

画像を介した自然言語表現の同義性判別

Discriminating Synonymous Expressions by Using Images

那須川 哲哉^{*1} 吉川 克正^{*1} 鈴木 祥子^{*1} 森田 千明^{*2}
Tetsuya Nasukawa Katsumasa Yoshikawa Shoko Suzuki Chiaki Morita

^{*1} 日本アイ・ビー・エム株式会社 東京基礎研究所
IBM Research – Tokyo, IBM Japan, Ltd.

^{*2} 日本アイ・ビー・エム株式会社 システム・テクノロジー開発製造
STG Laboratory, IBM Japan, Ltd.

Identification of word similarity plays important role in natural language processing. It supports wide varieties of applications such as text mining for aggregating synonymous expressions and machine translation for finding equivalent terms in different languages. Whereas state of the art of word similarity identification uses textual corpus as the knowledge resource, we propose a method to use images associated to each expression. After producing a list of candidate synonymous expressions for a query expression based on textual resources, images representing each candidate synonymous expression are compared with images representing the query expression for evaluating the synonymity.

1. はじめに

自然言語処理においては、同義性の判定が重要な役割を果たす。単一言語内であれば、同義表現を集約し表現の多様性を吸収することで、テキストマイニングなどに応用することができ、複数言語間の同義表現は、訳語として翻訳に応用できる。

近年、画像とテキストの融合が進み、画像検索など、言語表現と画像とが実用規模で結び付けられるようになった。この環境を活かし、本稿では、言語表現の同義性を各表現に紐付いた画像の類似性を用いて判別する手法とその有効性を示す。

2. 自然言語処理と画像処理の融合

自然言語処理の分野においては、1990年代からコーパスベースの処理手法が盛んになり、処理するテキストの量が増えるにつれて、解析精度が向上してきた。しかし、テキストの量を無限にすれば精度が100%になるわけではない。構文解析の精度はコーパスサイズ100万文程度で、格解析の精度はコーパスサイズ数千万文程度で飽和する傾向が見られる[喜連川2011]。これは、テキストに記述された情報のみを利用した自然言語処理の限界を示唆していると考えられる。

自然言語処理においては、テキストに記述された言語表現の意味を適切に扱う必要がある。言語表現の意味は、我々を取り巻く外界の状況や個々の経験と結び付いており、その状況や経験を全てテキストのみで適切に記述することができない。このことから、テキスト以外の情報を利用して言語表現の意味を扱うことで、自然言語処理の精度を向上できる可能性がある。

筆者らは、1980年代後半から、テキストの内容をイメージ化して解釈する手法を検討し、その実現に取り組んだ[那須川1990]が、1990年前後では、イメージ化するためのデータや計算機能力の問題などから、実用化には程遠い状況であった。

近年、画像や動画のデジタル化が進み、その解析技術が向上した結果、自然言語処理に画像や動画の情報を活用する試みが本格的に始まりつつある。例えば、ネット上の検索エンジンでは画像を言語表現で検索できるようになっており、多様な言

語表現が画像と結び付いている。このような環境を利用して、語義の曖昧性解消[Barnard 2003; Barnard 2005]や格解析[Bergsma 2011a]、訳語推定[Bergsma 2011b]に画像を活用したり、また、言語表現の参照先を画像中で特定 [Krishnamurthy 2013] したり、言語表現と動画中の動作を結び付け [Regneri 2013; Yu 2013] したりする試みが報告されている。

従来の自然言語処理の限界を超えるために、画像処理を融合させる試みには大きな可能性が感じられる。ただし、このような試みはまだ始まったばかりであり、実用化への課題は多い。例えば、[Bergsma 2011b] の訳語推定の実験対象は、物体を示す表現に限定され、実用性が低い。

3. 実システムにおける自然言語の同義性判別

本稿では、実用的な自然言語処理という観点で、膨大なテキストデータから有益な知見を導き出すテキストマイニング[那須川2006]のための同義性判別を考える。営業報告や不具合事象報告、コールセンターにおける対応記録といった多様なデータを対象として、そこに記述された多様な内容の分布傾向や増減傾向から有用な気付きを得るのがテキストマイニングの目的であり、特定の内容を示す表現を把握し、集約するために同義性判別が必要となる。その際、汎用的な同義性よりも、分析対象データに特化した同義性判別の重要性が高くなる。

