

# Исследование применимости методов обработки естественного языка к задачам поиска и сравнения изображений машиностроительных чертежей

К.Н. Фигура<sup>1</sup>

<sup>1</sup> ФГБОУ ВО «Братский государственный университет», 665709, Россия, г. Братск, ул. Макаренко, д. 40

## Аннотация

Проведенные в работе исследования показывают, что применение технологии дескрипторов особых точек в чистом виде к задаче сравнения и поиска чертежей является неэффективным. Выявлено, что основной причиной этому служит наличие в чертежах большого количества идентичных элементов (рамки, основная надпись, выносные линии, элементы шрифтов и др.). Для решения данной проблемы предложено использование метода tf-idf (term frequency-inverse document frequency), широко известного в технологии обработки естественного языка. В исследовании вместо векторов слов, применяемых в оригинальной методике tf-idf, использовались дескрипторы особых точек изображений, вычисленных по алгоритмам ORB и BRISK. В результате исследования получены следующие выводы: 1) показана высокая эффективность предлагаемого подхода для поиска копии изображения-запроса в базе данных. Так, для всех изображений, предложенных для поиска и имеющих свои полные аналоги в базе данных, было выявлено наличие копий. 2) Количество выявленных изображений, являющихся модификациями изображения-запроса, разнится и зависит от алгоритма нахождения особых точек и дескрипторов. Так, при использовании ORB максимальное количество выявленных модифицированных аналогов составило 60 %, при использовании BRISK – 80 % от всех аналогов изображения, находящихся в базе данных. 3) Предлагаемый подход показывает ограниченную эффективность для нахождения изображений, которые можно отнести к тому же классу, что и изображение-запрос (например, чертеж экскаватора, бульдозера, автомобильного крана). Здесь максимальное количество ложных определений достигло 60 %.

**Ключевые слова:** обработка естественного языка, метод tf-idf, поиск изображений, анализ изображений, распознавание образов, цифровая обработка изображений.

**Цитирование:** Фигура, К.Н. Исследование применимости методов обработки естественного языка к задачам поиска и сравнения изображений машиностроительных чертежей / К.Н. Фигура // Компьютерная оптика. – 2022. – Т. 46, № 4. – С. 590-595. – DOI: 10.18287/2412-6179-CO-1030.

**Citation:** Figura KN. Investigation of the applicability of natural language processing methods to problems of searching and matching of machinery drawing images. Computer Optics 2022; 46(4): 590-595. DOI: 10.18287/2412-6179-CO-1030.

## Введение

Поиск изображений по их семантическому содержанию, степени схожести является современной, актуальной задачей. Данной тематике посвящено достаточно большое количество исследований [1–5]. Можно выделить несколько подходов, наиболее часто применяемых для решения задачи поиска изображений.

1. Метод дескрипторов особых точек. Наиболее часто для решения различных задач компьютерного зрения (распознавание образов, создание панорам, вычисление трехмерных реконструкций, дополненная реальность и др.) используется технология дескрипторов особых точек. Суть данной технологии можно свести к поиску некоторых особенностей изображения, т.е. таких точек изображения, окрестности которых можно отличить от окрестностей любых других точек, и вычисления для данных точек числовых характеристик (дескрипторов), которые их описывают.

Таким образом, большинство алгоритмов, предназначенных для решения данной задачи, состоят из двух частей: детектора, отвечающего за поиск ключевых точек, и дескриптора, предназначенного для вычисления числовых характеристик особых точек. В задачах поиска и сравнения изображений дескрипторы применяют для вычисления метрик, на основе которых выдвигается суждение о степени сходства изображений.

Разработано большое количество алгоритмов для определения дескрипторов особых точек, их описание и сравнение можно найти в обзорах [6–8].

2. Структуры дескрипторов, представленных в виде графов. Для тех случаев, когда сопоставления отдельных дескрипторов недостаточно для сравнения изображений, применяют методы объединения дескрипторов в определенные структуры при помощи графов. Далее для определения степени сходства изображений сопоставляются уже графы, а не отдельные дескрипторы [9–11].

