ビューベーストアプローチに基づく 移動ロボットナビゲーション

松本吉央^{*1} 稲葉雅幸^{*2} 井上博允^{*3}

View-Based Approach to Robot Navigation

Yoshio Matsumoto^{*1}, Masayuki Inaba^{*2} and Hirochika Inoue^{*3}

Recently, view-based or appearance-based approaches have been attracting the interests of computer vision research. Based on a similar idea, we have proposed a view-based navigation method using a model of the route called the "View Sequence." It contains a sequence of frontal views along a route memorized in the recording run, and the recognition of the environment is realized based on the matching of the current view and memorized view sequence. In this paper, we discuss the required characteristics of the view for the view sequence, and evaluate our former method of generating views. Then we confirm that the stereo disparity satisfies the requirements of the view sequence through an experiment, and the disparity view sequence is applied for outdoor navigation. The experimental results indicate such views other than normal camera images can be utilized for our view-based navigation method.

Key Words: Robot Navigation, Robot Vision, View-Based Approach

1. はじめに

移動ロボットの研究において、環境を認識しロボットを誘導することは最も基本的な課題である。最近の多くの移動ロボット研究ではセンサとして視覚を用いて外界のセンシングを行っており、そのロボットの環境認識の方法は

- シーンの中の特定的特徴（ランドマーク）を用いる、モデルベーストアプローチ
- シーン全体の見え方を用いる、ビューベーストアプローチに大別できる。後者は近年注目されているアプローチであり、我々がこれまでに提案した“ビューシーケンス”[1]も後者に分類される。

本論文では、まずはじめに“ビューベーストアプローチ”についてその特徴を簡単に述べる。次に、ビューベーストアプローチに基づく移動ロボットのための経路表現手法としてこれまでに我々が提案した“ビューシーケンス”についてその特徴を述べ、用いる画像（ビュー）が満たすべき性質について明らかにする。さらに、我々の提案手法におけるこれまでのビューシーケンスの生成手法について、この観点から考察を行う。最後に、これまでに用いて来た濃淡画像以外にもその性質を満たす画像（具体的には視差画像）があることを実験的に示し、屋外環境で

の経路誘導への応用を試みる[2][3]。視差画像は、環境の三次元的な形状にのみ依存する“見え方”であるため、屋外でも明るさの変化に影響されにくい経路誘導が実現できる。以上、本論文では実験を通して、ビューシーケンスが濃淡画像以外にも適用可能である一般性の高い手法であることを示す。

2. ビューベーストアプローチ

2.1 ビューベーストアプローチによるパターン認識

人工知能の分野では、“記憶に基づく推論（Memory-Based Reasoning）”と呼ばれる手法が研究されている[4][5]。記憶に基づく推論では、いくつかの回答の中から最も適切なものを選択するような問題において、類似した問題の回答は同じになると仮定し、たくさんの事例の記憶を用いて適切な回答を導き出す。“記憶に基づく方法（Memory-Based Approach）”とは、このように認識のために他のアプローチと比較してより直接的に記憶を利用する手法の総称であり、膨大な記憶と単純な処理により現実世界の複雑な問題を解こうというアプローチである。これをパターン認識に適用したものが、“見え方に基づく方法”（View-Based Approach, Appearance-Based Approach, 以後ビューベーストアプローチ）である[6]。

パターン認識は基本的には、あらかじめ蓄えられた物体のモデルと、入力画像から何らかの処理により抽出された特徴との比較・照合により実現される。従来の三次元物体の認識はモデルと入力特徴との照合に幾何学的な特徴を利用する、“幾何学的特徴に基づく方法（Model-Based Approach）”と呼ばれるものが一般的であった。このアプローチでは、まず二次元画像か

原稿受付 2000年11月22日

^{*1} 奈良先端科学技術大学院大学情報科学研究科

^{*2} 東京大学工学系研究科

^{*3} Nara Institute of Science and Technology

^{*2} The University of Tokyo

らエッジやコーナなどの形状特徴を抽出し、これをもとに三次元的形状を抽出する。この三次元形状をあらかじめ用意してある三次元モデルから作られた特徴との間で照合する。このアプローチでは物体の三次元的な回転などの変化に容易に対処できるという点は優れているものの、二次元画像から幾何学的特徴を精度よく抽出することや、そこから三次元形状を復元すること、また三次元形状のモデルを作成することは現実には容易ではない。

これに対して、ビューベースアプローチでは、二次元の見え方画像がそのまま利用される。三次元物体はその取り得る二次元の見え方をあらかじめすべて記憶され、入力画像と比較することで認識される。このアプローチでは、画像をより直接的に利用することが特徴であり、また認識対象のモデルは画像を記憶することで作成できるため、従来の三次元的な形状モデルを用いる認識手法と比較して容易である。したがってこのアプローチは、ノイズが多い実画像から複雑な形状を持つ物体を認識できることが期待できるが、膨大な画像データの記憶が必要となるため、そのままでは記憶量や計算量の観点からみてあまり現実的ではなかった。

