

УДК 004.932.2

Исследование перцептивных хеш-функций изображений

Рудаков И. В.¹, Васютович И. М.^{1,*}

[*MONOmah_V@mail.ru](mailto:MONOmah_V@mail.ru)

¹МГТУ им. Н.Э. Баумана, Москва, Россия

В данной статье проведен анализ существующих методов поиска изображений по содержанию с использованием перцептивных хешей и возможности по их применению для поиска подобных изображений. Рассмотрены следующие методы генерации перцептивных хешей: хеш по среднему с применением низкочастотного фильтра, хеш-функция на основе гистограммы цветов, хеш-функция на основе дискретного косинусного преобразования. Исследования проводились по трем направлениям: скорость работы метода, способность эффективно различать изображения, устойчивость к трансформациям изображений (масштабирование, поворот, горизонтальное отражение). По результатам исследования был создан алгоритм, представляющий собой комбинацию методов на основе цветowych гистограмм и дискретного косинусного преобразования, применение которого при решении задачи поиска похожих изображений дает наиболее точные и стабильные результаты.

Ключевые слова: поиск изображений по содержанию, перцептивная хеш-функция

Введение

Стремительное развитие портативной электроники, в частности смартфонов и планшетов, и широкая доступность стабильного широкополосного доступа в интернет и к социальным сетям дали возможность пользователям производить и обмениваться большим объемом мультимедиа-информации. Причем объемы передаваемой и хранимой информации с течением времени только увеличиваются [1]. Все это приводит как к тому, что пользователям необходимо каким-то образом управлять своими коллекциями фотографий, видео- и аудиозаписей, так и к тому, что интернет-сервисам и контент-провайдерам необходимо справляться со все растущим объемом хранимой и передаваемой информации как можно более эффективно.

С другой стороны, по тем же причинам все более остро встает проблема контроля за передаваемой в сети информацией – отслеживание нарушений авторских прав, обнаружение и предотвращение злоупотребления передачей данных (спам).¹

Среди постоянно увеличивающегося количества способов поиска похожих изображений существует группа относительно универсальных и эффективных алгоритмов. Перцептивные хеш-функции – семейство алгоритмов, которые производят

так называемые «отпечатки» на основе отдельных признаков различных форм мультимедиа-информации [2].

Поэтому актуальным является исследование работы различных перцептивных хеш-функций при решении задачи нахождения коэффициентов подобия изображений и разработка метода, устойчивого к различным модификациям изображений.

1. Перцептивные хеш-функции изображений

1.1 Постановка задачи

Целью данной работы является исследование алгоритмов вычисления перцептивных хеш-функций при решении задачи поиска похожих изображений, а именно:

- Изучение скорости работы алгоритмов.
- Изучение способности алгоритмов различать изображения.
- Изучение устойчивости алгоритмов к различным видам трансформации изображений.

На основе полученных данных необходимо определить границы применимости рассматриваемых алгоритмов и разработать метод, который является одновременно устойчивым к трансформациям изображений и обладает приемлемой скоростью работы и способностью различать изображения.

Перцептивные хеш-функции были созданы для установления равенства восприятия мультимедийного контента. Подобные функции были разработаны для различных типов мультимедиа – изображений, аудио- и видеозаписей.

В применении к изображениям данные методы относятся классу методов, выделяющих низкоуровневые признаки/формы.

Перцептивные хеш-функции извлекают определенные признаки из мультимедиа-объекта и вычисляют хеш на основании этих признаков. Полученный хеш является своего рода отпечатком изображения. При сравнении изображений значения хешей сравниваются с использованием специальных функций. Эти функции вычисляют коэффициент различия или сходства между двумя значениями перцептивных хешей. Заключение о схожести изображений затем составляется на основании выбранного порогового значения.

Наибольший интерес [3] представляет исследование следующих алгоритмов:

1. Низкочастотный фильтр и хеш по среднему.
2. Хеш на основе гистограммы цветов.
3. Хеш на основе дискретного косинусного преобразования.

