

全方向ステレオシステム (SOS) のステレオ情報を用いた空間イベントの取得

棚橋 英樹†, 山本 和彦‡, 島田 大輔‡, 丹羽 義典†

† (財) ソフトピアジャパン HOIP 研究室/JST

‡ 岐阜大学工学部

あらまし: 人の行動を追跡・理解する上で人だけでなく人も含めた環境の情報を取得する必要がある。これまで、我々は観測点を中心とした全ての方向のカラー画像と3次元情報を同時に、リアルタイムで取得可能な全方向ステレオシステム (SOS) を開発した。

本稿では、本システムを用いて空間の複数のイベントを同時に取得する方法について報告する。空間イベントとしては、複数の移動体 (人物) の抽出とした。空間イベントを取得する方法としてカラー画像を用いて、あらかじめ取得した背景画像との差分により取得する方法があるが、照明条件の変化や背景と人物との色により安定に抽出できない場合がある。そこで、全方向ステレオシステムから得られる全方向のステレオ情報を用いて照明条件の変化に左右されにくい複数のイベントの同時取得を行った。

キーワード: 全方向ステレオ, イベント抽出, 背景差分, 全方位センサ

Acquisition of Events in the Space using Stereo Information of Stereo Omni-directional System (SOS)

Hideki TANAHASHI †, Kazuhiko YAMAMOTO ‡, Daisuke SHIMADA ‡ and Yoshinori NIWA †

† HOIP, Softopia Japan / JST

‡ Faculty of Engineering Gifu Univ.

Abstract: It is necessary for tracking and recognition of human motion to acquire not only human information but also environment information around a human. We have been developed a Stereo Omni-directional System, named SOS. This system can take all directional views and depth maps from an observation at the same time, in real-time.

In this paper, we describe the method of acquisition of events in the space by using the SOS. The events are moving objects. We extract the moving object regions in the space at the same time. In order to extract the moving object region from images, the subtracting background image method using color information has been proposed. However, in this method, which uses the color, the extraction result of moving object region is greatly influenced by lighting conditions and also the environmental condition. So, in order to extract the moving object region, we used the subtracting background image which uses the depth map calculated from the all direction stereo pair images acquired from SOS.

Keywords: All Directional Stereo, Event Extraction, Subtracting Background Image, Omni-directional Sensor

1. はじめに

近年の情報通信技術・コンピュータ技術の飛躍的な発展により、人間と情報関連機器のインターフェイスの高度化が図られつつある。これらの多くは、背景から人間を抽出し、顔の認識やジェスチャの認識など人間のみに焦点をあてた研究である。より高度化されたインターフェイスを実現するためには、人間とその周辺の環境を含めた大局的な観点から、人間の状況を理解する必要がある。こうした技術を我々は実環境情報処理技術⁽¹⁾と位置づけ、人間センシング技術や環境センシング技術の研究を行っている。

実環境における広範囲な情報を取得する方法として、カメラを回転させる手法^(2,3,4)、双曲面等の回転体ミラーを用いた全方位画像センサ^(5,6,7,8)を用いる手法やこれらのセンサを複数組み合わせる3次元情報を取得する方法^(9,10,11)や、取得したい環境の周囲に複数のカメラを配置する手法^(12,13)が提案されている。カメラを回転させる方法は、1台もしくは複数のカメラを回転し、観測点を中心とした画像や3次元情報を得る手法である。高い分解能の画像情報や3次元情報が得られるが、全ての方向の情報を取得するために時間を要するため、動的な環境では有効とは言えない。また、曲面ミラーを用いた全方位センサを用いる手法は、双曲面ミラーや円錐ミラー等、光学的工夫により観測点を中心とした全周囲の画像を1台のカメラで取得する方法である。リアルタイムに全方位画像が取得できるが、観測点を中心とした360度視野の全方位の画像を1台のカメラで取得するため空間分解能が低い等の問題がある。また、曲面ミラーを用いているため、センサと対象の位置関係により空間解像度が一定とはならない。高解像度の画像を取得するために様々な提案^(14,15)がされているが、空間解像度が一定ではない問題に関しては今後の課題と考えられる。これらの手法は、観測点を中心とした周囲360°の画像や3次元情報を得ることが可能であるが、対象の位置がセンサより上方にある場合や下方にある場合など、センサと対象の位置によって死角が生じる。一方、