3. Машинное обучение. Для поиска изображений по их контекстному содержанию активно используются различные методы машинного обучения. Данные методы применяются для классификации, генерации и сравнения изображений. Область машинного обучения для задач компьютерного зрения и распознавания образов является одной из наиболее динамично развивающихся [12–15].

Все приведенные выше методики обладают своими достоинствами и недостатками, но в целом все они могут достаточно эффективно справляться с задачей поиска изображений.

Тем не менее на сегодняшний день отсутствуют исследования, посвященные задаче поиска изображений, содержащих большое количество одинаковых, стандартных элементов. Примером такого рода изображений являются машиностроительные чертежи, оформленные в соответствии с Единой системой конструкторской документации [16]. Автоматизация процессов поиска чертежей, определения аналогов, выявление заимствований является насущной задачей при осуществлении этапа научно-исследовательской работы при подготовке проектов на любом машиностроительном предприятии. При этом наличие большого количества одинаковых элементов в семантически разных изображениях чертежей может значительно снизить эффективность применения известных методов сравнения и поиска изображений.

Поэтому цель данного исследования – выявление степени эффективности применения комбинированного подхода на основе технологии дескрипторов особых точек и метода tf-idf (term frequency-inverse

document frequency), ранее не применявшегося для сравнения изображений, содержащих большое количество идентичных элементов к поиску и сравнению изображений машиностроительных чертежей.

### 1. Применение технологии дескрипторов особых точек для определения степени схожести машиностроительных чертежей

Одним из наиболее эффективных методов, предназначенных для поиска и сравнения изображений, является метод дескрипторов особых точек. Данный метод известен достаточно давно и широко освещен в научной литературе [6–8, 17]. Его суть заключается в нахождении особых точек изображения (углы, участки окружностей и т.д.) и вычислении для данных точек векторов значений, которые их описывают (собственно дескрипторы). Разработано большое количество алгоритмов, предназначенных для решения данной задачи: Scale-Invariant Feature Transform (SIFT) [18], Speeded-Up Robust Features (SURF) [19], Binary Robust Independent Elementary Features (BRIEF) [20], Oriented FAST and Rotated BRIEF (ORB) [21], Binary Robust Invariant Scalable Keypoints (BRISK) [22], Features from Accelerated Test (FAST) [23] и др.

Для нашего исследования были выбраны распространенные дескрипторы FAST, BRISK и ORB. Для вычислительных экспериментов использовалась библиотека OpenCV 4.4.0 [24].

Особые точки, найденные на тестируемых изображениях алгоритмами FAST, BRISK и ORB, приведены на рис. 1 (обозначены окружностями).

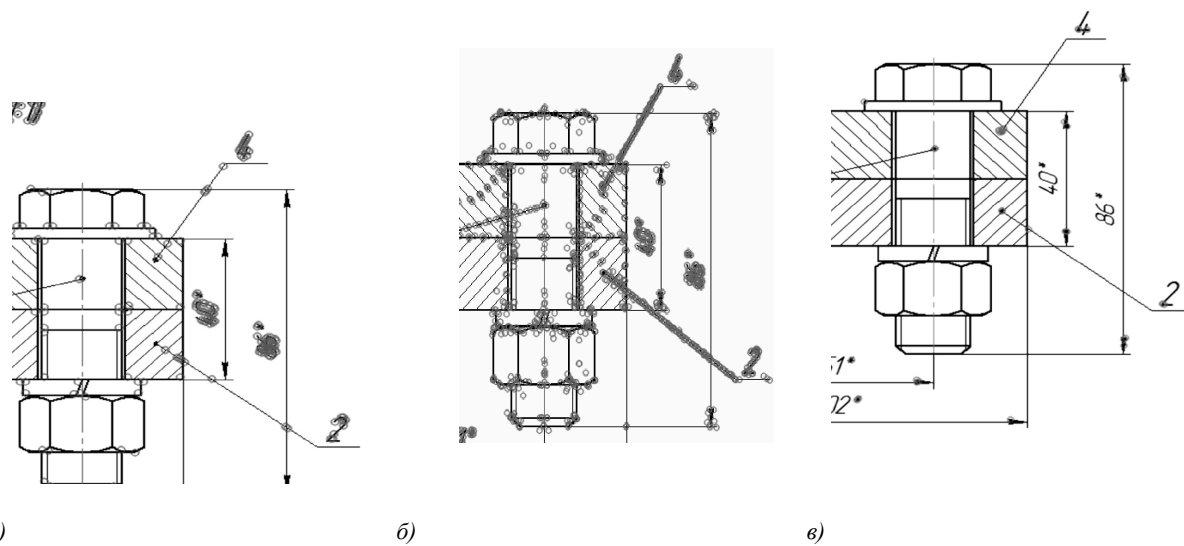


Рис. 1. Особые точки, найденные на изображениях чертежей: а) алгоритм FAST, б) алгоритм BRISK, в) алгоритм ORB