1.2 Низкочастотный фильтр и хеш по среднему

В изображениях высокие частоты обеспечивают детализацию, а низкие частоты определяют структуру изображения. Так, детализованное изображение высокого разрешения содержит большое количество высоких частот. Уменьшенная копия такого изображения практически целиком состоит из низких частот. Таким образом, можно построить алгоритм вычисления хеша изображения:

1. Перевести изображение в оттенки серого.

Изображение преобразуется из 64 трехцветных пикселей в 64 цвета.

2. Уменьшить изображение.

При уменьшении изображение до квадрата 8×8 пикселей происходит фильтрация высоких частот и выделяется структура изображения.

3. Вычислить среднее значение от полученных 64 цветов:

$$average = \frac{\sum_{i=1}^n v_i}{n}$$

4. Построить хеш на основе среднего значения цвета:

$$h_i = \begin{cases} 1, v_i > average \\ 0, v_i \leq average \end{cases}$$

Полученный таким образом хеш является устойчивым к масштабированию изображений и к изменению яркости и контраста. Кроме того, ввиду своей простоты, данный метод обладает очень низкой вычислительной сложностью.

1.3 Хеш на основе гистограммы цветов

Одним из простейших методов получения отпечатка изображения является построение гистограммы цветов. Гистограмма цветов – гистограмма уровней насыщенности изображения. Она позволяет оценить количество и разнообразие оттенков изображения, а также общий уровень яркости.

В дискретном пространстве цветов, определенном некоторым количеством цветовых осей (например, красный, зеленый, синий), гистограмма получается путем дискретизации цветов изображения и подсчетом количества пикселей, имеющих каждый конкретный цвет [4].

Данный метод часто используется в CBIR-системах [4] и является достаточно быстрым и простым в реализации. Поиск изображений с использованием гистограмм цвета устойчив к масштабированию (при нормализации гистограммы), аффинным преобразованиям.

Однако, этот метод не слишком хорошо показывает себя в ситуациях, когда разные изображения имеют очень похожую цветовую гамму, например, рисунки.

1.4 Хеш на основе дискретного косинусного преобразования (DCT)

Дискретное косинусное преобразование представляет собой разновидность преобразования Фурье. Графическое изображение можно рассматривать как совокупность пространственных волн, причем оси X и Y совпадают с шириной и высотой картинки, а по оси Z откладывается значение цвета соответствующего пикселя изображения. Дискретное косинусное преобразование позволяет переходить от пространственного представления изображения к его спектральному представлению и обратно.

Данный метод основан на таком важном свойстве дискретного косинусного преобразования, как «уплотнение энергии», что позволяет обнаружить ключевые особенности изображений [2].

Существует восемь различных вариантов DCT. Наиболее распространенный вариант – DCT-2 и именно его часто и называют просто DCT:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot \cos\left(\frac{(2n+1) \cdot k \pi}{2N}\right), k = \overline{0, N-1}$$

где $x[n], n = \overline{0, N-1}$ – последовательность -точек сигнала.

Двумерное дискретное косинусное преобразование эквивалентно последовательному применению DCT по каждому измерению:

$$X[k_1, k_2] = \sum_{n_1=0}^{N-1} \left(\sum_{n_2=0}^{N-1} x[n_1, n_2] \cdot \cos\left(\frac{(2n_2 + 1) \cdot k_2 \pi}{2N}\right) \right) \cdot \cos\left(\frac{(2n_1 + 1) \cdot k_1 \pi}{2N}\right). \quad (1)$$

Алгоритм построения перцептивного хеша с использованием DCT:

1. Перевести изображение в оттенки серого.
2. Уменьшить изображение до квадрата 32×32 пикселя.
3. Выполнить дискретное косинусное преобразование для полученного изображения;
4. Выделить из полученной матрицы матрицу 8×8 начиная со второй строки и столбца;
5. Вычислить среднее значение элементов матрицы;
6. Построить хеш из полученной матрицы путем сравнения элементов со средним.

