

質感認知の情報学の進展と将来

小松英彦

The Present and Future of *Shitsukan* Research

Hidehiko KOMATSU

Shitsukan means perception of materials and surface qualities of natural and man-made objects based on visual, somatosensory and auditory sensory signals. *Shitsukan* plays important roles in our daily life, but systematic attempts to understand the mechanisms of *Shitsukan* have emerged only recently. In this article, I will introduce current situations where the problems of *Shitsukan* are actively challenged by the cooperative efforts across different fields such as technological science, psychophysics and neuroscience.

Key words: *Shitsukan*, material perception, surface reflectance, texture perception, light field

質感認知とは、視覚、触覚、聴覚などの感覚刺激をもとに、物体の材質や表面あるいは内部の状態を知覚し認識する機能である。われわれは日常生活のあらゆる場面で質感認知を行っている。例えば、われわれは物体の材質が木でできているか金属でできているかを理解したり、表面がつるつるしているかざらざらしているかを容易に見分けることができる。また、地面が濡れているか乾いているかを判断したり、魚の鮮度を見分けるのも質感認知である。そして質感認知で得られた情報を、物を触るときの手指の動きや道を歩くときの歩行運動の制御に役立てたり、物を買うか買わないかといった行動を行うための意思決定に使ったりしている。

質感認知を考える場合、対象物の特殊性に注目したアプローチと、対象物を超えた普遍性に注目するアプローチの両方が重要と考えられる。例えば、化粧品と漆のそれぞれについて、別のマーケットが存在し、それぞれの質感を追究している専門家がいます。一方、化粧品の質感も漆の質感も、ある光学特性をもつ素材によって生み出される質感という意味では、共通の土俵で考えることができる。質感認知の情報学は、後者のさまざまな物が生み出す質感をできるだけ共通の言語で理解し、一般性のある科学として取り

扱うことをめざす学問分野であるといえるだろう。もちろんその場合でも、個別の対象に目を向けることは重要である。ある対象物においてどのような仕組みで固有の質感が生み出されるかを知ることから、普遍性のある原理を導いたり、新しい原理の発見につながると考えられるからである。

個別の対象についての質感の追究は人類の歴史の中でずっと行われてきたが、その中から普遍性を取り出し体系づける試みが始められたのは比較的最近のことと思われる。特に今世紀に入って、PCの高速化やメモリーの大容量化により、視覚質感認知の研究が心理物理学や脳科学の分野でも進められるようになった。そのような状況の中で、世界的に質感をめぐる研究が活発に行われている。2010年にはわが国で文部科学省科学研究費補助金新学術領域研究の「質感認知の脳神経メカニズムと高度質感情報処理技術の融合的研究」(略称：質感脳情報学)が始まり、工学、心理物理学、脳科学の研究者が連携して質感認知の研究に取り組んでいる*1。また2012年にはヨーロッパ7か国が、EUの研究費で共同で質感の研究や若手の育成を進めるPRISMという組織を構成して、領域融合的な活動を始めた*2。このほか学術雑誌でも質感の特集号が組

自然科学研究機構・生理学研究所感覚認知情報研究部門 (〒444-8585 岡崎市明大寺町字西郷中38) E-mail: komatsu@nips.ac.jp

*1 新学術領域研究「質感脳情報学領域」ホームページ <http://shitsukan.jp/>

*2 PRISM ホームページ <http://prism-network.eu/>

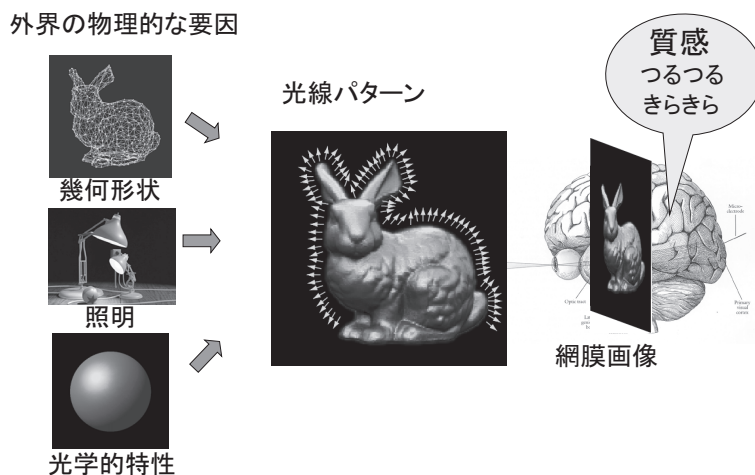


図1 質感認知に関わる外界の3つの物理要因。

まれたり、学会の分科会として質感を取り上げることが増えており、質感の研究は加速しているといえる。本稿では、このように多様な分野が相互に刺激を与えながら質感認知の情報学的研究が進んでいる状況を紹介したい。

1. Lytro の衝撃

個人的な思い出から始めさせていただきたい。2年あまり前に行われた質感脳情報学領域の班会議で、九州大の長原さんがポケットから小さな箱のようなものを出して、「これはLytroというライトフィールドカメラです」と説明を始めた。一見おもちゃのようなこのカメラは、普通のデジタルカメラができないことができる。撮影した画像の中のどこかに触れると、それまでピントがはずれてぼけていたその部分の画像がシャープにピントを結び、これまでピントの合っていた別の場所の画像がぼける、つまり撮影後の画像に焦点を当てなおすことが可能なのである*³。私の頭の中でそれまでぼんやりとしかつかめていなかったコンピューショナルフォトグラフィーの分野のイメージが、そのときくっきりと焦点を結んだ。