しかし、近年の計算機の性能向上は著しく、画像処理専用プロセッサの開発も進んでいる。これらの計算機環境の変化によりビューベースアプローチが現実味を帯びてきており、パターン認識の分野での研究が盛んに行われている。例えば Nayar ら [7] は、物体の見え方（姿勢や光源位置をパラメータとした多数のサンプル画像）を固有空間上の多様体として表現し、三次元物体の認識を効率良く実現する。ビューベースアプローチに基づく手法を提案し注目されている。

2.2 ビューベースアプローチによる位置の認識

ビューベースアプローチにより移動ロボットのための位置の認識を実現する場合、特定のランドマークを用いるのではなく、シーン全体の見え方を用いることになる。手順としては、まずはじめに画像を記憶するための教示を行ったあとで、そこで記憶した画像と現在の画像の比較・対応付け（マッチング）を行う。ここで、画像データは情報量が多いため、そのままでは記憶やマッチングの効率が悪い。そこで、(1) どのような画像を記憶し、(2) どのようなマッチングを行うか、という点がキーとなる。

Horswill [8] が開発した屋内案内ロボット Polly は、あらかじめ環境内の認識したい場所（各部屋の前など）で見える画像をあらかじめ記憶しておき、ロボットが自律移動しているときに見える画像が記憶した画像とのテンプレートマッチングで一致したら、その場所にいと認識した。このマッチングには 16×12 画素という低解像度の画像を用いている。前田ら [9] は、全方位画像をフーリエ変換したデータをマッチングに用いる手法を提案した。ロボットはあらかじめグリッド状の既知の位置において多数の全方位画像を記憶する。フーリエ変換により、全方位画像の回転はマッチングには影響を及ぼさなくなるため、ロボットの姿勢にかかわらず最も近い画像を効率よく探索することができる。用いるデータを低周波数領域に限定することで、画像を直接用いるよりもデータ圧縮をして記憶量や計算量も削減している。また、固有空間法を移動ロボットのシーンの認識

に適用した前田らの研究 [10] では、ロボットが得た周囲の画像は固有空間に投影され、あらかじめ記憶した固有空間上の画像とマッチングが取られる。この研究では一度のマッチングだけでロボットの位置を認識することをせず、一つの位置で複数の方向の観察を行ったり、移動しながら観察を行い、その複数の観察結果を統合するという、アクティブビジョンの枠組の中でビューベースアプローチによる効率のよいシーンの認識を目指しており興味深い。しかし位置認識の精度は記憶した位置から 50[cm] 離れると 50% 程度であり、現状では十分な精度が得られていない。また、小島ら [11] は進行方向のシーンの時系列画像を記憶し、その時系列画像の DP マッチングによりロボットの位置を認識した。この研究では、「時間」という制約をマッチングを用いることによって、安定なマッチングを目指すという方向をとっている。また、西村ら [12]、岩佐ら [13] はそれぞれ全方位画像を利用し、カメラの姿勢（垂直な軸周りの回転）に不変な特徴量を定義し、位置の認識に用いている。しかし、これらの手法では逆にロボットの姿勢を検出できないので、それだけでは移動ロボットの誘導に用いることはできない。つまり、誘導については他の方法を用いることを前提としている。これらのビューベースアプローチによる移動ロボットのシーンの認識の研究に共通する特徴は、特定のランドマークを用いるのではなく、シーン全体の見え方を用いるために、人工的なランドマークを敷設したり、画像中からランドマークを抽出するという処理が必要なくなるという点である。

また、Pomerleau [14]、Meng [15]、Zhu [16]、Crespi [17] はそれぞれ通路の見え方をニューラルネットやメモリーベースで学習することにより経路誘導を行う手法を開発している。しかし、これらの研究では「経路に沿って走行する」ことだけを目的として学習が行われており、ロボットの位置を知ることができない。そのため、これらの学習を用いた手法は、ビューベースアプローチに基づく移動ロボットの位置の認識には含めない。また、Ishiguro ら [18] はパノラミック・ビューによる経路表現を用いた“T-Net”と呼ばれる地図表現を提案しているが、その走行では画像中の局所領域をランドマークとしてトラッキングしており局所的な画像の変化に弱い点、厳密にはビューベースなナビゲーション手法ではない。

3. ビューシーケンスによる経路誘導

3.1 ビューシーケンス

ここまでに述べたビューベースのパターン／位置の認識の研究では、それが記憶したもののどれと一致するか、を認識することが目的であった。これに対して、経路誘導のための位置の認識には以下のような違いがある。

- (1) 実時間性が不可欠である。
- (2) 自己位置だけでなく、障害物検出や走行制御を行う必要がある。
- (3) ロボットの移動は連続的であるので、記憶のすべてを探索対象とする必要はない。

これらを考慮した上で、我々はこれまでに移動ロボットのための経路表現として“ビューシーケンス”を提案した [1]。ビューシーケンスとは、経路に沿ってロボットが進むときに見る環境

Fig. 1 Conceptual image of processing View Sequence

の見え方(ビュー)を, ある間隔で連続的に記憶した画像の列である。ロボットは, まずはじめにオペレータの操縦により指示走行を行い, 経路に沿って得られるビューを自動的に記憶することで, ビューシーケンスによる経路表現を作成する。ロボットが自律走行を行うときには, 記憶したビューシーケンスの中の画像と走行時に得られる画像のマッチングを実時間でを行い, ロボットの現在位置や障害物を認識し, ステアリングを制御する。ビューシーケンスによる経路誘導のイメージを Fig. 1 に示す。本手法は,