複数のカメラを配置する手法は、空間の周囲にカメラを配置し、複数の視点から得られた多視点画像から3次元シーンと動的イベントの検出が可能である。この手法ではシーン全体の映像と3次元情報の取得が可能であり、対象シーンの任意の視点からの画像を生成可能であるが、あらかじめ複数のカメラを空間の周囲に配置する必要がある。

こうした問題を鑑み、観測点を中心とした全ての空間のカラー画像と3次元情報を同時に得ることが可能で、また容易に移動可能な全方向ステレオシステムを開発した^(16,17)。全方向ステレオシステムは、センサを中心とした全ての方向のカラー画像と3次元情報が同時にリアルタイムで取得が可能である。本システムを用いることにより、複数の空間イベントが同時に取得可能であり、またそのイベントの3次元空間位置の取得や3次元的な解析も可能である。

本論文では、我々が開発した全方向ステレオシステムを用いて人間を中心とした環境における複数の空間イベントの取得を行ったので報告する。

2. 全方向ステレオシステム (SOS)

全方向ステレオシステム (Stereo Omnidirectional System : 以下 SOS) は、システムを中心とした3次元空間内の全方向のカラー画像と距離画像を同一時刻にリアルタイムで取得可能なシステムである。図1に本システムの外観

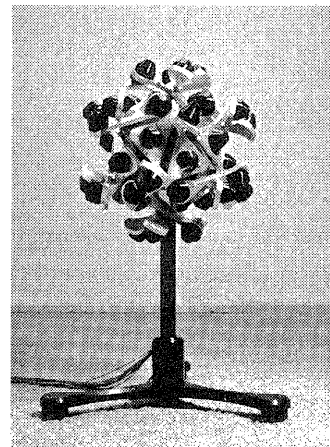


図1 全方向ステレオシステム (SOS) 外観

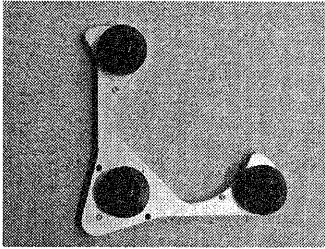


図2 ステレオユニット

を示す。本システムは、60個のカメラを用いて正20面体の各面上に3眼ステレオユニットを配置した。これによって3次元空間上の全ての方向におけるカラー画像と3次元情報を同一時刻にリアルタイムで取得することが可能である。また、同じ特性を持つユニットを正20面体の各面上に配置することにより、3次元空間を均等に分割し、高解像度の情報の取得が可能である。

本システムでは3個のカメラ対からなるステレオユニット（図2）を基本とし、このユニットを正20面体の各面に配置した。ユニット上の各カメラは、同一平面上に光軸が平行で、一つを基準カメラとして上下と左右の2つのステレオペアを構成するようにカメラを配置した。これによって、エピポーラ線が画像に対して水平、垂直であるため、ステレオペアにおける対応点探索を高速に行うことが可能である。また、この2つのステレオペアを用いることにより、それぞれのステレオペアから得られる信頼性の評価値を用いて統合することにより、計測結果の信頼性を上げることが可能となる。

ところで、ステレオ処理においては、カメラから遠く離れた3次元情報を精度良く求めるためには、ステレオペアを構成する2個のカメラ間隔（ベースライン長）を長くする必要がある。しかし、ベースライン長を長くするとステレオユニットが大きくなり、そのユニットをそのまま正20面体の面内に配置すると、面の一边の長さは大きくなり、システムの全体のサイズを大きくせざるを得ない。システムのサイズが大きくなると、隣り合った面のどちらかから最初に見え始める点までの距離（システムの中心からの距離）が大きくなり、

