

仮想環境での物体の見え方変動モデルの獲得と 実環境での未知の見え方へ対応可能なロボット視覚に関する研究

Construction of Variable Appearance Model for Objects in a Virtual Environment
and Object's Appearance Prediction for Robot Vision System in a Real Environment

萩原良信*¹

Yoshinobu Hagiwara

稲邑哲也*²

Tetsunari Inamura

*¹国立情報学研究所

National Institute of Informatics

*²国立情報学研究所/総合研究大学院大学

National Institute of Informatics/The Graduate University for Advanced Studies

If a lighting condition for capturing is different from the one for the reference images, the robot should measure the parameters of lighting condition or use a huge image database which covers so many lighting conditions in order to recognize the objects. However, it is difficult for real robots to create such a huge database due to high cost for preparing many objects under various lighting conditions. In this paper, we propose a novel approach to construct a huge image database with variable appearance in a virtual environment and apply this database to object recognition in real environments. Even if an image of an unknown object under unknown lighting condition is given, the robot estimates lighting condition filter that converts a reference image into the real image based on a known image in the environment and a huge image database.

1. はじめに

家庭やオフィスなどの人間が生活する複雑な環境における物体認識は、照明と物体、ロボットとの相対的な位置・姿勢の関係性によって生じる多様な見え方の変化に対応する必要がある。さらに近年では、これらの見え方の変化に頑健であるのみならず、その見え方の変化を感知して物体の様々な状態を推定し、その情報を人との円滑なコミュニケーションに役立てる事が期待されている。

従来、物体の種類毎に分類された膨大な画像のデータベースに基づく認識手法 [Nagahashi 07, Han 06] は、その物体が何であるかというカテゴリの認識において成果を上げてきた。これらの手法は、SIFT(Scale-Invariant Feature Transform)[Lowe 99], SURF(Speeded Up Robust Features)[Bay 08] などの画像特徴を用いる事で照明条件の変化に頑健な入力画像とデータベース内の参照画像との照合を実現している。しかしながら、これらの画像特徴は、環境の照明条件やその位置関係による見え方の変化を省いて物体を認識するため、人間が知覚しているような物体の見え方の変化を知る事が困難である。例えば、蛍光灯に照らされた赤りんごと夕日に照らされた青りんごの識別などである。

このような場合、ロボットが画像変換フィルタを適用して撮像画像を通常の照明条件下の画像に変換する方法が考えられる。従来、物体の画像間の特徴の対応から局所的な色変換を計算し、その投票によって全体の色変換を推定する方法 [Drew 02] や、異なる照明条件下のシーンの画像中の色の組から color flow を学習して色変換を行う手法 [Miller 01] が提案されている。しかしながら、人間の生活環境で動作するサービスロボットにこれらの手法を応用するには、未知環境下における多種多様な物体や参照光の獲得が課題となる。

このため、多様な照明条件や位置関係で撮像された物体の見え方を記憶した大規模データベースを獲得する必要がある。様々な角度、位置、照明条件で撮像された物体画像の大規模データベースから、ロボットによる画像変換フィルタの設計が期待できる。しかしながら、実環境において、これらの見え方



図 1: 既知物体の見え方変動モデルの獲得

に関する経験をロボットが得るには、膨大な数の物体と照明器具、各種の機器を操作する技術員、天候や時刻による長期間の観測が要求され、そのコストの高さが問題となる。

そこで本研究では、これらの見え方に関する経験をロボットが仮想環境において効率的に獲得し、この経験を実環境のロボットの物体認識に応用するアプローチを提案する。具体的には、まず、仮想環境においてロボットが多様な物体を多様な時刻や場所において観測して得られた物体の画像データベースを見え方変動モデルとして獲得する。次に、この見え方変動モデルから物体の見え方の変化を計算した画像変換フィルタを設計し、この画像変換フィルタを新たに観測された新規物体の画像に適用する事で未知の見え方を予測して新規物体を認識する。また、仮想環境で獲得された物体の見え方を計算する画像変換フィルタを実環境の物体認識に応用した実験についても述べる。本アプローチを効率化して実現するため、仮想環境において光源と物体、ロボットを多種多様に設定・設置して物体の画像を撮像する機能を備えた SIGVerse [Inamura 10] を本研究の開発と実験に用いた。

