

## 科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 5 月 24 日現在

機関番号：12102

研究種目：挑戦的萌芽研究

研究期間：2011～2012

課題番号：23650081

研究課題名（和文） シミュレーション駆動パターン認識に基づくステレオ視覚の実現

研究課題名（英文） Study of stereo vision based on simulation driven pattern recognition

## 研究代表者

福井 和広 (Fukui Kazuhiro)

筑波大学・システム情報系・教授

研究者番号：40375423

研究成果の概要（和文）：CG 技術により生成した視差ラベル付きステレオ左右画像を学習データとして視差クラス識別器を構成することで、従来の左右対応付けに基づくステレオ視覚と同程度以上の性能を有するステレオ視覚を実現した。また生成したステレオ左右画像セットを国内外の研究者が自由に使えるように New Tsukuba CG Stereo Dataset としてインターネット上に公開した。

研究成果の概要（英文）：We demonstrated that a stereo vision system can be constructed with multiple static pattern classifiers of disparity, which are learnt by a lot of synthesized stereo image pairs. The dataset of the synthesized stereo image pairs has been opened on the Internet, so that many researches in the world can utilize the dataset to advance their researches.

## 交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	2,700,000	810,000	3,510,000

## 研究分野：画像認識

科研費の分科・細目：情報学，知覚情報処理・知能ロボティクス

キーワード：ステレオ視覚，シミュレーション，パターン認識，CG 画像，New Tsukuba CG Stereo Dataset

## 1. 研究開始当初の背景

ステレオ視覚はロボット視覚の基本技術として現在でも精力的に研究が進められているが、ステレオマッチングに基づく基本原理は提案当初から変わっておらず、左右画像間で画素あるいは局所領域の対応付けが不可欠である。良く知られているように、対応付けは、隠れ、左右画像の輝度の違い、キャリブレーション時からのカメラ配置のズレ等の影響を受け易く、計算コストも大きかった。しかしながら、ロボットが最終的に必要な情報は、移動時に障害物に衝突するか否かを反映した、自身に最も近い点の情報（最大視差）であることを考えると、従来ステレオ視覚では、この最近点以外の取得も試みてお

り、必要以上に難しい問題に取り組んでいたと言える。このような背景から、左右画像の対応付けに基づく視差検出の幾何学的な問題を、ステレオパターンに対して最大視差クラスを出力するパターン識別器の学習問題に置き換えて、ステレオ視覚を構築することを試みる。

## 2. 研究の目的

本研究では、左右画像間の対応付けに基づく従来のステレオ視覚とはまったく異なるアプローチ，“シミュレーション駆動パターン認識”に基づくステレオ視覚の実現可能性を検証する。提案法では、ステレオ左右画像の局所領域を連結したパターンに対して、そ

の領域内の最大視差クラスを出力する識別器を統計学習し、これを多数用意することでステレオ視覚を実現する(図1).

