

Т.А. Паромова, старший преподаватель
И.Я. Зеленева, канд. техн. наук, доц.
Н.В. Луценко, старший преподаватель
С.А. Кленачов, магистр
Запорожский национальный технический университет, г. Запорожье, Украина
irina.zeleneva@gmail.com

Исследование схожести изображений с помощью перцептивных хэш-алгоритмов и расстояний Хэмминга

Рассмотрены результаты исследования количественной оценки схожести изображений с использованием перцептивных хэшей и расстояния Хэмминга при сравнении некоторого эталона с анализируемым изображением.

Ключевые слова: коллекции изображений, бинаризация, перцептивный хэш-алгоритм, расстояние Хэмминга.

Введение

В оцифровка и хранение большого объема визуальных материалов не является проблемой с технической точки зрения. В то же время особую актуальность приобретает задача обеспечения эффективного содержательного доступа к релевантной информации в электронных коллекциях изображений.

Подобная задача может возникнуть, например, при реализации сервиса поиска изображений по заданному образцу (Google Images, TinEye и т.п.), либо для поиска и фильтрации дубликатов изображений, которые занимают ресурсы сервера на веб-портале, предоставляющем пользователям свободно загружать различные фото и видео. Еще пример: одна из современных форм работы музеев, библиотек и архивов – формирование электронных коллекций изображений, в том числе и видео, с целью сохранения культурного наследия и его пропаганды за счет обеспечения удаленного доступа. Не менее важен эффективный поиск изображений для повышения достоверности диагностики при проведении компьютерных исследований различных заболеваний, где изображение объекта должно максимально совпадать с соответствующими формой и размерами изображенного объекта при сокращении вычислительных затрат [1, 4].

Таким образом, поиск изображений по образцу является одной из самых востребованных поисковых операций в настоящее время. Реализация такого поиска может быть выполнена различными способами. Одним из наиболее эффективных и универсальных методов является использование перцептивного хэш-алгоритма [2, 5].

Целью данной работы является получение количественной оценки схожести изображений с использованием перцептивных хэш-функций.

В связи с этим, в работе также рассмотрены исследования на основе расстояния Хэмминга,

в результате чего определено, при каком значении расстояния Хэмминга можно утверждать, что изображения похожи.

1. Способы поиска и сравнения изображений

В современном информационном пространстве существует немало способов поиска изображений в электронных коллекциях. До последнего времени традиционным способом решения этой задачи являлся поиск визуальной информации, опирающийся на индексирование текстовых описаний, ассоциированных с изображением. При очевидной необходимости организации такого доступа, данный подход представляется недостаточным. Неоднозначность соответствия между визуальным содержанием и текстовым описанием снижает показатели точности и полноты поиска. Некоторые изображения вообще трудно описать словами (очевидный пример – абстрактные картины) [1, 8].

В связи с этим возникает проблема организации доступа к современным электронным коллекциям изображений с использованием целого комплекса средств – как текстовых описаний, так и характеристик визуального содержания (простейших, типа цветовой гаммы, и более сложных, связанных с распознаванием образов). Текстовое описание и визуальная поисковая информация дополняют друг друга, обеспечивая возможность разностороннего поиска. Поиск может выполняться итеративно: сначала на основе ключевых слов, как более быстрый способ, затем среди отобранного множества материалов – более трудоемкий поиск с использованием визуальных характеристик.

Существует множество методов распознавания объектов на изображении. Выбор конкретных методов обусловлен особенностями объекта, который требуется распознать. Часто бывает, что

задача распознавания ставится неформальным образом, т.е. свойства искомого объекта задаются без строгих математических параметров. Для решения такой задачи необходимо сформулировать свойства нужного объекта и создать устойчивый алгоритм для обнаружения множества объектов, соответствующих заданным параметрам. Для решения поставленной задачи необходимо найти, обобщить и сформулировать в математических терминах эмпирические наблюдения, то есть формализовать параметры искомого объекта. Главная трудность состоит в том, что описать все свойства практически невозможно, и эти свойства могут соответствовать не всем объектам искомого множества. Поэтому в процессе математической формализации допускаются упрощения, которые в результате снижают качество алгоритма и понижают степень точности. В итоге можно сказать, что при решении задачи распознавания необходимо найти оптимальное соотношение сложности вычислений и желаемой точности.