Значения косинусов из формулы (1) для ускорения работы алгоритма можно рассчитать заранее с учетом того, что обрабатываемое изображение будет иметь размеры 32×32 пикселя, то есть $N=32$. Тогда, формулу (1) можно переписать следующим образом:

$$X[k] = \sum_{n=0}^{N-1} C[k, n] \cdot x[k], n = \overline{0, N-1}$$

где $C[n, m]$ – элементы матрицы DCT, которые рассчитываются следующим образом:

$$C[k, n] = \cos\left(\frac{(2n + 1) \cdot k \pi}{2N}\right), n, k = \overline{0, N-1} \quad (2)$$

Тогда, имея изображение I , представленное в виде двумерного массива, можно получить его двумерное косинусное преобразование следующим образом:

$$DCT(I) = C \cdot I \cdot C'$$

где C – матрица DCT, рассчитанная по формуле (2), а C' – транспонированная матрица C .

Данный алгоритм является наиболее вычислительно сложным среди всех представленных. Но его работу можно ускорить путем заблаговременного вычисления значений косинусов (матрица C) и выполнения преобразования только для требуемой части матрицы (шаг 4 алгоритма).

На основании исследованных алгоритмов создан комбинированный метод вычисления хеш-функции с использованием дискретного косинусного преобразования и цветовой гистограммы. Коэффициент подобия при использовании комбинации алгоритмов вычисляется по следующей формуле:

$$EP(I, M) = EP_{DCT}(I, M) \cdot \eta + EP_{HIST}(I, M) \cdot (1 - \eta)$$

где I, M – сравниваемые изображения; EP_{DCT}, EP_{HIST} – коэффициенты подобия, вычисленные с использованием DCT и гистограммы, соответственно; η и $(1 - \eta)$ – «веса» вычисленных коэффициентов.

Коэффициент η был подобран эмпирическим путем. В ходе тестирования значения η изменялись в диапазоне от 0.2 до 0.8 с шагом 0.2. В результате, было выявлено, что при значении коэффициента равном 0.8 наблюдается наилучший баланс между устойчивостью к рассматриваемым преобразованиям и способностью различать изображения.

2. Проведение исследования

Работа перцептивных хеш-функций может быть оценена различными способами. Тестирование и сравнение различных алгоритмов на различных множествах входных данных и параметров самих алгоритмов не только позволит сделать выводы об эффективности хеш-функций для конкретных областей применения, но и определить «узкие места» в алгоритмах и провести их оптимизацию.

Исследования проведены по трем направлениям:

- Исследование скорости вычисления перцептивной хеш-функции.
- Исследование способностей хеш-функции по различению изображений.
- Исследование устойчивости хеш-функции к различного рода модификациям изображений.

В исследованиях использовались два различных набора изображений, которые были получены сайта Wikimedia Commons (<http://commons.wikimedia.org>):

- Набор случайных изображений различной тематики. Данный набор состоит из 20 изображений хорошего качества – среднее разрешение 3514×2875 пк, средний размер 5.2 МБ.
- Набор изображений одной тематики (утки, плавающие в пруду). Данный набор также состоит из 20 изображений хорошего качества – среднее разрешение 3480×2520 пк, средний размер 3.3 МБ.

2.1 Исследование скорости вычисления перцептивной хеш-функции

Скорость расчета хеш-функции особенно важна, когда необходимо обработать большое количество изображений.

Для данного исследования использовался набор случайных изображений. Результаты исследования показаны на рисунке 1.

