

ディープニューラルネットワークを用いた 物体表面の光沢感認識

重川 裕和^{1,a)} 岡谷 貴之^{1,b)}

概要: 本研究は、1枚の画像から物体表面の反射特性を推定する問題を扱う。これまでいくつかの関連研究が行われてきたものの、人と同等の水準で、この課題を達成することはできていない。鍵となるのは、画像からどのような特徴を抽出すべきかであり、この観点から、学習によって、より適切な特徴の獲得が期待できる、ディープニューラルネットワーク (DNN) を用いた。具体的には、物体形状が既知、周囲の照明環境が未知である条件における、異なる反射特性クラスの識別課題を考えた。実験結果から、2段以上の畳み込みニューラルネットワーク (CNN) を用いた場合、人の視覚に匹敵する識別性能が得られることが分かった。また、同心円状で中心が明るく (暗く) 周囲が暗い (明るい) という特異な特徴が学習された。

1. はじめに

本研究では、単一の画像からそこに写る物体表面の質感を画像認識の枠組みで推定する問題を考える。質感とは、光沢感、透明感、滑らかさ、温かみ等の、人間が視覚や触覚を用いて物体表面から知覚する一連の感覚を指す。質感は物体の材質に最も関わりが深い、視覚的な知覚に限れば、表面反射特性から得られる情報がより重要である。故に問題は、表面反射特性そのもの、あるいはそれに関する何らかの情報を画像から推定することと定義できる。

物体の見えは、物体の表面反射特性のみならず、形状や照明条件に大きく影響を受けるため、表面反射特性を認識することは、他の要因についての事前知識を持たない限り、大変難しい問題であるように見える。それでも人は、視覚情報のみを用いても、かなり正確に物体表面の質感を推定することが可能である。

そのため、単一画像のみを用いて表面反射特性を推定するための研究がこれまでいくつか行われてきた [1], [6], [7], [12], [13]。コンピュータビジョンにおいて重要なのは、如何に良い特徴量を画像から抽出するかである。これまで手で設計した特徴量を用いた研究が行われてきたが [8], [10], [17], それらを用いた認識性能はどれも人間には遠く及ばないものであり、特徴を設計しては試すというやり方には限界がある。また、心理物理や神経科学の分野では、生物の視覚でなされている処理を解明しよう

とする研究がなされており、[13] のような興味深い成果も幾つか得られているが、これらの研究で全てを解明できたとは言えない。

本研究では、ディープニューラルネットワーク (DNN) を反射特性認識に適用した。近年、DNN が一般物体認識をはじめ、音声認識や自然言語処理などの様々な分野で従来手法を圧倒する性能を示すことが報告されている [3]。高い性能を示す鍵は、DNN が与えられた課題に応じた適切な特徴量を自動的に学習で得ることができる点にあると考えられ、このことは、画像からどういった特徴を取り出せば良いのか分からない質感の認識には特に魅力的である。我々の知る限り、DNN を反射特性認識に適用したとの報告は未だ無かった。

本研究では、数ある質感の中から光沢感を選び、物体形状が既知でかつ周囲の照明環境が未知の条件でこれを認識する問題を考える。実験のために、球の物体の様々な照明条件下における画像を市販のレンダラーを用いて生成した。識別対象とするクラスは、画像をレンダリングする過程で、表面反射特性に関する幾つかのパラメータを変化させることで作り出し、テスト用の画像各々が、用意したクラスのどれに属するかを識別できるよう、生成した学習用データセットを用いて DNN を訓練した。本研究では、幾つかの DNN の候補の中から、画像認識分野でこれまで数々の良い成果をあげている畳み込みニューラルネットワーク (CNN)[4], [9] を用いることにした。

以下の問に答えるため、幾つかの実験を行った。

- (i) CNN を用いて、照明未知の条件下で生成された 1 枚の画像から表面反射特性を認識することは可能か。

¹ 東北大学

Tohoku University

a) omokawa@vision.is.tohoku.ac.jp

b) okatani@vision.is.tohoku.ac.jp

- (ii) CNN により学習される特徴はどのようなものか.
- (iii) CNN の構造や層の数は認識精度にどう影響するか.