Самое простое и очевидное решение данной проблемы — это побайтовое сравнение двух файлов. Если поиск дубликатов осуществляется среди большого количества файлов, то сравнение каждого с каждым по его содержанию становится очень ресурсоемкой задачей, требующей больших вычислительных затрат. Существует немало программных продуктов, ускоряющих решение данной задачи [9], но и это не дает удовлетворительных результатов, т.к. сравнение изображений непосредственно друг с другом занимает слишком много процессорного времени и в случае использования даже на минимальных хранилищах, где количество изображений немногим больше сотни, абсолютно неэффективно. Кроме того, в отличие от других типов данных, таких как архивы, документы, либо бинарные файлы, изображения могут быть каким-либо образом изменены (например, иметь различный размер, водяной знак, либо иметь разное качество в случае использования таких алгоритмов сжатия с потерями, как JPEG), хотя при этом быть «одинаковыми».

Таким образом, более эффективными представляются методы поиска изображений не по самим исходным изображениям, а по данным, полученным в результате предварительной обработки этих изображений.

2. Анализ хеш-функции как способа сравнения изображений

Хэш-функции (или функции свертки) представляют собой функции преобразования, которые позволяют получить «отпечаток» фиксированной длины для исходных данных. Классические криптографические хэш-алгоритмы (например, MD5 или SHA-1) работают таким образом, чтобы для различных исходных данных, как мало бы они не отличались друг от друга, в результате

получались максимально отличные хэш-значения [2].

Перцептивные хэши — это другая концепция по сравнению с криптографическими хэш-функциями (MD5 и SHA1). В криптографии каждый хэш является случайным. Данные, которые используются для генерации хэша, выполняют роль источника случайных чисел, так что одинаковые данные дадут одинаковый результат, а разные данные — разный результат [2, 3]. Из сравнения двух хэшей SHA1 на самом деле можно сделать только два вывода: если хэши отличаются, значит, данные разные; если хэши совпадают, то и данные, скорее всего, одинаковые (поскольку существует вероятность коллизий, то одинаковые хэши еще не гарантируют совпадения данных). В отличие от этого, перцептивные хэши можно сравнивать между собой и делать вывод о степени различия двух наборов данных. Именно это свойство и используется для определения степени схожести изображений. Поскольку перцептивные хэш-алгоритмы описывают класс функций для генерации сравнимых хэшей, то характеристики изображения используются для генерации индивидуального (но не уникального) отпечатка, и эти отпечатки можно сравнивать друг с другом.

Таким образом, задача сравнения изображений сводится к вычислению хэш-значений этих изображений, и вычислению расстояния Хэмминга между ними. Чем меньше расстояние Хэмминга, тем более похожи данные изображения.

Алгоритмы вычисления перцептивного хэша обладают одинаковыми базовыми свойствами, поэтому картинки можно изменять в размере, менять соотношение сторон и даже незначительно менять цветовые характеристики (яркость, контраст и т.д.), но они всё равно совпадают по хэшу. Для получения хэша имеется несколько распространенных алгоритмов. Один из простейших перцептивных хэш-алгоритмов отображает среднее значение низких частот. В изображениях высокие частоты обеспечивают детализацию, а низкие частоты показывают структуру. Большое детализированное изображение содержит много высоких частот. В очень маленьком изображении нет деталей, так что она целиком состоит из низких частот. Самый простой способ избавиться от высоких частот — уменьшить изображение. Для сравнения разных хэшей подсчитывается количество разных битов в них (расстояние Хэмминга). Нулевое расстояние означает, что это, скорее всего, одинаковые изображения (или вариации одного изображения). Дистанция 5 означает, что изображения в чём-то отличаются, но в целом всё равно довольно близки друг к другу. Если дистанция 10 или больше, то это, вероятно, совершенно разные изображения [3, 5].

Итак, существуют различные алгоритмы вычисления перцептивных хэшей, которые различаются между собой чувствительностью к опре-

деленным типам искажений: изменению размера изображения, изменению соотношения сторон, цветовых характеристик (яркость, контраст, цветовая гамма), наложению водяных знаков и т.д.

3. Исследование схожести изображений с помощью перцептивных хэш-алгоритмов и расстояний Хэмминга

В данных исследованиях рассмотрены алгоритмы вычисления перцептивных хешей Average Hash, рHash, dHash [3]. Они сравнивались по простоте реализации, производительности и достоверности результатов сравнения изображений. Лучшие характеристики имеет dHash – алгоритм, который отличается простотой в реализации, имеет высокую скорость работы и более высокую точность по сравнению с первыми двумя вариантами. В то время как aHash основывается на среднем значении и рHash на частотных паттернах, dHash отслеживает градиент изображения.