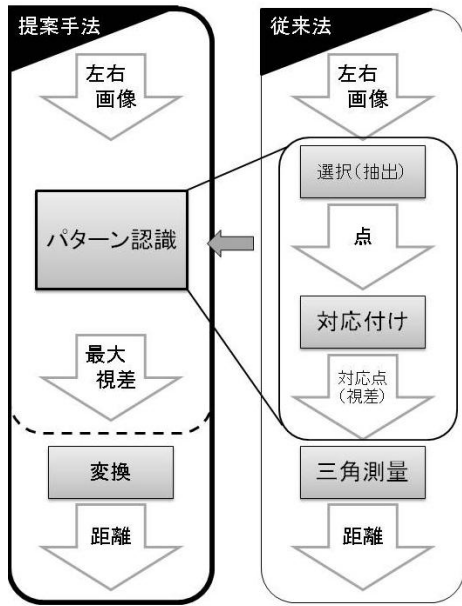


図 1



図 2

### 3. 研究の方法

#### (1) 視差ラベル付ステレオ画像の生成:

シミュレーション駆動パターン認識に基づくステレオ視覚を実現するためには、様々なシーンおよび撮影条件(照明条件、カメラパラメータ)における多様かつ大量の視差ラベル付きのステレオ左右画像ペアが必要となる。これらを実画像シーンから収集することは、多大な労力や多様性の確保の点で極めて困難である。本研究ではコンピュータグラフィック技術に基づいて精巧な3次元シーンを構築し、これを用いたシミュレーションにより、図2のような視差ラベル付きのステレオ左右画像ペアを大量に生成する。後で述べるように、提案法の有効性の検証は、ステレオ視覚の研究分野において標準的に使われている公開データベース Tsukuba stereo dataset (Head and Lamp) を用いて行う。そこで、この Head and Lamp シーンを忠実に再現した3次元シーン(図3)を作成し、カメラ位置・視点を動的に変化させながら多様なステレオ左右動画像を生成する。このステレオ左右動画像セットを New Tsukuba CG Stereo Dataset と名付ける。

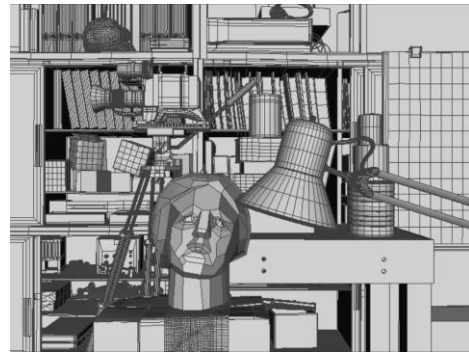


図 3

#### (2) 識別特徴の抽出:

左右ステレオ画像ペアに対して統計的パターン認識を直接に適用することも可能であるが、より高い識別性能を得るためには、学習画像から視差クラス識別に有効な特徴を抽出する必要がある。本研究では、従来ステレオ視覚において有効性が確認されている AD-Census コストと呼ばれる特徴量を用いる。画素  $(u,v)$  の視差  $d$  に対する AD-Census コスト  $c(u,v;d)$  は以下の式で定義される。

$$c(u,v;d) = f(C_{census}(u,v;d), \alpha_c) + f(C_{AD}(u,v;d), \alpha_A)$$

ここで、 $C_{census}()$  と  $C_{AD}()$  は、それぞれ局所パターン構造の不一致度と濃淡情報の差異を表している。また  $f(c, \alpha) = 1 - \exp(-c/\alpha)$  である。コストは画素レベルではなく、領域レベルで行った方が安定に求まるので、小領域  $(9 \times 7)$  ピクセルを設定して、この領域の全画素に対して上記コストを求める。図4は AD-Census コストマップを求める流れを示している。まず左右画像が与えられると、視差クラスに対応する2つの小領域  $(9 \times 7)$  ピクセルを左右画像に設定し、両者から上記 AD-Census コストマップ  $(9 \times 7)$  ピクセルを計算する。このようにして、最大視差  $D$  個分の AD-Census コストマップが求まる。これらを連結したパターン  $(9 \times D \times 7)$  ピクセルを学習のための Disparity pattern パターンとする。本研究では  $D=16$  とする。

New Tsukuba CG Stereo Dataset の最初の 261 フレームはオリジナルの Tsukuba Stereo Dataset と同じ視点となっているために、評価ではこれらのフレーム画像を除いた残りの 1700 フレームからランダムに 80,000 の左右画像ペアを抽出し、これらから視差ラベル付きの AD-Census コストマップを生成する。

#### (3) 識別器の学習:

左右画像ペアから求めた AD-Census コストマップをベクトル化する。次にこれらの学

習ベクトルを用いて、左右画像ペアをD個のクラスに分類する識別器を構築する。図4にこの大まかな流れを示す。まず先に述べたように入力された左右画像からAD-Census、さらにはDisparity mapを生成する。次に重判別分析により特徴抽出のための15次元(=16-1)の判別空間を生成し、この判別空間へ射影によりデータの次元削減を行う。識別では、各視差クラスの平均を参照ベクトルとする最近傍法を用いる。以上の流れは各小領域毎に視差を求める手順であるが、画像全体に対しては小領域を画像全体に渡ってスライドさせながら上記処理を行う。