Как видно из результатов экспериментов, большое количество особых точек, найденных алгоритмами, приходится на элементы чертежей, свойственные всем чертежам, оформленным в соответствии с ЕСКД (т.е. рамка, таблица основной надписи, выносные линии, элементы шрифтов и т.д.). Также можно отме-

тить, что разные алгоритмы находят на одних и тех же чертежах разные особые точки и разное их количество.

После определения особых точек на тестируемых изображениях и вычисления дескрипторов было осуществлено сравнение дескрипторов с целью выясне-

ния степени схожести изображений. Сравнение дескрипторов было реализовано путем их попарного сопоставления и вычисления для каждой пары нормы Хэмминга. Полученные результаты свидетельствуют о том, что применение технологии дескрипторов особых точек в чистом виде не подходит для решения задачи поиска и сравнения изображений машиностроительных чертежей. Причиной низкой эффективности данного метода является наличие в семантически разных чертежах большого количества идентичных элементов. И, как видно из проведенных экспериментов, наименьшее расстояние между дескрипторами характерно именно для таких особых точек, даже для идентичных изображений. Это в конечном счете приводит к неправильной идентификации изображения.

**2. Применение методов обработки естественного языка для поиска и сравнения изображений машиностроительных чертежей**

Для эффективного поиска среди изображений, содержащих большое количество стандартных, идентичных элементов, не относящихся к семантическому содержанию изображения, необходимо исключить из рассмотрения или уменьшить степень влияния дескрипторов особых точек, определяемых для данных элементов.

Для решения этой задачи были применены методы, хорошо известные в области обработки естественного языка. Для поиска текстовых документов используется векторная модель представления текста. В данном случае текстовые документы представляются в векторном виде, отображающем частотное распределение слов в тексте. Далее для каждого слова определяется его «вес», т.е. степень значимости слова. Чаще всего для этих целей используется метод *tf-idf* (term frequency-inverse document frequency) [25]. В данном методе частота термина (или слова) *w* в документе *d* определяется как:

$$tf_{w,d} = \frac{n_w}{\sum_i n_i},$$

здесь в числителе представлено число появлений термина *w* в документе *d*, а в знаменателе – общее число слов в документе *d*. Обратная частотность документа может быть определена как:

$$idf_{w,d} = \log \left( \frac{|D|}{\sum \{d : w \in d\}} \right),$$

где *|D|* – общее число документов, *d* – число документов из *D*, содержащих терм *w*.

Перемножение величин *tf<sub>w,d</sub>* и *idf<sub>w,d</sub>* дает вес каждого отдельного слова в корпусе документов *D*. Чем чаще терм *w* появляется в документе *d*, тем выше значение *tf<sub>w,d</sub>* и, соответственно, выше значение

*tf<sub>w,d</sub>idf<sub>w,d</sub>*. Но чем чаще терм *w* появляется во всех документах корпуса *D*, тем значение *tf<sub>w,d</sub>idf<sub>w,d</sub>* ниже. Применяя данную идею к задаче поиска изображений, мы можем снизить влияние дескрипторов, относящихся к стандартным, одинаковым элементам изображений, встречающихся в большинстве чертежей.

