

画像内容に基づく検索技術に 対する期待と現実

北本 朝展（きたもと あさのぶ）

国立情報学研究所 / 総合研究大学院大学

<http://agora.ex.nii.ac.jp/~kitamoto/>

画像検索とは？

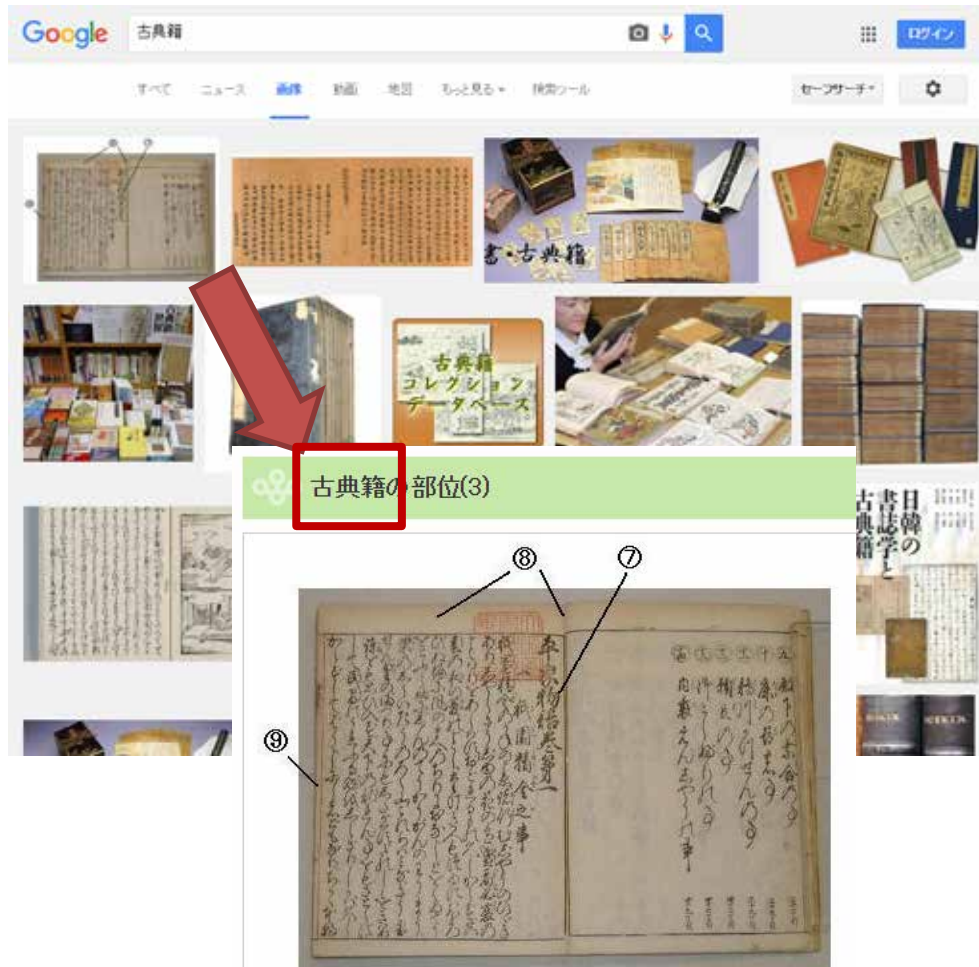
- 画像データベースの中から、利用者が求める画像データに到達するための技術。
- **入力**：キーワード（タグ）、テキスト、画像
- **技術**：インデックス、特徴表現、類似度（距離）定義、フィードバック...
- **出力**：画像、テキスト
- 網羅的に見るのと比べ、十分効率的か？

画像検索の主な方式

1. **テキストを入力**し、画像に付与された統制語やタグなどを検索。
2. **テキストを入力**し、画像に関連付けられたテキストを検索。
3. **画像を入力**し、画像から抽出したエンティティを検索。
4. **画像を入力**し、類似した画像を検索。
5. **テキストを入力**し、画像の自動アノテーションで得られたテキストを検索。