例えば、米国の政府機関である National Highway Traffic Safety Administration (NHTSA)¹ が収集し公開している自動車不具合情報²数十万件のデータを分析する上では、*gasoline* と *fuel* と *gas* が基本的には燃料という点で同じ内容を示していることを把握していれば、多様な表現で記述される燃料切れや燃料漏れなどの状況を集約することができ、特定の車種や特定の環境で燃料切れや燃料漏れを発生する問題に気付き易くなる。こういった表現の多様性は、データに依存し、一般的な定義が通用するとは限らない。例えば、NHTSA の自動車不具合情報のデータにおいては、*liftgate* という表現が *hatchback* (あるいは *liftback*) の後方ドアの意味で使われていることが多い。

連絡先: 那須川哲哉, 日本アイ・ビー・エム株式会社 東京基礎研究所, nasukawa@jp.ibm.com

¹ <http://www.nhtsa.gov/>

² <http://www-odi.nhtsa.dot.gov/downloads/>

Wikipedia によれば、この使い方は誤用となっているが、分析対象データにおいて、実際に後方ドアの意味で使われている以上、テキストマイニングのためには、後方ドアの同義語として扱えるようにすることが望ましい。

従って、実システムにおいては、対象データにおける同義性の判別が重要であり、一般的な同義表現辞書を予め用意することなく、与えられたデータにおける表現の使われ方に応じて同義性を判別する必要がある。

自然言語処理における同義性判別手法としては、対象となる表現 W に対し、データの各所に出現している W の周辺文脈表現 (例えば係り受け関係にある表現) の分布の類似性を比較する手法[Hindle 1990; Lin 1998; Gasperin 2001]が存在する。これは、特定の表現パターン[Yu 2002]や特定のデータタイプ[Wu 2003]に依存する手法よりも、多様なデータに比較的適用しやすく、汎用性が高い。この周辺文脈の分布類似性を比較する手法は、異なる言語のデータを対象に、周辺文脈情報を対訳辞書で翻訳することで、異言語での同義表現(すなわち訳語)抽出に適用できる[Rapp 1995; Fung 1997; Fung 1998; Rapp 1999]。

周辺文脈の分布類似性に基づく同義性の判別は、汎用性が高い反面、周辺文脈の表現の曖昧性や多様性などによるノイズを受けやすく、高い精度を出すのが難しい。書き手による表現の特徴を考慮[Murakami 2004]したり、計算方法を工夫[Andrade 2010]したりすることで多少なりとも精度を向上させることができるが、出力結果をそのまま、人手による確認抜きで、使えるレベルには至っていない。それが実現する機能は、同義表現抽出機能や訳語抽出機能というよりも同義表現抽出支援機能や訳語抽出支援機能というのが相応しい。例えば、前出の NHTSA 自動車不具合情報データと国交省の自動車不具合情報データを用いて訳語を抽出した実験では、第 1 候補が正しい訳語となる割合は 3 割程度にすぎなかった[那須川 2009]。上位 5 候補の中に正しい訳語が存在するのが 6 割程度であり、上位 20 候補まで広げると 8 割から 9 割程度に向上するといった結果であった。ただし、訳語にも同義表現が存在するため、正しい訳語が複数得られる場合がある。そのような場合、テキストマイニングにおいては、同義性を吸収して分析する効果を生じる。そのため、テキストマイニングという観点では、自動的に出力された訳語の結果が、上位候補 N 語中に複数の正しい訳語が存在するというレベルであっても、人手による修正を加えずに適用して、有益な気付きにつながり得るという結果[海野 2010]が確認されている。

4. 画像を介した自然言語表現の同義性判別

前述の自動車不具合情報データを対象として、自然言語処理で同義表現や対訳表現の候補を抽出し、各表現に紐付いた画像の特徴量を比較することで、同義表現の抽出精度を向上させる仕組みを構築し、有効性を調査した。

基本的には、図 1 のように、対象となる表現 W_0 に対する同義表現や対訳表現の候補集合 $\{S_j : j=1, \dots, m\}$ を作成し、対象となる表現 W_0 に関連する画像集合 $\{P_i : i=1, \dots, n\}$ と、各候補表現 S_j に関連する画像集合 $\{Q_{jk} : k=1, \dots, n'\}$ の特徴量を比較した。結果として、 W_0 に関連する画像の特徴量と類似性の高い特徴量を持つ画像に関連する候補表現ほど W_0 との同義性が高いと判断するようにした。下記実験では、関連画像の取得には Google Image 検索²を利用した