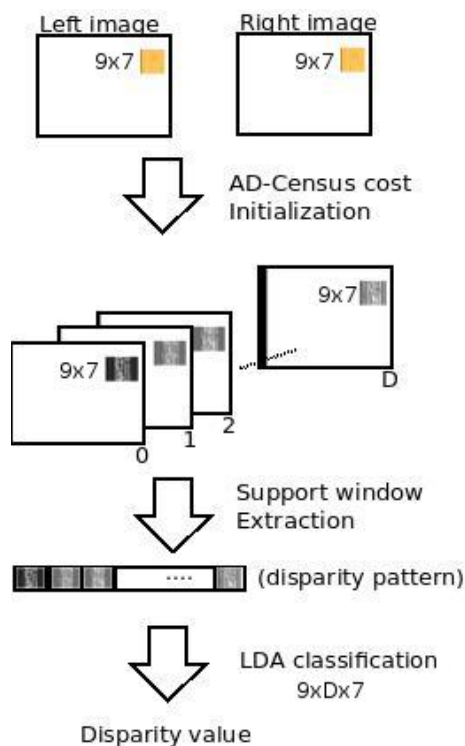


図 4

(4) 提案法の有効性の検証：

提案法が視差をどの程度の精度で検出できるかを評価する。このために、先に述べた、実画像の左右画像と正解視差マップで構成される Tsukuba stereo dataset (Head and Lamp) に対して、従来法と提案法を適用して、得られる両者の視差マップについて比較・考察する。

(5) 視差信頼度マップの生成：

隠れがある場合には、左右画像の対応付けが出来ないために、正しい視差が求まらない。そこで得られた視差に対する信頼度マップを生成する。信頼度マップは識別器が最大類似度を与えた視差クラス  $C_1$  とそれ以外の視差クラス  $C_n$  の関係を調べることで推定できる。小領域内に隠れが無い場合、視差の変化

は緩やかと考えられ、2番目、3番目に高い類似度を与える視差クラスは、視差クラス  $C_1$  に近い  $C_1 \pm T$  ( $T=1 \sim 3$ ) になると予測される。逆にこのようにならない場合には、隠れが生じている可能性が高いと考える。

4. 研究成果

本研究で得られた研究成果は以下の通りである。

(1) 提案するステレオ視覚を構成する識別器を学習するために、ステレオ視覚の研究分野では標準的に使われている公開データベース Tsukuba Stereo dataset (Head and Lamp) を3次元シーンとして忠実に再現した。図5の左列はステレオ右画像、右列はそれに対応する視差マップを示している。図6は Head and Lamp と同じシーンの異なる4種類の照明条件下におけるシミュレーション画像を示している。この New Tsukuba CG Stereo Dataset を用いて大量かつ多様な左右ステレオ画像ペアを生成した。



図 5



図 6



(2) New Tsukuba CG Stereo Dataset を用いて、提案するシミュレーション駆動パターン認識に基づくステレオ視覚を学習した。Tsukuba Stereo Dataset を用いて、提案法と従来法の視差マップを比較した。その結果、表 1 に示すように、提案法は従来法と同等以上の性能を実現できることを確認した。図 7 上右は正解視差マップ、下左は従来法の視差マップ、下右は提案法の視差マップを示している。特にテキストが少ない領域(緑の枠内)では不安定な従来法(左枠)に対して、提案法(右枠)では安定した視差が得られた。

表 1

Method	隠れなし	全体	不連続領域
従来法	10.7%	12.63%	<b>15.35%</b>
提案法	<b>9.16%</b>	<b>10.91%</b>	21.91%