2. 関連研究

視覚に関する心理物理学や神経科学では、人の視覚がどのように物体表面の質感を捉えているのかを解明するための研究 [7], [12], [13] が行われている. 最も大きな成果に, Motoyoshi らの研究 [13] がある. 彼らは, 人の光沢感知覚は, 画像の輝度ヒストグラムの歪度に強い相関があることを示し, さらに, 輝度ヒストグラムの歪度を計算する生体視覚モデルを提案した.

また, コンピュータビジョンの分野では, 本問題に対して 2 種類の方法が研究されている. 1 つは物理ベースの手法であり, 問題を照明と反射特性の同時推定とみなし, 単一の画像からこれを行おうとするものである [11], [16]. この手法の鍵は, 照明と反射特性に関する事前知識を最大限活用することであり, これらの方法の成否は, 事前知識の精度に依存する.

もう 1 つは, 機械学習をベースとした手法であり, 本研究もこの枠組みに属する. Dror ら [6] は, 画像から幾つかの単純な特徴を抽出し, SVM を用いて, それらの特徴と, 反射特性との間の相関を学習する手法を提案している. 彼らの手法は, 照明未知, 形状既知の条件下で, ある程度正確な反射特性クラスの推定を行えた. また, [15] に示される枠組みを用いて, Abe ら [1] は物体表面の質感属性に関する比較情報を基にしたランキング学習を用いた自然画像の質感推定法を示している.

本研究が扱う問題は, 材質認識とも強い関連性がある. 例えば Liu ら [10] は, ベイズ推定の枠組みを適用し, 多くの特徴の候補の中から, 材質認識に有効な特徴を選択することを考えた. 彼らは, SIFT と色ヒストグラム, Edge-slice が最も良い組み合わせであると報告している. また, Hu ら [8] はエッジ方向の輝度値の分散を特徴に用いて, Large Margin Nearest Neighbor による次元削減と分類を行う手法を提案し, Liu らを超越する性能を示した. さらに, Timofte と van Gool は, [5] で提案された BIF (Basic Image Features) と呼ばれる特徴と, 新しいプーリング手法とをあわせて用いることで, さらに高い性能を実現した [17].

上述の研究により, ある程度高い性能を実現することはできたが, それらの認識精度は, 人の視覚には大きく劣っており, 人と機械の視覚能力には, 依然大きな隔りがある.

3. 手法

3.1 データセット作成

実験では, 様々な自然照明の下でレンダリングした CG 画像を用いた. それらの画像を生成するため, Autodesk 社の Maya を用いた. より具体的には, レンダラーに Maya の内部レンダラーでありフォトリアルな画像が生成可能な



図 1 レンダリングに用いた自然照明画像の例. 上段: Interior. 下段: Nature.

mental ray を, シェーダとして, mental ray と相性が良い mia material を使用した. 後述するように, mia material 内に実装されている反射特性に関する数個のパラメータを変化させることで, 反射特性クラスを生成した. また, 生成する画像のサイズは, 128×128 ピクセルとした.

3.1.1 照明

照明用画像には, Dosch Design 社が有償で提供しているものの中から Interior と Nature V1 の 2 種類のデータセットを用いた. それぞれテーマに沿った様々なシーンの HDR 画像のセットとなっている. この 2 つのデータセットから 40 枚のシーン画像を選び, 合計 80 枚を照明用画像とした. 図 1 に画像例を示す. 以降, これら 2 種類の照明用データセットを, それぞれ Interior, Nature と呼ぶ. 我々は, DNN の学習に必要な大規模な画像セットを作るため, レンダリングする際, 生成画像の座標系から見て照明を, y 軸周りに $0, 72, 144, 216, 288$ 度, また, x, z 軸周りに $-30, -15, 0, 15, 30$ 度回転させることで, 異なる照明条件の画像の数を増やした. ここで, x, z 軸周りに関しては, 自然画像の上の方が明るい傾向があるという自然統計を崩さないようにするため, 変化は小さなものとした. この処理によって, 1 つの照明画像につき, $5 \times 5 \times 5 = 125$ 枚の画像を生成することができる.