2. 本アプローチの概要

2.1 既知物体の見え方変動モデルの獲得

図 1 に既知物体の見え方変動モデルの獲得の概念図を示す。図 1 中央のロボットは、図 1 左に示す時刻 12:00 から 18:00

連絡先: 萩原良信, 国立情報学研究所, 東京都千代田区一ツ橋 2-1-2, yhagiwara@nii.ac.jp

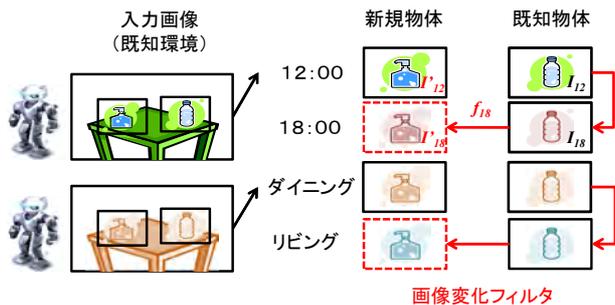


図 2: 画像変換フィルタによる新規物体の見え方の予測



図 3: 未知の見え方における新規物体の認識

への時間の経過に伴う太陽光による変化, ダイニングからリビングへの移動に伴う照明器具による変化において既知物体を観察する. この観察において, ロボットは, 異なる時刻や場所において既知物体の画像を撮像し, 図 1 右のように見え方の異なる物体の画像を時刻や場所と関連付けて見え方変動モデルを獲得する.

2.2 画像変換フィルタによる新規物体の見え方の予測

図 2 に既知物体の見え方変動モデルから計算された画像変換フィルタを用いた新規物体の未知の見え方の予測の概念図を示す. 本研究では, 屋外の時間の経過に伴う見え方の変化, 屋内の場所の移動に伴う見え方の変化を扱うが, ここでは, 前者の時間の経過に伴う見え方の変化を例として説明する. まず, 図 2 左上の入力画像のように既知物体と新規物体が同じ画像内に存在した場合, 図 2 右上のように既知物体の見え方変動モデルから入力画像の見え方 (ここでは, 時刻 12:00) が予測される. 次に, この既知物体の基準画像 I_{12} と予測したい時刻における既知物体の参照画像 I_{18} から見え方の変化を計算する画像変換フィルタ f_{18} を得る. 最後に, 時刻 18:00 における新規物体の未知の見え方の画像 I'_{18} は, 画像変換フィルタ f_{18} と時刻 12:00 における新規物体の基準画像 I'_{12} から式 (1) で計算される.

$$I'_{18} = f_{18}(I'_{12}) \quad (1)$$

同様に, 場所の移動による見え方の変化を計算する画像変換フィルタを構築し, 新規物体の異なる場所における見え方の変化を予測した画像データを得る事ができる.

2.3 未知の見え方における新規物体の認識

図 3 に予測された画像データを用いた未知の見え方における新規物体の認識の概念図を示す. 図 3 左上の入力画像は, 時刻 18:00 における未知の見え方で新規物体を撮像したもので



図 4: 屋外における時間経過による物体の見え方の変化

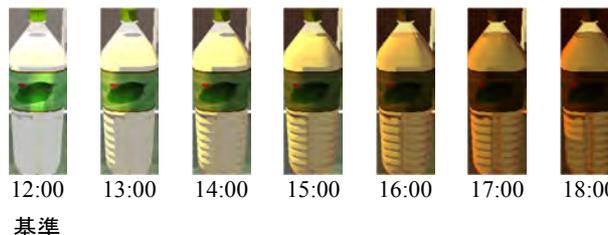


図 5: 既知物体の時間経過における見え方変動モデル

ある. 図 3 右上の時刻 12:00 を基準として予測した新規物体の画像データを用いて, 時刻 18:00 の未知の見え方において新規物体を認識する. 同様に, ダイニングを基準としてリビングでの新規物体の見え方を予測し, 場所の移動による見え方の変化についても述べる.

3. 仮想環境における評価と実環境への応用

3.1 時間の経過による見え方の変化に対する実験

本実験では, 実環境で想定される代表的な見え方の変化の一つである時間の経過による物体の見え方の変化を扱う. 屋外環境における太陽光による物体の見え方の変化を仮想環境に再現した実験を行った.