ライトフィールドカメラはライトフィールド、つまり光の場そのものを記録するカメラである。質感認知においては図1に示すように、外界の物理的な要因つまり物体の幾何形状、光学的特性と照明環境が相互作用して作り出す網膜画像の特徴の分析とその脳内処理、そしてその結果生み出される質感知覚の関係が分析される。しかし外界の物理的な要因が相互作用して生み出すのは光線の場であり、網膜画像はその限られた一部を切り取ってきたものに過ぎない。光線の一本一本は、そのたどってきた経路で出会った

物の素材や表面の情報を伝えている。光線の場がどのような性質をもつのか、またどのようにそれらの情報を取り出せばよいのかについて理解が進めば、質感認知の研究に新たな展望が開けるものと期待される。コンピューショナルフォトグラフィーの分野は、まさにそのような方向から質感認知の研究の発展にきわめて重要な役割を果たすと考えられる¹⁾。

すでにAdelsonとBergenは、ライトフィールドを記述する関数として、プレノプティック関数(plenoptic function)という概念を1991年に提唱している(plenopticは「完全な」を意味するplenusとopticsを合成した言葉である)²⁾。プレノプティック関数はある環境である時刻 t におけるすべての光線の場を $P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$ という7次元で表す関数で、 (V_x, V_y, V_z) は環境内の任意の点の座標、 θ, ϕ はその点を通る光の方向、 λ は波長を表す。彼らはこの関数から視覚系がさまざまな情報を取り出すときに、これらのパラメータ空間内で微分を計算することが有効であることを主張し、それらの微分演算子が実際に初期視覚皮質で見いだされているさまざまな特徴抽出細胞のフィルター特性と類似していることを述べている。視覚系の情報処理を網膜像の元になる光の場の構造にさかのぼってとらえる考え方はまだ一般的なものとはいえないが、視覚の原理から出発する考え方は、この分野に将来大きなインパクトを与える可能性を秘めていると思われる。

視覚における質感認知の研究の基本的な問題は、素材の認識と表面状態の認知である。以下では、このそれぞれについての研究の展開を眺めてみたい。

*³ Lytro ホームページ <https://www.lytro.com/>

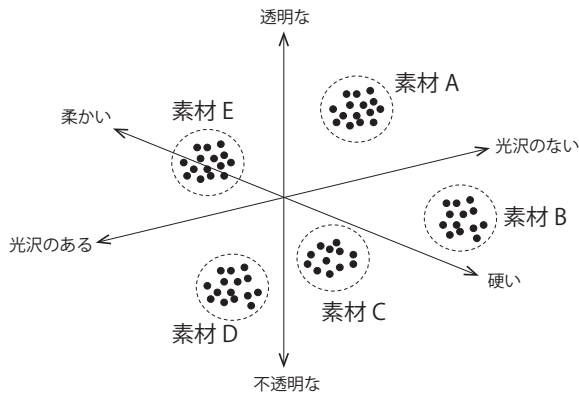


図2 さまざまな性質を軸とする多次元空間による素材の表現。図では三次元で描いているが、実際にはもっと高次元。

2. 素材の認識

われわれはさまざまな素材に共通したイメージを抱いている。例えば、金属は冷たく硬い、布は柔らかく温かい、などである。木のぬくもりのある家というようなフレーズのコマーシャルが意味をもつのは、素材に対して共通したイメージが存在することによる。さまざまな素材に対してヒトがどのようなイメージをもち、それが素材間でどのように似ていたり異なっていたりするののかという問いは、質感の基本的な問題であるが、まともにそのような問題に取り組んだ研究は意外に乏しい。イメージを定量化するひとつの方法は言語を媒介にするもので、物の性質や受ける印象を表す形容詞がどの程度当てはまるかを数値で表すというやり方である。FlemingらはFlickr Material Databaseの画像130枚を用いて、9つの性質（光沢、透明、色鮮やかさ、粗さ、硬さ、冷たさ、脆さ、自然さ、可愛らしさ）について、被験者に評価をしてもらうという方法により素材の性質の表現を調べた³⁾。その結果、被験者間で評価値の相関は0.6~0.8程度あり、被験者は各素材についてかなり共通したイメージを抱いていることが裏付けられた。9つの性質のそれぞれと評価値の関係を示すパターンは、各刺激がどのようにイメージされたかを表しているが、このパターンは同じ素材の画像間では類似しており、異なる素材の間では異なっていた。別の見方をすると、図2に模式的に示すように、それぞれの性質を軸にとった9次元の空間の別々の場所に、異なる素材の画像が雲のように集まって位置づけられるということである。

われわれはある素材を見たときに、それが光沢があるか透明であるかといった視覚的な判断だけでなく、硬いか柔らかいか、あるいは冷たいか温かいといった触覚的な性質についても判断できる。Flemingらの実験でも、画像を

見せて評価を行う性質の中には、硬さや冷たさといった触覚的な性質が含まれていた。このように質感は、見ただけで触った感じがわかるといった多感覚的な特性をもっている。これはおそらく生後の発達過程においてヒトが身の回りの物を見て触るという経験を積み重ねてきた学習の要素が大きいと想像されるが、視覚刺激に含まれる特徴と触覚刺激に含まれる特徴の関係が、生得的な知識として脳の中にある程度埋め込まれている可能性も否定できない。質感についての発達研究はまだあまり進んでいないのが現状であるが、近年活発に進められつつあり、後で述べるように興味深い成果が現れつつある。