表1 SOS システム仕様

撮像素子	1/3 インチカラーCMOS イメージセンサ
解像度	644(H)×484(V)
焦点距離	2.9mm
画角	96.6° (H)×71.9° (V)
ベースライン	約90mm
アレー部外径	約27cm
アレー部重量	4.5Kg

システムの中心近くで近接死角を生じる。このため全ての方向が見え始めるシステムの中心からの距離を小さくするためには、多面体の大きさを小さくする必要がある。本システムでは、こうした相反する条件を満たすため、十分な3次元情報の精度を持つステレオユニットを、互いに隣り合うステレオユニットを入れ子状態に配置することによってカメラのベースラインと面のサイズを独立させる構成を考案した⁽¹⁸⁾。更なるサイズの限界を求めるために3次元CADを用いてカメラ毎に視野角のシミュレーションを行い、各ユニットが互いに隣り合うユニットのカメラの視野を遮らないようにユニットを配置した。こうしたユニットの配置により、システムから遠く離れた3次元情報を精度良く取得可能でありながら、近接死角を小さくすることが可能となった。

表1にシステムの仕様を示す。本システムでは、水平画角96.6度、垂直画角71.9度のレンズを使用することによって中心から400mm以上の空間は死角がなく、画像と3次元情報が得られる。各ユニットからは、1枚のカラー画像と2枚の白黒画像からなるステレオ画像が取得され、全方向の20枚のカラー画像とステレオ画像を1セットとして、15セット/秒でメモリユニットに転送する。その後、パソコン上に画像を転送し、ステレオ処理を行う（図3）。

本システムの特徴を以下に示す。

1. センサを中心とした空間全ての方向のカラー画像と3次元情報を取得できる。

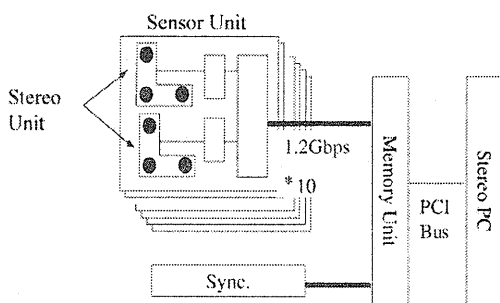


図3 システム構成図

2. リアルタイムで取得可能である。
3. 方向による解像度の差がない高分解能の情報が得られる。
4. 高精度の3次元カラー情報が得られる。

3. 空間イベントの抽出

本稿では、開発したSOSを用いて、3次元空間における空間イベントの抽出及び追跡について検討した。実環境は、静的な環境と動的なイベントから構成されるため、あらかじめ静的な環境の画像を背景情報として持っておき、現在の画像との差分を取ることで空間イベントの抽出を行った。空間イベントは3次元空間における複数の移動体（人）の同時抽出である。

3.1 視差画像の取得

SOSからは、20面体の各面の方向に対して2つのステレオペアが得られる。これらの画像から、視差画像の評価は、奥富ら⁽¹⁹⁾が開発したマルチベースラインステレオ法をベースに行う。入力画像は初めに、レンズ歪みを補正し、エッジ画像を生成する。その後、基準画像から領域を切り出し、その画像を上と左の画像それぞれにずらしながら画像間の相関を取る。相関演算には、SAD (Sum of Absolute Difference) を用いた。水平、垂直の2つの方向から得られた相関値から、相関の高い方のずらし量をその画素における視差とした。

画像間で相関演算を行う時、テクスチャが存在

しない領域に対しては、誤対応が生じる。そのため、相関マスクの領域内におけるテクスチャの分布を、対応の確からしさとして用いることにより、誤対応の点を除いた。これら一連のステレオ処理には、ステレオビジョンソフトウェアライブラリ^(20,21)を用いた。