- 経路上の自己位置, 進行方向および障害物を認識できる
- 実時間で走行制御を行うことができる
- 一回の記録走行により容易に経路指示が行える

という特徴がある。

記憶したビューと現在のビューなど, 2枚のビューを比較しそれらがどのくらい似ているかを計算する処理を, ビューのマッチングと呼ぶ。これは Fig. 2 に示すように, ビューの中心部の縦長の領域をテンプレートとしたブロックマッチングで, マッチングエラー e は以下の式で定義される。

$$e(u) = \sum_{x=s}^{w-s} \sum_{y=0}^h |I_1(x, y) - I_2(x+u, y)| \quad (1)$$

$$e = \{\min e(u) \mid -s \leq u < s\} \quad (2)$$

ただし w, h はビューの縦横のサイズ, $I_1(x, y), I_2(x, y)$ は二つのビュー I_1, I_2 における (i, j) の位置の画素の輝度値, s は水平方向の探索範囲である。このときテンプレートのサイズは $(w-2 \cdot s) \times h$ となる。ここでは前述の通り $w = 32, h = 32, s = 8$ としたため, テンプレートのサイズは 16×32 また探索範囲は $-8 \leq u < 8$ となる。相関演算 LSI を搭載した画像処理ボード [19] ではこの処理をハードウェアで高速に実行することができ, その結果として, マッチングエラー e (ブロック間誤差, この値が小さいほど相関が高い) およびそのときの水平方向の変位 u が得られる。テンプレートが縦長であるのは, ロボットの回転や並進によりビューの変位が横方向には大きくなる可能性があり探索範囲が必要であるのに対して, ロボットの移動が平面上であるので縦方向には(理想的には)発生せず, 縦方向には探索範囲が必要ないからである。

なお, ビューシーケンスにおいてはロボットの行動として,

Fig. 2 Matching method between two view images

Fig. 3 Matching error of views in image space

- 前進
- コーナーリング

の二つを用意し, その繰り返しによりロボット誘導を行う [1]. このうちのコーナーリングにおいては記憶したビュー 1 枚のみと現在の画像のマッチングを取ることで行動が制御される。そのためビューシーケンスにおいて特徴的なビューの遷移というものは起こらない。そこで, 以下では前進行動のみを対象としてビューシーケンスの特性について議論する。

3.2 画像空間上でのビューシーケンス

画像を n 次元の画像空間 (n は画素数) 上のベクトル $X = (x_1, x_2, \dots, x_n)$ とみなす。Fig. 3 は画像空間上の画像の配置を模式的に三次元で示したものであるが, 実際には空間は三次元ではなく n 次元である。すると, 二つの画像 X, Y 間の(ユークリッド)距離 d は,

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (3)$$

となる。この計算を簡略化するために, Nayar らは固有空間法 [7] を用いることで画像空間の次元を下げている。この方法ではマッチングの精度をあまり落とさずに次元を下げる事が可能であるため, 記憶量やマッチングの計算量が少なくて済む。これに対し, 本研究では画像の次元は下げずに, 画像間の距離として

$$d' = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n| \quad (4)$$

を用い, これを 2 枚の画像のマッチングエラーと呼ぶことにする (n はテンプレートの画素数)。

このマッチングエラーをある探索範囲で計算しその最小値を求めるものが式 (1), (2) のテンプレートマッチングであり, この計算は単純でありハードウェアで高速に実行することが可能

Fig. 4 View Sequence in image space

になっている [19] ことが利点である。また、探索範囲を持つことで画像が多少ずれていても正しい画像間の距離を求めることができるだけでなく、そのずれもマッチングの結果として得られるため、ロボットの制御に用いることができる。さらに、生の画像を記憶しているためにマッチした画像の差分をとることで、容易に変化領域を抽出することも可能となる。ただし画像の大きさは記憶量や計算量に大きく効いてくるため、慎重に決定する必要がある。

ロボットにビューシーケンスを与えるためには、一度経路に沿ってロボットを移動させる必要がある。この走行は教示走行と呼ばれる。教示走行時のビューの様子を Fig. 4 の上に示す。ロボットは i 枚目のビュー M_i を記憶したところとする。ロボットが移動すると、そこで得られるビュー V は徐々に M_i から離れて行く。このマッチングエラーを計算しながらある経路上を移動して、閾値を超える M_{i+1} を新たなビューとして記憶する。これにより、画像空間上ではビューは等間隔に並ぶことになり、全体としてある走行経路は画像空間上では Fig. 4 下のようなビューシーケンスとして表現される。

3.3 マッチングエラーの推移

ロボットの移動距離に比例して、マッチングエラーが直線的に増加すると仮定すると、教示走行中のマッチングエラーの推移は、Fig. 5 のようになる。図中の $P1, P2 \dots$ は経路中のビューを記憶した位置、 $e1, e2 \dots$ は直前に記憶したビューとその位置でのマッチングエラーである。マッチングエラーの変化の仕方（傾き）は環境に依存して異なるため、ロボットの移動距離や移動時間の点では等間隔とはならない。