В проведенном исследовании нами был применен следующий подход. Во-первых, для набора изображений, содержащего 1000 чертежей в формате *jpg*, были вычислены дескрипторы. Вычисление дескрипторов производилось по алгоритмам *ORB* и *BRISK*. Затем для выявления наиболее типичных образцов из множества дескрипторов была применена кластеризация по методу *k*-средних. Таким образом, визуальным аналогом слов в нашем случае являются центры классов, выделенных посредством алгоритма *k*-средних. После этого для каждого изображения, входящего в тестовый набор, были рассчитаны гистограммы визуальных слов, т.е. значения того, как часто в изображении появляется дескриптор, соответствующий ближайшему центру определенного класса. Затем на основе гистограмм слов были рассчитаны показатели *tf* и *idf* для каждого изображения в наборе тестовых изображений. Расчет производился по следующей методике:

$$tf = \frac{\sum_{i=1}^n [imh_i > 0]}{|d|},$$

где *imh* – визуальная гистограмма для данного изображения, *imh<sub>i</sub>* – количество дескрипторов в изображении, соответствующих ближайшему центру *i*-го класса, *n* – количество центров классов, на которое было разбито пространство дескрипторов тестируемых изображений (в нашем исследовании *n* было принято равным 100), *|d|* – количество дескрипторов в данном изображении. В данном случае сумма положительных значений в гистограмме визуальных слов делится на общее количество дескрипторов данного изображения.

Для определения величины *idf* использовалась следующая зависимость:

$$idf = \log \left( \frac{|D|}{\sum_{i=1}^n [imh_i > 0]} \right),$$

где *|D|* – общее количество дескрипторов во всем тестовом наборе изображений. Здесь общее количество дескрипторов делится на сумму положительных значений в гистограмме визуальных слов.

После этого для каждого изображения был рассчитан показатель *tf-idf*, представляющий собой произведение *tf* на *idf*.

Для поиска наиболее схожих изображений среди находящихся в тестовом наборе для изображения-запроса рассчитывался показатель  $tf-idf$  по выше описанной методике. Затем вычисленный показатель  $tf-idf$  изображения-запроса сравнивался по метрике косинусного расстояния с показателями  $tf-idf$  всех изображений, находящихся в базе данных. Набор базовых изображений представляет собой коллекцию черте-

жей, оформленных в соответствии с ЕСКД. В наборе можно выделить несколько классов машин (экскаваторы, бульдозеры, автомобильные краны). С целью проверки эффективности предлагаемого подхода в плане поиска аналогов и заимствований некоторые чертежи были модифицированы (поменяны местами виды, изменена основная надпись, удалена рамка, изменены конструктивные элементы).