3.1.2 光沢感

mia material 中の 2 つのパラメータを変化させることで, 光沢感に関するパラメータの違いによるクラスを 5 種類用意した. 1 つは, 鏡面反射と拡散反射の割合 (refl_gloss), もう 1 つは, 入射光に対する反射光の角度依存性である. 後者は, 90 度におけるその割合 (brdf_90_degree_refl) を 1.0 に固定し, 0 度における割合 (brdf_0_degree_refl) を, 0.0 から 1.0 まで変化させた. この値が 1.0 に近いほど, より金属らしい光沢感を実現することができる. 我々は, 光沢感に関する 5 つのクラスに対応するパラメータの組として, (0.2,0.8), (0.8,0.8), (0.5,0.5), (0.8,0.2), (0.2,0.2) を選んだ. 図 2 に, レンダリングした画像の例を示す.

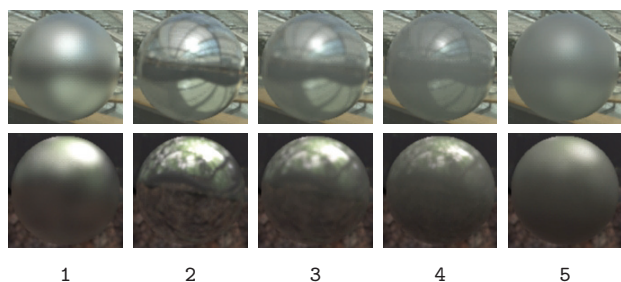


図 2 異なる光沢感パラメータを持つ 5 クラスの画像例. 各クラス ID を画像下に示す. 上段は *Interior* を, 下段は *Nature* の照明画像を基にレンダリングした画像である.

3.2 識別課題

未知の照明を基にレンダリングした画像に対し, 光沢感に関するクラス分類を行う. そのため, 照明画像セットを, 学習用とテスト用に分ける. 光沢感に関しては, *Interior* と *Nature* からそれぞれ 32 枚を学習用, 残りの 8 枚をテスト用とした. つまり, 各クラス中に含まれる画像は, 学習用が $8000 (= 32 \times 2 \times 125)$ 枚, テスト用が $2000 (= 8 \times 2 \times 125)$ 枚となる.

3.3 畳み込みニューラルネットワーク

実験では, cuda-convnet^{*1}を用いて CNN による学習及びテストを行った. 本稿では, ネットワークを簡潔に記述するため, 以下の表記を用いる. まず, 層の種類に関して, 畳み込み層を C , プーリング層を P (本実験では, 最大プーリングを使用する), 同一マップ内における正規化層を N , マップ全体にわたっての正規化層を \bar{N} とする. これらの層は, 全て局所受容野を持っており, そのサイズと隣接局所受容野との間隔を示すために, それぞれ下付き, 上付き文字を用いる. 例えば, C_{17}^2 は, x, y 軸それぞれの方向に 2 ピクセル間隔をとって, 入力層に適用される 17×17 ピクセルの大きさのフィルタを持つ畳み込み層であることを示している. 以降の実験では, 全ての畳み込み層で, フィルタの数を 32 個に, また, 1 段目の畳み込み層のフィルタサイズは, 17×17 ピクセルとした.

活性化関数に一般的なシグモイド関数ではなく, 修正線形関数 $f(x) = \max(0, x)$ を用いることで NN の学習をより高速に行うことができるとの報告がある [9]. この活性化関数を用いた場合, 続く処理で正規化を行う層を介す必要があり, これに従い, 用いる全ての畳み込み層の活性化関数にこの関数を採用し, 正規化層を畳み込み層とあわせて使用した. 図 3 に, 2 つの畳み込み層を持つ CNN の例として, $C_{17}^2, P_5^2, N_4^1, C_3^1, P_3^2, N_2^1$ の模式図を示す.

CNN の最上位層は, 5 つのユニットから成り (後述するシーンについての 2 クラス分類実験の場合は 2 つ), この層と 1 つ手前の層との間は全結合により結ばれている. こ

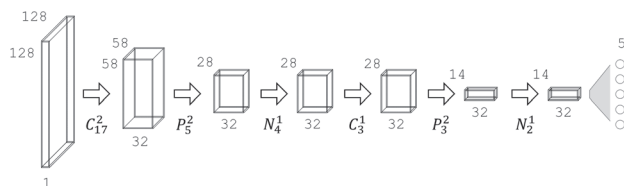


図 3 実験で用いる CNN の一例.