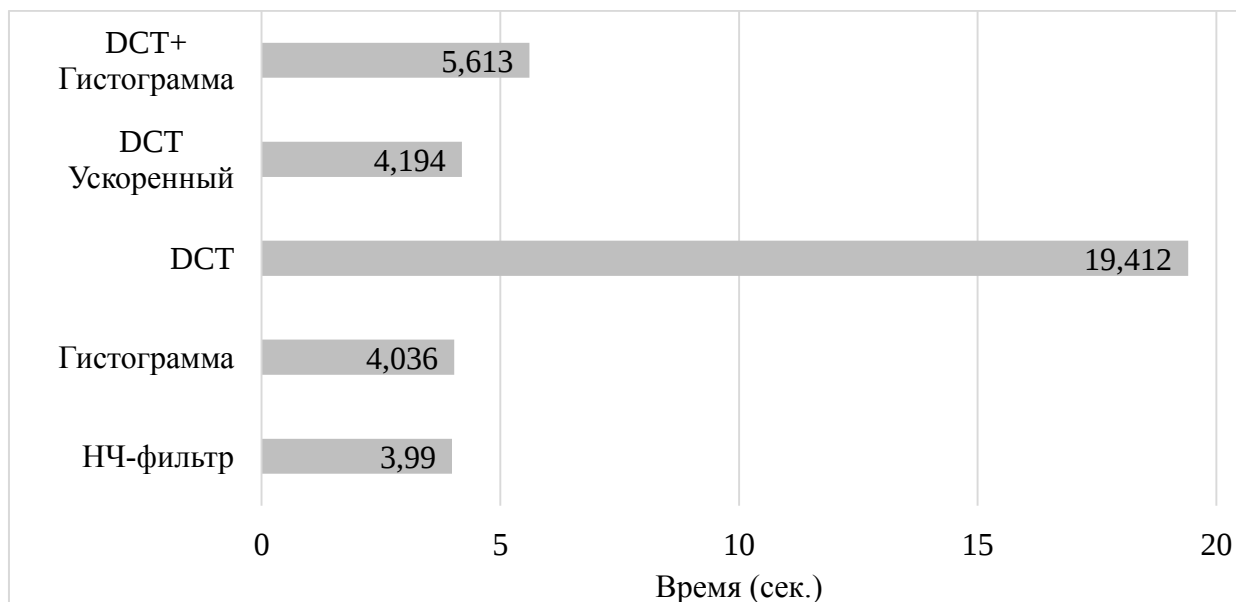


Рисунок 1. Результаты измерений скорости работы алгоритмов

По результатам данного исследования можно заключить, что скорости работы алгоритмов практически равны и не являются определяющим фактором при выборе алгоритма для выполнения поиска подобных изображений.

2.2 Исследование способностей хеш-функции по различению изображений

При сравнении двух различных изображений, идеальная перцептивная хеш-функция должна всегда приводить к коэффициенту подобия изображений равному 0.5 [2].

Интерес также представляет зависимость распределения коэффициентов подобия от рассматриваемого множества изображений. Так, если входными данными будет множество похожих изображений (например, все рассматриваемые изображения представляют собой горный пейзаж), для функции достичь «идеального» коэффициента подобия может быть очень сложно. Поэтому данное исследование проводилось на двух множествах изображений, описанных выше. Для выполнения исследования каждая из хеш-функций была использована для вычисления коэффициента подобия для каждой возможной пары изображений из тестируемого множества.

Значения полученных коэффициентов представлены в таблице 1.

Таблица 1. Результаты оценки хеш-функций по способностям различения изображений

	НЧ-фильтр	Гистограмма	DCT	DCT + Гистограмма
Набор случайных изображений				
Среднее	0.476	0.202	0.493	0.492
Максимум	0.809	0.375	0.656	0.719
Минимум	0.180	0.007	0.313	0.328
Набор изображений одной тематики				
Среднее	0.481	0.181	0.503	0.502
Максимум	0.867	0.299	0.719	0.719
Минимум	0.125	0.010	0.328	0.344

По результатам эксперимента видно, что хеш-функция на основе дискретного косинусного преобразования наряду с комбинированной функцией обладает лучшими возможностями по различению изображений среди представленных.

2.3 Исследование устойчивости хеш-функций к модификациям изображения

Для тестирования устойчивости хеш-функций используются различные преобразования изображений.

Горизонтальное отражение. При горизонтальном отражении, двоичное представление изображения меняется практически полностью, в отличие от семантических изменений и изменения восприятия отраженного изображения человеком. Именно поэтому было проведено тестирование хеш-функций на предмет устойчивости к таким модификациям. Результат тестирования представлен в таблице 2, на рисунке 2 показано распределение коэффициентов подоби

Таблица 2. Коэффициенты подоби

	НЧ-фильтр	Гистограмма	DCT	DCT + Гистограмма
Среднее	0.291	0.000	0.471	0.096
Максимум	0.578	0.000	0.578	0.113
Минимум	0.031	0.000	0.094	0.019