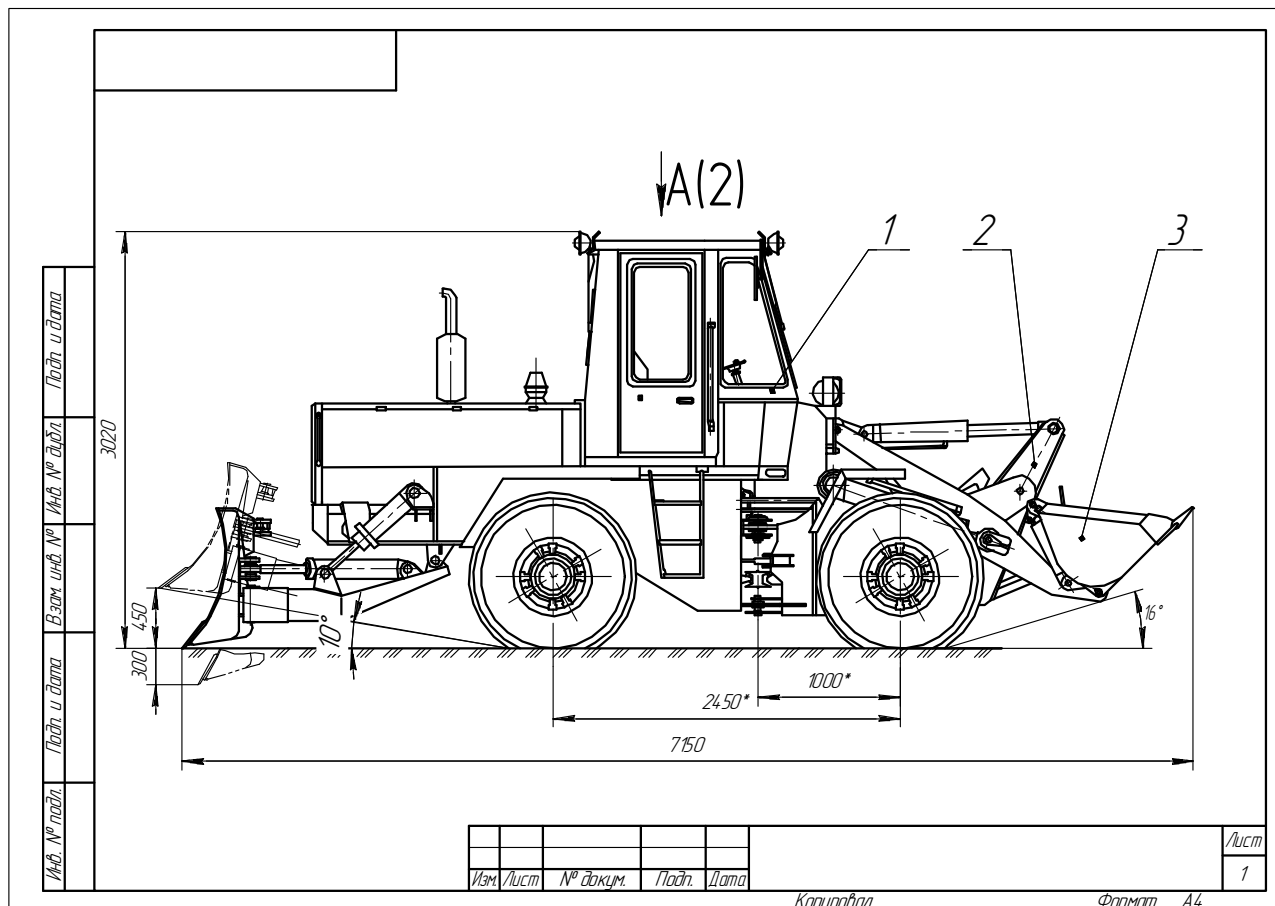


Рис. 2. Пример изображения-запроса

Пример изображения-запроса приведен на рис. 2. На рис. 3 приведены результаты поиска изображений в базе данных. Здесь показаны усредненные значения количества найденных изображений для каждого типа модификации, нормированные к 100 %. Всего было использовано 100 тестовых изображений-запросов, для каждого из которых было подготовлено восемь модификаций различного типа и две полных копии. Поиск каждый раз осуществлялся по всему набору изображений, состоящему из 1000 экземпляров.

### Заключение

Полученные результаты позволяют сделать следующие выводы.

1) Показана высокая эффективность предлагаемого подхода для поиска копий изображения-запроса в базе данных. Так, для всех изображений, предложенных для поиска и имеющих свои полные аналоги в

базе изображений, во всех 100 тестовых случаях наличие копий было выявлено.

2) Количество выявленных изображений, являющихся модификациями изображения-запроса, различается и зависит от алгоритма нахождения особых точек и дескрипторов. Так, например, при использовании ORB максимальное количество выявленных модифицированных аналогов составило 60 %, при использовании BRISK – 80 % от всех аналогов изображения, находящихся в базе данных.

3) Предлагаемый подход показывает ограниченную эффективность для нахождения изображений, которые можно отнести к тому же классу, что и изображение-запрос (например, чертеж экскаватора, бульдозера, автомобильного крана). Основным критерием, служившим для оценки данного показателя, являлось количество изображений техники того же класса, что и изображение-запрос, но не являющихся

его копиями или модификациями, возвращаемых в результате поискового запроса. Здесь максимальное количество ложных определений достигло 60 %.

В целом можно отметить, что задача контекстного поиска изображений, содержащих большое количество идентичных элементов, является сложной, требующей дальнейшей проработки и исследований. Ограничения описанного в работе подхода в области

нахождения объектов подобного класса, но не являющихся непосредственными аналогами, возможно решить путем комбинации предлагаемой методики и подхода, основанного на применении графов. Кроме того, в качестве направления дальнейших исследований можно указать выяснение эффективности других методов повышения релевантности поисковых запросов, отличных от *tf-idf*, например, Okapi BM25 [26].

