

УДК 004.08

ВЫБОР ТЕХНОЛОГИИ ХЕШИРОВАНИЯ ДЛЯ ПРОВЕРКИ УНИКАЛЬНОСТИ ИЗОБРАЖЕНИЙ В СЕРВИСЕ МУЗЫКАЛЬНЫХ ОБЛОЖЕК

Серпинский Роман Эдуардович,

Студент группы ИУК5-21М

Калужский филиал Московского государственного технического университета имени Н.Э. Баумана

serpinskiyrea@student.bmstu.ru

Буракова Мария Сергеевна,

Ассистент кафедры ИУК5 КФ

Калужский филиал Московского государственного технического университета имени Н.Э. Баумана

m.burakova@bmstu.ru

Аскеров Салех Теймур оглы,

Студент группы ИУК5-21М

Калужский филиал Московского государственного технического университета имени Н.Э. Баумана

askerovst1@student.bmstu.ru

Аннотация

В данной работе проводится сравнительный анализ четырёх алгоритмов перцептивного хеширования изображений – aHash, pHash, dHash и wHash. Изучаются особенности работы каждого алгоритма, точность определения визуального сходства и скорость генерации хеша. Экспериментальная часть включает тестирование алгоритмов на наборе изображений с различными искажениями (поворот, изменение яркости, размытие и др.). Особое внимание уделяется выбору наилучшей технологии хеширования для сервиса проверки уникальности музыкальных обложек. На основании полученных данных сформулированы рекомендации по применению алгоритмов в реальных условиях.

Ключевые слова: перцептивное хеширование, проверка уникальности изображений, музыкальные обложки, aHash, pHash, dHash, wHash, сравнение алгоритмов, хеш-функции, визуальное сходство.

CHOOSING A HASHING TECHNOLOGY TO VERIFY THE UNIQUENESS OF IMAGES IN THE MUSIC COVER SERVICE

Roman E. Serpinski,

Student of the IUK5-21M group

Kaluga Branch of the Bauman Moscow State Technical University
serpinskiyrea@student.bmstu.ru

Maria S. Burakova,

Assistant of the Department of IUK5 CF
Kaluga Branch of the Bauman Moscow State Technical University
m.burakova@bmstu.ru

Saleh T. Askerov,

Student of group IUK5-21M
Bauman Moscow State Technical University (Kaluga Branch)
askerovst1@student.bmstu.ru

ABSTRACT

This paper provides a comparative analysis of four perceptual image hashing algorithms – aHash, pHash, dHash and wHash. The features of each algorithm, the accuracy of determining visual similarity, and the hash generation rate are studied. The experimental part includes testing algorithms on a set of images with various distortions (rotation, brightness variation, blurring, etc.). Special attention is paid to choosing the best hashing technology for the service to verify the uniqueness of music covers. Based on the data obtained, recommendations on the use of algorithms in real conditions are formulated.

Keywords: perceptual hashing, image uniqueness verification, music covers, aHash, pHash, dHash, wHash, algorithm comparison, hash functions, visual similarity.

Введение

Современные цифровые платформы, работающие с визуальным контентом, нуждаются в быстрых и надёжных способах определения уникальности изображений. Особенно это актуально для специализированных сервисов, таких как маркетплейсы визуального искусства, где визуальное оформление играет важную роль. Для защиты от дублирующего контента и нарушения авторских прав необходимо внедрять алгоритмы, способные эффективно находить визуально схожие изображения, даже если они были частично изменены.

Перцептивное хеширование – это подход, при котором изображение преобразуется в компактный хеш-код, сохраняющий его визуальные особенности [2]. Это позволяет быстро сравнивать изображения между собой, опираясь не на точное совпадение пикселей, а на их общее визуальное содержание. Существует несколько алгоритмов перцептивного хеширования, каждый из которых реализует свой способ анализа изображения.

В данной работе рассматриваются и сравниваются четыре популярных алгоритма:

aHash (Average Hash) – использует усреднение яркости пикселей,

pHash (Perceptual Hash) – анализирует частотные компоненты изображения,

dHash (Difference Hash) – оценивает относительные различия между соседними пикселями,

wHash (Wavelet Hash) – применяет вейвлет-преобразование для извлечения текстурных признаков [1].