次に、記憶したビューシーケンスを用いて自律走行する場合を考える。マッチングエラーの増加のしかたが経路の前後方向で同じだと仮定すると、Fig. 5 のようにして記憶したビューシーケンスでは Fig. 6 のようにマッチングエラーが推移すると期待できる。ここでマッチングエラーとしているのは、すべてのビューと各位置でのビューのマッチングエラーのうちの最小値である。ただし、最小値を探すために常にすべてのビューをマッ

Fig. 5 Matching error in recording run

Fig. 6 Matching error in autonomous run

チングに用いる必要はない。

なぜなら、ロボットの移動によるビューの変化は連続的であるとみなせるからである。例えば Fig. 6 の $P1$ 付近にいることが分かっているときには、1 枚目のビューと 2 枚目のビューだけを用いて、その小さい方がマッチングエラーだとみなせる。ロボットが連続的に移動していき、 $e1'$ よりも $e2'$ の方が小さくなったならば 2 枚目のビューと 3 枚目のビューだけを用い、 $e2'$ と $e3'$ を比較するといった具合である。

3.4 ビューの探索範囲の局所性

このように、ビューのマッチングにおいて探索範囲を局所的（2 枚のみ）に限定することには、以下のような理由がある。

- マッチングの計算時間を、ビューシーケンスの長さに関わらず一定にし、ロボットの環境認識の実時間性を保証する。
- Fig. 4 のように、画像空間内でビューシーケンスが偶然接近しても、シーンを誤認識しないようにする。

ここで、近傍のビューだけをマッチング対象とするだけで十分であることを保証するのが、各ビューの間隔が画像空間中で等しいという性質である。もし、一定の走行距離ごとにビューを記憶したとすると、経路中に似たシーンが続いてビューの変化が少ない場所があると、似たビューが連続して記憶されてしまう。この場合、自律走行時の自己位置の認識（ビューシーケンス中のどこにいるか）という、マッチングの結果が不定になる。また、そのままロボットが進んでビューに変化がある場所に出たときにも、探索範囲をどこまでにするか決まらないので、正しい遷移ができないことになる。

画像空間中で等間隔にビューが並んでいる場合には、そもそも似たビューが記憶した中に連続して存在しないので、自己位置が不安定になることがない。そのため、マッチング対象は今最もマッチしているビューとその次のビューの 2 枚だけで十分であるといえるのである。ただし、見え方の変化が少ない場所では、1 枚のビューに対応する経路上の距離が長くなり、（特定の場所である行動を行いたいという場合の）位置決め精度は低

Fig. 7 Locations where view were memorized in a dark corridor

下する。これは、画像情報のみから自己位置を認識していることによる限界であるといえる。これまでの実験において、通常の環境でのビューの記憶はほぼ等距離間隔であることが分かっている。

しかし画像の変化の仕方が一定でない環境では、ビューを記憶する間隔もばらつく。Fig. 7は、夜間、ほとんど照明がなく真っ暗な廊下環境において経路を教示した実験の結果である。画像を記憶した場所が黒い点で示されているが、その間隔は等距離ではなくばらついていることが分かる。画像を集中して記憶しているのは照明のもれている部屋の近辺で、それ以外の場所では通常よりも広い間隔（最大およそ8[m]）で画像を記憶している。しかし画像空間中でのビューの距離が等しいので、ビューの探索を安定して行うことができ、このように極端な環境でも自律走行が可能であった。

3.5 ビューに必要な性質

Fig. 5, Fig. 6のグラフはマッチングエラーがロボットの移動距離に比例して直線的に増加すると仮定しているが、実際にはマッチングエラーはどのような増加をすれば、ビューシーケンスのマッチングは成功するかをここで考察する。

Fig. 8 左は、マッチングが成功する場合の典型的なビューシーケンスを示している。自律走行時のビュー V は、記憶したビュー M_{i-1}, M_i, M_{i+1} の近傍を移動している。このとき、 e_i, e_{i+1} を比較するだけで、 V が M_i, M_{i+1} のどちらに近いかを決定することができる。この場合、ロボットの移動距離（画像を記憶した位置からの距離）とそのビューを用いたマッチングエラーとの関係は、単調増加になる。

次に、Fig. 8 右はマッチングが失敗する場合の典型的なビューシーケンスを示している。自律走行時のビュー V は、全体としては記憶したビュー M_{i-1}, M_i, M_{i+1} の近傍を移動している。しかし、 e_i, e_{i+1} を比較するだけでは V が M_i, M_{i+1} のどちらに近いかを決定することができない。この場合、ロボットの位置とそのビューを用いたマッチングエラーとの関係は、全体としては増加傾向であってもノイズが大きく、単調増加にはなっていない。マッチングが単調増加にならないと、二つのマッチングエラーの比較によってロボットの位置を推定することができない。また教示走行時にも、実際にはロボットがほとんど移動していないのに新たなビューを記憶してしまうといった問題も起きることになる。

