実空間強化とヒューマン・ロボットインタフェースの研究
— 対話型レンジファインダの構成 —

山下 雅也* 佐藤 信 坂根 茂幸
中央大学大学院理工学研究科

Augmented reality and human-robot interface
– Implementation of an interactive range finder –

Masaya YAMASHITA* Shin SATO Shigeyuki SAKANE
Chuo University Graduate School of Science and Engineering

Abstract - A prototype system of human-robot interface using projector-based augmented reality is under development. The paper describes a subsystem which allows a conbination of interactive and active range finders to extract 3D information of the robot task environment.

Key Words: Augmented reality, human-robot interface, DigitalDesk, tracking vision, range finder

1 はじめに

AR(augmented reality) はコンピュータの生成する仮想的な世界と実世界とを融合し、その実空間で人間にコンピュータとのインタラクションが可能なシステムを構成する。 Wellner らが構成した DigitalDesk [1] もその試みの一つであり、机の上にさまざまな記号やパターンをプロジェクタで投影し、それに対する人間の手の動作認識というサイクルを繰り返して、柔軟なインタラクションが可能となることを示した。そこには、コンピュータの画面の前に座ってキーボードやマウスで入出力する従来の GUI とは異なる、実世界指向のインタフェースの可能性を見ることができる。

プロジェクタを用いる AR には、HMD 等の特別な機器を身体に装着することが不要で、また、複数の人間が融合環境を共有できるという利点がある。筆者らは、この情報投影とそれに対する人間の行動認識を基礎とする枠組みを参考に、人間とロボットとの協調作業における自然な相互インタラクションを構成することを目指したインタフェースシステムの開発を行なっている。これまでに、(1) 仮想操作パネルを用いた情報入力、実画像の投影による作業教示の機能 [2]、さらに (2) 実空間に直接マークを投影する指さしポインタ機能 [3] を構成してきた。(1) は机の上に投影した仮想のメニューボタンなどを利用して情報入力を実現した。(2) では人間の指示した場所をシステムが推定し実空間にマークを投影する。これにより、オペレータは HMD やモニタディスプレイを見ることなく、システムの推定状態を監視できる。

ロボットの作業教示では、システムが作業環境の 3 次元情報を入力する機能を持つとそれに基づく環境理解が可能となり、オペレータの負担を軽減し、より効果的な作業教示が可能になる。本報告では、このヒューマン・ロボットインタフェースの一環として構成している対話型レンジファインダ機能 [4] について述べる。

2 従来の研究

従来、ロボットの作業環境の 3 次元データを入力するために様々な手法、システムが開発されてきた。その一つの例として投光型レンジファインダがある。投光型レンジファインダは通常、スポット光、スリット光、空間コード化したパターンなどの特徴図形を投影し、3 角測量により計測する。しかし、従来は特殊な光学装置が必要であった。たとえば、スリット光レンジファインダではレーザー光をシリンドリカルレンズを通してスリット状に変換し、ミラーを振って対象物に走査する。ミラーの偏光角度から投影情報を計算し、これとカメラに移る像の位置から距離を測定する [5]。このシステムでは、スリット光を自由に変化させることはできない。そのため、対象となる全空間に対してスリット光を投影するため不要な情報まで読みとる可能性や、対応点問題の誤対応を起こす可能性もある。

3 レンジファインダ機能の構成

3.1 対話型レンジファインダ

対話型レンジファインダにおける「対話」には、オペレータとレンジファインダ間の対話、および、複数のレンジファインダ同士の対話、の 2 重の意味を持たせている。例えば、前者は、環境モデルを対話的にオペレータが実環境下で作成する環境モデリング、また、後者は、作業環境の状況に応じてレンジファインダ同士が協調して働き 3 次元データを取得するような役割が挙げられる。本論文では、後者の場合について構成したシステムについて述べる。