最近、Flemingは質感研究についてのレビューの中で、質感に関する問題が、大きくはクラス分け(classification)と評価(estimation)の2つの処理に分けられるという考えを述べている⁴⁾。このうちクラス分けが素材の認識に対応すると考えられる。クラス分けで問題になるのは、クラス間の境界線を引くこと、そして個々のサンプルがどのクラスに属するかを判断することである。多様なバラエティーに富む素材の画像に対してどのように境界線を引けばよいのかというのは難しい問題である。ヒトはかなり素早く画像から素材を見分けることができる。Sharanらはインターネット上のFlickrの画像から10種類の素材を含む画像各100枚を選んでデータベース(Flickr Material Database)*⁴⁾を作り、これを利用して素材識別の実験を行ったところ、40msの短時間提示でもチャンスレベルより有意に高い成績で正しく識別できることを示した⁵⁾。

どのようなアルゴリズムを用いればさまざまな素材の画像を正しく分類できるのだろうか。コンピュータービジョンの分野では、画像中の物体やシーンを正しくラベル付けする問題について多くの研究がなされてきた。それらの研究においては、画像を輝度勾配の情報についての局所特徴量であるSIFT(scale invariant feature transform)やHOG(histograms of oriented gradients)とよばれる特徴の集まりとして記述し、それを元に物体カテゴリを識別する統計的学習手法が発達してきた⁶⁾。Flickr Material Databaseの画像について、Sharanたちはこれらを含む局所特徴量を用いて素材カテゴリの識別を行うことを試み、55.6%までの正答率を得ることに成功している⁷⁾。これは多様な素材の認識の正答率としてはかなり高いものであるが、まだヒトの識別性能には及ばない。その理由として、シーンや物体カテゴリに比べて素材の見えの変化が大きい可能性も考えられるが、素材画像データベースの規模が小さく、十分に

*⁴⁾ Flickr Material Database ホームページ <http://people.csail.mit.edu/cehu/CVPR2010/FMD/>

機械学習の識別器の訓練ができないことによる可能性もある。

近年はウェブ上に膨大な画像データが蓄積されており、それを利用した物体やシーン画像の統計的機械学習の研究が発展している。例えばHaysとEfrosは、一部が物体による遮蔽などで欠損したシーンの画像を、ウェブ上に存在する画像を使って、画像に付されたラベルを参照することなしにきわめて自然に見える完全なシーンに補完することに成功した⁸⁾。また、大規模なデータを用いて学習させた多層のニューラルネットが物体認識において高い性能を示すことが、一昨年からにわかに注目を集めるようになった⁹⁾。さらに、同じ年にNgたちのグループは、数十億のパラメータをもつ12層の大規模ニューラルネットを構築し、YouTubeの画像1000万枚を16コアのPC1000台からなるPCクラスターで教師なしで学習させたところ、ヒトの顔やネコ、あるいは人体といった高次元物体カテゴリーの画像の検出器が自動的に生成されたことを報告し、話題になった¹⁰⁾。このように、人が設計した特徴量ではなく、大規模なサンプルデータを用いることにより、高い能力をもつ識別器が統計的機械学習によって作られることが物体認識の分野で示されており、素材の識別に関してもこのような方向の研究の発展が注目される。また、計測技術の向上によって生み出される大量の高次元データから意味のある情報を効率よく抽出する手法の数理工学的研究が、スパースモデリングや圧縮センシングなどをキーワードとして近年急速に発展している⁵⁾。このような分野とうまく連携を取ることも、質感認知の情報学の発展には重要であると考えられる。

3. 素材識別の脳内処理

大脳皮質における視覚情報処理は視覚野とよばれる脳領域でなされるが、視覚野はいくつもの領域から構成される。目から入った視覚情報は網膜で神経の信号に変換され、視神経を通して外側膝状体で中継された後、大脳皮質の一番後ろに位置する第一次視覚野(V1ともよばれる)に伝えられる(図3)。一次視覚野(V1)で処理された信号は、その前方に隣接するV2野を経て前方の視覚領域に伝えられていく。これらの経路は頭頂連合野に向かう背側経路と下側頭皮質に向かう腹側経路に解剖学的に区別され、両者は機能的にも違う役割を果たしていることが示されている。背側経路は空間や運動に関連した情報処理を行うの

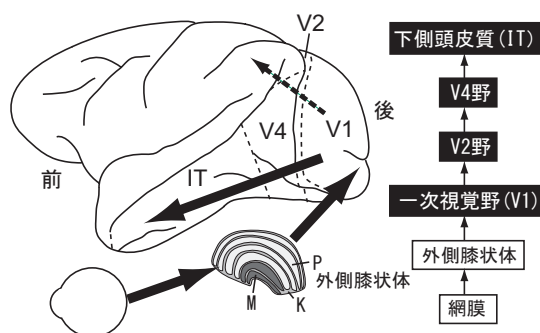


図3 サルの視覚神経系の模式図。目から外側膝状体で中継された視覚情報は脳の一番後ろにある第一次視覚野(V1)に伝えられる。大脳皮質は階層的に構成されたいくつもの視覚領域から成り、2つの経路に分かれて視覚情報は処理される。1つの経路は頭頂連合野に向かう背側経路(破線矢印)であり、もう1つは下側頭皮質に向かう腹側経路(実線矢印)である。右の線でつないだブロックは腹側経路に含まれる脳領域を示す。