このように、ビューシーケンスとして利用可能なビューは、マッチングエラーの変化がロボットの移動により単調増加するもの、ということができる。ただし、マッチングエラーがいつまでも単調増加することはあり得ない。例えば廊下においてあ

Fig. 8 Typical good and bad views

る位置から 100[m] 先の画像と 101[m] 先の画像をそれぞれマッチングに用いても、100[m], 10[m] という距離とマッチングエラーの大小との相関関係は期待できない。このため、ビューシーケンスとして利用可能なビューの条件は、

- ビューのマッチングエラーが、ある区間で単調増加することと表現できる。この単調増加する区間のうちに新たなビューを記憶する必要がある。この区間が長いほど画像を記憶する間隔が長くできて、効率がよい。また、上の条件を満たすセンサ情報であれば、通常の画像そのものでなくてもビューシーケンスとして利用することができる。さらに、そのままではビューシーケンスには適さないセンサ情報でも、前処理により利用可能な情報に変換することが可能な場合もあり得る。

一般に、ビューがこの単調性を持つかどうかには、環境およびセンサにおける以下の要素がかかわってくる。

- 環境の形状の複雑さ
- 環境のテクスチャの複雑さ
- センサの空間分解能
- センサの精度、S/N 比

このため、単純に例えばテクスチャの複雑さだけにより単調性の有無を判断することはできない。実際にはその環境においてビューシーケンスを取り込み、マッチングエラーの変化を調べてみないと、そのビューがその環境に適用できるかどうか分からない。コンピュータビジョンにおけるビューベースの物体認識の研究 [6] [7] でも同様で、やはりテクスチャだけで、ある物体の認識が可能であるかどうか、判断できるものではない。ただし、我々は屋内の様々な環境（廊下、広いホール、エレベータなどを含む）においてビューシーケンスを用いた長距離の経路誘導実験を行っており [20]、その有効性を確認している。

4. ビュー生成手法の評価

ビューベースの認識では、ビューに含まれる情報量が多い方が安定した認識が実現できる。しかし本研究ではこれまで、記憶およびマッチングの効率を上げるために画像の解像度を落としてきた。ただし、縮小画像を作る際には間引くのではなく、平滑化を行うことで画像の質の低下を抑える工夫をしている。ここではこれらのビューの生成手法を、前章の議論を踏まえて評

価する。

4.1 ビューの解像度

前章で述べた通り、ビューのマッチングエラーはロボットの移動距離に応じて単調増加する必要がある。さらに、できるだけ長い間隔で記憶できるように、エラーが長い区間で単調増加することが望ましい。エラーの性質が変わらないのであれば、解像度は低い方が記憶量やマッチングの計算の点で有利である。画像の記憶量は画素数に比例するので、解像度を n とすると $O(n^2)$ である。またテンプレートマッチングの計算量は [テンプレートの画素数 $\times$ 探索回数] によって決まり、画素数が $O(n^2)$ 、探索回数は探索範囲を横方向だけなので $O(n)$ で、全体では $O(n^3)$ ということになる。

解像度とマッチングエラーの関係を調べるため、ビューの解像度を変えながらマッチング実験を行った結果を **Fig. 9** に示す。ロボットはある位置で1枚のビューを記憶する。このグラフは、ロボットは徐々に移動しながらその場で得られるビューと記憶したビューのマッチングを行った結果のマッチングエラーをプロットしたものである。元の画像は VRAM 上の 480×480 の領域を利用し、 160×160 、 96×96 、 32×32 、 16×16 のビューをすべての点を用いた平滑化により作成した。マッチングエラーはビューの画素数で割り、1画素当たりのマッチングエラーとして比較する。「マッチングエラーの滑らかさと単調増加する区間の長さ」の観点でこのマッチング結果を比較すると、 160×160 および 96×96 は滑らかに単調増加しており、ほとんど差はない。 32×32 にしても、まだ滑らかさに問題は見られないが、 16×16 ではやや滑らかでない部分が出てくる。

ビューの解像度を 16×16 まで落とせば、より長い距離のデータを表現することができるようにはなるが、その反面 (1) 滑らかでない部分が出て来る、(2) ステアリング制御量の分解能が下がる、(3) 障害物検出のための差分画像の解像度が下がる、という欠点が出てくる。現在のビューは $32 [\text{pixel}] \times 32 [\text{pixel}] \times 8 \text{bit}$ で1枚当たり $1 [\text{kB}]$ のデータ量を持ち、例えば $1 [\text{GB}]$ のハードディスクには百万枚記憶することができる。仮にビューの記憶間隔を $1 [\text{m}]$ とすると、このハードディスクはおよそ $1,000 [\text{km}]$ の経路を表現することができることになる。これは屋内環境における経路表現としては十分な量であり、上記の欠点を考慮するとこれ以上解像度を下げる必要はなく、妥当なサイズである。また、このサイズは本研究で用いている相関演算 LSI がサポートするテンプレートマッチングの探索画像のサイズにも一致しており、計算が最も簡単に行えるという利点もある。