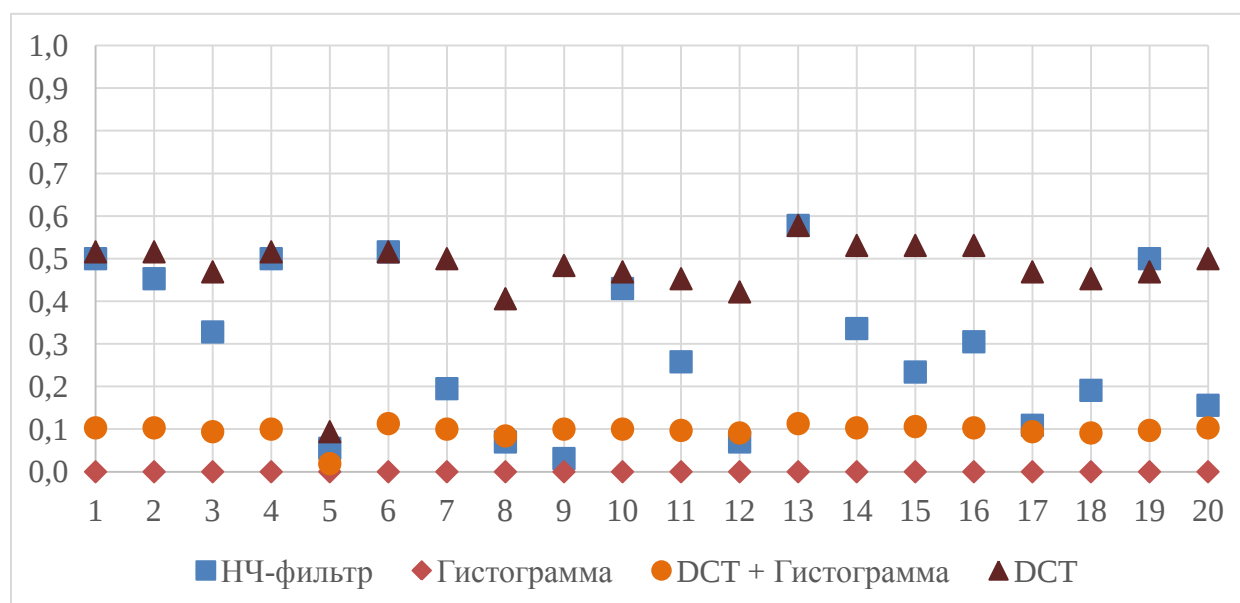


Рисунок 2. Коэффициенты подоби

Можно заметить, что функции, основанные на выделении форм на изображении непригодны для поиска изображений, полученных горизонтальным отражением. В свою очередь, метод с использованием цветowych гистограмм показывает практически полное совпадение изображений.

Уменьшение размера (масштабирование). При проведении данного исследования изображения пропорционально уменьшались до ширины 1024 пикселя и сравнивались с исходным. Результаты эксперимента представлены в таблице 3 и на рисунке 3.

Таблица 3. Коэффициенты подобия масштабированных изображений

	НЧ-фильтр	Гистограмма	DCT	DCT + Гистограмма
Среднее	0.023	0.000	0.003	0.000
Максимум	0.133	0.000	0.016	0.006
Минимум	0.000	0.000	0.000	0.000