本システムでは、プロジェクタを用いて情報投影を行なうので、その機能を拡張すれば各種の光パターンの投影に

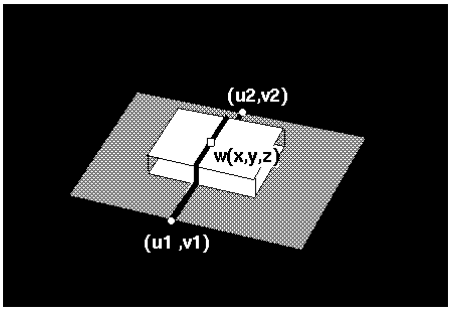


Fig.1: Camera image of using light-striped range finder

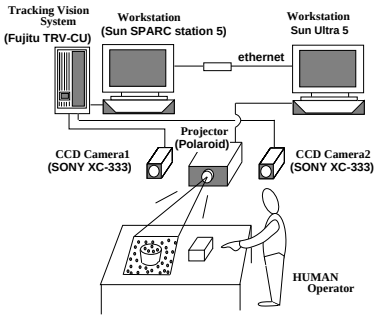


Fig.2: Hardware of the randomdot stereo range finder

基づくアクティブレンジファインダの構成が可能になる。例えば、（1）点図形の投影によるスポット光レンジファインダ、（2）多数点の投影によるランダムドットステレオ[6]、（3）線図形の投影によるスリット光投影レンジファインダなどのレンジファインダが構成可能である。従って、オペレータは多様なレンジファインダ方式を用いて、対象物の形状や状況に応じて選択するシステム化（仮想操作パネルの利用）、そして、指さしポインタ機能と結合した部分的な3次元情報の入力システム化への拡張が可能となる。

本報告では、上記の (2) と (3) のレンジファインダが協調して働くシステムの構成について述べる。具体的には上記の方式 (2) のランダムドットステレオにより、ラフな物体形状を抽出する。その結果に基づいて (3) のスリット光投影レンジファインダで適切な位置／方向にスリット光を投影し、より詳細な形状情報を求める。

3.2 ランダムドットの投影によるレンジファインダ

ステレオ視における重要な課題は左右両眼の対応点探索問題であるが、ランダムドットパターンの投影により誤対応を減らすことができる。対象物の測定面が一様で画像特徴のない面であっても、その面に模様をつけることに相当するのでエッジ画像に基づくステレオ法と比較して多くの対応点の情報が得られる利点がある。相関演算を用いるレンジファインダ方式なので、トラッキングビジョン装置 [7] との相性が良い。

3.3 局所領域におけるスリット光レンジファインダ

近年のプロジェクション機器の性能の向上により、環境に対する複雑なイメージの比較的高輝度での投影が可能となった。このプロジェクタを用いることにより、従来の投影パターンの異なるレンジファインダ方式が可能となる。そこで、本システムでは、スリット光の投影（従って計測）範囲の局所化と、スリット光の投影角度が可変なレンジファインダ方式の構成を試みた。

図1はプロジェクタ画像面上の座標値 $(u_1, v_1), (u_2, v_2)$ を結ぶ線分を投影スリット光とし、対象物上の座標値 $w = (X, Y, Z)^T$ にスリット光に投影したイメージ場面である。初めに

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} u_2 - u_1 \\ v_2 - v_1 \\ u_1 v_2 - u_2 v_1 \end{pmatrix} \tag{1}$$

とする。ここでプロジェクタ画像面上における直線の式は

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + t \begin{pmatrix} a \\ b \end{pmatrix} \tag{2}$$

となり、射影カメラモデルの関係からプロジェクタの射影カメラ行列を P_p とするとさらに

$$\lambda \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{3}$$

となる。ここで (2)(3) より λ, t を消去して

$$m^T R w = -m^T T \tag{4}$$

となる。ただし、

$$m = \begin{pmatrix} -b \\ a \\ c \end{pmatrix}, R = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}, T = \begin{pmatrix} p_{03} \\ p_{13} \\ p_{23} \end{pmatrix} \tag{5}$$

である。以上の式とカメラの撮影したスリット光の位置座標、そしてカメラの射影カメラの校正データから投影されたスリット像の3次元座標が得られる。

4 システム構成と実験

4.1 システムの構成

本システムのハードウェア構成を図2に示す。描画用コンピュータ (Sun Ultra 5) でランダムドットのパターンを X-window を用いて描画し、接続した液晶プロジェクタ (POLAROID, COLORVIEW LIGHT) により実空間に投影する。2台の CCD カメラ (SONY XC-333) でこのパターンを投影した環境の画像を入力し、トラッキングビジョン (Fujitsu TRV-CU) により相関演算を行なっている。上位のプログラムは EusLisp[8] を使用し、2値画像処理には画像処理ソフトウェアパッケージ SPIDER[9] を利用した。