の層の出力は, 入力画像が, 各クラスに属する事後確率を算出するため, softmax 関数によって正規化される. CNN の学習は, 正しいラベルと予測されたラベルとの間で定められる交差エントロピー誤差関数を, 逆誤差伝播法によって最小化を図ることで進む. つまり, 学習が進むにつれて, 初期値をランダムとして始めた各畳み込み層のフィルタや, 最上位層の全結合層の重みパラメータが, 次第に各識別課題に最適化されていくことになる. 性能比較を容易にするため, 各 CNN の学習は, 学習用画像セットを 150 回反復させたところで打ち切った.

3.4 比較実験

3.4.1 SIFT-BoF

CNN と比較するため, 物体認識分野においてもベースラインとして採用されている bag-of-features ベースの手法を採用した. 具体的には, 画像をグリッドに区切り密にサンプリング (サンプリング幅は, 3 ピクセルに設定) した SIFT 特徴量を抽出 (各キーポイントにつき, 17×17 サイズの近傍領域から抽出) し, それら抽出された SIFT 特徴量を 2000 個の visual words に量子化する (これらの visual words は, k-means クラスタリングにより前もって用意する). このようにして, 1 枚の画像を BoF 特徴表現に変換し, これに対し, 線形 SVM を用いてクラス分類を行う.

データセット中の画像は, 照明に使用されたシーンの一部が, 球体の背景となっている. この部分は, 物体の光沢感と関係無いため, BoF ベースの手法の認識精度に悪影響を及ぼすことが考えられる. 実際に, 背景でない部分のみから SIFT-BoF を行った場合の方が, 画像全体を用いた場合の認識精度よりも高い性能を示した. そのため, 後述する結果には, 前者の性能を示している. 一方, CNN に関しては, 両者の認識精度における明確な違いが見受けられなかったため, 画像全体を用いている.

3.4.2 人間の視覚

もう 1 つの比較対象として, 人間の視覚による光沢感認識に関する同様の実験を行った. 被験者は, 20 から 25 歳までの 4 名の男性である. 実験を開始する前に, テストセットの各クラスの画像から, *Interior*, *Nature* 各々について 8 枚ずつ, つまり, 計 $80 (= 5 \times 2 \times 8)$ 枚ランダムに選択する. 本実験では, この 80 枚の画像をテストセットとして用い

*1 <http://code.google.com/p/cuda-convnet/>でソースコードが公開されている. 詳細は, [9] を参照.

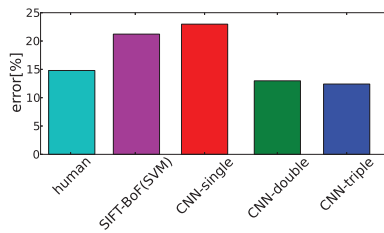


図 4 光沢感認識に関する 5 クラス分類の各手法の誤認識率. CNN-single, -double and -triple は、各々 1,2,3 段の CNN を指す.

る. 実験は、初めに被験者に 3 分間学習用画像セットを自由に閲覧させることで、5 クラスの違いを十分に理解させた後、事前に用意した 80 枚のテストセットをランダムに 1 枚ずつ見せ、それがどのクラスに属する画像であるかを答えさせる手順で行った. 回答時間は無制限とし、被験者には、回答中も自由に学習用セットを閲覧することを許可した. なお実験では画像を提示するのに、EIZO FlexScan SX2762W LCD を用いた.

4. 結果

4.1 認識精度

図 4 に、光沢感認識に対する異なる手法の誤認識率を示す. 比較する手法は、人間の視覚, SIFT-BoF(SVM), また、1, 2, 3 段の CNN (各々、用いられている畳み込み層の数が 1, 2, 3 個の CNN を指す) の 5 種類である. なお、CNN に関しては、各段の CNN ごとに、最も低い誤認識率を示したネットワーク構造のみについてその値を示している. 各々の構造は、 $C_{17}^1 P_{32}^8 N_2^1$, $C_{17}^2 P_5^2 N_4^1 C_3^1 P_3^2 N_2^1$, $C_{17}^2 P_3^2 N_9^1 C_4^1 N_9^1 P_3^1 C_3^1 N_9^1 P_3^2$ である. 結果から、1 段の CNN は、SIFT-BoF(SVM) に匹敵する性能を示したものの、人の視覚には及ばないことが分かる. また、3 段の CNN は、本実験中最高の認識精度を達成した. 2 段の CNN もこれに並ぶ性能であり、これら 2 つは、人の視覚の性能より優れている.