に対し、腹側経路は物体認識に関係する処理を行っている⁶⁾。

このうち、腹側経路が質感認知に関係することがいくつかの研究で報告されている。平松と郷田らは、素材のCG画像をヒトに見せたときに、視覚野のどの部分が活動し素材の種類を見分けているかを、機能的磁気共鳴画像法(機能的MRI; fMRI)を用いて調べた¹¹⁾。その結果、V1から紡錘状回に至る腹側経路の領域において、素材の種類によって局所的な活動分布のパターンが変化し、素材を見分けていることがわかった。次に彼らは、これらの活動が素材についてのどのような情報を表現しているのかを、2通りの仮説に基づいて調べた。1つの仮説は、ニューロン活動が素材間での比較的単純な画像特徴の違いを反映しているというものである。そこで、各素材の画像ごとに局所の方位、空間周波数成分の大きさ、および平均的な輝度と色についての計20個の画像特徴を求め、素材による脳活動の違いが画像特徴の違いとどの程度相関しているかを調べた。もう1つの仮説は、ニューロン活動が素材間での知覚される印象の違いを反映しているというものである。知覚される印象を定量化するために、〈光沢あり-光沢なし〉〈透明-不透明〉〈単調-複雑〉〈冷たい-温かい〉〈硬い-柔らかい〉など12の形容詞対について、各画像を5段階で被験者に評価してもらった。そして、このように定量化した素材による印象の違いが、素材に対する脳活動の違いとどの程度相関しているかを調べた。すると、V1, V2といった初期視覚野の脳活動は、画像特徴と高い相関を示す一方、知

⁸⁾ 新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成領域」ホームページ <http://sparse-modeling.jp/>

⁶⁾ 大脳皮質視覚野における情報処理の概説 <http://www.nips.ac.jp/scinfo/kaisetsu.htm>

覚的な印象とは相関が低いことがわかった。一方、腹側経路の高次領野である紡錘状回付近の活動は、逆に知覚的な印象と相関が高く、単純な画像特徴とはあまり相関がみられなかった。この結果は、腹側視覚経路において素材識別が行われていること、また腹側経路での処理の過程で素材の印象が形成されることを示している。また興味深いことに、知覚的印象を評価するときに用いた形容詞には、視覚的な形容詞だけでなく〈冷たい-温かい〉〈硬い-柔らかい〉といった触覚的な形容詞も含まれるが、それらだけを取り出して解析を行っても同様な結果が得られた。このことは、腹側高次視覚野では、視覚と触覚の情報が統合された多感覚的な形で素材が表現されている可能性を示している。郷田らは最近、ヒトで用いたのと同じ素材画像をサルに見せて、fMRI を用いて脳活動を調べる実験を行った結果、ヒトと同様に腹側視覚経路が素材を見分ける活動を示し、また画像特徴から知覚的な印象への変換がこの経路で起こっていることを見いだしている¹²⁾。

生体において質感認知がどのように行われているかを知ることは、質感認知の生物学的な仕組みを明らかにするという意義のほかに、生体のすぐれた質感認知の機能の仕組みを知ることによって新しい情報処理アルゴリズムを生み出す上でのヒントを得るという側面がある。視覚野の単純型細胞に類似した特徴検出細胞と、複雑型細胞に類似した空間的な不変性をもつ細胞を、階層的に複数回積み上げたニューラルネットはその一例であり、パターン認識にすぐれた能力をもつことが示されている¹³⁾。

素材認識に深く関係すると考えられる画像処理は自然テクスチャーの識別である。布、毛、石、樹皮、皮革のように、多くの素材は特有の自然テクスチャーをもっている。これらのテクスチャーの識別を行うアルゴリズムについても、多くの工学的研究がなされている。その中で、生体における素材認識機構を考える上で参考になると思われるのは、テクスチャー処理の初期段階でガボールフィルター様の特徴検出器で画像を局所の方位と空間周波数の成分に分解するという、視覚野の処理と同様の処理を行うアルゴリズムである。

サルやネコなどの哺乳動物の一次視覚野 (V1) には、特定の方位と空間周波数の刺激に選択的に反応する細胞が存在し、そのうち単純型細胞とよばれる細胞の受容野は、ガボール関数によく似た空間構造をもつことが知られている。一方、複雑型細胞とよばれる細胞は、位相が 90 度ずつづれた複数の単純型細胞の出力を組み合わせたものとして表される。したがって、素材画像の自然テクスチャーは、これらの細胞によって局所の方位と空間周波数に分解

されて表現されることになる。生体における素材識別の仕組みを考える上で重要なことは、V1 で表現されている局所の方位や空間周波数の信号をどのように利用すればよいのかということになる。これについて重要な手がかりを与えると考えられるのは、自然テクスチャーを合成するために開発されたアルゴリズムである。Portilla と Simoncelli が開発したアルゴリズムでは、テクスチャー画像をガボール様のフィルターで局所の方位と空間周波数の信号に分解した後、それらの成分の空間位置間の相関、異なるスケールのフィルター出力間の相関、異なる方位のフィルター出力間の相関などを元画像について計算し、各統計量が元画像と同じ値をもつように新しい画像を合成していくことによって、元画像と同じように知覚される自然テクスチャーを合成することに成功した¹⁴⁾。素材テクスチャーの識別を問題にしているときに、どうしてテクスチャーの合成の話をするのか疑問に感じられるかもしれないが、それは以下のような理由による。例えば 2 本の松の木の樹皮を比較してみると、同じようなテクスチャーをもち、いずれもまぎれもない松の木の樹皮であるが、細部を比較すると、ピクセルの明るさや細かいパターンは両者で全く異なっている。つまりテクスチャー知覚において重要なことは、テクスチャー画像がもつ何らかの統計的な性質をヒトは検出しており、それが同じであればヒトには同じテクスチャーとして知覚されるということである。Portilla と Simoncelli のアルゴリズムでは、元の自然テクスチャー画像をガボール様のフィルターに通したときの出力を出発点として、それらを組み合わせて作られる高次統計量を元画像に一致させることにより、ヒトの目に同じように知覚される自然テクスチャーを合成することに成功していた (図 4)。したがって、このアルゴリズムで用いられた高次統計量が視覚野の処理において計算されている可能性が考えられるということである。