図 4 に仮想環境の屋外の異なる時刻において既知物体の見え方の変化をロボットが観察する様子を示す. 画像中央のテーブルの上に対象となる物体を配置し, テーブルに向かって設置されたロボットのカメラで物体の画像を撮像する. 物体の画像を撮像する度に, シミュレーション環境内の時刻を進め, 時間の経過に伴って図 4(a),(b) の昼, 夕のように光源の位置と色温度を変化させて照明条件が実際の太陽光に似た変化をするように設定した. このとき, シミュレーション環境内での時刻は, 時刻 12:00 の正午から時刻 18:00 の日の入りまでとし, 30 分間隔で撮像した合計 13 枚の画像を時刻と関連付けて既知物体の見え方変動モデルとして記憶した.

図 5 に本実験において記憶した 13 枚の画像から一時間毎の 7 枚の既知物体の画像を示す. 図 5 の左の時刻 12:00 から右の時刻 18:00 まで既知物体の明るさや色合いが異なり, 見え方が変化している.

図 6 に画像変換フィルタによる新規物体の未知の見え方の予測の概要を示す. 図 6 には, 時刻 12:00 を基準の時刻として, 時刻 17:00 のにおける新規物体の未知の見え方を予測する例を示している. まず, 基準となる時刻 12:00 の既知物体と新規物体の画像が得られているとする. このとき, 既知物体の時間経過による見え方の変化は観察済みであり, 新規物体の異なる時刻における見え方は未知である. 次に, 既知物体の見え方変

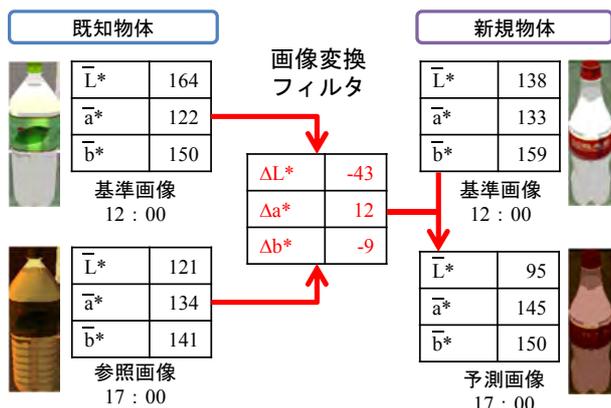


図 6: 画像変換フィルタによる新規物体の未知の見え方の予測



図 7: 新規物体の異なる時間における見え方の予測画像

動モデルの時刻 12:00 と 17:00 の画像における見え方の差から画像変換フィルタを設計する。ここで、画像変換フィルタの計算には、 $L^*a^*b^*$ (1976 CIE $L^*a^*b^*$ 色空間)を用いた。まず、画像 T の座標位置 (x, y) における各画素の L^* , a^* , b^* の値 $T_{L^*}(x, y)$, $T_{a^*}(x, y)$, $T_{b^*}(x, y)$ から平均値 \bar{L}^* , \bar{a}^* , \bar{b}^* を計算する。図 6 左の画像と表は、既知物体の時刻 12:00, 17:00 の画像と、画像から計算した \bar{L}^* , \bar{a}^* , \bar{b}^* の値を示している。さらに、基準となる時刻に得られた基準画像の $L^*a^*b^*$ の平均値と異なる時刻に得られた参照画像の $L^*a^*b^*$ の平均値の差から画像変換フィルタ ΔL^* , Δa^* , Δb^* を得る。図 6 中央の表は、基準の時刻を 12:00, 異なる時刻を 17:00 とした場合の画像変換フィルタを示している。最後に、時刻 12:00 に得られた新規物体の基準画像の各画素の L^* , a^* , b^* の値に画像変換フィルタの ΔL^* , Δa^* , Δb^* の値を加算し、時刻 17:00 の新規物体の未知の見え方を予測した画像を構築する。

同様の処理を全ての時刻について実行し、時刻 13:00 から時刻 18:00 までの新規物体の未知の見え方を予測した画像データが図 7 である。ここでは、時刻 12:00 から 18:00 までに 30 分毎に予測した 13 枚の画像から一時間毎の 7 枚の画像のみを示している。図 7 の左の時刻 12:00 の基準画像から右の時刻 18:00 の画像まで時間の経過に伴った新規物体の見え方の変化が確認できる。