3.2 空間イベントの抽出

我々は、これまでSOSから得られる全方向カラー画像を用いて、背景のカラー画像との差分を取ることで、イベント領域を抽出し、その領域に対して3次元情報を計算することにより空間におけるイベント抽出を行った⁽²²⁾。しかしながら、カラー画像を用いた背景差分では、抽出した結果は、光源変化の影響や環境に左右される。例えば、空間イベントが背景の色と同じ場合、カラー画像を用いた背景差分では抽出が困難となる。本稿ではSOSから得られたステレオ画像から視差画像を生成するプロセスの並列化を図り、得られた視差画像を用いてあらかじめ取得した背景視差画像との差分処理により空間イベントの抽出を行うことにより、カラー画像を用いた背景差分より背景に依存しにくい空間イベントの抽出を行った。

しかしながら、ステレオ画像から得られる距離情報は光源の全体的な変化にはあまり左右されないが、光源のちらつきによって得られる対応点が異なり、そのまま背景との差分を取るとノイズとなって現れる。そのため、取得した距離情報に対してメディアンフィルターを施すことにより空間における誤対応の除去を行った。また、背景画像に関してはこの空間のメディアンフィルターを施すとともに時間軸方向にもメディアンフィルターを施すことによりイベント取得の安定化を図った。

SOSからは、同期の取れた全方向の2つのステレオペアを構成するステレオ画像対が20枚得られる。得られたステレオ画像対から視差画像を計算し、得られた各視差画像と背景視差画像とをそれぞれ差分処理を行う。その後、閾値処理により2値画像を作成し、その画像に対して、縮小、膨張処理を行い、マスク画像を作成した。このマス

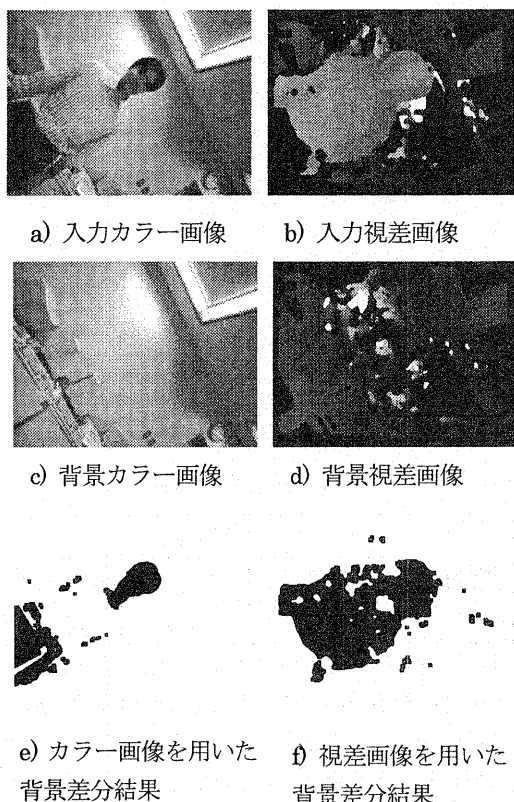


図4 カラー画像と視差画像による背景差分結果

ク領域に対して視差画像から3次元情報を取得した。なお、これらの処理は、 160×120 の解像度で行った。図4に背景差分にカラー画像を用いた場合と視差画像を用いた場合の結果を示す。図4 a), b)は同じ時間でのカラー画像とステレオ画像から計算した視差画像である。視差画像は、距離の近い点ほど白く表示しており、黒い部分は背景と判断された点を示す。また、図4 e), f)は、それぞれ背景差分を行った結果を2値化し、縮小、膨張処理を行ったマスク画像である。カラー画像を用いた背景差分では、背景と同一色の領域が抽出できないが視差画像を用いることにより、領域全体を抽出できているのがわかる。