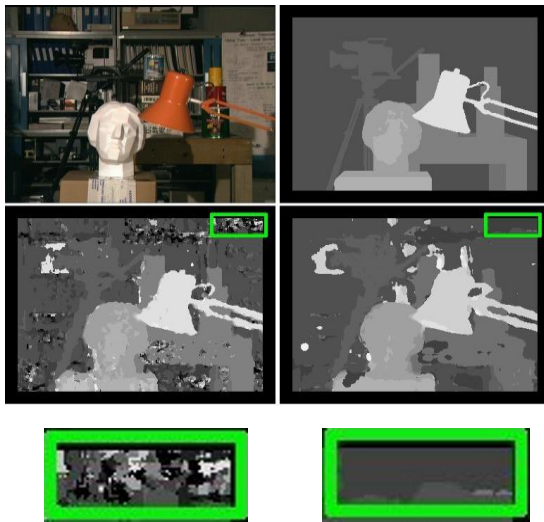


図 7

(3) 学習データ数を変化させて性能評価を行った。その結果、400 程度の学習データ数で従来法と同程度になり、1000 個の学習データでほぼピーク性能に達することが確認できた(図 8)。

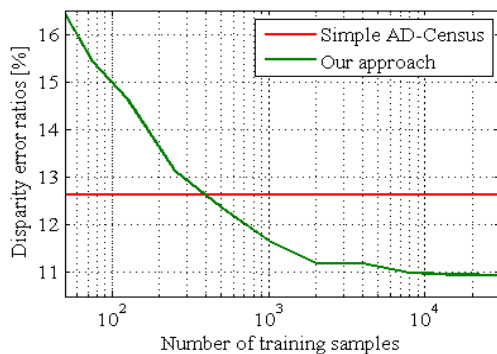


図 8

(4) 識別器が出力する各視差クラスの類似度セットから、図 9 に示すような視差信頼度マップを生成することができた。



図 9

(5) 研究成果をパターン認識・コンピュータビジョン分野のトップコンファレンスの一つである International Conference on Pattern Recognition (ICPR) 2013, および ICPR に併設されたワークショップ等にて発表し、大いに注目を集めた。

(6) New Tsukuba CG Stereo Dataset をインターネット上で一般公開した。5 月中旬の時点で多数の世界的な有力研究機関からのダウンロードがあり、その数は約 300 に達している。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- ① Martin Peris Martorell, Atsuto Maki, Sara Martull, Yasuhiro Ohkawa, Kazuhiro Fukui, "Towards a Simulation Driven Stereo Vision System", International Conference on Pattern Recognition (ICPR) 2012, pp. 1038-1042, 2012, 査読有。
- ② Sarah Martull, Martin Peris Martorell, Kazuhiro Fukui, "Realistic CG Stereo Image Dataset with Ground Truth Disparity Maps", ICPR2012 workshop TrakMark2012, pp. 40-42, 2012, 査読有
- ③ Sarah Martull, Martin Peris, Kazuhiro Fukui, "Realistic CG Stereo Image Dataset With Ground Truth Disparity Maps", 電子情報通信学会 PRMU 研究会, vol. 111, No. 430, pp. 117-118, 2012, 査読無。

[学会発表] (計 3 件)

- ① Martin Peris Martorell, Atsuto Maki,

Sara Martull, Yasuhiro Ohkawa, Kazuhiro Fukui, "Towards a Simulation Driven Stereo Vision System", International Conference on Pattern Recognition (ICPR) 2012年11月13日, つくば国際会議場 (茨城県)

② Sarah Martull, Martin Peris Martorell, Kazuhiro Fukui, "Realistic CG Stereo Image Dataset with Ground Truth Disparity Maps", ICPR2012 workshop TrakMark2012, 2012年11月11日, つくば国際会議場 (茨城県)

③ Sarah Martull, Martin Peris, Kazuhiro Fukui, "Realistic CG Stereo Image Dataset With Ground Truth Disparity Maps", 電子情報通信学会 PRMU 研究会, 2012年2月9日. 東北大学(宮城県)

[その他]

New Tsukuba Stereo Dataset ホームページ  
<http://www.cvlab.cs.tsukuba.ac.jp/dataset/tsukubastereo.php>

本研究課題で開発した CG ステレオ画像データセット (New Tsukuba CG Stereo Dataset) のダウンロードサイトである.

## 6. 研究組織

### (1) 研究代表者

福井 和広 (Fukui Kazuhiro)  
筑波大学・システム情報系・教授  
研究者番号：40375423