図 5 は、人の視覚及び、各 CNN による識別結果の混同行列を示したものである. この図から分かることは、1 段の CNN では、クラス 1 とクラス 5、クラス 3 と 4 間の区別を行う事が困難である一方、CNN を 2 段、3 段とすると、人間と同じ用にどのクラスも高い水準で識別できていることが分かる.

4.2 学習された特徴

図 6 は、各段の CNN を用いて光沢感認識課題を行った際に、1 層目の畳み込み層におけるフィルタを示したものである. 各フィルタ画像は、輝度の範囲が [0:255] となるように正規化している. これらから、CNN の段数に依らず、同心円状でかつ、その中心が明るく周囲が暗いものと、反対に、中心が暗く、周囲が明るい 2 種類のフィルタが学習されているのが分かる. これらのフィルタは、各々異なるスケ-

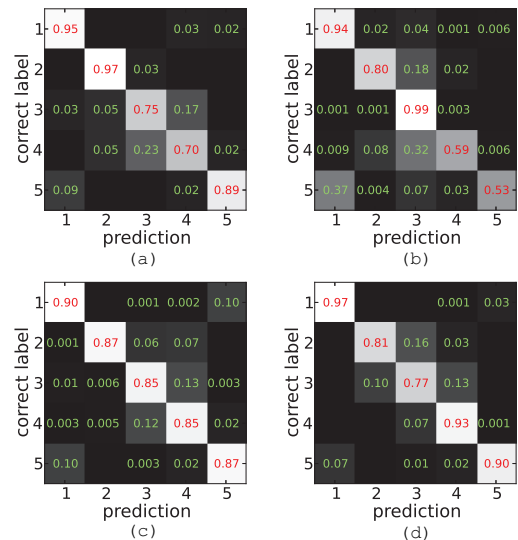


図 5 各手法で光沢感認識を行った際の混同行列. (a) 人の視覚. (b) CNN(1 段). (c) CNN(2 段). (d) CNN(3 段).

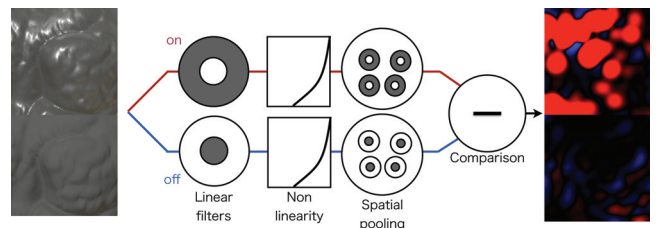


図 7 Motoyoshi らが提案した人間の視覚における光沢感知モデル [13]. はじめに、オン中心細胞とオフ中心細胞を用いて入力画像をフィルタリングし、それらの出力を非線形関数にかけた後、プーリング処理を行う. 最終的に、これら 2 つのチャンネルからの出力を統合することで、物体表面の光沢度を知ることができる.

ルに最適化されたフィルタのようにも見える.

これらのフィルタは、生体視覚システムにおける初期段階で用いられている受容野に類似している. それらは、中心が明るい方がオン中心、もう一方がオフ中心細胞と呼ばれており、入力画像のコントラスト検出に用いられていると考えられている. これらのオン中心、オフ中心細胞は、Motoyoshi らが提案した、人の視覚における光沢感認識のための神経モデル [13] においても重要な役割を果たしている. 彼らのモデルは、人が感じる光沢感と画像の輝度ヒストグラムの歪度が高い相関を示したとの心理物理実験の結果を基に作成したものであり、このことは、人は、光沢感知覚にその歪度を利用していることを示唆している. 図 7 は、画像の各局所領域の輝度ヒストグラムの歪度を求める彼らのモデルの模式図である.