3. 3 空間イベント統合

3. 2節より空間中の各方向画像からイベント

領域が取得できるが、イベントによっては、一つの方角だけでなく複数の方向にまたがる場合がある。例えば SOS のレンズ特性から各カメラの中心から 1 (m) 離れた水平、垂直の視野範囲は、それぞれ $2.92 \text{ (m)} \times 1.45 \text{ (m)}$ である。そのため、各方向から得られた3次元情報を、ユニット間の位置関係を利用して一つの座標系に変換し、統合することにより、複数の方向で観測されたイベントの統合及びイベントのセグメンテーションを行った。各ユニット間の3次元位置関係は、CAD から得られたパラメータを用いた。

各ユニットから得られた視差画像には光源のちらつきやステレオ処理における誤対応によるノイズ等により、イベント領域以外も抽出される。また、対象の距離によっては空間的に重なりが生じる。そのため、空間内におけるイベント領域のセグメンテーションにそのまま3次元情報を用いるのは困難である。そのため、対象とする空間を $N_x \times N_y \times N_z$ 個のボクセル領域に分割し、そのボクセルを用いて、各ユニットからの3次元情報の統合と、ノイズの除去や空間内のイベントのセグメンテーションを行った。

3. 2節より得られた各ユニットのイベント領域の視差画像から3次元情報を取得する。得られた3次元情報から各ユニットの3次元位置情報を用いて、グローバル座標系に変換し、対応するボクセルに投票を行う。得られたボクセル空間で、各ボクセルの連続する領域を統合することによってイベント領域のセグメンテーションを行う。次に、この統合された領域の重心及び領域の大きさによりノイズの除去を行う。また、ここで抽出対象の幾何的拘束（例えば人間の追跡では身長、大きさなど）を導入し、イベント抽出の安定性を図った。

4. 実験

3次元空間における動的イベントとして部屋の中を移動する複数の人物の抽出を行った。図5に実

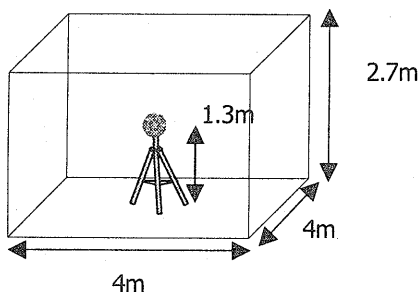
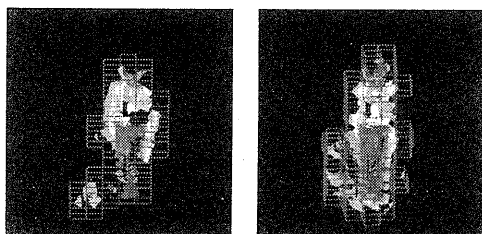


図5 実験環境



a) 視差画像を用いた背景差分結果 b) カラー画像を用いた背景差分結果

図6 背景差分結果の比較

験環境を示す。実験は、周囲をパーティションで仕切り、 $4\text{m} \times 4\text{m}$ の範囲である。また、床面から天井までは 2.7m で、SOS の中心は、ほぼ実験環境の中心に配置した。

イベント統合に用いたボクセルの数は、 $4\text{m} \times 4\text{m} \times 3\text{m}$ の3次元空間を N_x , $N_y \times 41$ 分割、 $N_z=31$ 分割として、1つのボクセルの表現している領域は、 $10\text{cm} \times 10\text{cm} \times 10\text{cm}$ の正方領域とした。背景差分の結果からの閾値は実験的に求め、固定とした。また、抽出対象の幾何拘束として本稿では、ボクセルの連続する領域の高さを 1m 以上とした。

視差画像を用いた背景差分の有効性を示すため、この SOS の中心から約 1m の円周上を移動する1人の人物の追跡を行った。図6に視差画像を用いた背景差分から抽出した人物とカラー画像を用いた背景差分から抽出した人物を3次的に統合した画像を示す。どちらの結果もボクセルにより1つの領域として統合されているのがわかる。しかしながら、カラー画像を用いた差分では、影に

