, pp.705-711, 2006.
- [11] G. J. Brostow and R. Cipolla, "Unsupervised bayesian detection of independent motion in crowds", CVPR, Vol.1, pp.594-601, 2006.
- [12] A.M. Cheriyyadat, B.L. Bhaduri and R.J. Radke, "Detecting Multiple Moving Objects in Crowded Environments with Coherent Motion Regions", IEEE Computer Society Workshop on Perceptual Organization in Computer Vision, 2008.
- [13] B. Leibe, E. Seemann and B. Schiele, "Pedestrian detection in crowded scenes", CVPR, Vol.1, pp. 875-885, 2005.
- [14] A. C. Davies, J. H. Yin and S. A. Velastin, "Crowd monitoring using image processing", Electron. Commun. Eng. J., Vol.7 pp.37-47, 1995.
- [15] S.-Y. Cho, T. W. S. Chow and C.-T. Leung, "A neural-based crowd estimation by hybrid global learning algorithm", IEEE Trans. Syst. Man, Cybern. B,

- Vol.29, pp.535-541, 1999.
- [16] N. Paragios and V. Ramesh, "A mrf-based approach for real-time subway monitoring", CVPR, Vol.1, pp.1034-1040, 2001.
  - [17] C. S. Regazzoni and A. Tesei, "Distributed data fusion for real-time crowding estimation", Signal Process., vol. 53, pp. 47.63, 1996.
  - [18] D. Kong, D. Gray and H. Tao, "A Viewpoint Invariant Approach for Crowd Counting", IEE Int. Conf. on Pattern Recognition, Vol.3, pp.1187-1190, 2006.
  - [19] A. N. Marana, L. F. Costa, R. A. Lotufo, "Estimating crowd density with Minkowski fractal dimension", Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, vol.6, pp.3521-3524, 1999.
  - [20] 奥富正敏, 金出武雄: "複数の基線長を利用したステレオマッチング", 電子情報通信学会論文誌, vol.J75-D-II, no.8, pp.1317-1327, 1992 .
  - [21] Umeda, K., et al., "Subtraction Stereo -A Stereo Camera System That Focuses On Moving Regions -", Proc. of SPIE-IS&T Electronic Imaging, Vol.7239 Three-Dimensional Imaging Metrology, 723908, 2009.
  - [22] 梅田和昇, 寺林賢司, 橋本優希, 中西達也, 入江耕太: "差分ステレオ - 運動領域に注目したステレオ視 - の提案", 精密工学会誌, Vol.76, No.1, 2010.
  - [23] "人物検知センサ", 三菱プレジジョン技報, Vol.3, pp.19, 2008.