本アプローチの有用性を評価するため、異なる時刻における新規物体の認識において、予測画像有り無しの場合の画像照合の相関値の比較した。ここで、画像照合には、RGB 各値の二乗平均誤差に基づく単純なテンプレートマッチングを用いた。結果のグラフを図 8 に示す。横軸が時刻、縦軸は画像照合の相関値、●が予測有、▲が予測無の結果を示している。基準

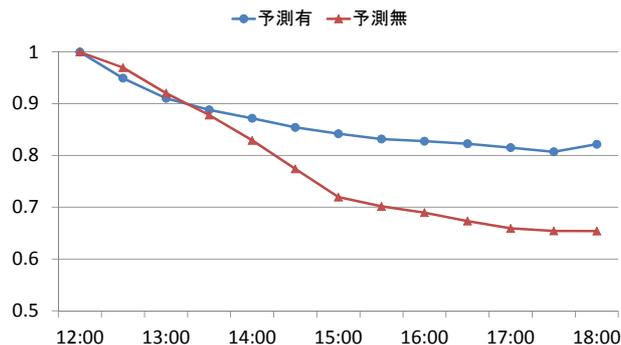


図 8: 未知の見え方における新規物体の画像照合の相関値

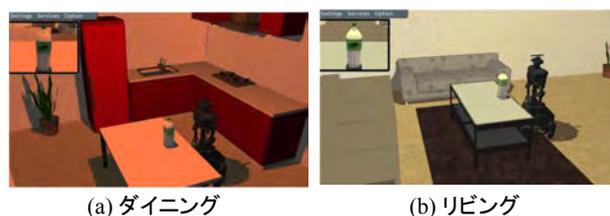


図 9: 屋内における場所の移動による物体の見え方の変化

の時刻 12:00 では、二つの相関値に大きな差は見られないが、時間が経過した夕方時刻 18:00 では、予測無に比べて予測有の結果が高い相関値を示している。この結果から、仮想環境において獲得した既知物体の見え方変動モデルが新規物体の未知の見え方において認識の確度を高める事が示唆された。

3.2 場所の移動による見え方の変化に対する実験

次に、屋内におけるロボットの場所の移動による物体の見え方の変化について実験を行った。また、本実験では、仮想環境において見え方変動モデルを獲得し、これを実環境の物体の見え方の変化に応用する。

図 9 に屋内の照明条件の異なる部屋における既知物体の見え方の変化をロボットが観察する様子を示す。図 9(a) がダイニング、(b) がリビングの様子であり、それぞれ白熱灯と蛍光灯の色温度を環境の光源として設定した。

まず、図 10(a),(b) のように仮想環境において白熱灯と蛍光灯の照明条件の下で見え方の異なるペットボトルの画像を撮像した。また、図 10(c) は、実環境の白熱灯の照明条件で撮像したマグカップの画像である。このとき、実環境の蛍光灯の照明条件におけるマグカップの見え方は未知である。本実験では、



図 10: 仮想環境と実環境で撮像した画像

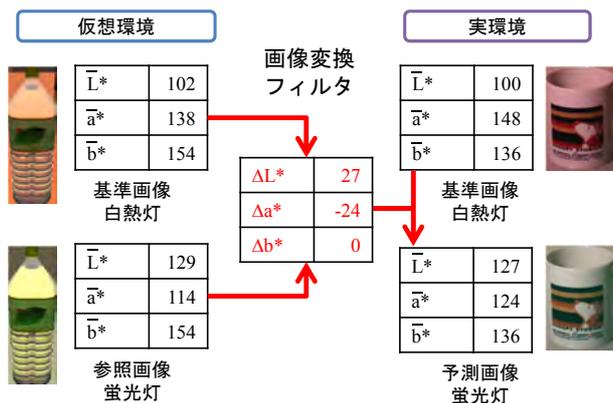


図 11: 仮想環境の見え方変動モデルによる実環境の見え方の予測

表 1: 予測の有無による画像照合の相関値

	画像照合の相関値
予測 無し	0.75
予測 有り	0.80

仮想環境で得られた図 10(a),(b) の二枚の画像を見え方変動モデルとして画像変換フィルタを設計し、これを実環境の白熱灯で得られた図 10(c) の画像に適用する事で実環境の蛍光灯におけるマグカップの見え方を予測して認識する。