近年、このような可能性を生理学的に検証する試みが始まっている。Freeman たちは Portilla と Simoncelli の統計量をすべて用いて合成した自然テクスチャー画像と、ガボール様フィルターの出力のみを使って合成した画像の 2 種類の画像を用いて、それらをサルに見せたときの V1 と V2 野のニューロン活動を比較した¹⁵⁾。その結果、V1 では 2 種類の画像に対する応答に差がなかったのに対して、V2 では前者に対して有意に強い応答が得られた。V1 で応答に差がなかったことは、V1 がテクスチャー画像に含まれる局所の方位と空間周波数のみを検出しているという考えと一致した結果である。一方、V2 において自然テクスチャー画像でより強い応答が得られたことは、V2 におい

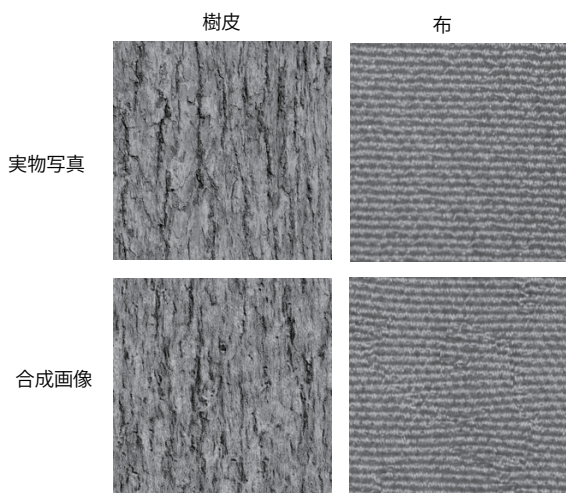


図4 素材テクスチャーの合成. Portilla と Simoncelli のテクスチャー合成アルゴリズムによる樹皮 (左) と布 (右) の合成例. 上が実物で下が合成した画像.

て Portilla と Simoncelli のアルゴリズムで用いられた高次統計量の計算の少なくとも一部が行われている可能性を示すものである. Freeman らの実験では個々のニューロンがどのような統計量を検出しているかについては示されていないが, 最近岡澤らはこの問題に答えを与える実験を V2 の次の段階であるサル V4 野で進めている¹⁶⁾.

これらの研究は, 素材識別の問題を超えて, 腹側経路における視覚情報処理の研究に新しい発展をもたらす可能性がある. これまで多くの研究によって, 下側頭皮質のニューロンは顔や体あるいは物体画像に含まれる複雑なパターンに対して選択性をもつことが明らかにされてきた¹⁷⁾. しかし, V1 で取り出された方位や空間周波数の情報から下側頭皮質でみられるそのような複雑な選択性が, 視覚皮質のどのような処理で生み出されるのかについては, ほとんどわかっていなかった. 素材テクスチャーに関する最近の研究の発展は, 対象としている画像特徴がはっきりしており, 出発点となる V1 の選択性との関係も明瞭である. そのため, これまで謎のまま残っていた中間段階での処理が, これらの研究をきっかけにして進む可能性があると考えられる.

4. 表面状態の認知

次に, 質感認知のもうひとつの側面である表面状態の認知に目を転じたい. 視覚による表面状態の認知に関して最も多く研究がなされてきたのは, 光沢についてであろう. 光沢は物体表面の反射特性と密接に関わっている. 光沢は実用的な重要性が高いため, 工学的に物体の反射特性を特定の測定条件に限定して定量化する試みはさまざまに行わ

れてきたが, 一般性のある形で反射特性を記述する方法は, Nicodemus らが用いた双方向反射率分布関数 (bidirectional reflectance distribution function; BRDF)¹⁸⁾ として, 照明光と反射光をそれぞれすべての方向の組み合わせについて計測する方法であり, 今日一般的な方法として用いられている.

図1をもう一度見ていただきたい. 網膜画像は, 物体の幾何形状, 光学的特性と照明という3つの要因が相互作用してできる光の場が, 網膜に結像することによって生じることを述べた. 質感認知はこの矢印を逆にした過程として考えることができ, 網膜画像を元にその元になる物体の光学的特性を推定する機能と考えることができる. いま光学的特性として物体表面の反射に限定して考えると, これは画像から BRDF を推定する問題ということになる. このように, 画像を元にその原因となる物理的な要因を推定する問題は, 工学分野で研究されてきた. 従来, 3つの物理的要因のうち2つを既知として残りの1つを推定する研究が行われてきた¹⁹⁾ が, 近年, より制約の少ない条件で画像から外界の要因を推定する試みが行われている. そのひとつのきっかけは, さまざまな素材の BRDF や自然照明環境の計測が行われ, それらの性質の分析が進んだことが挙げられるであろう. BRDF や自然照明環境のもつ統計的な性質を逆問題を解くにあたっての拘束条件として利用することにより, 物体の反射特性を画像から推定することが成功を収めつつある^{20,21)}. これらの研究は現段階では脳の視覚系が行っている情報処理とどのように関係するのかは不明であるが, 光沢知覚においても自然な照明環境のもつ統計的な性質が重要であることが示されており^{22,23)}, 工学的研究で用いられているアルゴリズムが知覚のメカニズムを理解する上で手がかりを与える可能性がある.