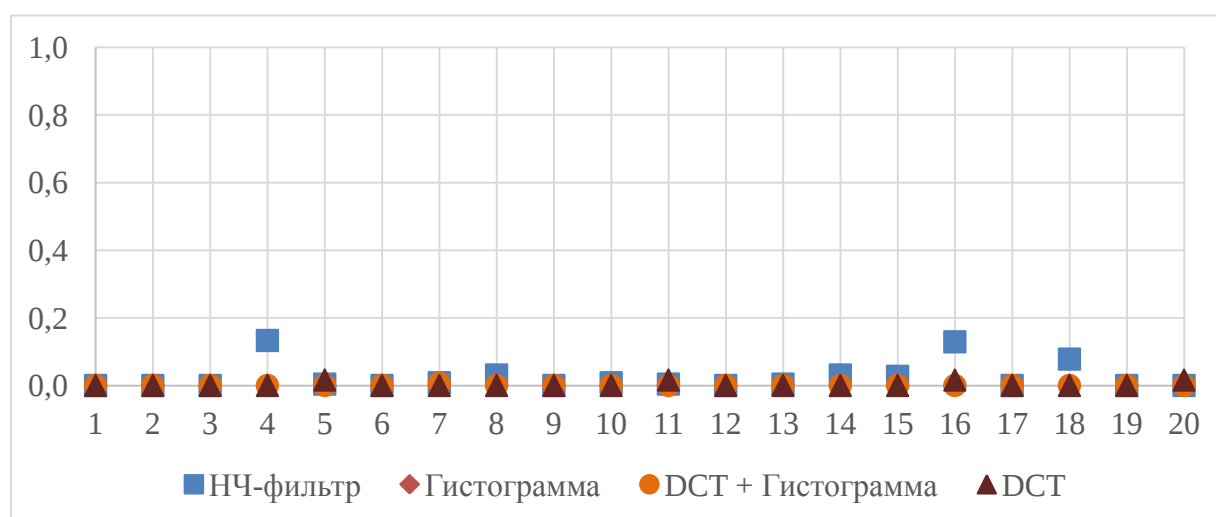


Рисунок 3. Коэффициенты подобия масштабированных изображений\

Из результатов исследования видно, что все представленные алгоритмы устойчивы к масштабированию изображений.

Поворот. При проведении данного исследования изображения были повернуты на 5° с применением бикубической интерполяции. При таком виде преобразования наилучшие результаты получены при использовании хеш-функции на основе гистограммы цветов (Таблица 4). На рисунке 4 показано распределение коэффициентов подобия в данном эксперименте.

Таблица 4. Коэффициенты подобия повернутых изображений

	НЧ-фильтр	Гистограмма	DCT	DCT + Гистограмма
Среднее	0.095	0.026	0.213	<i>0.131</i>
Максимум	0.207	0.087	0.375	<i>0.244</i>
Минимум	0.008	0.000	0.109	<i>0.031</i>