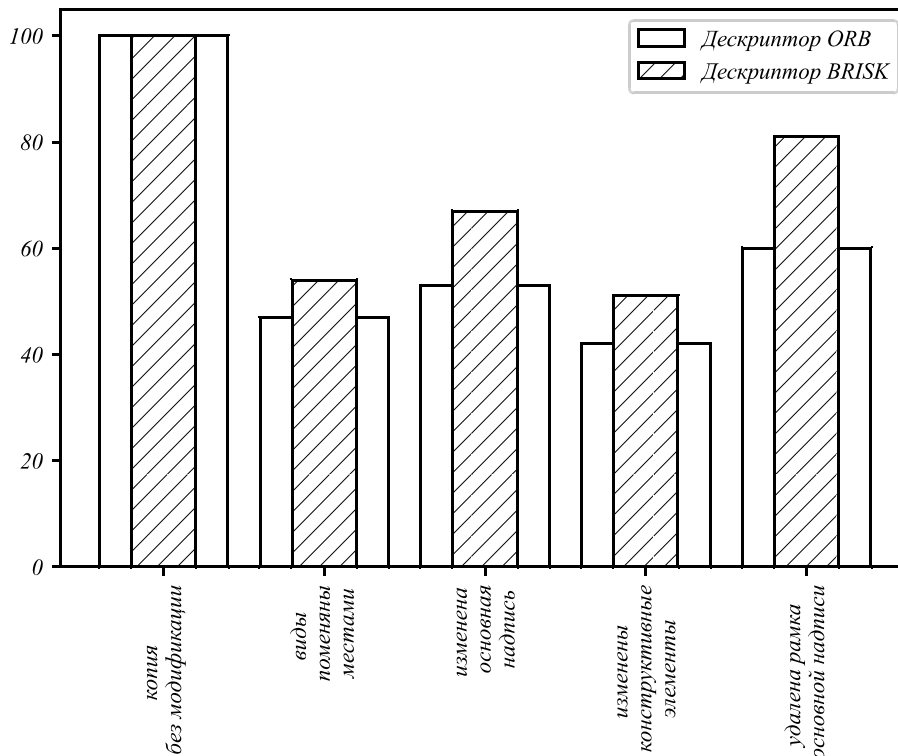


Рис.3. Результаты поиска изображений в базе данных

Предлагаемый подход показал свою эффективность при поиске аналогов и модифицированных копий чертежей в базе данных и может использоваться в качестве основы для разработки систем графического антиплагиата [27–29]. Кроме того, предлагаемая методика позволяет производить поиск без предварительной обработки чертежей. Также следует отметить, что описанный в данном исследовании подход может быть применен к изображениям любого семантического содержания.

### References

- [1] Ahmed KT, Ummesafi S, Iqbal A. Content based image retrieval using image features information fusion. *Inf Fusion* 2019; 51: 76-99. DOI: 10.1016/j.inffus.2018.11.004.
- [2] Duan G, Yang J, Yang Y. Content-based image retrieval research. *Phys Procedia* 2011; 22: 471-477. DOI: 10.1016/j.phpro.2011.11.073.
- [3] Tzelepi M, Tefas A. Deep convolutional learning for Content Based Image Retrieval. *Neurocomputing* 2018; 275: 2467-2478. DOI: 10.1016/j.neucom.2017.11.022.
- [4] Haji MS, Alkawaz MH, Rehman A, Saba T. Content-based image retrieval: a deep look at features prospectus. *Int J Comput Vis Robot* 2019; 9(1): 14-38. DOI: 10.1504/IJCVR.2019.098004.
- [5] Rana SP, Dey M, Siarry P. Boosting content based image retrieval performance through integration of parametric & nonparametric approaches. *J Vis Commun Image Represent* 2019; 58: 205-219. DOI: 10.1016/j.jvcir.2018.11.015.
- [6] Mouats T, Aouf N, Nam D, Vidas S. Performance evaluation of feature detectors and descriptors beyond the visible. *J Intell Robot Syst* 2018; 92: 33-63. DOI: 10.1007/s10846-017-0762-8.
- [7] Mukherjee D, Wu QMJ, Wang G. A comparative experimental study of image feature detectors and descriptors. *Mach Vis Appl* 2015; 26(4): 443-466. DOI: 10.1007/s00138-015-0679-9.
- [8] Saha SK, Xiao D, Frost S, Kanagasigam Y. Performance evaluation of state-of-the-art local feature detectors and descriptors in the context of longitudinal registration of retinal images. *J Med Syst* 2018; 42(2): 57. DOI: 10.1007/s10916-018-0911-z.
- [9] Ma J, Jiang X, Fan A, Jiang J, Yan J. Image matching from handcrafted to deep features: A survey. *Int J Comput Vis* 2021; 129: 23-79. DOI: 10.1007/s11263-020-01359-2.
- [10] Zakharov AA, Zhiznyakov AL, Titov VS. A method for feature matching in images using descriptor structures. *Computer Optics* 2019; 43(5): 810-817. DOI: 10.18287/2412-6179-2019-43-5-810-817.
- [11] Zakharov AA, Barinov AE, Zhiznyakov AL, Titov VS. Object detection in images with a structural descriptor