ウェブ画像検索

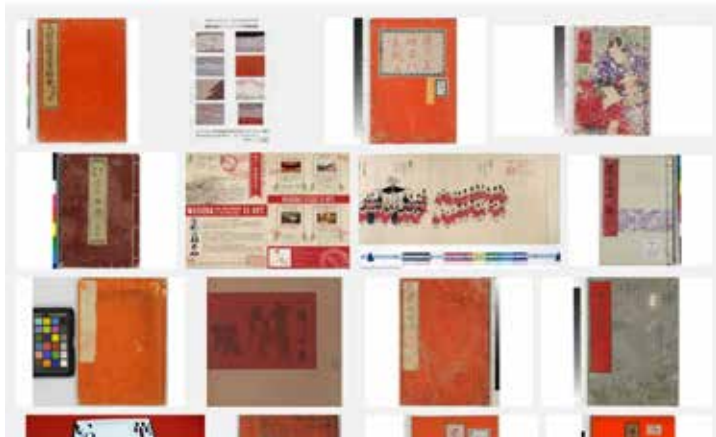
Google画像検索（標準）



- 画像の周辺に存在するテキストを検索する。
- **画像そのものを分析しているわけではない。**
- 簡単な画像解析オプションはある。

<http://www.library.pref.osaka.jp/site/osaka/kosho-bui.html>

Google画像検索（色）

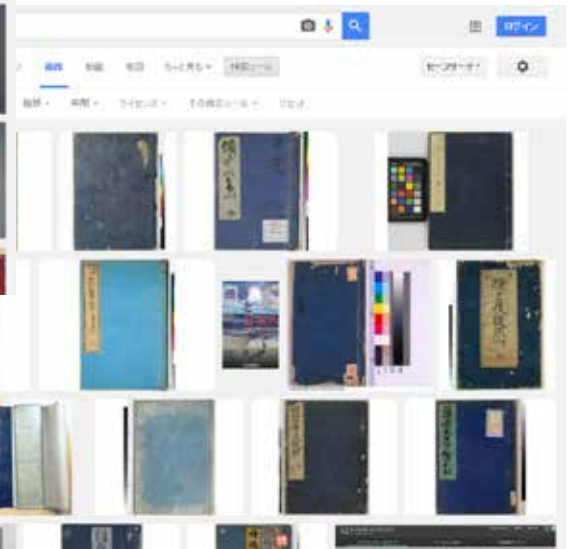


緑



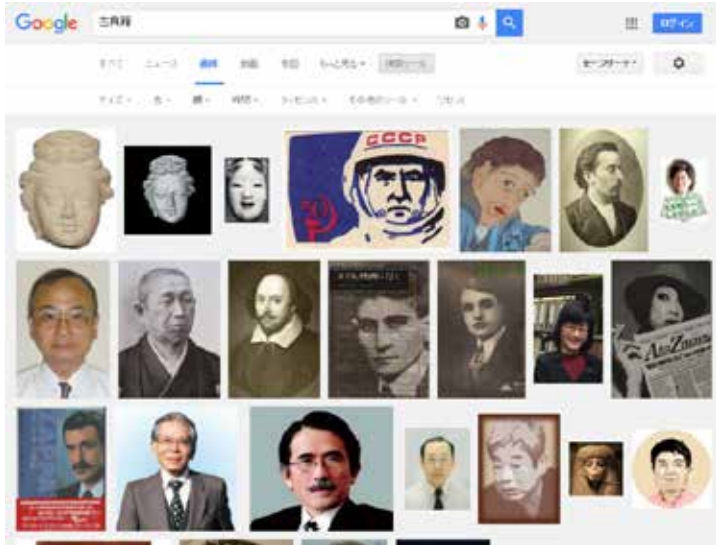
青

赤



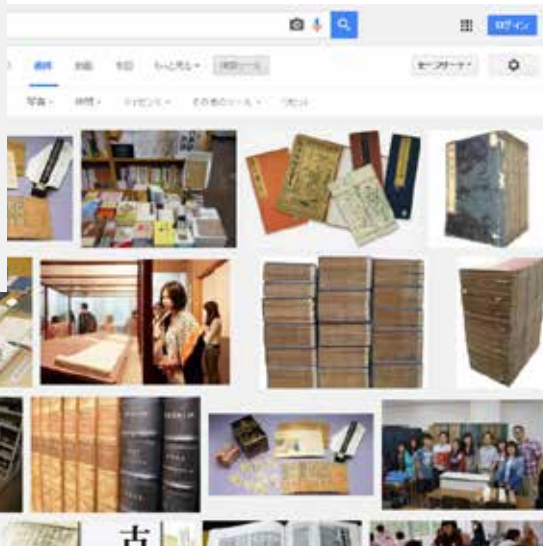
検索オプションに色を指定すると、色に応じて検索結果が変化する。

Google画像検索（種類）



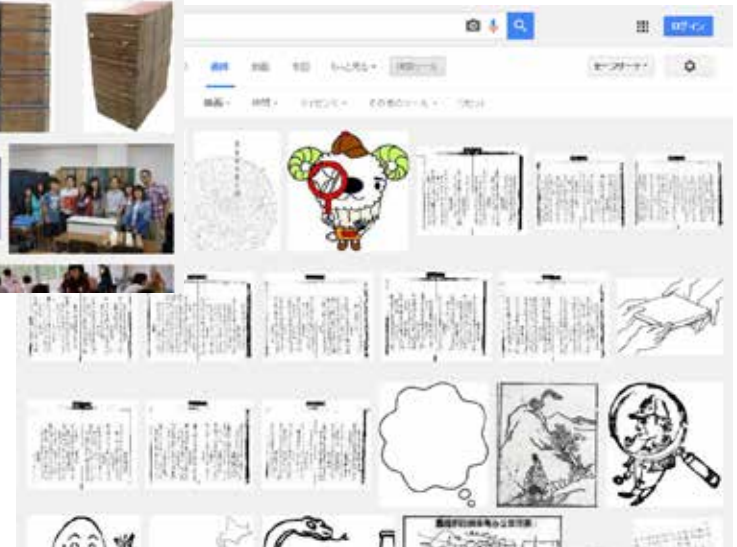
顔

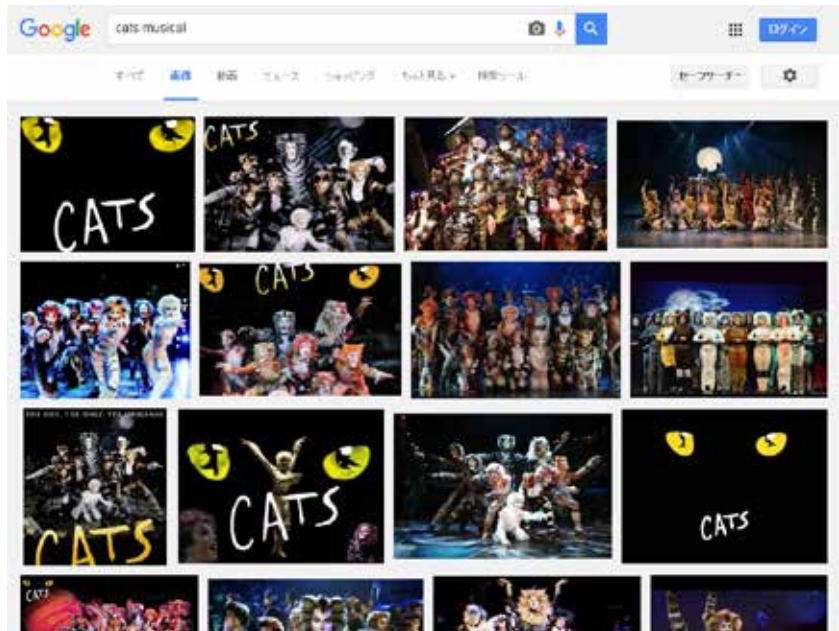
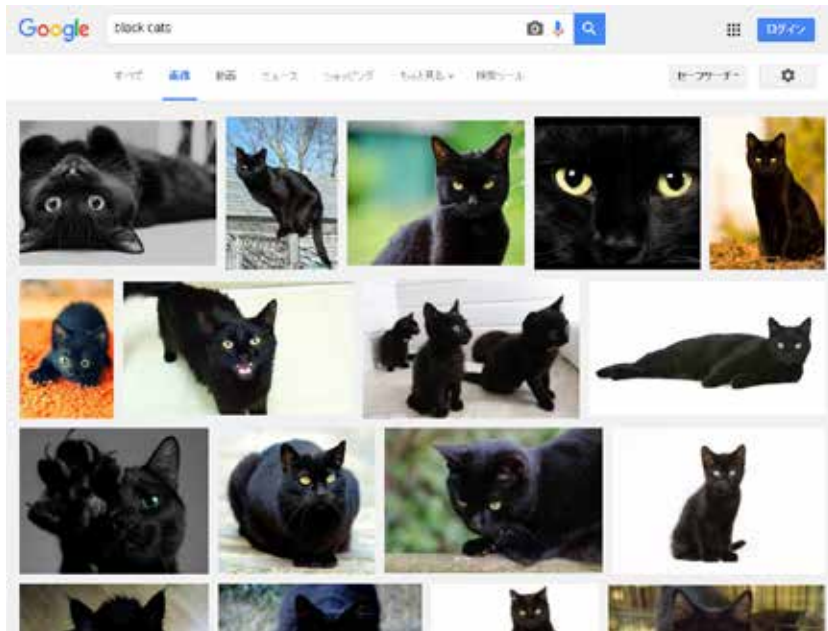
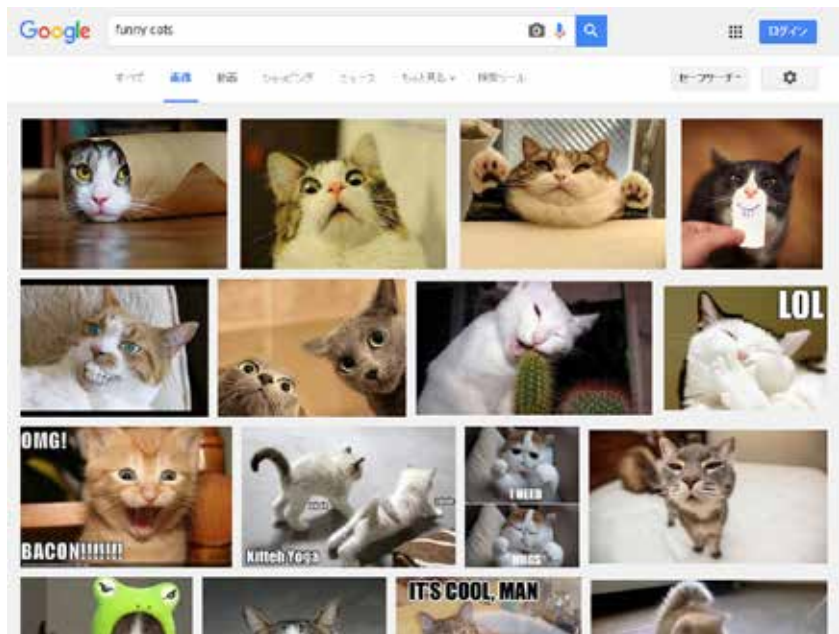
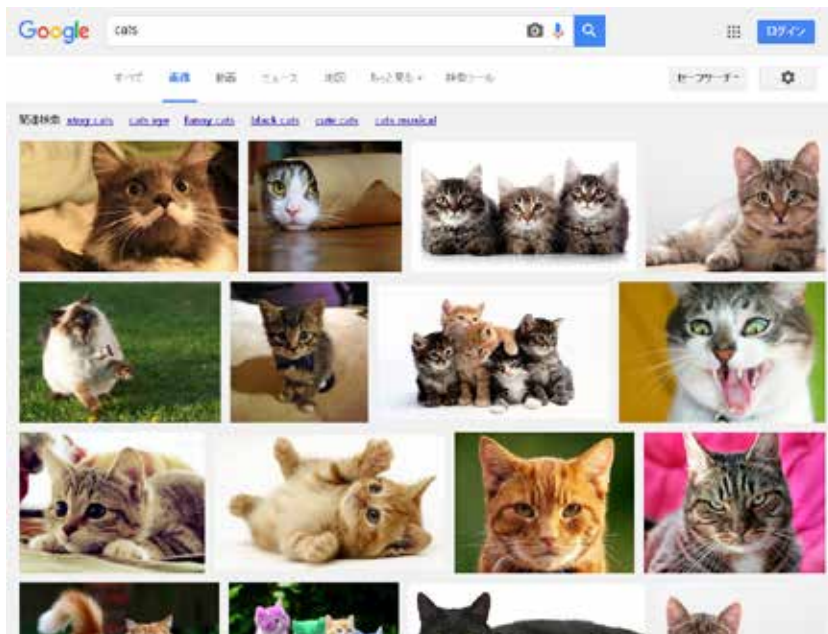
写真