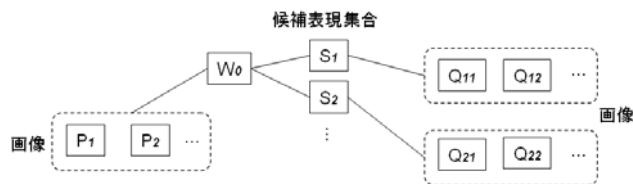


図 1: 対象表現 W_0 とその同義/対訳表現候補と画像の関係

4.1 日本語自動車不具合情報データでの同義性判別

自動車のフロントガラスに関して、同義語である「フロントガラス」「フロントグラス」「フロントウインドウ」「フロントウインドウ」と、周辺文脈の分布類似性を用いたテキストベースの処理で同義語の候補として抽出され易い「ワイパー」及び「リアガラス」との同義性を評価した。各表現に対して 5 つの関連画像を用いた、画像の特徴量抽出と類似度算出は下記ステップの通り。

1. surf[Bay 2008] によりモノクロ画像で keypoints を抽出
2. 同義語候補の画像の keypoints を k-means でクラスタリング。クラスタ個数は 500 を選択 (visual words の抽出)
3. フロントガラスの画像と、各同義語候補の画像の keypoints にクラスタ ID (visual words) を付与
4. 各画像に対して visual words のヒストグラムを作成し、これを画像ベクトルとして利用
5. コサイン類似度を算出

最終的に得られた類似度上位の 25 表現を表 1 に示す。表中最上段の行は、「フロントガラス」の 4 番画像と「フロントウインドウ」の 5 番画像との類似度が最も高かったことを示している。結果的に、「フロントガラス」の同義語の画像が類似度上位に並び、本手法の同義語判別への有効性が確認できた。非同義語の「ワイパー」で最も類似度の高い画像は 43 番目であった。

表 1: フロントガラスと類似性の高い関連画像を持つ表現

| 同義語画像番号 (フロントガラス) | 候補語 (同義語 or 関連語) | 候補語画像番号 | コサイン類似度 |
|----------------------|---------------------|---------|----------|
| 4 | フロントウインドウ | 5 | 0.374818 |
| 5 | フロントウインドウ | 1 | 0.351583 |
| 5 | フロントウインドウ | 5 | 0.351583 |
| 2 | フロントウインドウ | 5 | 0.350032 |
| 5 | フロントグラス | 4 | 0.348452 |
| 3 | フロントグラス | 3 | 0.337632 |
| 5 | フロントグラス | 3 | 0.331811 |
| 5 | フロントウインドウ | 5 | 0.329496 |
| 5 | フロントグラス | 1 | 0.328216 |
| 2 | フロントグラス | 4 | 0.313787 |
| 1 | フロントグラス | 3 | 0.313047 |
| 2 | フロントウインドウ | 3 | 0.311591 |
| 2 | フロントウインドウ | 3 | 0.311591 |
| 5 | フロントグラス | 2 | 0.297334 |
| 1 | フロントグラス | 5 | 0.295604 |
| 2 | フロントグラス | 5 | 0.293101 |
| 1 | フロントグラス | 3 | 0.292100 |
| 3 | フロントグラス | 1 | 0.292100 |
| 3 | フロントグラス | 2 | 0.288572 |
| 4 | フロントグラス | 5 | 0.286588 |
| 4 | フロントグラス | 2 | 0.285401 |
| 1 | リアガラス | 2 | 0.285138 |
| 1 | フロントウインドウ | 5 | 0.282935 |
| 2 | フロントグラス | 3 | 0.282829 |
| 2 | フロントグラス | 3 | 0.281119 |