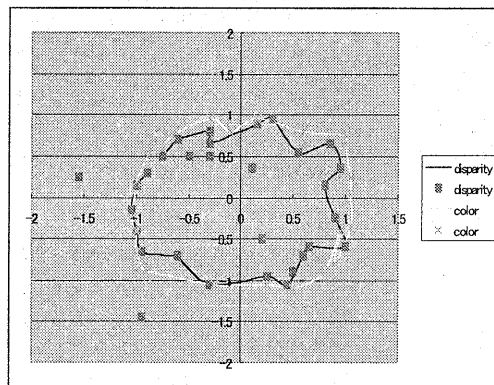


図7 抽出された人物の軌跡

表2 軌跡の位置誤差

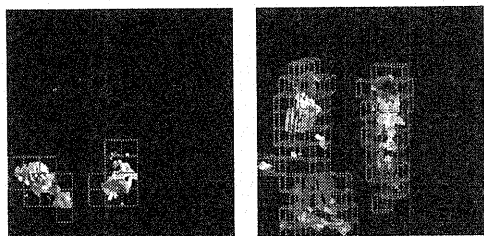
	平均誤差	誤差分散
視差画像背景差分	0.014	0.0003
カラー画像背景差分	0.021	0.0007

よって背景の部分も取得されている。影による背景部分の影響を小さくするために背景差分に用いる閾値を大きくすることも考えられるが、背景色の影響を大きく受けることとなる。また、図7に2つの手法から抽出した領域の重心を追跡した結果を示す。また、表に半径 1m の円との2乗平均誤差と誤差分散を示す。図7上の点は、人物の領域以外にも取得された領域である。どちらも精度よく人物領域が追跡されているが、位置誤差を比較すると視差画像を用いた背景差分の追跡結果の方が平均誤差、誤差分散とも小さい。これは、前述した人物による影等の領域も同じ領域として抽出されているためと考えられる。

最後に2人の人物を抽出した結果を図8に示す。図8a), b)とも同じ画像で視点を変化させた画像である。2人の人物がそれぞれ別の領域に分離されている。また、高解像度の3次元カラー情報も取得されているため、容易に任意視点から画像が生成できる。

5. まとめ

本稿では、我々が開発した全方向ステレオシステム(SOS)から得られた3次元情報を用いて、空間内におけるイベントの抽出を行った。SOSは、



a) 上面透視画像 b) 鳥瞰透視画像

図8 複数人物の抽出例

観測点を中心とした全方向のカラー画像とステレオ画像対がリアルタイムで得られるため、水平だけでなく上下も含めた空間内の複数のイベントを同時に取得可能である。

本稿では、SOS から得られたステレオ画像対から視差画像を取得し、背景の視差画像と差分を取ることによりイベントの抽出を行った。本手法は、カラー画像を用いたイベント抽出より環境に左右されにくい。

今後、全ての方向視野が取得可能な SOS の特性を生かし、複数のイベントの追跡とカラー画像と3次元情報を用いることにより安定なイベント抽出をするシステムの構築を図る予定である。