線画

検索オプションに種類を指定すると、種類に応じて検索結果が変化する。



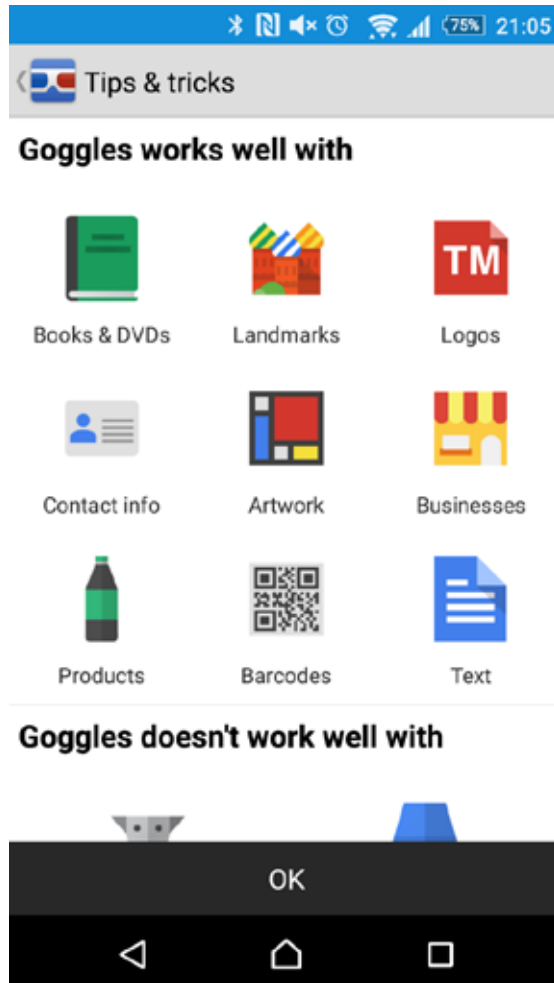


画像内容に基づく検索

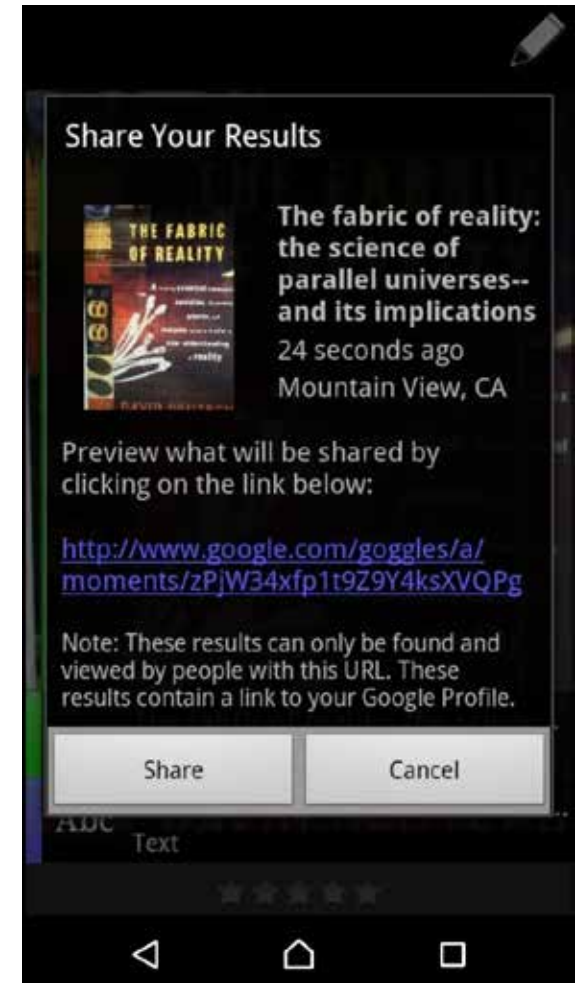
画像内容に基づく検索

- 画像外のメタデータではなく、画像内に含まれる内容分析を重視する手法。
 1. 同じものを探す（エンティティ）。
 2. 同じ種類のものを探す（カテゴリ）。
 3. 似ているものを探す（距離、類似度）。
- 1と2はある程度使われているが、3は期待の割にはまだブレークしていない。
- タグ付けによらない検索機能を実現する。

同じものを探す



Google Goggles



同じ画像・似た画像を探す



Corporate and Enterprise Image Tracking



Image tracking for our corporate clients who need continuous and automatic tracking of