¹ <http://www.mlit.go.jp/jidosha/carinf/rcl/index.html>

² <https://www.google.co.jp/imgph>

表2: 自然言語処理による対訳表現候補の抽出結果(類似順)

| engine | transmission | injury | fire |
|--------|------------------|---------|----------|
| エンジン | ギヤ | エアバッグ | 配線 |
| ガasket | ギア | 事故 | 火災 |
| 信号待ち | シフト | バック | エンジンルーム |
| ディーラー | オートマチックトランスミッション | 衝突 | 煙 |
| 信号 | シフトチェンジ | エア | ボンネット |
| カバー | マニュアル | サイドブレーキ | 熱 |
| クーラント | マニュアル車 | 助手席 | エンジンルーム内 |
| 後方 | クラッチ | 怪我 | 火 |
| シリンダ | ミッション | 車 | 原因 |
| オイル | チェンジ | エアバッグ | 電気配線 |
| ヘッド | シフトレバー | 壁 | 駐車場 |
| 液晶画面 | D | ブレーキ | 電気 |
| 再始動 | Dレンジ | 前面 | 車両火災 |
| 鏡 | 変速 | レンジ | 白煙 |
| 写し鏡 | レンジ | 衝突時 | 白 |
| 写し | トランスミッション | | |

4.2 米日自動車不具合情報データからの訳語獲得

NHTSA 自動車不具合情報データと国交省の自動車不具合情報データを用いて、周辺文脈の分布類似性をベースに英日の翻訳対を抽出 [那須川 2009] して得られた表 2 の結果を各表現の関連画像の類似性で並べ替える実験を行った。

画像の特徴量抽出と類似度算出は下記ステップの通り。

1. sift[Lowe 2004] によりモノクロ画像で keypoints を抽出
2. 日本語(訳語)候補画像の keypoints を k-means でクラスタリング. クラスタ個数は 500 を選択(visual words の抽出)
3. 英語画像と、(訳語)候補語画像の keypoints にクラスタ ID(visual words)を付与
4. 各画像に対して visual words のヒストグラムを作成し、これを画像ベクトルとして利用
5. コサイン類似度を算出

最終的に得られた類似度によって、表 2 の各対訳表現を並べ替えた結果を表 3~表 6 に示す。(下線の候補語が正解)

表3: engineの対訳表現候補の画像類似度によるソート結果

| 対象表現画像番号 | 候補語 | 候補語画像番号 | コサイン類似度 |
|----------|-------|---------|----------|
| 9 | エンジン | 2 | 0.793757 |
| 10 | ヘッド | 9 | 0.790767 |
| 9 | 写し | 3 | 0.742549 |
| 9 | エンジン | 4 | 0.734717 |
| 9 | 写し鏡 | 10 | 0.731012 |
| 4 | シリンダ | 10 | 0.726663 |
| 4 | ディーラー | 5 | 0.724024 |
| 9 | ヘッド | 7 | 0.722495 |
| 7 | エンジン | 9 | 0.718701 |

表4: transmission 対訳表現候補の画像類似度でのソート結果

| 対象表現画像番号 | 候補語 | 候補語画像番号 | コサイン類似度 |
|----------|------------------|---------|----------|
| 5 | トランスミッション | 1 | 0.574225 |
| 4 | トランスミッション | 1 | 0.565086 |
| 4 | トランスミッション | 6 | 0.537392 |
| 5 | ギヤ | 7 | 0.536524 |
| 4 | オートマチックトランスミッション | 4 | 0.534691 |
| 4 | チェンジ | 7 | 0.528256 |
| 3 | トランスミッション | 3 | 0.523765 |
| 5 | トランスミッション | 2 | 0.522819 |
| 5 | トランスミッション | 8 | 0.518128 |
| 3 | チェンジ | 7 | 0.516999 |

表5: injury の対訳表現候補の画像類似度によるソート結果

| 対象表現画像番号 | 候補語 | 候補語画像番号 | コサイン類似度 |
|----------|---------|---------|----------|
| 3 | ブレーキ | 7 | 0.853467 |
| 1 | 助手席 | 5 | 0.746171 |
| 1 | 怪我 | 5 | 0.721599 |
| 1 | 助手席 | 7 | 0.706157 |
| 1 | エア | 5 | 0.704402 |
| 1 | バック | 10 | 0.704122 |
| 3 | サイドブレーキ | 5 | 0.701945 |
| 5 | エアバッグ | 7 | 0.696534 |
| 7 | 衝突 | 7 | 0.694013 |
| 1 | エアバッグ | 4 | 0.689688 |