При реализации этого алгоритма первым шагом является уменьшение размера изображения, но в данном алгоритме используется не квадратная, а прямоугольная матрица изображения размером 9x8 (в общем случае N + 1 x N).

Таким образом, изображение переводится в градации серого, после чего для каждой строки вычисляется разница между следующим и предыдущим пикселем, и таким образом, в результате получается матрица размером 8x8 пикселей, которая представлена на рисунке 1.

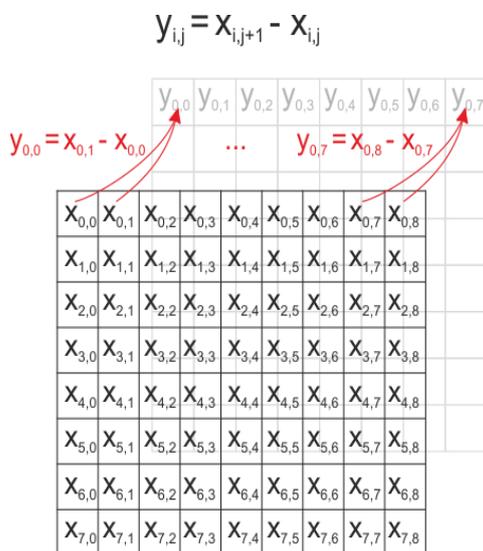


Рисунок 1. – Матрица значений алгоритма dHash

Затем с помощью операции бинаризации выполняется радикальное уменьшение количества информации, содержащейся на изображении. В процессе бинаризации исходное полутоновое изображение, имеющее некое количество уровней яркости, преобразуется в черно-белое изображе-

ние, пиксели которого имеют только два значения – 0 и 1 [6, 7].

Следующим шагом выполняется расчет хэша: если значение текущего пикселя меньше предыдущего ($y[i, j] > 0$), значение хэша принимается 1, в противном случае 0. В исходном коде эти два шага так же объединяются в один, и значение хэша получается из разницы значений соседних пикселей [8].

Далее, на этапе сравнения полученного хэша изображения с хэшем эталона используется расстояние Хэмминга, как мера различия объектов одинаковой размерности.

Расстояние Хэмминга определяется числом позиций, в которых соответствующие символы двух слов одинаковой длины различны. В более общем случае расстояние Хэмминга применяется для строк одинаковой длины любых q-ичных алфавитов и служит метрикой различия (функцией, определяющей расстояние в метрическом пространстве) объектов одинаковой размерности.

Пусть $u=(u_1, u_2, \dots, u_n)$ – двоичная последовательность длиной n.

Число единиц в этой последовательности называется весом Хэмминга для вектора u и обозначается как $w(u)$. Например: $u=(1\ 0\ 0\ 1\ 0\ 1\ 1)$, тогда $w(u)=4$.

Пусть u и v – двоичные слова длиной n. Число разрядов, в которых эти слова различаются, называется расстоянием Хэмминга между u и v и обозначается как $d(u, v)$. Например: $u = (1\ 0\ 0\ 1\ 0\ 1\ 1)$, $v = (0\ 1\ 0\ 0\ 0\ 1\ 1)$, тогда $d(u, v) = 3$. Таким образом, расстояние между двоичными последовательностями u и v равно весу их поразрядной суммы, т. е.

$$d(u, v) = w(u+v). \tag{1}$$

Вероятность того, что слово v будет принято за u, равна

$$P_{ou} = p^{n-d(u, v)} \times q^{d(u, v)}, \tag{2}$$

где p – вероятность правильной передачи бита сообщения; $q=1-p$ – вероятность ошибки.

Задав линейный код, т. е., определив все 2^k кодовые слова, можно определить расстояние между всеми возможными парами кодовых слов, минимальное из них называется минимальным кодовым расстоянием d_{min} .

Нетрудно показать, что расстояние между нулевым кодовым словом и одним из кодовых слов, входящих в порождающую матрицу, равно d_{min} (согласно определению, строки порождающей матрицы линейного блочного кода сами являются кодовыми словами данного кода).

Тогда минимальное кодовое расстояние d_{min} равно минимальному весу Хэмминга для всех строк порождающей матрицы кода [8].

Для получения количественной оценки совпадения изображений методом перцептивных хешей был разработан программный комплекс, работающий в соответствии с алгоритмом, представленным на рисунке 2.