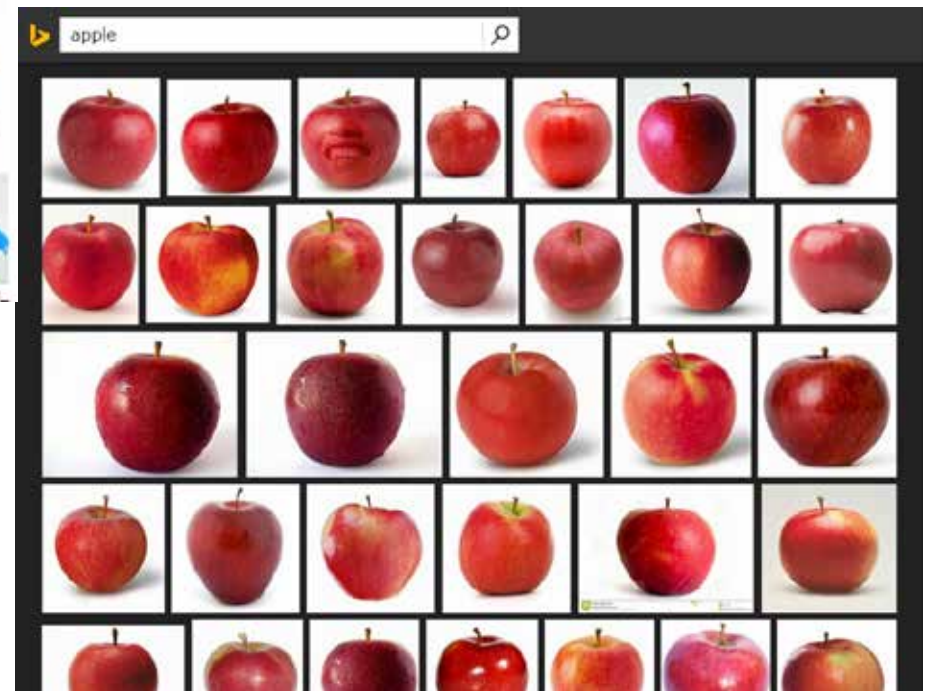
<http://www.tineye.com/>



画像の類似性



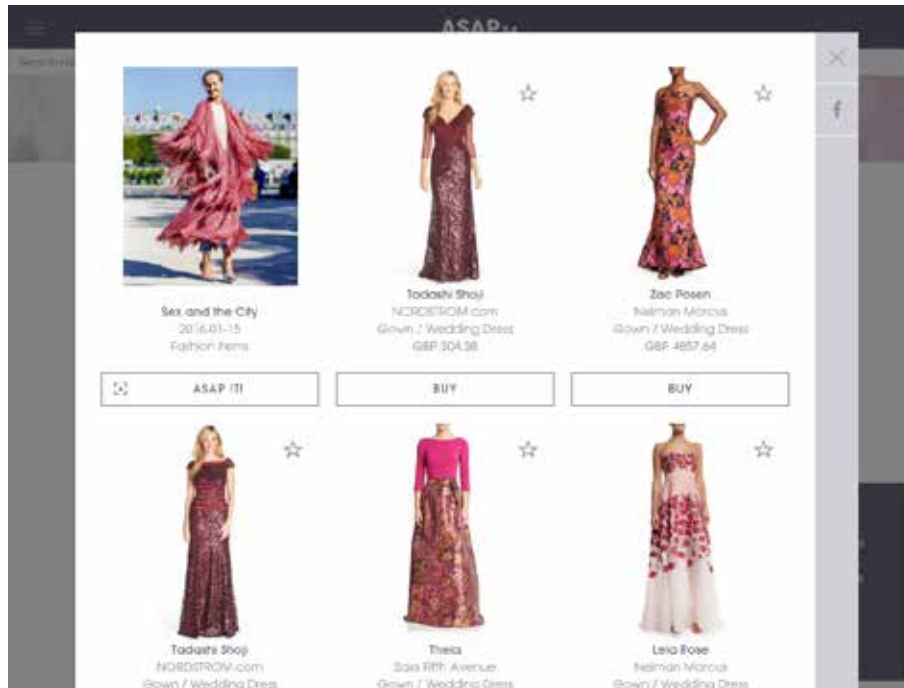
一枚の画像を選んで類似画像を検索すると、色が考慮されるため、（赤い）りんごだけになる。



Appleで検索した結果。りんごのAppleとAppleのロゴが混ざっている。単語だけでは区別できない。

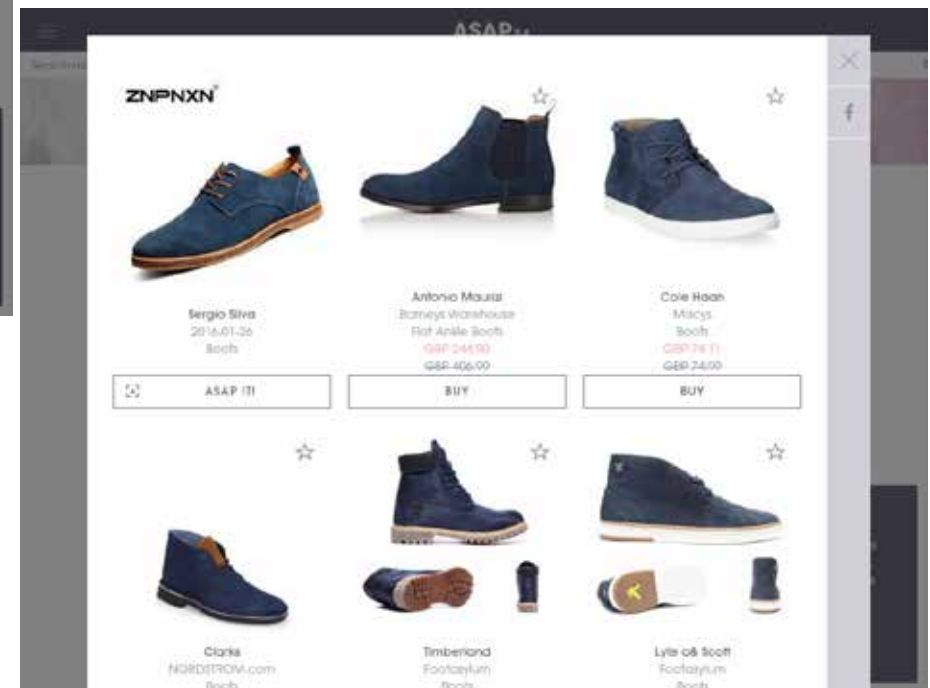
ショッピングへの応用

町で素敵な人を見つけたら、スマホで撮影し、似た服や靴を検索し、オンラインショッピングへ。



色や形で検索すると、
キーワードでは表現できない要素で探せる。

<https://www.asap54.com/>



デジタル・シルクロード・プロジェクトでの取り組み



東洋文庫貴重書 デジタルアーカイブ

<http://dsr.nii.ac.jp/toyobunko/>

- 237冊の書籍70,898ページをデジタル化。
- 研究コミュニティで必須の書籍を選定。
- キャプションや目次などは人手で入力。
- 全文テキストOCR入力（誤りは未訂正）。