Сравнение моделей

Для оценки эффективности алгоритмов перцептивного хеширования изображений (aHash, pHash, dHash, wHash) были выделены два ключевых критерия:

Точность определения визуального сходства изображений

Скорость генерации хешей [5]

1. Точность. Под точностью понималась способность алгоритма определять, что модифицированное изображение (например, с поворотом, шумом или изменением яркости) является визуально схожим с оригиналом. Для оценки качества использовались следующие метрики:

Precision — доля найденных совпадений, которые действительно являются модифицированными копиями;

Recall — доля всех модифицированных изображений, которые были корректно распознаны как дубликаты;

F1-score — гармоническое среднее между Precision и Recall, отражающее общий баланс между точностью и полнотой [3].

2. Скорость вычислений. Для каждого алгоритма измерялось:

Среднее время генерации хеша одного изображения (в миллисекундах);

Все измерения проводились в идентичных условиях на локальной машине с использованием языка программирования Python и библиотеки ImageHash [6].

Экспериментальная база

Для тестирования был использован набор изображений, DataSet.

Для каждого изображения автоматически создавалась модифицированная версия с помощью искажений: поворот, изменение яркости, размытие, контраст, шум.

Сравнение производилось попарно между оригиналом и его искажённой версией.

Все алгоритмы тестировались на одних и тех же данных, с единым порогом расстояния Хэмминга (≤ 5), при котором изображения считались схожими [4].

Решение

В рамках проведённого исследования была рассмотрена задача выбора подходящего алгоритма перцептивного хеширования для реализации функции проверки уникальности изображений в специализированном сервисе музыкальных обложек. Были протестированы и сравнены четыре популярных алгоритма: aHash, pHash, dHash и wHash.

В качестве тестового датасета использовались две выборки изображений: первая папка содержала оригинальные музыкальные обложки, вторая — модифицированные версии тех же изображений (например, с изменением яркости, контраста, масштабированием, небольшими поворотами и другими преобразованиями).

Для каждого изображения из оригинальной папки был рассчитан хэш с использованием каждого из четырёх алгоритмов. Затем выполнялось сравнение с хэшами изображений из второй папки с целью определения степени схожести. На основании порогового значения рассчитывалось, удалось ли алгоритму корректно выявить соответствие (т.е. определить, что модифицированное изображение является вариантом оригинала) или отличить нерелевантные изображения.

Precision (Точность):

Доля корректно найденных дубликатов среди всех предсказанных совпадений.

$Precision = (TP / (TP + FP)) \times 100\%$

(TP (True Positives) — количество верно обнаруженных дубликатов.)

(FP (False Positives) — количество ложных срабатываний.)

Recall (Полнота):

Доля реальных дубликатов, которые алгоритм смог обнаружить.

$Recall = (TP / (TP + FN)) \times 100\%$

(FN (False Negatives) — количество пропущенных дубликатов.)

F1-score:

Гармоническое среднее precision и recall. Оптимизирует баланс между ними.

$$F1=2 \times (\text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})) \times 100\%$$

Таблица 1 – Результаты точности и скорости алгоритмов

Алгоритм	Precision (%)	Recall (%)	F1-score (%)	Сред. время (мс)
aHash	94.00	64.00	75.91	0.83
pHash	98.00	54.00	69.52	4.09
dHash	91.00	54.00	67.67	0.78
wHash	95.00	62.00	74.83	6.24

Заключение

В рамках проведённого исследования была рассмотрена задача выбора подходящего алгоритма перцептивного хеширования для реализации функции проверки уникальности изображений в специализированном сервисе музыкальных обложек. Были протестированы и сравнены четыре популярных алгоритма: aHash, pHash, dHash и wHash. Сравнение проводилось на основе ключевых метрик качества – precision, recall и F1-score, при этом особое внимание уделялось устойчивости алгоритмов к визуальным искажениям, таким как поворот, изменение яркости, контраста и размытие.