Рисунок 2 – Алгоритм сравнения в системе фильтрации изображений

Данный программный комплекс составляет основу автоматизированной системы сравнения и

фильтрации изображений, предложенной авторами для информационного освещения в социальных сетях различных мероприятий культурно-рекламного характера.

4. Результаты исследований

Исследования проводились на пяти выборках (V1, V2...V5) по девять изображений каждая. Выборки были сформированы из реальных видеоматериалов, освещающих события пяти разных мероприятий рекламного и культурного характера. Каждая выборка содержала изображение-эталон, шесть изображений, похожих на эталон и относящихся к данному мероприятию, а также три изображения, которые не проходили по критерию схожести и могли не относиться к данному мероприятию.

Изображение-эталон обозначим как ethalon. Похожие изображения обозначим как p1, p2...p6. Изображения, которые не похожи, обозначим n1, n2, n3.

Таким образом, каждая выборка изображений представлена тремя подмножествами $E=\{ethalon\}$, $P=\{p1,p2, p3, p4, p5, p6\}$, $N=\{n1, n2, n3\}$. В каждом подмножестве содержатся заранее известные элементы.

В ходе исследований определены перцептивные хэши для каждого изображения из подмножеств и найдено расстояние Хэмминга между элементом подмножества E и элементами подмножеств P и N.

Результаты определения расстояния Хэмминга для изображений из первой выборки представлены в таблице 1.

Таблица 1 – Результат исследований для первой выборки

Изображение	Перцептивный хэш	Расстояние Хэмминга
ethalon	fff0000b6bc00	-
p1	d7d70290bc3400	11
p2	303d7f7df0f4000	15
p3	ffff10302c2400	10
p4	ffff0043ee7c20	9
p5	20dfdf0a0606080c	16
p6	3c047d1166246e	20
n1	c0c4a0433300000	26
n2	143e0e1707100000	26
n3	8027160e104800	26

Гистограмма по результатам для первой

выборки представлена на рисунке 3.

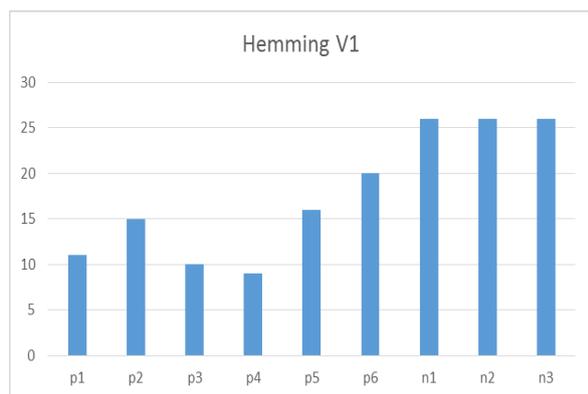


Рисунок 3. – Гистограмма результатов исследований первой выборки

По результатам исследования расстояния Хэмминга для изображений первой выборки видно, что максимальное расстояние Хэмминга для похожих изображений составляет 20, а минимальное расстояние для не соответствующих эталону изображений составляет 26.

Аналогичные исследования были проведены для второй, третьей, четвертой и пятой выборок. Результаты определения расстояния Хэмминга для изображений из второй выборки: максимальное значение для похожих равно 23, а минимальное расстояние для не соответствующих эталону изображений составляет 28. Для третьей выборки, соответственно 23 и 28, для четвертой 23 и 25, для пятой 23 и 28.

Результаты исследований представлены по всем пяти выборкам на рисунке 4.