表6: fire の対訳表現候補の画像類似度によるソート結果

| 対象表現画像番号 | 候補語 | 候補語画像番号 | コサイン類似度 |
|----------|----------|---------|----------|
| 8 | 火災 | 10 | 0.736875 |
| 7 | 火災 | 10 | 0.735284 |
| 1 | 駐車場 | 2 | 0.725225 |
| 9 | 火災 | 7 | 0.714769 |
| 1 | エンジンルーム内 | 4 | 0.703265 |
| 7 | 火 | 9 | 0.700451 |
| 8 | 火 | 9 | 0.694985 |
| 8 | 煙 | 2 | 0.689051 |
| 8 | 配線 | 4 | 0.67684 |
| 9 | 火災 | 10 | 0.673375 |

injury の対訳である「怪我」に関しては、表 2 で 8 番目だった類似度が、表 5 では 3 番目と、高まっている。しかし、第 1 候補にはなっていないため、類似度の指標を、はずれ値に対して頑健とされる L1 距離に変更して比較した。sift で出したヒストグラム特徴量から各画像の L1 距離を測定し、小さい順にソートした。その結果、表 7 の通り、「怪我」が第 1 候補となった。

表7: injuryの対訳表現候補のL1画像距離によるソート結果

| 対象表現画像番号 | 候補語 | 候補語画像番号 | L1 距離 |
|----------|-------|---------|----------|
| 1 | 怪我 | 5 | 1.00106 |
| 1 | エアバッグ | 10 | 1.046223 |
| 7 | 事故 | 10 | 1.050087 |
| 1 | 衝突 | 5 | 1.053474 |
| 1 | バック | 10 | 1.054752 |

5. まとめと今後の課題

自然言語表現の同義性判別に画像を用いる手法を示し、実問題のデータを処理した結果を通じて、その有効性を示した。

画像処理の計算量を考えると、画像と紐付いた全ての表現を比較し、全ての表現の組合せの中から同義表現を抽出するのは現実的でない。実用上で重要なのが分析対象データにおける同義性であることから、分析対象のテキストデータのみを最初に用いて自然言語処理の範囲内で候補を抽出し、比較対象を限定してから画像処理を適用することで、全体的な処理効率を

高めることができる。すなわち、自然言語処理と画像処理を融合させることで、効率良く自然言語処理の精度向上を実現しようというのが本手法の狙いである。

イメージ検索で多様な表現と結び付いた画像を確認すると、物体以外でも、日本語と英語で類似した画像と結び付いている表現は意外に多い。例えば、*odor* と「臭い」の画像を比較すると、鼻をつまむイメージに共通性が見受けられる。また、*philosophy* と「哲学」においても、各表現と結び付いた画像の中にギリシャ哲学の絵画など共通な画像が見受けられる。しかし、*greeting* と「挨拶」のように、文化的な違いから異なる画像が結び付いているケースや、*trouble* と「問題」のように、多義性も影響して、画像に共通性がないケースも見られる。

今後は、大規模な実験を通じて、本手法が有効である表現と有効でない表現の違いを調査し、その判別の自動化を検討したり、より効果的な画像類似性判別手法を実現したりする必要がある。本稿で紹介した実験の範囲でも、人間にとって似ていると感じられる画像の類似度が低かったり、逆に似ていないようでも類似度が高かったりするケースが散見された。実運用上は、同義性判別結果を画像と共に提示するインターフェースなど、活用方法を工夫する仕組みも重要である。