Алгоритм aHash продемонстрировал сбалансированную работу – при высокой точности (94%) он обеспечил неплохой показатель полноты (64%). Это делает его пригодным для использования в системах, где важно быстрое сравнение изображений и допустим некоторый уровень пропуска схожих вариантов. Алгоритм особенно хорошо справляется с простыми случаями, однако становится менее надёжным при сложных искажениях.

Алгоритм pHash показал наилучшую точность (98%), что говорит о его способности избегать ложных совпадений, однако при этом его recall составил всего 54%, что указывает на склонность пропускать изменённые копии изображений. Это делает его подходящим для систем, где критически важно избегать ошибок первого рода, но можно допустить частичный недообнаружение дубликатов.

Алгоритм dHash оказался самым лёгким и быстрым, однако его показатели точности и полноты были ниже остальных (precision – 91%, recall – 54%). Он хуже справляется с изображениями, содержащими сложную структуру и изменения. Тем не менее, его высокая скорость может быть востребована в условиях, где приоритетом является производительность, а не максимальная точность.

Алгоритм wHash показал наилучший общий баланс между точностью (95%) и полнотой (62%). Он оказался наиболее устойчив к искажениям изображений и может считаться наиболее надёжным вариантом среди протестированных. Несмотря на более высокую вычислительную нагрузку, его использование может быть оправдано в системах, где важно качество поиска и минимизация ошибок.

Таким образом, выбор алгоритма должен основываться на специфике проекта и его приоритетах – таких как скорость обработки, устойчивость к искажениям и допустимый уровень ошибок. В рамках моего проекта, ключевым критерием является высокая точность обнаружения схожих изображений, поскольку система должна максимально надёжно выявлять повторяющийся визуальный контент и предотвращать загрузку дубликатов. С учётом этого, наилучшим выбором для реализации задачи является алгоритм wHash, который показал стабильную работу при различных визуальных искажениях и

продемонстрировал наилучший баланс между точностью и полнотой. Этот алгоритм обеспечивает высокую надёжность при проверке уникальности и наиболее соответствует требованиям проекта.

Список литературы:

1. Charikar M. S. Similarity Estimation Techniques from Rounding Algorithms // Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC). – 2002. – Pp. 380–388. – DOI 10.1145/509907.509965.
2. Jégou H., Douze M., Schmid C. Product Quantization for Nearest Neighbor Search // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, No. 1. – Pp. 117–128. – DOI 10.1109/TPAMI.2010.57.
3. Gong Y., Lazebnik S. Iterative Quantization: A Procrustean Approach to Learning Binary Codes for Large-scale Image Retrieval // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2013. – Vol. 35, No. 12. – Pp. 2916–2929. – DOI 10.1109/TPAMI.2012.193.
4. Datta R., Joshi D., Li J., Wang J. 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.
5. Mikolajczyk K., Schmid C. A Performance Evaluation of Local Descriptors // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27, No. 10. – Pp. 1615–1630. – DOI 10.1109/TPAMI.2005.188.
6. Torralba A., Fergus R., Freeman W. T. 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2008. – Vol. 30, No. 11. – Pp. 1958–1970. – DOI 10.1109/TPAMI.2008.128.

References:

1. Charikar M. S. Similarity Estimation Techniques from Rounding Algorithms // Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC). – 2002. – Pp. 380–388. – DOI 10.1145/509907.509965.
2. Jégou H., Douze M., Schmid C. Product Quantization for Nearest Neighbor Search // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, No. 1. – Pp. 117–128. – DOI 10.1109/TPAMI.2010.57.
3. Gong Y., Lazebnik S. Iterative Quantization: A Procrustean Approach to Learning Binary Codes for Large-scale Image Retrieval // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2013. – Vol. 35, No. 12. – Pp. 2916–2929. – DOI 10.1109/TPAMI.2012.193.
4. Datta R., Joshi D., Li J., Wang J. 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.
5. Mikolajczyk K., Schmid C. A Performance Evaluation of Local Descriptors // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27, No. 10. – Pp. 1615–1630. – DOI 10.1109/TPAMI.2005.188.

6. Torralba A., Fergus R., Freeman W. T. 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. - 2008. - Vol. 30, No. 11. - Pp. 1958-1970. - DOI 10.1109/TPAMI.2008.128.