我々の知る限り、上述したようなオン中心、オフ中心細胞と同様の同心円状のフィルタが NN によって学習されたとの報告はこれまでにない. これらのフィルタは、CNN により物体認識を行った場合 [9] や、自然画像に対してスパース

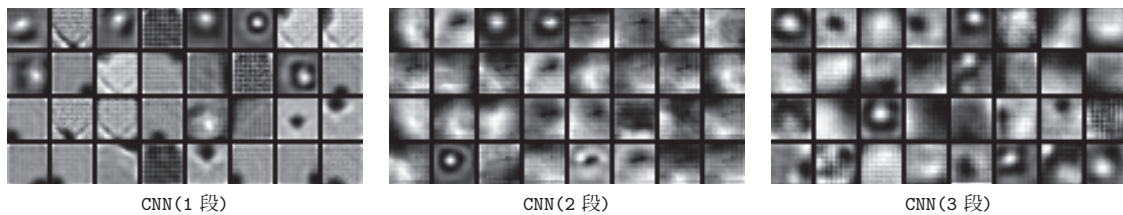


図 6 光沢感認識を行った際に 1 層目の畳み込み層で学習されたフィルタ. CNN の段数を問わず、同心円状で、中心が明るく周囲が暗いものと、反対に、中心が暗く周囲が明るいものが学習されている.

コーディングを用いた場合 [14] に学習されることが知られているガボールフィルタのようなものとは大きく異なっている.

CNN により、オン中心、オフ中心細胞に似たフィルタがなぜ学習されたのかを調べるために、追加実験を行った. 先に述べたように、*mia material* の 2 つのパラメータ (1 つは鏡面反射と拡散反射の比率, もう 1 つは, “金属らしさ” に影響するパラメータ) を変化させることで, 光沢感の 5 つのクラスを定めている. この追加実験では, これら 2 つのパラメータ各々に関する 2 クラス分類を行った. 具体的には, クラス 1, 2 間 (鏡面反射と拡散反射の比率) と, クラス 2, 4 間 (“金属らしさ”) についてである. 図 8 は, 5 クラス分類を行った際と同じ 2 段の CNN によってこれら 2 つの分類課題を行った際に学習されたフィルタである. 誤認識率は, 各々 12.1% と 6.95% であった. 他の構造を持つ CNN を用いた場合にも各々の識別課題において, 同様なフィルタが学習されたこと, また, 他のクラス対に関する識別課題, 則ち, クラス 4, 5 間 (鏡面反射と拡散反射の比率) と 1, 5 間 (“金属らしさ”) 各々について 2 クラス分類を行った際も図 8 と同様の傾向が得られたことを述べる. この結果から, 鏡面反射と拡散反射の比率に関して 2 クラス分類を行うと, オン中心, オフ中心細胞に似たフィルタが学習される一方, “金属らしさ” に関して分類を行った場合は, オフ中心細胞に似たフィルタのみが学習されることが分かる. このことから, 図 6 に見られるオン中心細胞に似たフィルタは, 鏡面反射と拡散反射との比率に関して区別する際に, また, オフ中心細胞に似たフィルタは, 2 つのパラメータに関する分類両方にとって重要な役割を果たしていると考えられる.

更に比較のため, 同じ 2 段の CNN により, 光沢感に関する 5 クラス分類を行ったときと同じデータセットを用いて, 照明に関する 2 クラス分類, すなわち, 画像が *Interior* と *Nature* どちらの照明を基に合成されたものであるかの分類実験を行った. この 2 クラス分類課題においては, 1 段の CNN でも 95 % 以上の認識精度を示した. 学習されたフィルタは, 図 8 に示した通り, エッジのようなフィルタのみであり, 同心円状のものは全く学習されなかった.

以上の結果から, CNN が, オン中心, オフ中心細胞に似

たフィルタを学習するのは, 光沢感認識特有であると考えられる. 光沢感に関する認識を行った際, これらのフィルタが, CNN の構造や様々なパラメータを変えても安定して学習されたことは注目すべきである. この傾向は, 畳み込み層における活性化関数に修正線形関数の代わりに, シグモイド関数を用いた場合にも当てはまることが確認できた. しかし, この場合, 物体認識認識における報告 [9] と同様に学習速度が著しく減少し, また, 認識精度も落ちる結果となった.