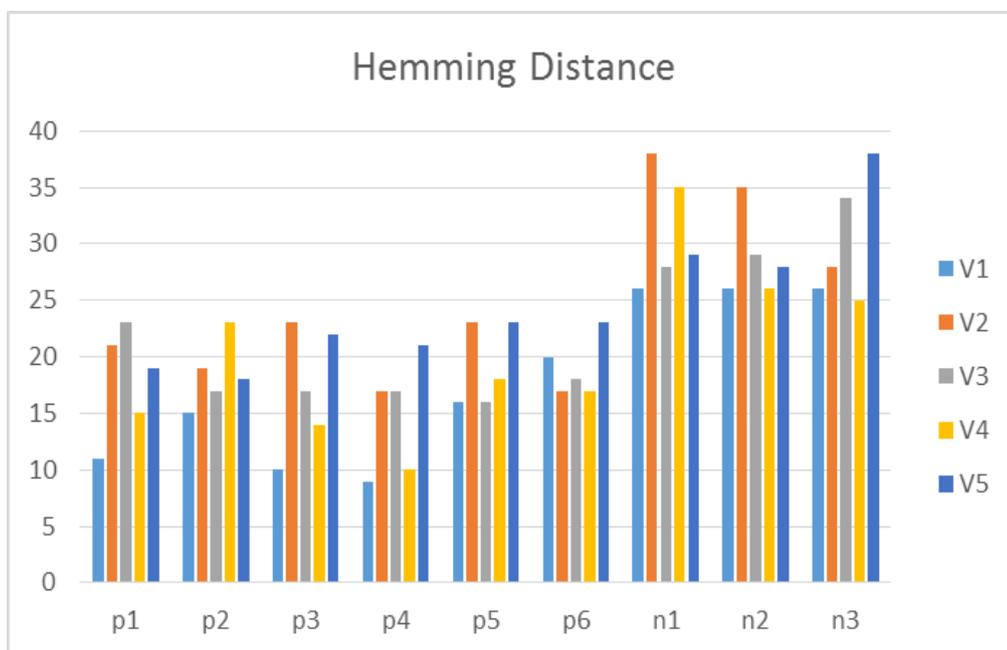


Рисунок 4 – Гистограмма результатов исследований по всем выборкам

На основе полученных данных можно сделать вывод, что максимальное значение расстояния Хэмминга для элементов подмножества P всех выборок составляет 23, а минимальное значение расстояния Хэмминга для элементов подмножества N составляет 25.

Таким образом, изображения можно считать похожими, то есть попадающими в множество изображений P, если расстояние Хэмминга меньше или равно 23.

5. Практическое приложение результатов исследований

Апробация результатов исследований была

проведена с помощью разработанной автоматизированной системы на основе алгоритма, приведенного на рис.2. В данной системе пользователь – владелец мероприятия – выдает разрешение на доступ к своему информационному каналу в системе другим пользователям, которые находятся в ограниченном рамках данного мероприятия пространстве. Главный пользователь ведет фотообзор с места события, к нему подключаются другие, и отправляют свои изображения в главный информационный канал. Для подтверждения истинности, изображения от наблюдателей сравниваются с эталонным изображением владельца канала, снятым с ближайшей точки.

Для начала работы в системе нужно зарегистрироваться и создать аккаунт-профиль. С

помощью этого профиля можно создавать мероприятия и делать фотообзор.

Система обеспечивает контроль изображений из альтернативных источников данных для объективной оценки мероприятий. Для этого с помощью геолокации ограничивается зона, из которой можно передавать изображения в канал организатора мероприятия. Размер зоны задает организатор. Альтернативные источники данных проходят отсеивание по зоне мероприятия. Если источник окажется в разрешенной зоне, он имеет право отправлять свои изображения, сделанные с камеры мобильного устройства, в медиа-канал организатора мероприятия.

Также применяется функция фильтрации от нежелательных изображений, которые не относятся к данному мероприятию. После съемки изображение проходит сравнение с эталоном организатора. Функция фильтрации реализуется с помощью сравнения изображений из альтернативных источников данных с эталоном, сделанным организатором мероприятия. При этом используются результаты данных исследований, позволяющие определить, при каком расстоянии Хэмминга можно утверждать, что изображения похожи и относятся именно к данному мероприятию.

Если изображение похоже на эталон, то есть попадает в коридор процентного соотношения схожести, то оно автоматически отправляется на сервер, и его можно будет увидеть в медиа-канале организатора на странице данного мероприятия.

Заключение

Как показали эксперименты, сравнение перцептивных хэшей достаточно эффективно, но в некоторых случаях недостаточно эффективно, поскольку не учитывает цветовую гамму, а также наличие на изображении «ключевых» объектов, несущих основную информационно-смысловую нагрузку. Важность для человека цветового восприятия изображения обуславливает значение методов поиска визуальной информации на основании схожести с образцом по цветовым характеристикам. Возможно использование таких показателей, как средний или основной цвета, а также множества цветов; эти характеристики имеют смысл применять для локального индексирования областей изображения. Дальнейшие исследования авторов связаны с идеей метода цветовых гистограмм для индексирования и сравнения изображений.