また、データ量を削減するには画像を1枚単位で行うだけでなく、列として行うこと (例えば MPEG 圧縮) も考えられる。ビューシーケンスではそれぞれのビューは前後のビューと類似しており相関が高いため、この効果は大きいはずである。しかし、処理が複雑になり処理時間も増大すること、また画像へのランダムアクセスに制限が生じること、さらに上で述べたように 32×32 にしたことで十分データ量は小さくなっていること、を考慮し現在のところは画像列としての圧縮は行っていない。

4.2 平滑化の効果

ビューを生成するために解像度を下げるときには、画像を単に間引くのではなく、平滑化 (平均・縮小処理) を行う。これ

Fig. 9 Matching error with various view resolution

Fig. 10 Matching error with and without smoothing

は、ある領域に含まれる複数画素の輝度値の平均をとり、それをその領域を代表する一画素の輝度値とする処理である。平滑化には間引くことと比較して、

- (1) 画像の高周波成分を取り除く
 - (2) ビデオ信号にのっているノイズを取り除く
- という二つの効果がある。

間引きでは、サンプリング定理—サンプリング周波数が f_0 のとき、 $f_0/2$ 以下の周波数成分は正しくサンプリングできない—により説明される通り、点、線などの高周波数成分領域がある場合、同じ画像を用いても間引く位置のずれによって結果の縮小画像は大きく変化する。この現象は間引くときのサンプリング間隔が大きくなるほど顕著になり画像上のノイズとして現れるが、平滑化では画像の平均をとり高周波成分を取り除いてから縮小するので画像は安定する。また、ビデオ信号 (アナログ信号) に含まれる微小なノイズも、平均を取ることで減少させることができる。

この平滑化の効果を確かめるため、実際の廊下画像で間引きと平滑化の双方によって生成したビューと、これら2種類のビューを用いた場合のマッチングエラーの比較を **Fig. 10** に示す。ビューの解像度はどちらも 32×32 である。平滑化を行うことで画像上のノイズが減り、マッチングエラーの安定性が大きく改善されていることが分かる。

5. 視差画像を用いたビューシーケンス

前章で述べたように「ビューのマッチングエラーが、ある区間で単調増加する」ような画像特徴量を利用すればビューシーケンスとして成り立つ。これまでのビューシーケンスに基づく経路誘導では、ビューとして濃淡画像を用いていた。これは建物内の廊下環境をターゲットとしていたため「明るさなど、環境の見え方が時間によって大きく変化しない」という仮定が置けたため問題にはならなかった。しかし、屋外環境ではこのような仮定はできないため、このビューシーケンスはそのままでは屋外には適用できない。

そこで、本章では明るさの変化に影響されない「視差」を用いてビューを定義し、屋外での経路誘導実験を行う。ビューベースアプローチで認識を行う場合、用いる画像はできるだけ特徴抽出を行わず生に近い画像であることが望ましい。そこで、ビューを定義するときにはデータの抽象化（シンボル化）を行わず、またデータの信頼性を定義し重み付けに用いることで、マッチングの安定化を図る。

視差とは、ステレオカメラを用いて両眼立体視をしたときの、対象物体の画像中の水平位置の差である。平行ステレオの場合、視差が0であれば対象は無限遠に、また視差が大きくなればなるほど近くにあることになる。視差は距離によって決まるが、距離を記憶しマッチングに用いるのはビューベースアプローチとは言えない。距離は幾何学的な情報であり「見え方」とは異なる。また、視差から距離への変換は線形でなく、遠方の物体までの距離は大きな量子化誤差を含み、ノイズやキャリブレーション誤差にも弱い。これに対して視差は画像から直接得られる「見え方」であり、ビューベースアプローチに適した特徴量であると考えられる。

5.1 視差画像の生成

視差画像は、テンプレートマッチングを用いて生成する。視差画像の大きさは 32×1 で、画面上の中央に横1ラインに視差を検出するテンプレート領域を配置する。Fig. 11は屋外において取り込んだ画像（実際にはステレオ画像）とそこから生成した視差画像、および視差の信頼性である。視差画像は明るいほど大きな視差、つまり近い物体を示す。Fig. 11では通路の右側近くの木とやや左側の離れた位置にある木の部分が明るくなっているのが分かる。また、視差の信頼度には相関演算の信頼度 [21] を用いている。視差画像が水平1ラインであるのは、処理時間の制限によるものである。

5.2 視差画像のマッチング

記憶したテンプレート視差画像を $d_t(x)$ 、走行時の視差画像を $d_s(x)$ とすると、視差画像を用いたマッチングエラーは以下のように定義できる。

$$e(u) = \sum_{x=u}^{w-u} |d_t(x) - d_s(x+u)| \quad (5)$$

$$e = \{\min e(u) \mid -s \leq u < s\} \quad (6)$$

ここで、色相画像で行ったのと同様に、ここでも視差画像の信頼度 $r(x)$ を用いて以下のように定義し直す。

input image

stereo disparity

reliability of stereo matching

Fig. 11 Disparity View

Fig. 12 Result of basic matching experiment using stereo disparity

$$e(u) = \frac{\sum_{x=s}^{w-s} |d_1(x) - d_2(x+u)| \cdot \min(r_t(x), r_s(x+u))}{\sum_{x=s}^{w-s} \min(r_t(x), r_s(x+u))} \quad (7)$$