画像切り出し + 画像類似検索



図の部分の自動切り出し

国立情報学研究所 - デジタル・シルクロード・プロジェクト
『東洋文庫所蔵』貴重書デジタルアーカイブ

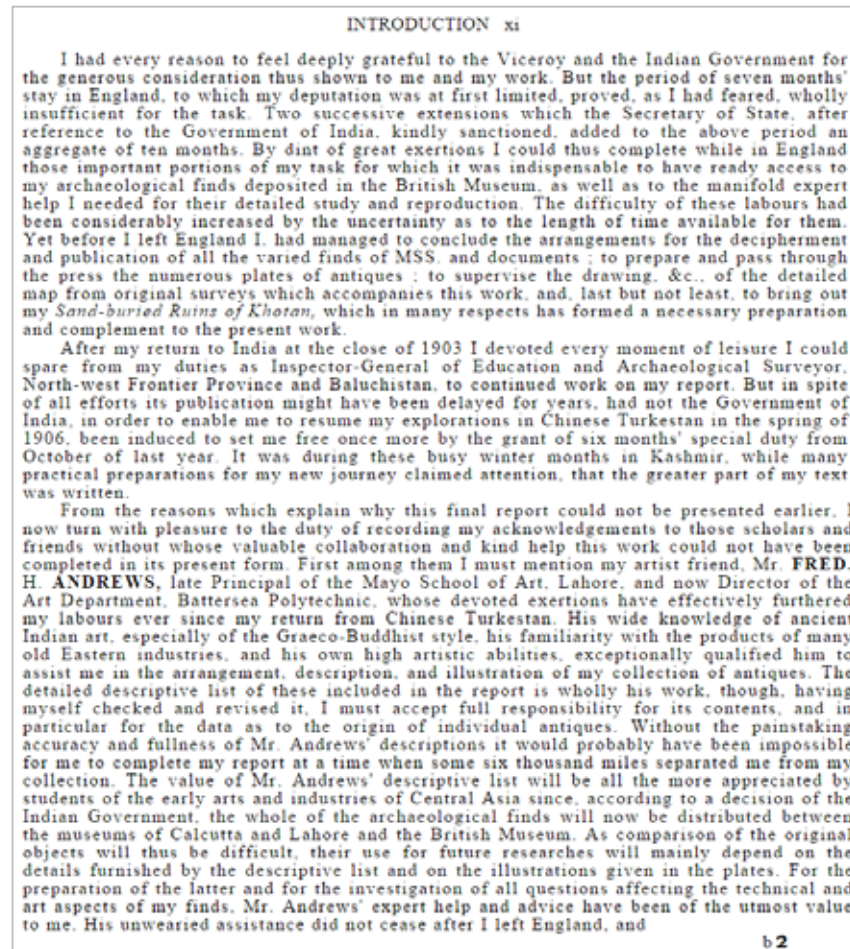
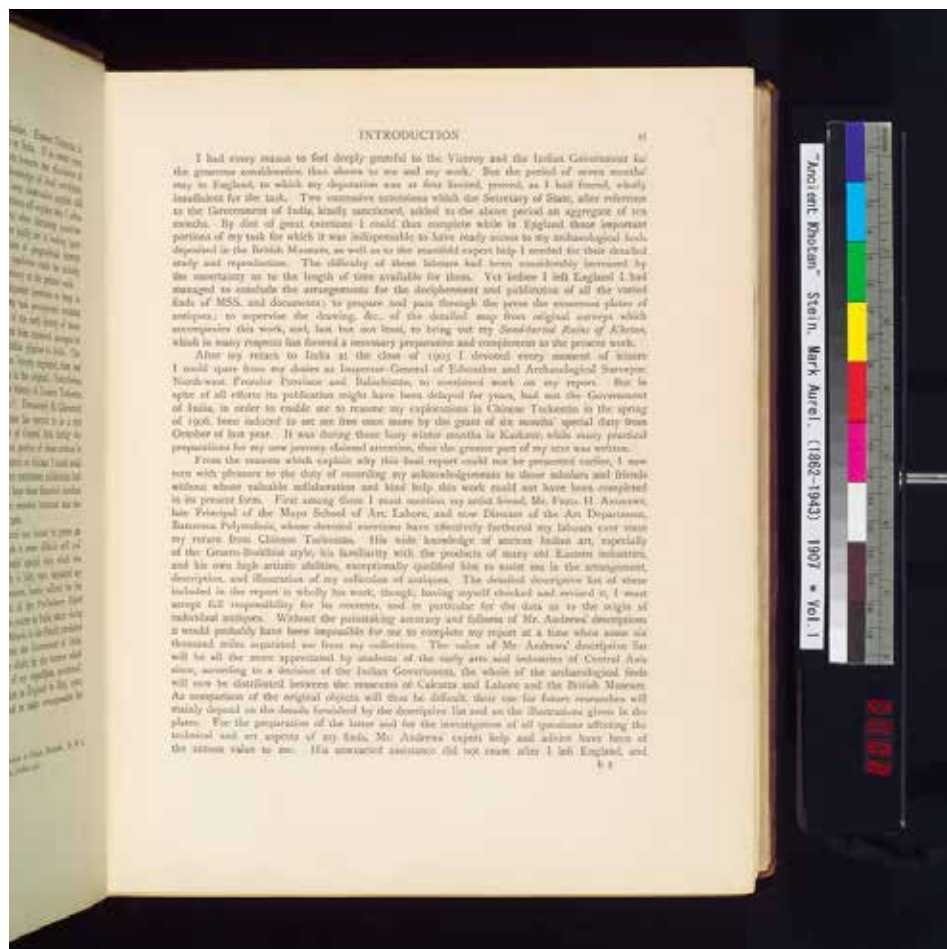
デジタル・シルクロード > 東洋文庫アーカイブ > 類似画像検索

画像検索結果 (類似検索) : 29 件 || Search in English

Query 1	1	2	3	4
VIII-5-B2-19/V-1 Page 0371	VIII-5-B2-19/V-1 Page 0377	La-56/V-1 Page 0141	T-VIII-5-A-8-3/V-3 Page 0173	VIII-5-B2-19/V-1 Page 0483
5	6	7	8	9
VIII-5-B2-19/V-1 Page 0371	VIII-5-B2-19/V-1 Page 0391	VIII-5-B2-8/V-2 Page 0287	LFe-21/V-1 Page 0258	VIII-5-B2-19/V-1 Page 0377
10	11	12	13	14
LFe-21/V-1 Page 0252	VIII-5-B2-19/V-1 Page 0385	VIII-5-B2-19/V-1 Page 0391	VIII-5-B2-19/V-1 Page 0385	LFe-21/V-1 Page 0261
15	16	17	18	19
VIII-1-E-37/V-1 Page 0008	LFe-21/V-1 Page 0263	LFe-21/V-1 Page 0264	VIII-5-B2-19/V-1 Page 0403	XII-4-2/V-5 Page 0026