光沢知覚の心理物理学的研究は以前より行われていたが, 近年, PC の高速化や CG のプラットフォームの整備により, さまざまな反射特性をリアルに模擬した刺激を比較的容易に作れる環境が整い, 研究が発展している. また BRDF は 6次元の関数であり, 1つのサンプルだけでも膨大なデータとなるが, 現実の素材は可能な BRDF 空間のごく一部を占めるに過ぎず, CG のプログラムでは少数のパラメータを用いて反射特性を近似してリアルな画像を作ることが一般に行われる. 図5A はよく用いられる3つのパラメータで反射をモデル化する方法の各パラメータを表したもので, 1つは方向性をもたない反射成分 (拡散反射), 第2は強い方向性をもつ反射成分 (鏡面反射), 第3は物体表面の微小な凸凹によって生じる鏡面反射の広がりである.

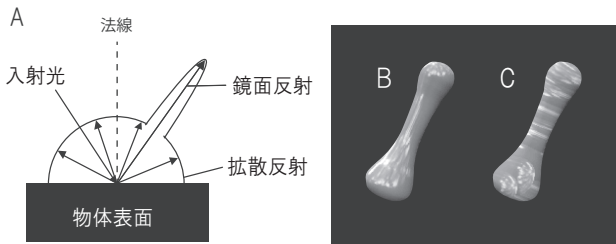


図5 A: 鏡面反射と拡散反射の模式図。拡散反射は入射した光がすべての方向に一樣に反射する成分、鏡面反射は入射光が法線方向をはさんで反対の方向に反射する成分。鏡面反射は方向性が高いが、物体表面の微小な凹凸により反射方向に広がりが生じる。B: 鏡面反射の強い物体画像のCG画像、C: Bの物体のハイライトを回転して作った画像、Cでは光沢感が減少している。

これらの物理的パラメーターが光沢知覚とどのように関係しているかを調べるために、Ferwardaらは上の3つの反射パラメーターのさまざまな組み合わせで作ったCG物体画像を被験者に提示し、光沢知覚の強さを説明する変数を求める実験を行った²⁴⁾。その結果、2つのパラメーターが知覚される光沢の強さと相関することを見いだした。1つは拡散反射率と鏡面反射率を非線形に組み合わせた変数であり、これは物体表面のハイライトの部分とそれ以外の部分の明るさのコントラストを表す。もう1つの変数は鏡面反射の広がり度で、これは物体表面に映りこんだ環境のイメージの鮮明さを表す量である。

5. 光沢知覚のメカニズム

光沢に関する脳活動を調べた研究では、素材識別と同様腹側視覚経路において光沢をもつ刺激に対する強い活動が見いだされた²⁵⁾。また、サルの下側頭皮質には光沢を見分ける細胞が存在しており、それらのニューロン集団の活動は、さまざまな光沢を系統的に表現していることが示されている²⁶⁾。これらの生理学的な研究は、視覚系が画像中のどのような情報を手がかりとして光沢を知覚しているかという問題については、今のところ答えを与えることはできていない。心理物理学的なアプローチにより、この問題についての研究が現在活発に進められている状況である。視覚系がどのように光沢の判断を行っているかという問題に関して、大きく2つの異なる考え方が存在する。1つの考え方は、光沢の検出はそれに関係する外界の物理的な要因、つまり光沢の場合であれば反射パラメーターの推定を行っているというものである。この考え方を逆光学説とよぶことにする。物体の画像は反射特性だけでなく、物体の幾何形状や照明環境の要因も混じって作られるので、反射特性を推定するためには幾何形状や照明環境の要因を差し引く必要がある。これは、色覚において色の恒常性の計算

を行うためには、照明光の影響を何らかの方法で差し引く必要があることと類似の問題設定である。画像からの反射特性の推定に関しては、前章で述べたように、工学的にはある程度成功しつつある。もう1つの考え方は、光沢の検出において視覚系が行っていることは、他の要因を分離して反射特性を推定することではなく、照明環境によらず反射特性と相関する画像特徴を取り出すことであるというものである。この考え方を、ここでは画像特徴説とよぶことにする。西田と新谷は、凸凹した表面のCG画像を用いた実験において、ヒトの光沢の判断が同じ反射特性でも表面の形状に依存して変化することを見いだすとともに、画像の輝度ヒストグラムが光沢の判断に関係することを示唆した²⁷⁾。その後、本吉らは、物体画像の輝度やその周波数サブバンドのヒストグラムの歪度が光沢知覚に重要であることを報告した²⁸⁾。逆光学説と画像特徴説は光沢知覚を舞台にして競い合っているが、これは視知覚の本質に関わる問題点をめぐる議論であり、質感認知の取り扱う問題に視知覚の本質的な問題が凝集していることを示すものであるといえる。

その後の研究で、輝度ヒストグラムの歪度だけでは光沢知覚が説明できない例が報告されているが、その1つは、ハイライトを画像中で回転すると輝度ヒストグラムは変化しないにもかかわらず知覚される光沢の強さが低下するというものである²⁹⁾(図5B, 5C)。これは、ハイライトが光沢知覚の手がかりとなるためには、物体表面の三次元形状によって生じる陰影とハイライトのパターンに整合性が必要であることを示唆している。光沢を中心とした質感研究の発達研究は中央大の山口らのグループが精力的に進めており、楊らは、ヒトには生後7~8か月で光沢のある物体とない物体を見分ける能力があらわれることを報告している³⁰⁾。興味深いことに、乳児はすでにこの段階で正しいハイライトと画像中で回転したハイライトを区別しており、このことは、自然環境におかれた外界の物体が作る画像の性質について、このような早い発達段階ですでに正しい知識をもっていることを示唆している。