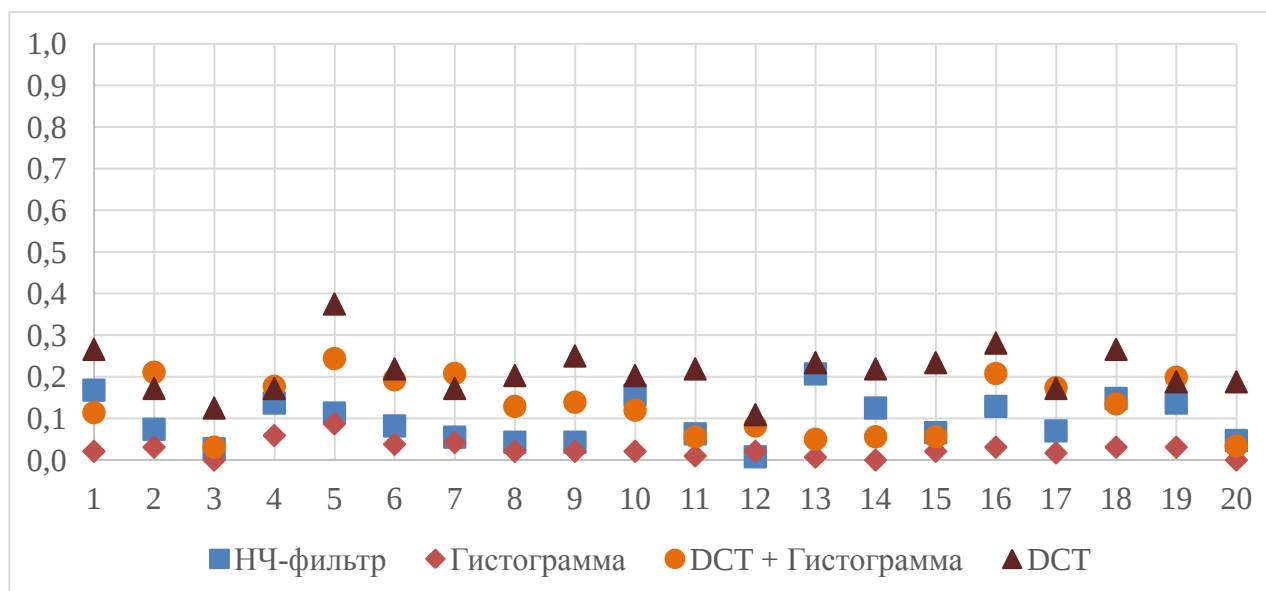


Рисунок 4. Коэффициенты подобия повернутых изображений

Поворот на 90 градусов. При проведении данного исследования изображения были повернуты на 90° по часовой стрелке с применением бикубической интерполяции. Как видно из таблицы 5, хеш-функция на основе гистограммы цветов является абсолютно устойчивой к данному виду преобразований. В свою очередь, алгоритмы на основе дискретного косинусного преобразования и низкочастотного фильтра считают такие изображения практически полностью различными. Комбинированный метод ввиду абсолютной устойчивости гистограммы показывает хорошие результаты – средний коэффициент различия изображений меньше 0.01. На рисунке 5 показано распределение коэффициентов подобия в данном эксперименте.

Таблица 4. Коэффициенты подобия повернутых на 90° изображений

	НЧ-фильтр	Гистограмма	DCT	DCT + Гистограмма
Среднее	0.433	0.000	0.482	<i>0.097</i>
Максимум	0.602	0.000	0.563	<i>0.116</i>
Минимум	0.070	0.000	0.313	<i>0.063</i>

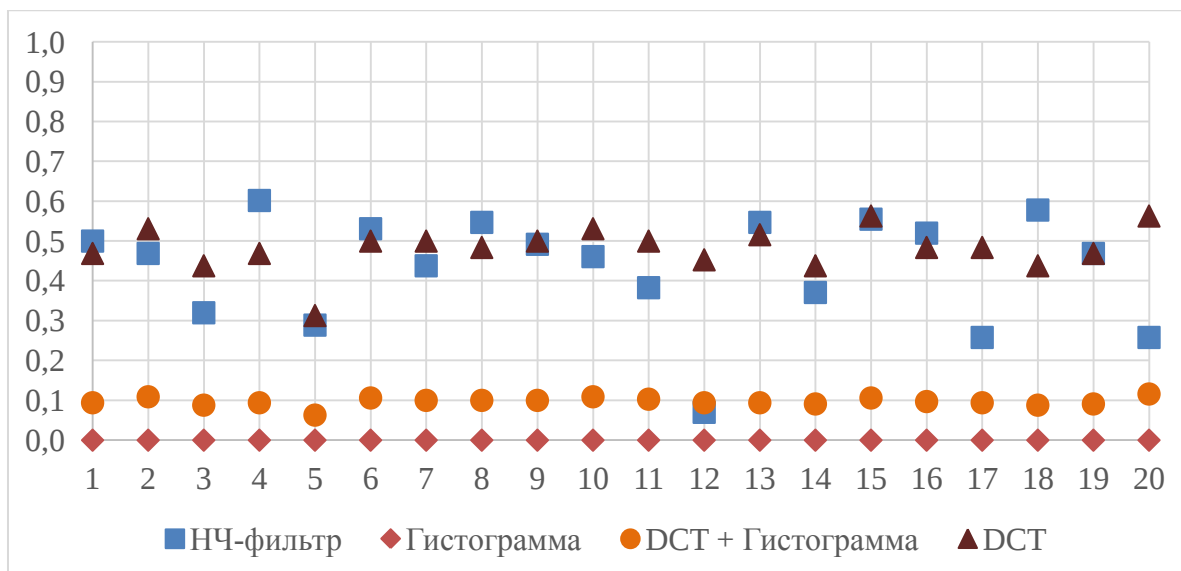


Рисунок 4. Коэффициенты подобию повернутых на 90° изображений

Заключение

Оптимизированный DCT-метод вычисления перцептивных хешей по результатам исследований показал себя как надежное средство поиска изображений по содержанию. Однако, он является неустойчивым к повороту и горизонтальному отражению изображения.