次に、屋外環境においてロボットの移動距離とマッチングエラーの関係を調べた。ロボットは Fig. 12 (左) の画像が見える位置でテンプレートを記憶し、そこから徐々に前方に移動しながらマッチングエラーを記録していく。Fig. 12 の右はこの実験の結果を示す。グラフ中の実線は視差の信頼度により重み付けを行った場合、点線は重み付けを行わなかった場合の結果であり、重み付けをした方がマッチングエラーの値は変動がやや小さくなった。濃淡画像のビューのマッチングエラーと比較して滑らかさの点はかなり劣っているが、全体としては 2.5 [m] までマッチングエラーは単調増加しているとみなせるため、ビューシーケンスに用いることができると判断した。

5.3 視差画像を用いた経路誘導

次に、屋外経路において視差画像のビューシーケンスを教示したところ、約 12 [m] の経路に対して 9 枚の画像を記憶した。記憶した視差画像のビューシーケンスと、その位置での写真、記憶した位置を Fig. 13 に示す。視差画像を並べてみると、写真に示した ①～⑤ の物体（木など）が、視差画像中で中心近くに現れ、徐々に両端へ向かって移動していくのが分かる。また、この教示データを用いて、自律走行を行ったときのマッチング

Fig. 13 Memorized Disparity View Sequence

エラーのグラフを Fig. 14 に示す。マッチングエラーは教示時の閾値 (6.0) を超えることなく正しく推移しており、視差画像ビューシーケンスを用いた経路誘導は成功した。

5.4 考察

以上の実験結果から、これまで用いてきた画像とはまったく異なるタイプの「見え方」情報を用いてもビューシーケンスが成り立つことが示された。このことはビューシーケンスの適用範囲を広げるものとして評価できる。

ただし、どのような環境ではどのような画像を用いればよいのかを評価する手段を確立するまでには至っていない。これは、3.5 節に述べた通り、ビューのマッチングエラーがロボットの移動に対して単調に増加するかどうかは、環境およびセンサの性質に依存するためである。実際に走行環境においてビューシーケンスを取り込み、マッチングエラーの変化を調べてみると、そのビューがその環境に適用できるかどうか分からないので、人間が使えるようなビューを定義して与えて試した、という段階である。

また、マッチングの基礎実験のグラフから分かるように、視差画像の生成自体が十分安定であるとは言いがたく、まだ改良の余地がある。また、処理時間が 7 [frame] かかっている点も移動ロボットにとって十分な処理速度が達成できているとはいえない。しかし、ステレオマッチングを高速に行う研究は現在盛んに行われており (例えば文献 [22])、将来的にはより密で安定した視差を得ることができれば、さらに安定したマッチング結果が得られることが期待できる。

6. おわりに

本稿では、ビューベースアプローチとビューシーケンスについて考察し、これまでのビューシーケンスの生成方法の評価を行った。また「視差」の画像がビューシーケンスの持つべき性質を満たすことを実験的に示し、屋外環境での経路誘導へ応用した。

モデルベースアプローチとビューベースアプローチは異

Fig. 14 Experimental result of matching in autonomous navigation

なる特性を持っているため、一概にどちらが優れているとはいえない。心理学には、人間が三次元物体を認識する際に、物体が見慣れていないような場合には三次元構造を考慮しながらモデルとの照合をとるが、よく見慣れた物体については二次元照合を行っている、という報告がある。また最近の生物学での研究には、蟻のナビゲーションにおいても、ビューシーケンスと同じように複数のビューを記憶し、順番にマッチングしながら位置を認識している (文献 [23] [24])、という報告もあり、ビューベースアプローチは生物の視覚の認識法を考える上でも興味深い。