5. 結論

形状既知, 周囲の照明環境未知の条件下における単一画像からの光沢感の認識に CNN を適用する実験を行った. その結果, CNN を用いて人間の視覚に匹敵する精度で, 光沢感の認識を行うことができた. また, 高い認識率を達成するには, 2 段以上の深い構造を持つ CNN である必要があることが分かった. さらに, 光沢感認識を CNN により行った場合, 生体視覚システムや, Motoyoshi らが提案した人の光沢感知覚モデルで重要な役割を担っているオン中心, オフ中心細胞に似た同心円状のフィルタが学習されることを発見した.

以上のフィルタに関する考察は, 1 層目の畳み込み層に関してのみについてのものである. 2 段以上の構造が人の視覚に匹敵する高い認識精度の達成に必要であることから, 2 段目以降の畳み込み層で得られた特徴についての解析が重要である. Motoyoshi らが提案したモデルは, 制限付きの条件下でのみ適用出来るものであり, 人の光沢感認識に関する全ての振る舞いを説明するまでには至っていないことが指摘 [2] されている. 従って, CNN のより上層に学習されている特徴の解析を行うことが, こうした隔たりを埋めることに繋がると考えられ, 今後の課題である.

参考文献

- [1] T. Abe, T. Okatani, and K. Deguchi. Recognizing surface qualities from natural images based on learning to rank. In *Proc. ICPR*, pages 3712-3715, 2012.
- [2] B. L. Anderson and J. Kim. Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11), 2009.
- [3] Y. Bengio, A. C. Courville, and P. Vincent. Unsuper-

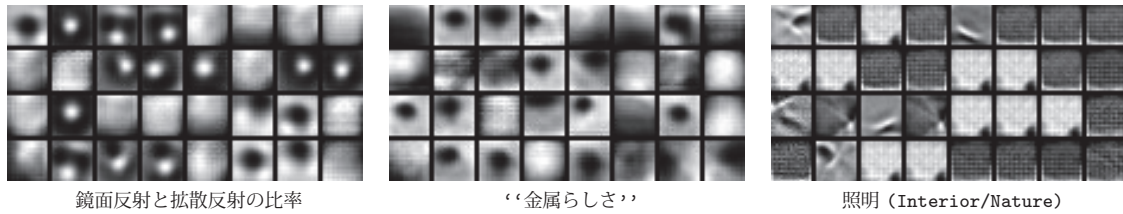


図 8 2 段の CNN を用いて 3 つの 2 クラス識別課題を行った際に学習されたフィルタ. 左から順に各々, 鏡面反射と拡散反射の比率, “金属らしさ”, 照明 (*Interior/Nature*) に関する分類を行った際に得られたものである.

vised feature learning and deep learning: A review and new perspectives. *CoRR*, abs/1206.5538, 2012.

- [4] D. C. Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In *Proc. CVPR*, 2012.
- [5] M. Crosier and L. D. Griffin. Using basic image features for texture classification. *International Journal of Computer Vision*, 88(3):447-460, 2010.
- [6] R. O. Dror, E. H. Adelson, and A. S. Willsky. Recognition of surface reflectance properties from a single image under unknown real-world illumination. In *Proc. the Workshop on Identifying Objects Across Variations in Lighting at CVPR*, 2001.
- [7] C. Hiramatsu, N. Goda, and H. Komatsu. Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *NeuroImage*, 57(2):482-494, 2011.
- [8] D. Hu, L. Bo, and X. Ren. Toward robust material recognition for everyday objects. In *Proc. BMVC*, pages 48.1-48.11, 2011.
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, 2012.
- [10] C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz. Exploring features in a bayesian framework for material recognition. In *Proc. CVPR*, 2010.
- [11] S. Lombardi and K. Nishino. Reflectance and natural illumination from a single image. In *Proc. ECCV*, 2012.
- [12] I. Motoyoshi. Highlight shading relationship as a cue for the perception of translucent and transparent materials. *Journal of Vision*, 10(9), 2010.
- [13] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson. Image statistics and the perception of surface qualities. *Nature*, 447(7141):206-209, 2007.
- [14] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607-609, 1996.
- [15] D. Parikh and K. Grauman. Relative attributes. In *Proc. ICCV*, 2011.
- [16] F. Romeiro and T. Zickler. Blind reflectometry. In *Proc. ECCV*, 2010.
- [17] R. Timofte and L. Van Gool. A training-free classification framework for textures, writers, and materials. In *Proc. BMVC*, pages 93.1-93.12, 2012.