Метод на основе гистограмм цветов обладает худшей среди представленных способностью различать изображения, что приводит к очень большому количеству ложноположительных результатов.

Реализованный комбинированный метод, использующий дискретное косинусное преобразование и цветовые гистограммы с указанными выше коэффициентами, обладает высокой скоростью вычисления хеша (сравнимой с алгоритмами на основе DCT и цветовых гистограмм по отдельности), имеет хорошие способности различать изображения и устойчив к различным представленным видам преобразований изображений.

Список литературы

1. Datta R., Joshi D., Li J., Wang J.Z. Image Retrieval – Ideas, Influences and Trends of the New Age // ACM Computing Surveys. 2008. Vol. 40, no. 2. P. 1-60. DOI: [10.1145/1348246.1348248](https://doi.org/10.1145/1348246.1348248)
2. Zauner C. Implementation and Benchmarking of Perceptual Image Hash Functions: Ph.D. thesis. University of Applied Sciences, Hagenberg, 2010.
3. Krawetz N. Looks Like It. Available at: <http://www.hackerfactor.com/blog/index.php?/archives/432-Looks-Like-It.html> , accessed 01.04.2015.
4. Swain M.J., Ballard D.H. Color Indexing // International Journal of Computer Vision. 1991. Vol. 7, no. 1. P. 11-32. DOI: [10.1007/BF00130487](https://doi.org/10.1007/BF00130487)

Analysis of Perceptual Image Hash Functions

I.V. Rudakov¹, I.M. Vasiutovich^{1,*}

[*MONOmah_V@mail.ru](mailto:MONOmah_V@mail.ru)

¹Bauman Moscow State Technical University, Moscow, Russia

Keywords: content-based image retrieval, perceptual hash function

Perceptual hash functions generate specific values on the basis of visual data of images that are called footprints or perceptual hashes. For similar and different images, such functions compute similar and different hashes, respectively. As a result, using the functions to compute differences or similarities between the hashes, it is possible to draw conclusion whether the relevant images are similar or not. Perceptual hash functions can be used to identify or verify image integrity.

This article analyzes existing methods of content-based retrieval by means of perceptual hash, and possibilities to use them to search for similar images. It considers the following methods to generate perceptual hashes: hash on average using low-pass filter, hash function based on histogram of colors, and hash function based on discrete cosine transformation. The article also investigates properties of hash function, which is a combination of methods based on discrete cosine transform and color histogram.

The studies were conducted in three areas: method speed, method of efficient distinguishing between images, resistance to image transformations (scaling, rotation, horizontal reflection).

All of considered algorithms have low computational complexity and, consequently, almost the same speed to calculate hashes. Hash functions based on discrete cosine transformations have shown the best ability to distinguish between the images. The color histogram-based function is the most resistant to transformation, but at the same time has the worst ability to distinguish between the images.

As a result of research the algorithm has been created, which is a combination of methods based on color histograms and discrete cosine transformation. Using this algorithm in solving the problem of search for similar images provides the most accurate and consistent results.

Further researches are supposed to expand a scope of considered algorithms to calculate the perceptual hashes.

References

1. Datta R., Joshi D., Li J., Wang J.Z. Image Retrieval – Ideas, Influences and Trends of the New Age. *ACM Computing Surveys*, 2008, vol. 40, no. 2, pp. 1-60. DOI: [10.1145/1348246.1348248](https://doi.org/10.1145/1348246.1348248)
2. Zauner C. *Implementation and Benchmarking of Perceptual Image Hash Functions*. Ph.D. thesis. University of Applied Sciences, Hagenberg, 2010.
3. Krawetz N. *Looks Like It*. Available at: <http://www.hackerfactor.com/blog/index.php?/archives/432-Looks-Like-It.html> , accessed 01.04.2015.
4. Swain M.J., Ballard D.H. Color Indexing. *International Journal of Computer Vision*, 1991, vol. 7, no. 1, pp. 11-32. DOI: [10.1007/BF00130487](https://doi.org/10.1007/BF00130487)