Flemingは最近のレビューの中で上記の2つの考え方(逆光学説と画像特徴説)を取り上げ、それぞれの問題点を指摘した後、彼が「統計的見えのモデル」(statistical appearance model)とよぶ第三のアプローチについての主張を展開している⁴⁾。それによると、視覚系が行っていることは、表面の性質が変わったときに画像がどのように変化するかについての統計的な規則性を把握し、新しいサンプルが与えられたときにその規則性を利用して特定の質感の変化として知覚するというものである。まだラフな構想

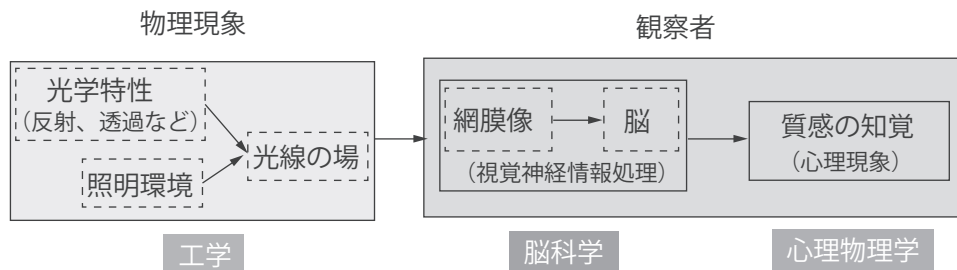


図6 質感認知の情報学にも関係する3つの学問分野（工学，脳科学，心理物理学）と，それらがおもに対象とする視覚質感認知の3つの段階（物理現象，脳情報処理，知覚）の概念図。

の段階であり，今後具体的な例に適用して肉付けが行われるとともに理論としての有効性が検証されていく必要があるが，新しい発展が期待される注目すべき考え方のように思われる。

表面状態の認知については光沢のみを取り上げてみてきたが，BRDFが取り扱えるのは物体表面での反射だけである。しかし，世の中に存在する素材は多かれ少なかれ光を透過する性質をもち，物体内部に光が侵入して散乱を受けたり，物体内部に存在する異なる媒質間の界面で反射したり，あるいは物体の他の面から物体外に出ていく。それらの光学過程によりライトフィールドは複雑に変化し，また波長分布にも変化が生じる。現実の物体がもつ多様な質感に対する認知を対象とするためには，それらの素材の光学的な特性の計測や解析がまず必要であり，それをめざした工学的な研究が進められている³¹⁾。本稿の最初に触れたコンピュータショナルフォトグラフィーの分野においても，このような現実世界で生み出される複雑なライトフィールドから情報を取り出す研究がさまざまに進められている。例えば長原らは透明物体を透過して観察される背景のひずみ方が透明物体の形状に依存することに注目して，ライトフィールドカメラが取得する多視点画像を解析することにより，背景によらず透明物体の形状を認識する興味深い研究を行っている³²⁾。心理物理学の分野においても，半透明感^{33,34)}や流体の粘度³⁵⁾など，光沢以外の質感を生み出す画像手がかりについての先駆的な研究が進められつつある。

質感認知がさまざまな産業において興味をもたれる大きな理由は，質感認知が嗜好や価値判断に密接につながっているからである。ヒトの嗜好や価値判断を分析して商品開発に結び付ける研究は感性工学の分野で多くなされてきたが，質感認知の情報学の発展により，ヒトの心が生み出す知覚の成り立ちの仕組みを明らかにすることが可能になる。一方で，質感認知を感性と結びつけていくためには，

感性工学でつちかわれた感性評価のさまざまな方法が参考になるものと思われる。

質感認知の情報学は，物と光の相互作用の仕方，それにより生じるライトフィールド，網膜画像の形成，そして脳内情報処理を経て知覚へと至る一連の流れを系統だてて解明することをめざしている。それにより，世界とわれわれがどのように関係しているのかという人類にとっての大きな問題を，より深いレベルで理解できるようになるだろう。それは，世界を構成する多様な事物がわれわれの心にもたらす豊かさの源泉に迫っていく，ひとつのメインルートとなるに違いない。工学と心理物理学と脳科学の連携を中心にすえ，諸科学が交流する場がさらに活発に広がってゆき，質感認知の情報学が発展することを願っている（図6）。

文 献

- 1) 日浦慎作：“コンピュータショナルフォトグラフィ理解のための光学系入門”，コンピュータビジョン最先端ガイド4—CVIM チュートリアルシリーズ—，八木康史・斉藤英雄編（アドコム・メディア，2011）pp. 23-52.
- 2) E. H. Adelson and J. R. Bergen: “The plenoptic function and the elements of early vision,” *Computational Models of Visual Processing*, ed. M. S. Landy and J. A. Movshon (MIT Press, Cambridge, 1991) pp. 3-20.
- 3) R. W. Fleming, C. Wiebel and K. Gegenfurtner: “Perceptual qualities and material classes,” *J. Vision*, **13**(8) (2013) 9:1-20.
- 4) R. W. Fleming: “Visual perception of materials and their properties,” *Vision Res.*, **94** (2014) 62-75.
- 5) L. Sharan, R. Rosenholtz and E. Adelson: “Material perception: What can you see in a brief glance?” *J. Vision*, **9**(8) (2009) 784.
- 6) 柳井啓司：“一般物体認識の現状と今後”，情報処理学会論文誌，コンピュータビジョンとイメージメディア，**48** (2007) 1-24.
- 7) L. Sharan, C. Liu, R. Rosenholtz and E. Adelson: “Recognizing materials using perceptually inspired features,” *Int. J. Comp. Vis.*, **103** (2013) 348-371.
- 8) J. Hays and A. A. Efros: “Scene completion using millions of photographs,” *Proc. of SIGGRAPH* (2007).
- 9) A. Krizhevsky, I. Sutskever and G. E. Hinton: “ImageNet classification with deep convolutional neural networks,” *Proc. of Neural Information Processing Systems (NIPS)* (2012).
- 10) Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S.