- based on graphs. *Computer Optics* 2018; 42(2): 283-290. DOI: 10.18287/2412-6179-2018-42-2-283-290.
- [12] Zheng L, Yang Y, Tian Q. SIFT meets CNN: A decade survey of instance retrieval. *IEEE Trans Pattern Anal Mach Intell* 2018; 40(5): 1224-1244. DOI: 10.1109/TPAMI.2017.2709749.
- [13] Alrahhhal M, Supreethi KP. Content-based image retrieval using local patterns and supervised machine learning techniques. 2019 Amity Int Conf on Artificial Intelligence (AI-CAI) 2019: 118-124. DOI: 10.1109/AICAI.2019.8701255.
- [14] Ali A, Sharma S. Content based image retrieval using feature extraction with machine learning. 2017 Int Conf on Intelligent Computing and Control Systems (ICICCS) 2017: 1048-1053. DOI: 10.1109/ICCONS.2017.8250625.
- [15] Saritha RR, Paul V, Kumar PG. Content based image retrieval using deep learning process. *Cluster Comput* 2019; 22: 4187-4200. DOI: 10.1007/s10586-018-1731-0.
- [16] GOST 2.001-2013 Unified system for design documentation (ESKD). General Provisions (as amended) 22 November 2013 [In Russian]. Source: <https://docs.cntd.ru/document/1200106859>.
- [17] Krasnabayeu YA, Chistabayeu DV, Malyshev AL. Comparison of binary feature points descriptors of images under distortion conditions. *Computer Optics* 2019; 43(3): 434-445. DOI: 10.18287/2412-6179-2019-43-3-434-445.
- [18] Lowe DG. Object recognition from local scale-invariant features. *Proc Seventh IEEE Int Conf on Computer Vision* 1999; 2: 1150-1157. DOI: 10.1109/ICCV.1999.790410.
- [19] Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Comput Vis Image Underst* 2008; 110: 346-359. DOI: 10.1016/j.cviu.2007.09.014.
- [20] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: Binary robust independent elementary features. In Book: Daniilidis K, Maragos P, Paragios N, eds. *Computer vision – ECCV 2010*. Berlin, Heidelberg: Springer; 2010: 778-792. DOI: 10.1007/978-3-642-15561-1\_56.
- [21] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. 2011 Int Conf on Computer Vision 2011: 2564-2571. DOI: 10.1109/ICCV.2011.6126544.
- [22] Leutenegger S, Chli M, Siegwart RY. BRISK: Binary robust invariant scalable keypoints. 2011 Int Conf on Computer Vision 2011: 2548-2455. DOI: 10.1109/ICCV.2011.6126542.
- [23] Rosten E, Drummond T. Machine learning for high-speed corner detection. In Book: Leonardis A, Bischof H, Pinz A, eds. *Computer vision – ECCV 2006. Part I*. Berlin, Heidelberg: Springer; 2006: 430-443. DOI: 10.1007/11744023\_34.
- [24] OpenCV 4.4.0. OpenCV 2020. Source: <https://opencv.org/opencv-4-4-0/>.
- [25] Roelleke T, Wang J. TF-IDF uncovered: a study of theories and probabilities. *Proc 31st Annual Int ACM SIGIR Conf on Research and Development in Information Retrieval* 2008: 435-442. DOI: 10.1145/1390334.1390409.
- [26] Whissell JS, Clarke CLA. Improving document clustering using Okapi BM25 feature weighting. *Inf Retrieval* 2011; 14: 466-487. DOI: 10.1007/s10791-011-9163-y.
- [27] Bruno A, Cattaneo G, Petrillo UF, Narducci F, Roscigno G. Distributed anti-plagiarism checker for biomedical images based on sensor noise. In Book: Battiato S, Farinella GM, Leo M, Gallo G, eds. *New trends in image analysis and processing – ICIAP 2017*. Cham: Springer International Publishing; 2017: 343-352. DOI: 10.1007/978-3-319-70742-6\_32.
- [28] Iwanowski M, Cacko A, Sarwas G. Comparing images for document plagiarism detection. In Book: Chmielewski LJ, Datta A, Kozera R, Wojciechowski K, eds. *Computer Vision and Graphics*. Cham: Springer International Publishing; 2016: 532-543. DOI: 10.1007/978-3-319-46418-3\_47.
- [29] Chen Y, Gan L, Zhang S, Guo W, Chuang Y, Zhao X. Plagiarism detection in homework based on image hashing. In Book: Zou B, Han Q, Sun G, Jing W, Peng X, Lu Z, eds. *Data science*. Singapore: Springer Singapore; 2017: 424-432. DOI: 10.1007/978-981-10-6388-6\_35.