参考文献

- [1] 松本吉央, 稲葉雅幸, 井上博允: “視野画像列を利用した経路表現に基づくナビゲーション”, 日本ロボット学会誌, vol.15, no.2, pp.236-242, 1997.
- [2] 松本吉央, 坂井克弘, 池田和徳, 稲葉雅幸, 井上博允: “ビューベースアプローチに基づく移動ロボットナビゲーション”, 第4回ロボティクスシンポジウム予稿集, pp.283-288, 1999.
- [3] Y. Matsumoto, M. Inaba and H. Inoue: “View-Based Approach to Robot Navigation,” Proc. of 2000 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1702-1708, 2000.
- [4] 北野宏明: “超並列人工知能”, 人工知能学会誌, vol.7, no.2, pp.224-262, 1992.
- [5] C. Stanfill and D. Waltz: “Toward Memory-Based Reasoning,” Communications of the ACM, vol.29, no.12, pp.1213-1228, 1986.
- [6] 村瀬洋: “古くて新しい画像認識法—固有空間による画像認識—”, 情報処理, vol.38, no.1, pp.54-60, 1997.
- [7] S.K. Nayar, H. Murase and S.A. Nene: “Learning, Positioning, and Tracking Visual Appearance,” Proc. of IEEE Int. Conf. on Robotics and Automation, pp.3237-3244, 1996.
- [8] I. Horswill: “Polly: A Vision-Based Artificial Agent,” Proc. of Int. Conf. on AAAI '93, pp.824-829, 1993.
- [9] 前田武志, 石黒浩, 辻三郎: “全方位画像を用いた記憶に基づく未知環境の探索”, 電子情報通信学会技術研究報告, PRU92-10, pp.73-80, 1995.
- [10] 前田佐嘉志, 久野義徳, 白井良明: “固有空間解析に基づく移動ロボットの位置認識”, 電子情報通信学会論文誌, vol.J80-D-II, no.6, pp.1502-1511, 1997.
- [11] 小島浩, 伊藤慶明, 岡隆一: “Reference Interval-Free 連続 DP を利用した移動ロボットの時系列画像による位置同定システム”, 電子情報通信学会論文誌, vol.J80-D-II, no.3, pp.724-733, 1997.
- [12] 西村拓一, 野崎俊輔, 岡隆一: “Non-monotonic 連続 DP によるスポットティングに基づく移動ロボットの時系列画像を用いた大域的位置の推定”, 電子情報通信学会論文誌, vol.J81-D-II, no.8, pp.1876-1884, 1998.
- [13] 岩佐英彦, 栗飯原述宏, 横矢直和, 竹村治雄: “全方位画像を用いた記憶に基づく位置推定”, 電子情報通信学会論文誌, vol.J84-D-II, no.2, pp.310-320, 2001.
- [14] D.A. Pomerleau: “ALVINN: An Autonomous Land Vehicle in a Neural Network”, Technical Report CMU-CS-89-107, CMU,

- 1989.
- [15] M. Meng and A.C. Kak: "NEURO-NAV: A Neural Network Based Architecture For Vision-Guided Mobile Robot Navigation Using Non-Metrical Models of the Environment," Proc. of Int. Symp. on Robotics Research, pp.750-757, 1993.
 - [16] Z. Zhu, H. Xi and G. Xu: "Combining Rotation-Invariance Images and Neural Networks for Road Scene Understanding," In Proc. IEEE Int. Conf. on Neural Networks, pp.1732-1737, 1996.
 - [17] C. Furlanello, B. Crespi and L.Stringa: "A memory based approach to navigation," Biological Cybernetics, vol.69, pp.385-393, 1993.
 - [18] H. Ishiguro, T. Miyashita and S. Tsuji: "T-Net for Navigating a Vision-Guided Robot in a Real World," Proc. of IEEE Int. Conf. on Robotics and Automation, pp.1068-1073, 1995.
 - [19] 井上博允, 稲葉雅幸, 森武俊, 立川哲也: "局所相関演算に基づく実時間ビジョンシステムの開発", 日本ロボット学会誌, vol.13, no.1, pp.134-140, 1995.
 - [20] Y. Matsumoto, K. Ikeda, M. Inaba and H. Inoue: "Visual Navigation using Omnidirectional View Sequence," Proc. of 1999 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp.317-322, 1999.
 - [21] 森武俊, 松本吉央, 稲葉雅幸, 井上博允: "相関値分布の分類に基づく追跡注視点生成", 日本機械学会ロボティクス・メカトロニクス講演会'95 講演論文集, pp.1076-1079, 1995.
 - [22] 岡田慧, 加賀美聡, 稲葉雅幸, 井上博允: "PC による高速対応点探索に基づくロボット搭載可能な実時間視差画像・フロー生成法と実現", 日本ロボット学会誌, vol.18, no.6, pp.896-901, 2000.
 - [23] S.P.D. Judd and T.S. Collett: "Multiple stored views and landmark guidance in ants," Nature, vol. 392, pp.710-714, 1998.
 - [24] M.V. Srinivasan: "Ants march as they march," Nature, vol. 392, pp.660-661, 1998.

松本吉央 (Yoshio Matsumoto)

1970 年 8 月 21 日生。1993 年 3 月東京大学工学部卒業。1998 年 3 月同大学大学院工学系研究科情報工学専攻博士課程修了。同年 4 月オーストラリア国立大学研究員。1999 年 1 月奈良先端科学技術大学院大学情報科学研究科助手, 2001 年 1 月同助教授。ロボットビジョンおよびヒューマンインタフェースの研究に従事。電子情報通信学会会員。(日本ロボット学会正会員)

井上博允 (Hirochika Inoue)

1942 年 7 月生。1965 年東京大学工学部卒業。1970 年同大学大学院博士課程修了。工学博士。同年電子技術総合研究所入所。知能ロボットの研究開発に従事。1977 年東京大学工学部機械工学科助教授。1984 年教授。現在, 機械情報工学科教授。ロボット全般, 人工知能, 情報システム工学の研究と教育に従事。(日本ロボット学会正会員)

稲葉雅幸 (Masayuki Inaba)

1958 年生。1981 年東京大学工学部機械工学科卒業, 1986 年東京大学大学院工学系研究科情報工学専門課程修了。工学博士。1986 年東京大学講師, 1989 年同助教授, 2000 年同教授, 大学院情報学環所属, 現在に至る。日本機械学会, 情報処理学会, 計測自動制御学会, 人工知能学会, ソフトウェア科学会各会員。(日本ロボット学会正会員)