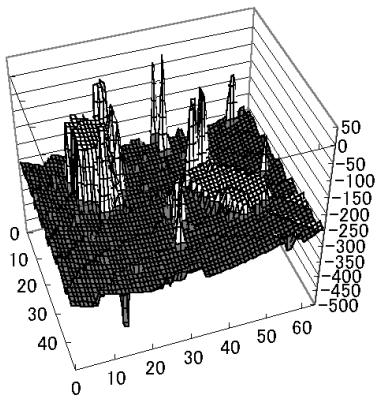


Fig.3: An example of 3D data extracted by the range finder

4.2 ランダムドットの投影による三次元情報の推定

ランダムドットを投影したシーンに対して、ステレオ視により 3 次元計測を行なう。右カメラ画像 (640×480pixel) において、32×32 pixel を 1 ブロックとする領域を画面縦横に 8pixel ずつ移動させたものに区分けする。横 69× 縦 49 の合計 3381 ブロックを各参照画像とし、ブロックサイズ 8×8、間引き率 4 で左カメラ画像との相関演算を行ない、対応する画像上の位置を求める。両カメラの対応点の位置と、予めキャリブレーションにより求めたステレオの変換行列に基づいて 3 次元位置を計算する。相関の精度はランダムドットの径の大きさと分布密度に依存するので、誤対応の発生頻度が少ないドットの径と密度の組合せを調べ、ドットの径を 10pixel(実空間で 12mm 相当)、ドット数を 2500 個とした。投影したドットパターンの実空間での大きさは約 920mm ×700mm の領域である。

図 3はこの方法で求めた作業環境の 3 次元データの例を示している。垂直方向の軸は Z 座標値である。水平方向の 2 軸は X,Y 座標値ではなく、測定データを格納する配列 (69×49) のインデックスの値そのままで示している。全画面にわたる相関演算と対応点の 3 次元座標の計算に要した時間は約 1.33 sec であった。

4.3 ランダムドットの認識結果の投影

図 4は上記のシステムにより抽出して物体上にラベル A,B, C を実空間に投影した場面を示している。具体的な処理は、ランダムドットレンジファインダシステムにより 3 次元位置を計測したデータ配列 (69×49) の高さ方向 -250mm 以上 200mm 未満の計測値を物体の上面と仮定し、2 値化する。この出力データに Edge Preserving Smoothing 処理 [10] を 2 回行ない、その抽出領域 (連結成分) にラベルをつける。その上で領域の重心となる配列要素の 3 次元座標値にプロジェクタから物体識別のラベルを投影している。

Fig.4: Projection of labels to the detected objects

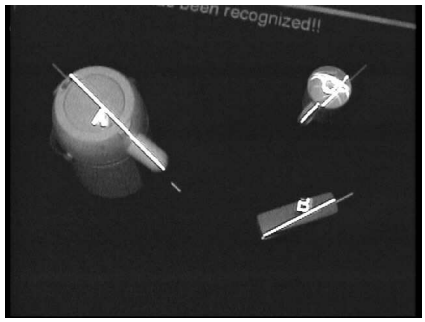


Fig.5: Projection of principal axis to the objects

4.4 スリット光レンジファインダの切替え

上で求めた連結成分の主軸方向に向いた外接長方形の頂点の絶対座標値 2 点をプロジェクタ画像面上に変換する。この 2 点を結ぶ線分上をスリット像の中点が移動するように、さらに、画像面上 2 点間の水平、垂直方向のベクトル成分のうち大きい値をスリット線分の長さとして設定する。これにより投影されたカメラ画像面上のスリット像の位置を検出し、3.3 節で述べた式からその絶対座標値を算出する。なお、スリット像の位置の検出方法は事前に撮影したスリット像 (8 × 8 pixel 縦、横、斜め 2 枚の計 4 枚) を参照画像とし、横方向について相関演算をし、最適となった部位を検出する。これを 8pixel ずつ上下にずらして全画面を探索した。

図 5はオブジェクトに対して推定した主軸方向を、プロジェクタで実物体上に投影した様子を示している。

図 6はスリットを画像内で探索し、特定したカメラ画像上の位置をマス目状に描画している様子である。特に (c) ではコップの取っ手の方向と垂直にスリットを描画しているため、細長い取っ手部分の検出可能性を高めている。また、ランダムドットレンジファインダでは上面が平面状でない物体を検出することが困難であったが、(c) においてコップの縁部分も検出されていることが示されている。

4.5 考察

スリット光投影レンジファインダの機能によりシステムが物体にラベルづけを行ない、そのラベルを実物体上に投影することにより、オペレータの作業教示を支援することが可能になる。例えば図 4の状態において、これによりオペレータがこのラベルを見て、「pick A」のように対象物を指定することが可能になる。さらに、切り替えたスリット光レンジファインダにより、ランダムドットレンジファインダでは判別の難しいコップの縁のような特徴的な部分を判別できる。コップの取っ手部分もスリット光が垂直に近くなるように走査できるため、従来の 1 直線方向のみのスリット光レンジファインダに比べ、縁部分の検出可能性を高めることが可能になった。