色が類似した画像を検索

文字認識



画像分析に基づくが、画像検索とは言わない。

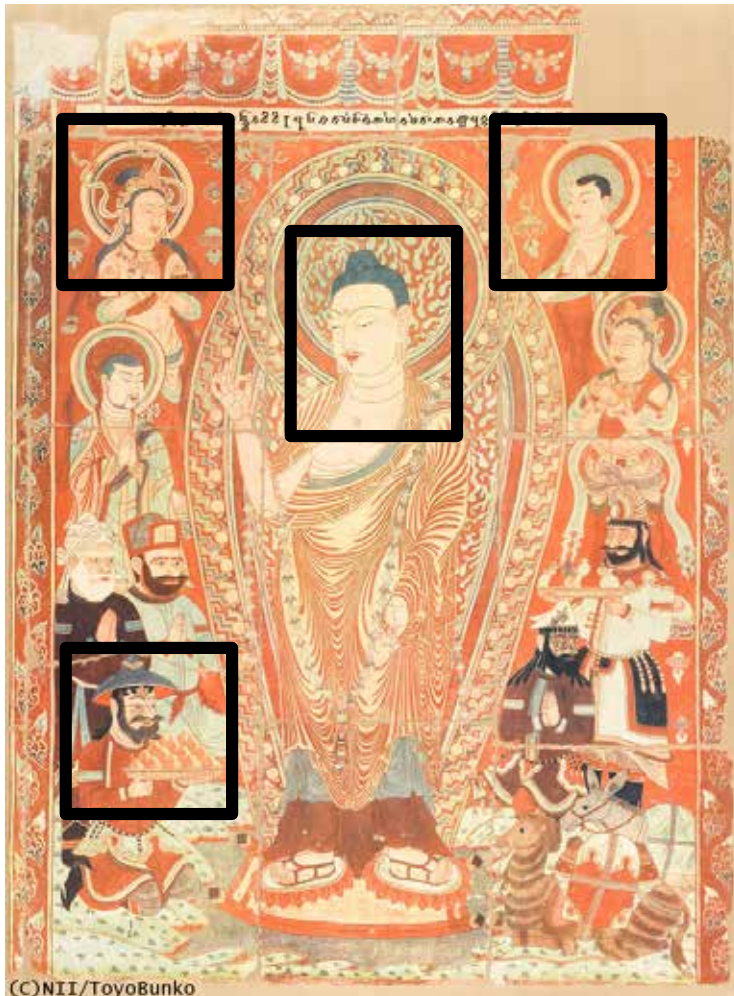
(人間の) 顔認識



画像処理ライブラリOpenCVの顔認識アルゴリズムは、古写真でもある程度は使える。

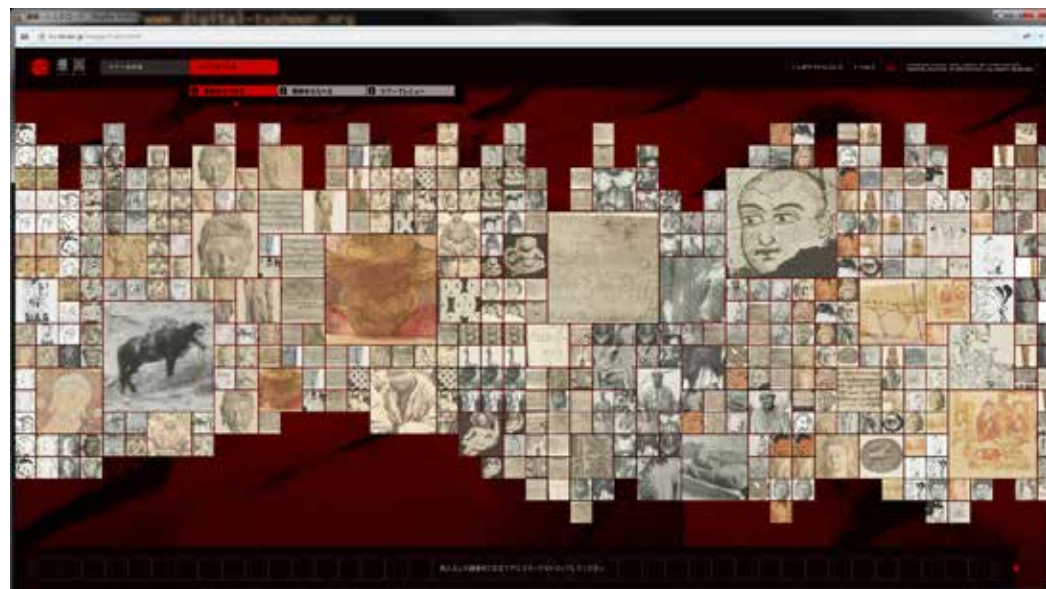
遷画：作品の断片化と再構成

<http://dsr.nii.ac.jp/senga/>



- 東洋文庫貴重書デジタルアーカイブのデジタル書籍の利活用。
- 画像の一部を断片化し、メタデータを付与して素材画像化。
- 素材画像を利用者が任意に再構成し、新たな作品として公開。

「遷画」の基本構造



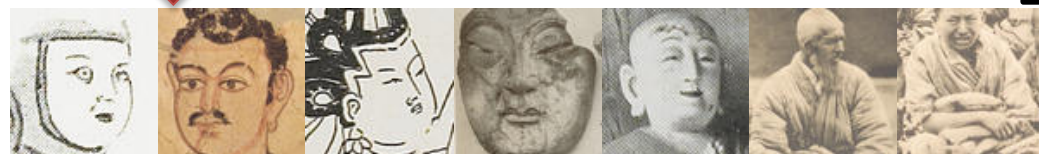
収集支援機能

テーマ検索
エリア検索
類似画像検索
協調画像検索

収集 (Collect)



並べ替え (Order)



共有

東洋文庫ミュージアム
ポストカード





COPYRIGHT © 2007, 2010, DIGITAL SILK ROAD PROJECT,
NATIONAL INSTITUTE OF INFORMATICS. ALL RIGHTS RESERVED.

<http://dsr.nii.ac.jp/senga/>

東洋文庫ミュージアム来館記念

2012年03月08日

ONICO

イケ面

画像の意味の文脈依存性

アニマル三昧



仲良し二人組み



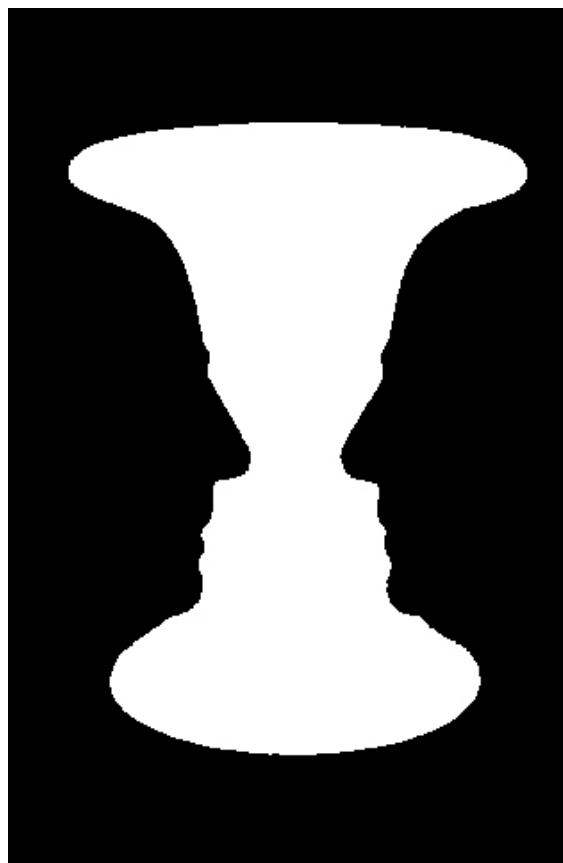
飛び出し注意



敦煌 鳥



画像の多義性



ルビンの壺

Wikipediaより

自動画像アノテーションとDEEP LEARNING（深層学習）

自動画像アノテーション

A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



A skateboarder does a trick on a ramp.



A dog is jumping to catch a frisbee.



A group of young people playing a game of frisbee.



Two hockey players are fighting over the puck.



A little girl in a pink hat is blowing bubbles.



A refrigerator filled with lots of food and drinks.



A herd of elephants walking across a dry grass field.



A close up of a cat laying on a couch.



A red motorcycle parked on the side of the road.



A yellow school bus parked in a parking lot.