参考文献

- [喜連川 2011] 喜連川優: 情報爆発のこれまでとこれから, 電子情報通信学会誌, Vol.94, No8, 2011.
- [那須川 1990] 那須川哲哉, 寺西克之, 崔鐘根, 伊東幸宏, 高木朗, 小原啓義: イメージの対象世界モデルを用いた文章理解, 第4回人工知能学会全国大会, pages 279-282, 1990.
- [Barnard 2003] Kobus Barnard, Pinar Duygulu, David Forsyth, Nando de Freitas, David M. Blei, Michael I. Jordan: Matching words and pictures, The Journal of Machine Learning Research archive, Volume 3, Pages 1107-1135, 2003.
- [Barnard 2005] Kobus Barnard and Matthew Johnson: Word sense disambiguation with pictures, Journal of Artificial Intelligence - Special volume on connecting language to the world archive, Volume 167 Issue 1-2, Pages 13-30, 2005.
- [Bergsma 2011a] Shane Bergsma and Randy Goebel: Using visual information to predict lexical preference. In Proceedings of Recent Advances in Natural Language Processing, pages 399-405, 2011.
- [Bergsma 2011b] Shane Bergsma and Benjamin Van Durme: Learning Bilingual Lexicons using the Visual Similarity of Labeled Web Images, In proceedings of International Joint Conference on Artificial Intelligence, pages 1764-1769, 2011.
- [Krishnamurthy 2013] Jayant Krishnamurthy, Thomas Kollar: Jointly Learning to Parse and Perceive: Connecting Natural Language to the Physical World, Transactions of the Association for Computational Linguistics, 1(2), pages 193-206, 2013.
- [Regneri 2013] Michaela Regneri, Marcus Rohrbach, Stefan Thater, Dominikus Wetzel, Bernt Schiele, Manfred Pinkal: Grounding Action Descriptions in Videos. Transactions of the Association for Computational Linguistics, 1, pages 25-36, 2013.
- [Yu 2013] Haonan Yu and Jeffrey Mark Siskind: Grounded language learning from videos described with sentences. ACL 2013, pages. 53-63, 2013.
- [那須川 2006] 那須川哲哉: テキストマイニングを使う技術/作る技術—基礎技術と適用事例から導く本質と活用法, 東京電機大学出版局, 2006.
- [Hindle 1990] Donald Hindle: Noun Classification From Predicate-Argument Structures. Proceedings of the 28th Annual Meeting of ACL, pp.268-275, 1990.
- [Lin 1998] Dekang Lin: Automatic Retrieval and Clustering of Similar Words, COLING - ACL, pp768-774, 1998.
- [Gasperin 2001] Caroline Gasperin, Pablo Gamallo, Alexandre Agustini, Gabriel Lopes, and Vera de Lima: Using Syntactic Contexts for Measuring Word Similarity In the Workshop on Semantic Knowledge Acquisition & Categorisation (ESSLLI 2001), 2001.
- [Wu 2003] Hua Wu and Ming Zhou: Synonymous Collocation Extraction Using Translation Information, Proceedings of the 41st Annual Meeting of ACL, pp.120-127, 2003.
- [Yu 2002] Hong Yu, Vasileios Hatzivassiloglou, Carol Friedman, Andrey Rzhetsky, and W. John Wilbur: Automatic extraction of gene and protein synonyms from MEDLINE and journal articles, Proceedings of the American Medical Informatics Association (AMIA) Symposium, 2002.
- [Rapp 1995] Reinhard Rapp: Identifying word translations in non-parallel texts, Proceeding ACL '95 Proceedings of the 33rd annual meeting on Association for Computational Linguistics, Pages 320-322, 1995.
- [Fung 1997] Pascale Fung and Kathleen McKeown: Finding terminology translations from non-parallel corpora. In Proc. of the 5th Annual Workshop on Very Large Corpora, pages 192-202, 1997.
- [Fung 1998] Pascale Fung and Lo Yuen Yee: An IR approach for translating new words from nonparallel, comparable texts, In Proc. 17th COLING, pages 414-420, 1998.
- [Rapp 1999] Reinhard Rapp: Automatic identification of word translations from unrelated English and German corpora, Proceeding of ACL '99 Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, Pages 519-526, 1999.
- [Murakami 2004] Akiko Murakami and Tetsuya Nasukawa: Term aggregation: mining synonymous expressions using personal stylistic variations. In Proceedings of COLING '04, pages 806-812, 2004.
- [Andrade 2010] Daniel Andrade, Tetsuya Nasukawa, and Jun'ichi Tsujii: Robust measurement and comparison of context similarity for finding translation pairs. In Proceedings of COLING '10, pages 19-27, 2010.
- [那須川 2009] 那須川哲哉, Andrade, D., 海野裕也, 村松祐希, 山本和英: 言語横断テキストマイニングのための翻訳対抽出, 言語処理学会第15回年次大会発表論文集, pages 108-111, 2009.
- [海野 2010] 海野裕也, 那須川哲哉: 言語横断テキストマイニング, 第24回人工知能学会全国大会, 2010.
- [Bay 2008] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool: SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, 2008.
- [Lowe 2004] David G. Lowe: Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, Volume 60 Issue 2, pages 91-110, 2004.