- Corrado, J. Dean and A. Y. Ng: "Building high-level features using large scale unsupervised learning," *Proc. of International Conference on Machine Learning (ICML)* (2012).
- 11) C. Hiramatsu, N. Goda and H. Komatsu: "Transformation from image-based to perceptual representation of materials along the human ventral visual pathway," *NeuroImage*, **57** (2011) 482–494.
 - 12) N. Goda, A. Tachibana, G. Okazawa and H. Komatsu: "Representation of the material properties of objects in the visual cortex of nonhuman primates," *J. Neurosci.*, **34** (2014) 2660–2673.
 - 13) K. Fukushima: "Neocognitron: "A self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, **36** (1980) 193–202.
 - 14) J. Protilla and E. P. Simoncelli: "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comp. Vision*, **40** (2000) 49–71.
 - 15) J. Freeman, C. M. Ziemba, D. J. Heeger, E. P. Simoncelli and J. A. Movshon: "A functional and perceptual signature of the second visual area in primates," *Nat. Neurosci.*, **16** (2013) 974–981.
 - 16) G. Okazawa, T. Tajima and H. Komatsu: "Natural texture selectivity of macaque V4 neurons examined by adaptive sampling," *Abst. for 43rd Soc. Neurosci. Meeting* (2013) 555.06.
 - 17) K. Tanaka: "Inferotemporal cortex and object vision," *Annu. Rev. Neurosci.*, **19** (1996) 109–139.
 - 18) F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg and T. Limperis: "Geometrical considerations and nomenclature for reflectance," *National Bureau of Standards Monograph 160* (1977).
 - 19) 佐藤洋一, 向川康博: "インバースレンダリング", 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM) **91** (2004) 65–76.
 - 20) F. Romeiro and T. Zickler: "Blind reflectometry," *European Conference on Computer Vision (ECCV)* (2010).
 - 21) S. Lombardi and K. Nishino: "Reflectance and natural illumination from a single image," *European Conference on Computer Vision (ECCV)* (2012).
 - 22) R. W. Fleming, R. O. Dror and E. H. Adelson: "Real-world illumination and the perception of surface reflectance properties," *J. Vision*, **3** (2003) 347–368.
 - 23) I. Motoyoshi and H. Matoba: "Variability in constancy of the perceived surface reflectance across different illumination statistics," *Vision Res.*, **53** (2012) 30–39.
 - 24) J. Ferwerda, F. Pellacini and D. P. Greenberg: "A psychophysically-based model of surface gloss perception," *Proc. SPIE*, **4299** (2001) 291–301.
 - 25) G. Okazawa, N. Goda and H. Komatsu: "Selective responses to specular surfaces in the macaque visual cortex revealed by fMRI," *NeuroImage*, **63** (2012) 1321–1333.
 - 26) A. Nishio, N. Goda and H. Komatsu: "Neural selectivity and representation of gloss in the monkey inferior temporal cortex," *J. Neurosci.*, **32** (2012) 10780–10793.
 - 27) S. Nishida and M. Shinya: "Use of image-based information in judgments of surface-reflectance properties," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.*, **15** (1998) 2951–2965.
 - 28) I. Motoyoshi, S. Nishida, L. Sharan and E. H. Adelson: "Image statistics and the perception of surface qualities," *Nature*, **447** (2007) 206–209.
 - 29) J. Kim, P. Marlow and B. L. Anderson: "The perception of gloss depends on highlight congruence with surface shading," *J. Vision*, **11**, No. 9 (2011) article 4, 1–19.
 - 30) J. Yang, Y. Otsuka, S. Kanazawa, M. K. Yamaguchi and I. Motoyoshi: "Perception of surface glossiness by infants aged 5 to 8 months," *Perception*, **40** (2011) 1491–1502.
 - 31) 向川康博: "反射・散乱の計測とモデル化", コンピュータビジョン最先端ガイド4—CVIMチュートリアルシリーズ, 八木康史, 斉藤英雄編 (アドコム・メディア, 2011) pp. 121–150.
 - 32) K. Maeno, H. Nagahara, A. Shimada and R. Taniguchi: "Light field distortion feature for transparent object recognition," *Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 2786–2793.
 - 33) R. Fleming and H. H. Bulthoff: "Low-level image cues in the perception of translucent materials," *ACM Trans. Appl. Percept.*, **2** (2005) 346–382.
 - 34) I. Motoyoshi: "Highlight-shading relationship as a cue for the perception of translucent and transparent materials," *J. Vision*, **10**, No. 9 (2010) article 6, 1–11.
 - 35) K. Maruya, T. Kawabe and S. Nishida: "Material from motion: Human perception of fluid properties from motion vector fields," *J. Vision*, **13**, No. 9 (2013) article 207.

(2014年2月24日受理)