Describes without errors

Describes with minor errors

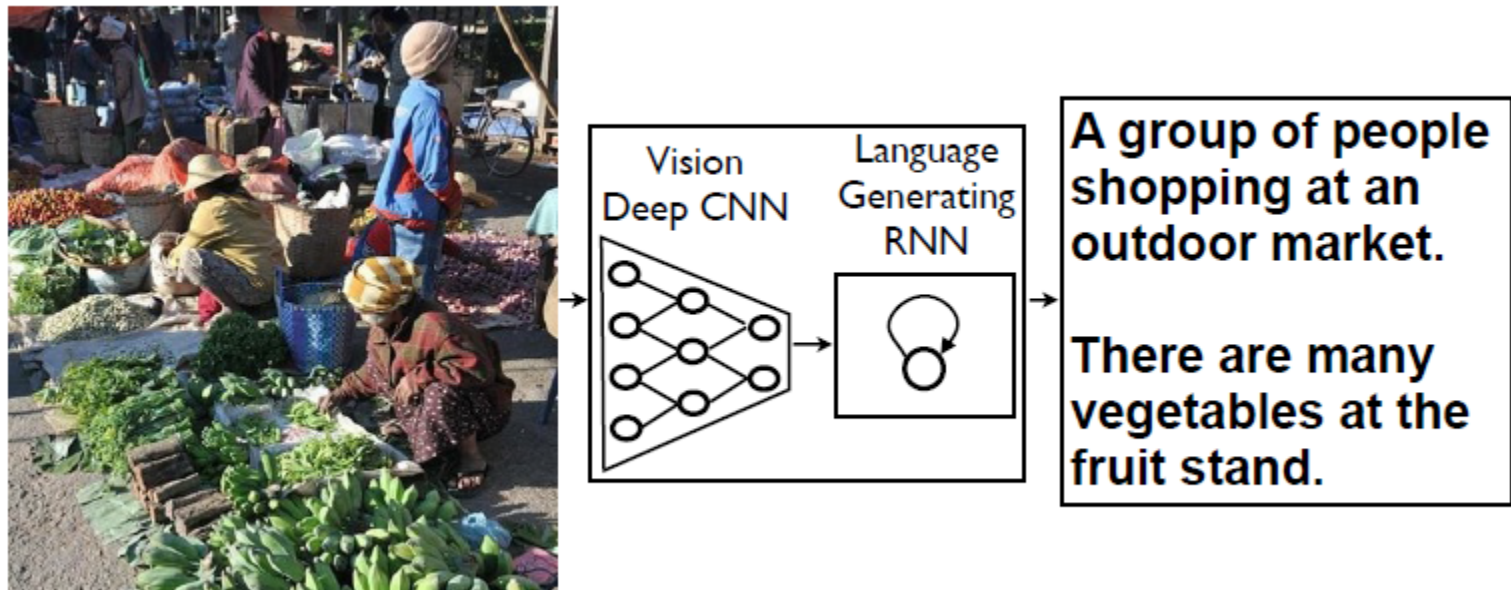
Somewhat related to the image

Unrelated to the image

Vinyals, et.al., Show and Tell: A Neural Image Caption Generator, CVPR 2015

ディープラーニング

- 画像とテキストを同じ空間に埋め込む深層学習（Deep Learning）を使って、画像キャプションを自動的に生成。

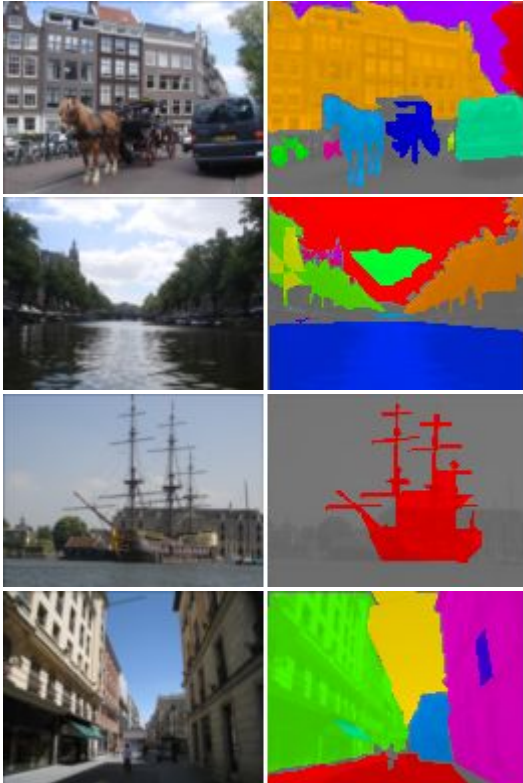


Vinyals, et.al., Show and Tell: A Neural Image Caption Generator, CVPR 2015

深層学習の大流行

- 画像分類問題では、人間に匹敵する分類精度を記録（SuperVision, ILSVRC 2012）
- 教師なし学習では、猫の顔に相当するパターンを自動獲得（Google, ICML 2012）
- 画像学習結果を流用し、「悪夢のような」アート作品を生成（DeepDream）
- 膨大な学習データと莫大な計算機パワーが必要。使うのはそう簡単ではない。

クラウドソーシング



http://labelme.csail.mit.edu/Release3.0/browserTools/php/mechanical_turk.php

2016/2/6

- 勝負の鍵：（高品質）**学習データが多ければ性能が向上する**（特に深層学習）。
- 通常：**ネット上の不特定多数の人間**に、タグ付け等の単純作業を低額で発注。
- 巨人：GoogleやFacebookは、**無料自社サイト**に良質な写真データを大量に集積。

ヒューマンコンピューテーション



<https://www.google.com/recaptcha/intro/index.html>

- **CAPTCHA** : HCの代表的な例。機械と人間の能力差をテスト。
- **reCAPTCHA** : OCR誤りという、機械に難しい問題を人間に解かせる。
- **くずし字版CAPTCHA** : 一般人の能力が機械以下だと意味ないが...

ゲームの活用（ESP Game）



Figure 1. Partners agreeing on an image. Neither of them can see the other's guesses.

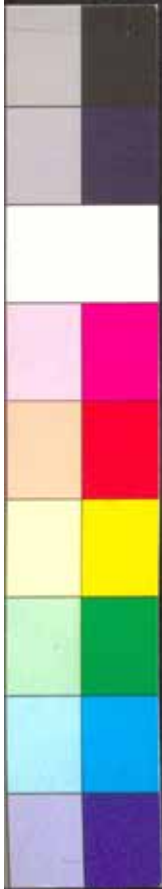
Luis von Ahn, et.al. Labeling Images with a Computer Game, Proc. CHI, 2004