図 11 に仮想環境の既知物体の画像から実環境の新規物体の見え方を予測した結果を示す。図 11 左が図 10 の仮想環境で白熱灯と蛍光灯の照明条件におけるペットボトルの見え方を撮像した画像である。まず、図 11 左の仮想環境で得られた白熱灯と蛍光灯の画像の差から図 11 中央の画像変換フィルタを計算した。次に、図 11 右上の実環境の白熱灯において得られたマグカップの画像に図 11 中央の画像変換フィルタを適用して図 11 右下の実環境の蛍光灯におけるマグカップの見え方を予測した画像を構築した。

表 1 は、実環境の蛍光灯の照明条件化におけるマグカップの認識において、この予測画像を用いた場合と用いなかった場合の画像照合の相関値の違いを示している。表 1 の一列目が予測画像の有無、二列目がその条件における画像照合の相関値である。表 1 において、予測画像を用いなかった場合の相関値 0.75 に対して予測画像を用いた場合には相関値が 0.80 まで向上している。この結果から、仮想環境において獲得した見え方に関する経験が実環境の物体認識において有用な結果を導き出す一つのモデルを示す事ができた。

4. おわりに

本稿では、ロボットが仮想環境において効率的に見え方に関する経験を獲得し、これを実環境のロボットの物体認識に応用するアプローチを提案した。本アプローチの最初の段階として、時間の経過と場所の移動による見え方の変化を仮想環境に再現し、異なる物体の見え方を予測して認識する実験を行った。実験では、実環境のロボットではコストが高くなる長期間の観察を仮想環境において実施し、一日の物体の見え方の変化

を複数の部屋において記憶し、見え方変動モデルを効率的に獲得した。さらに、見え方の変化を計算する画像変換フィルタを構築し、これを新規物体に適用する事で未知の見え方を予測して物体を認識する方法を示した。実験の結果から、本アプローチは物体認識に有用であるのみならず、その物体の未知の見え方を予測して認識に役立てる事が確認された。実環境における実験では、仮想環境において獲得した見え方に関する経験を実環境の物体認識に応用し、有用な結果を得る事ができた。実験において用いた仮想環境の光源のパラメータは色温度のみに基づく簡単なものであったが、この設定においても実環境の物体認識に成果があった事から、実環境の見え方に関するパラメータを獲得する事でさらなる効果が期待される。今後は、仮想環境において、さらに多様な時刻や場所で見え方の異なる膨大な物体の画像を獲得し、仮想環境における繰り返し試行によって画像処理の要素技術に関する最適化を行って本手法の効果を高めていきたい。

参考文献

- [Nagahashi 07] T. Nagahashi, H. Fujiyoshi, and T. Kanade: "Object Type Classification Using Structure-based Feature Representation", in *MVA2007: IAPR Conference on Machine Vision Applications*, pp.142-145, 2007.
- [Han 06] F. Han, Y. Shan, R. Cekander, H. S. Sawhney, and R. Kumar: "A two-stage approach to people and vehicle detection with HOG-based SVM", in *Proc. Of Workshop on Performance Metrics for Intelligent Systems*, pp.133-140, 2006.
- [Lowe 99] D.G. Lowe: "Object recognition from local scale-invariant features", in *Proc. of the 7th International Conference on Computer Vision*, pp.1150-1157, 1999.
- [Bay 08] H. Bay, T. Tuytelaars, and L. B. Gool: "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding*, Vol.110, No.3, pp.346-359, 2008.
- [Drew 02] M.S. Drew, Z.N. Li, and Z. Tauber: "Illumination color covariant locale-based visual object retrieval", *Pattern Recognition*, Vol.35, No.8, pp.1687-1704, 2002.
- [Miller 01] E.G. Miller and K. Tieu: "Color eigenflows: Statistical modeling of joint color changes", *Proc. 8th Int. Conf. on Computer Vision*, Vol.1, pp.607-614, 2001.
- [Inamura 10] T. Inamura et al: "Simulator platform that enables social interaction simulation -SIGVerse: SocioIntelliGenesis simulator-", in *IEEE/SICE International Symposium on System Integration*, pp.212-217, 2010.