### Сведения об авторе

**Фигура Константин Николаевич**, 1985 года рождения, в 2008 году окончил Братский государственный университет. В 2013 году защитил кандидатскую диссертацию по специальности «Машины, агрегаты и процессы», работает доцентом на кафедре математики, информатики и физики Братского государственного университета. Область научных интересов: машинное обучение, обработка графических изображений, методы поиска изображений по их семантическому содержанию. E-mail: [kfigura@internet.ru](mailto:kfigura@internet.ru).

ГРНТИ: 20.19.29

Поступила в редакцию 24 августа 2021 г. Окончательный вариант – 31 октября 2021 г.

---

# Investigation of the applicability of natural language processing methods to problems of searching and matching of machinery drawing images

*K.N. Figura<sup>1</sup>*

<sup>1</sup> *Bratsk State University,  
665709, Bratsk, Russia, Makarenko 40*

## *Abstract*

In this work it is shown that the application of the technique of local feature descriptors in its pure form to the task of searching and matching of drawings is ineffective. It is revealed that this is mainly due to the presence in the drawings of a large number of identical elements (frames, a title block, extension lines, font elements, etc.). It is proposed that this problem should be solved using a tf-idf (term frequency-inverse document frequency) method, which is widely known in natural language processing. In the study, instead of the word vectors used in the original tf-idf technique, descriptors of image feature points calculated using the ORB and BRISK algorithms were used. The study has led to the following conclusions: 1) the proposed approach offers high efficiency in finding a copy of the image-query in the database. Thus, copies of all images presented for search and having their full analogs in the database are revealed. 2) The identification rate of modified image-queries varies, depending on the algorithm used for finding keypoints and descriptors. So, the maximum percentage of identified modified analogs is 60% when using ORB and 80% when using BRISK - out of all image analogs in the database. 3) The proposed approach shows a limited efficiency in finding images that can be attributed to the same class as the image queries (for example, a drawing of an excavator, a bulldozer, or a truck crane). Here, the maximum proportion of false identification has reached 60%.

*Keywords:* natural language processing, tf-idf method, image retrieval, image analysis, pattern recognition, digital image processing.

*Citation:* Figura KN. Investigation of the applicability of natural language processing methods to problems of searching and matching of machinery drawing images. *Computer Optics* 2022; 46(4): 590-595. DOI: 10.18287/2412-6179-CO-1030.

---

## *Author's information*

**Konstantin Nikolaevich Figura**, born in 1985, graduated from Bratsk State University in 2008. In 2013 he defended his Ph.D. thesis in the specialty «Machines, Aggregates and Processes», works as an assistant professor of Mathematics, Informatics and Physics department of Bratsk State University. Research interests: machine learning, graphic image processing, content-based image retrieval. E-mail: [kfigura@internet.ru](mailto:kfigura@internet.ru).

---

*Received August 24, 2021. The final version – October 31, 2021.*

---