Список литературы

- 1 Огневой Г. Д. Методы и алгоритмы поиска изображений [Электронный ресурс] - Режим доступа: <http://www.bntu.by/news/67-conference-mido/1580-2014-11-23-09-36-26.html>
- 2 Кузин А. Генерация хешей MD2, MD4, MD5 с помощью OpenSSL [Электронный ресурс] - Режим доступа: <http://www.firststeps.ru/linux/r.php?18>.
- 3 Чикало О. В. Способ бинаризации медицинских изображений, © FindPatent.ru - патентный поиск, 2012-2016 [Электронный ресурс] - Режим доступа: <http://softxaker.ru/others/311-universalnyy-generator-heshey-hash-generator.htm>
- 4 <http://www.findpatent.ru/patent/253/2533876.html>
- 5 Ализар А. Как работает перцептивный хеш [Электронный ресурс] - Режим доступа: <https://habrahabr.ru/post/120562/>
- 6 Бинаризация изображений. Прикладные вопросы распознавания образов Распознавание образов для программистов [Электронный ресурс] - Режим доступа: <http://recog.ru/blog/applied/15.html>
- 7 Ализар А. Способы бинаризации [Электронный ресурс] - Режим доступа: <https://habrahabr.ru/post/128768/>
- 8 Федоров А. Бинаризация черно-белых изображений [Электронный ресурс] - Режим доступа: <http://it-claim.ru/Library/Books/ITS/wwwbook/ist4b/its4/fyodorov.htm>
- 9 Цветков А. А., Алгоритмы распознавания объектов [Текст] / Цветков А. А., Шорох Д. К., Зубарева М. Г., Юрсков С. В., Шуклин А. В., Хамуш А. Л., Ануфриев И. Б. // Технические науки: проблемы и перспективы: материалы IV междунар. науч. конф. (г. Санкт-Петербург, июль 2016 г.). — СПб.: Свое издательство, 2016. — С. 20-28.

Надійшла до редакції 15.10.2016

Т.О. ПАРОМОВА, І.Я. ЗЕЛЕНЬОВА, Н.В. ЛУЦЕНКО, С.А. КЛЕНАЧОВ

Запорізький національний технічний університет (Україна)

ДОСЛІДЖЕННЯ СХОЖОСТІ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ПЕРЦЕПТИВНИХ ХЕШ-АЛГОРИТМІВ І ВІДСТАНЕЙ ХЕММІНГА

Розглянуто результати дослідження кількісної оцінки схожості зображень з використанням перцептивних хешей та відстані Хеммінга при порівнянні деякого еталона з аналізованим зображенням.

Ключові слова: колекції зображень, перцептивний хеш-алгоритм, бинаризація зображень, відстань Хеммінга.

T.A.PAROMOVA, I.J. ZELENJOVA, N.V. LUTSENKO, S.A. KLENACHOV

Zaporizhya National Technical University (Ukraine)

RESEARCH OF THE IMAGES SIMILARITY USING PERCEPTUAL HASH ALGORITHMS AND HAMMING DISTANCES

In today's information environment it is becoming particularly urgent to ensure the effective meaningful access to relevant information from digital image collections. A similar problem may arise, for example, in the implementation of image search services on a given pattern (Google Images, TinEye, etc.), or for searching and filtering duplicate images that take up server resources on the web portal that provides users to freely upload a different photo and video. Thus, the image search is one of the most popular search operations at this time. Implementation of such a search may be performed in various ways. One of the most effective and universal method is the use of perceptual hash algorithm.

The aim of this work is to obtain a quantitative assessment of the similarity of images with the use of perceptual hash functions.

In this regard, such problem is discussed in the research on the basis of Hamming distance, whereby it is determined at what value of the Hamming distance can be argued that the images are similar.

There are many methods of recognition of objects in the image. The choice of specific methods depends on the characteristics of the object that you want to recognize. It often happens that the recognition problem is formulated informally, i.e. the desired object properties are set without rigorous mathematical parameters. Thus, more effective methods of image retrieval are represented not by the source images, but the data obtained as a result of pre-processing the images. For example, you can use perceptual hashes, compare them and draw conclusions about the degree of difference between the two data sets. This property is used to determine the degree of similarity between images.

In these studies three algorithms for calculating perceptual hashes are considered - Average Hash, pHash, dHash. They were compared on the simplicity of implementation, performance and reliability of the image comparison results. It has better characteristics dHash - algorithm, which is simple to implement, has high speed and higher accuracy compared with the first two embodiments.

For quantitative evaluation of image matching with the method of perceptual hashes it was developed a software system operating in accordance with the known image comparison algorithm. This software system is the basis of the automated comparison and image filtering system, proposed by the authors for the media coverage of various activities of cultural and advertising in the social networks. The research results allow us to determine at what distance Hamming it can be argued that the images are similar and relate it to this event.

Key words: *collection of images, perceptual hash algorithm, image binarization, Hamming distance.*