5 むすび

拡張 DigitalDesk を用いた柔軟なヒューマン・ロボットインタフェースシステムの一環として構成している、対話型レンジファインダの機能について述べた。ここでは、プロジェクタとトラッキングビジョンというハードウェア構成をそのまま光投影方式のレンジファインダに用いることができる。本報告では、ランダムドットレンジファインダとスリット光投影レンジファインダの協調システムの構成について述べた。

今後は、本機能で得られた 3 次元情報とこれまでに開発した諸機能を用いる作業教示へ発展させることが課題である。

参 考 文 献

1) P.Wellner, Interacting with paper on the DigitalDesk, Communications of the ACM, Vol.36, No.7, pp.87-96, (1993)

2) 寺島, 坂根: 拡張ディジタルデスクを用いるヒューマン・ロボットインタフェース, 日本ロボット学会誌, Vol.16, No.8, pp.1091-1098, (1998)

3) 佐藤, 寺島, 坂根: 実空間強化とヒューマンロボットインタフェース - 指さしポインタの構成 -, 第 16 回日本ロボット学会学術講演会, pp.23-24, (1998)

4) 山下, 佐藤, 坂根: 実空間強化とヒューマン・ロボットインタフェース - 対話型レンジファインダ機能の構成 -, 第 17 回日本ロボット学会学術講演会, pp.419-420, (1999)

5) 荒木、田中、藤野、佐藤: 高速 3 次元形状計測装置の試作, 電子情報通信学会論文誌 D,Vol. J71-D, No.10, pp.2059-2068,(1988)

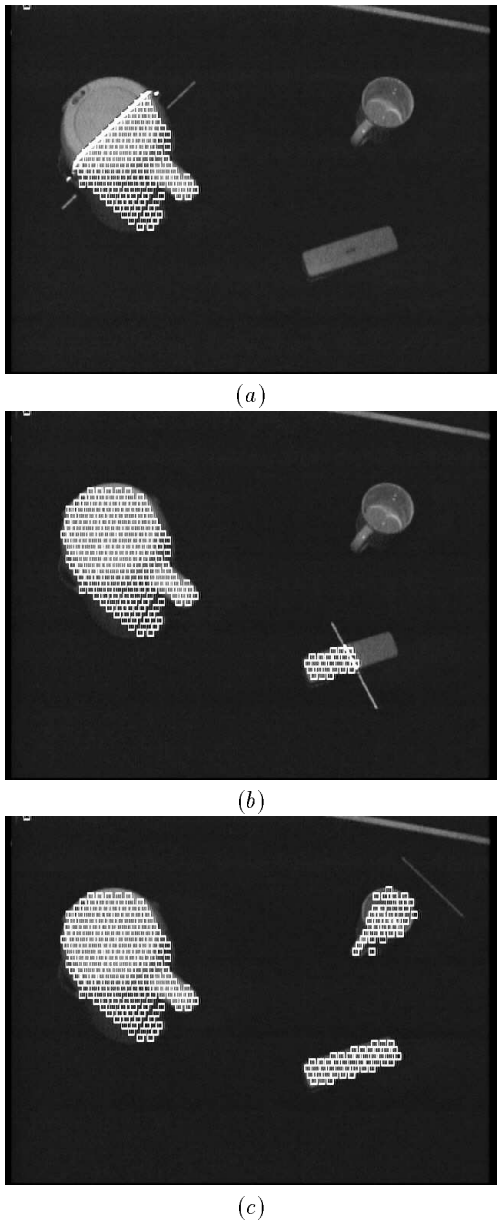
6) 橋本他: ランダムドットパターン投光ステレオを用いた物流ロボット視覚システムの開発、日本ロボット学会誌、 Vol.17, No.1, pp.48-49,(1999)

7) 森田、沢崎、内山、佐藤: カラートラッキングビジョン, 第 14 回日本ロボット学会学術講演会, pp.279-280, (1996)

8) 松井: 幾何モデリング機能を備えたマルチスレッド並列オブジェクト指向言語 Euslisp, 日本ロボット学会誌, Vol.14, No.5, pp.650-654, (1996)

9) 情報処理振興事業協会: 画像処理サブルーチン・パッケージ SPIDER-II, 情報処理振興事業協会 協同システム開発株式会社, (1986)

10) M.Nagao, T.Matsuyama:Edge preserving smoothing, CGIP, vol.9, No4, pp.394-407,(1979)



(a) (b) (c)

Fig.6: Experiments of 3D measurements using the lighted-striped range finder