

## 画像と言語を用いた質感情報表現のディープラーニング

研究代表者 岡谷 貴之（東北大学大学院情報科学研究科・教授）

研究分担者 山口 光太（東北大学大学院情報科学研究科・助教）

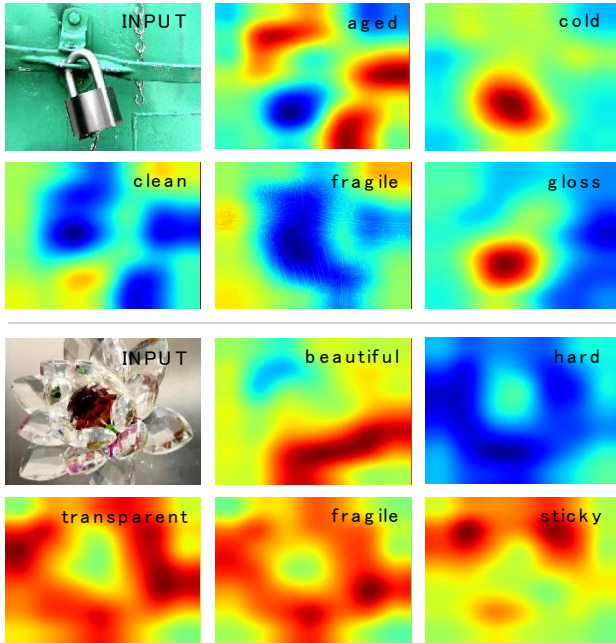


図1. 画像のどの部分がどういう形容詞で表現されるかの表示. 左上の画像に対し, 各形容詞の程度が強い場所が赤く, 弱い場所が青く表示されています.

### ○ 研究の背景

人は画像に写るシーンや物体を見て, 様々なことを理解し, 感じ取ることができます. 同じことをコンピュータで行えるようにすることは, 人工知能(AI)研究の基本課題です. 最近の研究の進展により, この課題は解決へ向けて大きく前進しました. しかしながら, 画像に写る物の質感を人と同じように理解することは, 未だに AI には難しいことです. 本研究では, 視覚による質感理解が人にとってたやすく, かつ日常の生活に欠かせないことを踏まえ, 質感を認識するシステムを実現しようとしています.

### ○ この研究の目指すもの

画像からそこに写る物の質感をコンピュータが認識するシステムの実現を目指し, 大規模な画像データを用いたディープラーニング(深層学習)により, 質感情報の表現空間を構築しようとしています. 最近の AI の進展は, 大量のデータを用いた機械学習に支えられています. 質感認識の難しさは, 同じような学習用のデータを得ることが容易でないことです. 質感とは, 人でもしばしば言語化が困難な, 言

葉(意味)を超えた心理的な表象であり, 個人差や主観の排除も難しいからです.

そこで本研究では, 物体カテゴリ認識など, 大量のデータを容易に準備できるタスクをディープニューラルネットに学習させ, そこで獲得される内部表現(≒入力画像に対する中間層の活性化パターン)をマイニングし, 質感を表現する特徴の空間を構築しようとしています. 核となるアイデアは, 多層のニューラルネットワークに質感認識の「代理タスク」を学習させ, 獲得される内部表現を元に, 質感の表現空間を構築することにより, 質感表現のマイニングを行うというものです.

言葉で表しにくいゆえに人と人, あるいは人と機械の間での伝達が難しかった質感を「数値化」し, 工業デザインの効率化(=製品の試作や実物目視の必要性を排除)や, ウェブの検索サイト等での情報取得や, E-コマースでの商品推薦の精度向上を狙います.

### ○ これまでに得られた成果

画像1枚からそこに写る物の質感を, いくつかの形容詞について認識するシステムを実現しました. 具体的には, 質感に関わる形容詞を予め選んでおき, 入力画像からその形容詞の尺度で人が受ける印象を再現します(図1). このシステムは, サンプル画像群に対し, 人が各形容詞の尺度を与えたデータを元に, 深層学習を用いることで実現しています.

また, ウェブの大量データを用いて, 質感を表現する語彙を自動的に発見するシステムを実現しました. インターネットで様々な商品を販売するサイトを対象に, そこに掲載されている商品の画像およびその紹介文(テキスト)を使います. 商品の画像の内容と, 商品の紹介文に登場する語彙との間の相関を調べることで, 画像の内容(=質感)をよく説明する語彙を選び出します. 画像の内容は, 上述の画像特徴を用いて取り出しています.

### ○ 関連する研究発表

1. Vittayakorn S, Umeda T, Murasaki K, Sudo K, Okatani T, Yamaguchi K, Automatic Attribute Discovery with Neural Activations, In Proc. European Conference on Computer Vision, 2016.
2. Liu X, Ozay M, Zhang Y, Okatani T, Learning Deep Representations of Objects and Materials for Material Recognition, Vision Sciences Society Annual Conference, 2016