参考文献

- (1) 山本和彦：“知的センシングに基づく実環境情報処理技術とその応用”，2000年度電気関係学会東海支部連合大会予稿集，pp.S-61—S-62 (2000)
- (2) K. B. Saraclik：“Characterizing an Indoor Environment with a Mobile Robot and Uncalibrated Stereo”，Proc. IEEE International Conference of Robotics and Automation，pp. 984-989 (1989)
- (3) H. Ishiguro and S. Tsuji：“Omni-directional stereo”，IEEE Trans. Pattern Analysis and Machine Intelligence，Vol. 14, No. 2, pp.257-262 (1992)
- (4) 石黒浩，山本雅史，辻三郎：“全方向視野の距離情報獲得”，信学論 (D-II)，Vol.J74-D-II, No.4, pp.500-508 (1991)
- (5) Y.Yagi：“Omnidirectional Sensing and Its Applications”，IEICE Trans. Information and Systems，Vol.E82-D, No.3, pp.568-579 (1999)
- (6) 西沢慶満，八木康史，谷内田正彦：“全方向視覚センサ COPIS を搭載した移動ロボットのための環境マップの生成と移動自由空間の推定”，日本ロボット学会誌，Vol.11, No.6, pp.868-874 (1993)
- (7) 山澤一誠，八木康史，谷内田正彦：“移動ロボットナビゲーションのための全方向視覚センサ HyperOmniVision の提案”，信学論 (D-II)，Vol.J79-D-II, No.5, pp.698-707 (1996)
- (8) Y. Onoe, N. Yokoya, K. Yamazawa and H. Takemura：“Visual Surveillance and Monitoring System Using an Omnidirectional Video Camera”，Proc. 14th ICPR, pp.588-592 (1998)
- (9) 佐川立昌，石黒浩，石田亨：“複数の全方向視覚センサを用いた人間の行動の実時間追跡”，第4回 SSII，pp.179-184 (1998)
- (10) 茶園篤，山澤一誠，横矢直和，竹村治雄：“全方向ステレオ視による3次元情報の獲得”，映像情報メディア学会誌，Vol.51, No.10, pp.1745-1753 (1997)
- (11) 川西隆仁，山澤一誠，岩佐英彦，竹村治雄，横矢直和：“六角錘ミラーを用いた全方向画像センサによる全周ステレオパノラマ動画の生成”，MIRU'98，Vol.I, pp.177-182 (1998)
- (12) S.Vedula, P.Rander, H.Saito, and T.Kanade：“Modeling, Combining, and Rendering Dynamic Real-World Events from Image Sequences”，Proc.4th VSMM, Vol.1, pp.326-332 (1998)
- (13) H.Saito, S.Baba, M.Kimura, S.Vedula, and T.Kanade：“Appearance-Based Virtual View Generation of Temporally-Varying Events from Multi-Camera Images in the 3D Room”，Proc.2nd 3DIM (1999)
- (14) H. Nagahara, Y. Yagi and M. Yachida：“Super-resolution From An Omnidirectional Image Sequence”，IEEE Conf. IECON2000, pp.2559-2564 (2000)
- (15) 山澤一誠，市川徹，寺沢征彦，竹村治雄，横矢直和：“全方向ビデオカメラを用いた行動記録システム”，「視覚情報メディアのためのパターン認識・理解」平成12年度報告書，pp.105-114 (2000)
- (16) 棚橋英樹，山本和彦，桑島茂純，丹羽義典：“全方向ステレオシステムの開発”，第6回 SSII，pp.145-150 (2000)
- (17) 桑島茂純，桑原裕之，山本和彦，棚橋英樹，

丹羽義典：“全方向ステレオカメラシステム”，O plus E, Vol.22, No.12, pp.1585—1591 (2000)

(18) 山本和彦, 桑島茂純, 棚橋英樹：“全方向ステレオ画像撮影装置及びステレオ画像撮影装置”，特願 2000-098812 (2000)

(19) M. Okutomi, T. Kanede, “A multi-baseline stereo”, *IEEE Trans. PAMI*, Vol.15, No.4, pp.353-363, 1993

(20) 桑島茂純：“世界初のPCベース・フルカラーステレオ・ビジョン・システム「COLOR TRICLOPS」”，映像情報, Vol.31, No.7, pp.17-22 (1999)

(21) Point Gray Research Inc. : “User's Guide and Command Reference · TRICLOPS Stereo Vision Software Development Kit(SDK) Version 2.2.2 - “

(22) 棚橋英樹, 島田大輔, 山本和彦, 丹羽義典：“全方向ステレオシステム (SOS) を用いた移動体追跡”，View2000, pp.67-72 (2000)