- 他者のタグと一致したら成功。他者を想像させ「合意できるタグ付け」を実現。

古典籍の画像検索

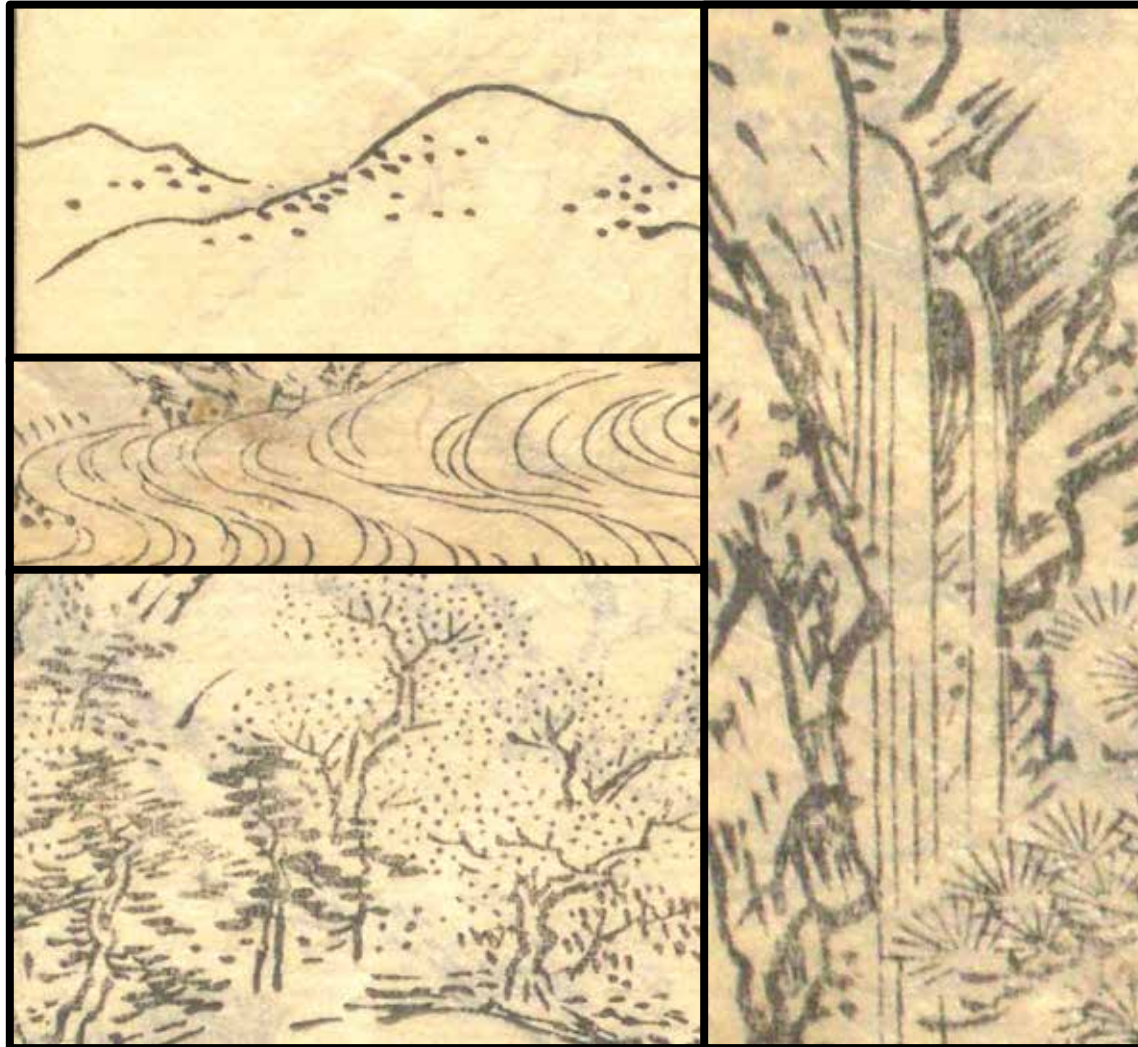
古典籍画像検索の可能性

- **タグ検索**：既入力データを学習データとして活用することで、タグ付け作業を支援しスピードアップ。
- **パターン検索**：文様などテキスト化しにくいパターンを自動的に抽出して検索。
- **イラスト（線画）検索**：自然画像（写真）に比べて抽象度が高い（詳細度が低い）ため、意味付けがより難しい。



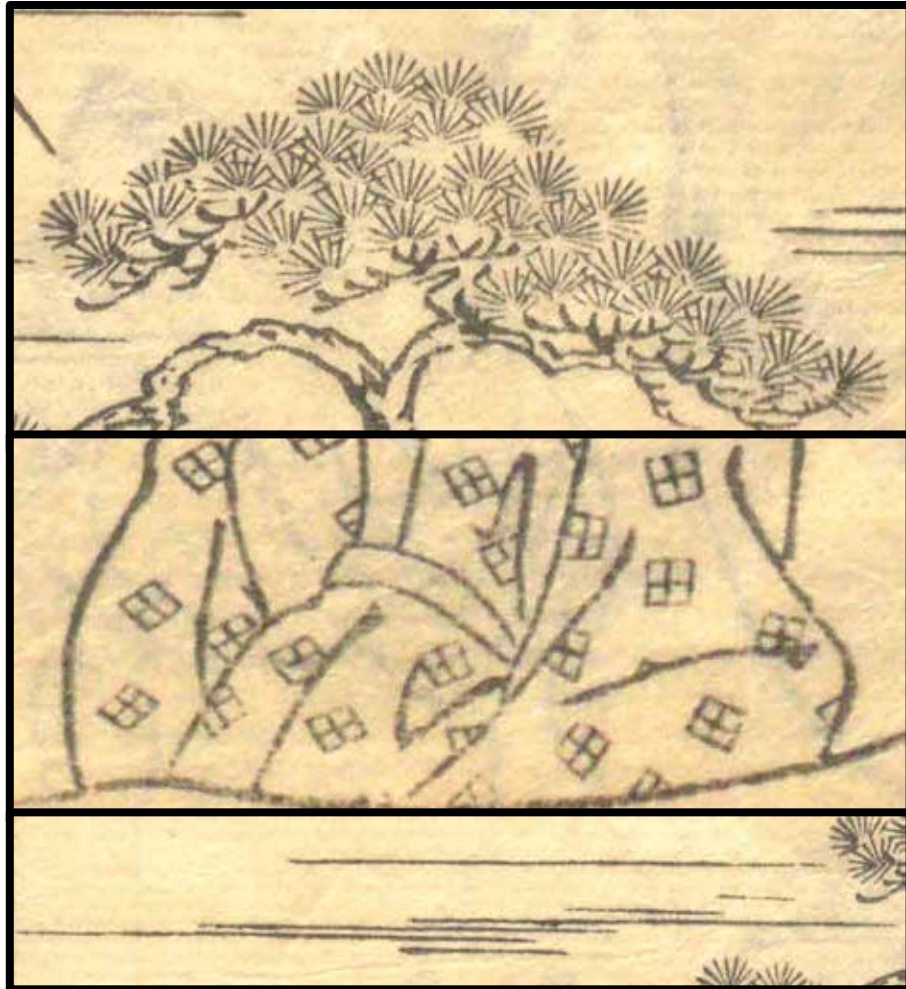
絵本磯駒松 <http://www2.dhii.jp/nijl/NIJL0007/049-0245/049-0245-00010.jpg>

抽象化された表現



- 山、川、滝、木。細部の省略による現実世界の高度な抽象化表現。
- 抽象度が高い概念は、自動タグ付けが非常に困難。

規則的なパターン



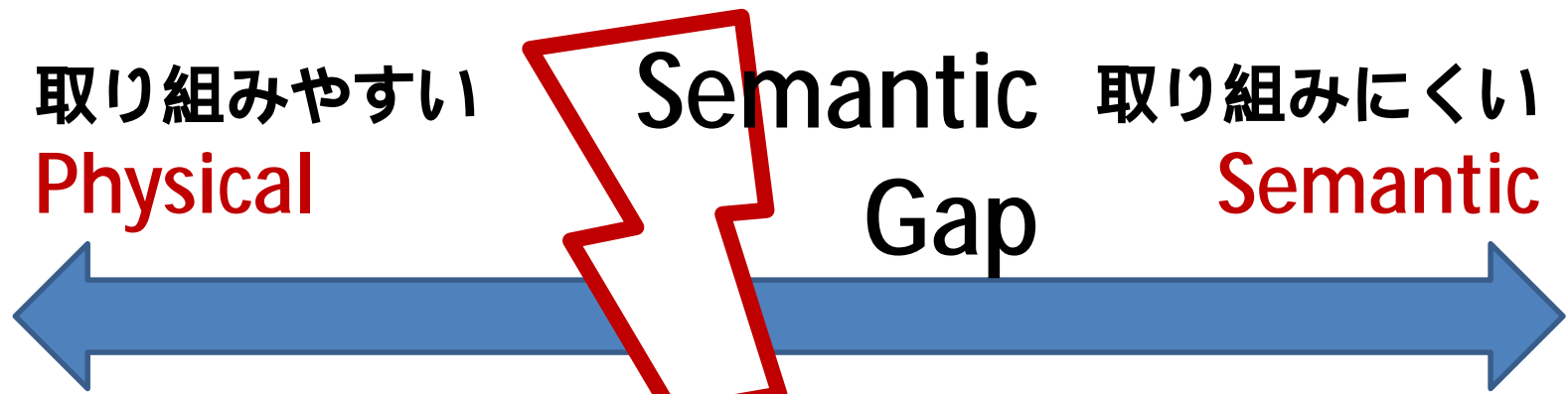
- 松の葉、着物の文様、奥行き感を出す水平線など。
- 抽象化と同時に、規則性が読み取れるパターンである。
- 検索したいパターンを特定し、データから学習させる。

まとめ

画像分析の難易度を考える



2方向の画像分析戦略



- **専門家の置換**：大量・高精度な作業を高速化・自動化する。
- 例：写本の同一性判定や特徴の分類。
- **専門家の補完**：少数の高度な判断を支援・効率化する。
- 例：画像のフィルタリングやタグ付け。

関連情報

- デジタル・シルクロード

- <http://dsr.nii.ac.jp/>

「デジタルデータ・ツールを活用した人文学研究」に関心があれば、**人文情報学**や**デジタル・ヒューマニティーズ**等の研究会にもご参加下さい。

- 人文科学とコンピュータ研究会

- <http://www.jinmoncom.jp/>

- 日本デジタル・ヒューマニティーズ学会

- <http://www.jadh.org/>