УДК 004.932.2

МЕТОДЫ ЭКСТРАКЦИИ ПРИЗНАКОВ ИЗОБРАЖЕНИЯ С ЦЕЛЬЮ ПОИСКА И РАСПОЗНАВАНИЯ ЛИЦ

METHODS FOR EXTRACTING FEATURES FROM AN IMAGE FOR SEARCH AND FACE RECOGNITION TASKS

Бахтияр Жумалыевич Жарлыкасов
аспирант
bakhtiyarzbj@gmail.com
ФГАОУ ВО «Российский государственный профессионально-педагогический университет», Екатеринбург, Россия

Калыбек Сапарович Мауленов
магистрант
k_maulenov@inbox.ru
РГП на ПХВ «Костанайский государственный университет имени А. Байтурсынова», Костанай, Казахстан

Bakhtiyar Zhumalievich Zharlykasov
Russian State Vocation Pedagogical University, Yekaterinburg, Russia

Kalybek Saparovich Maulenov
Kostanay State University named after A. Baytursynov, Kostanay, Kazakhstan

Аннотация. *Приведен анализ методов разреженного представления и преобразования цифровых изображений для экстракции и уменьшения исходного пространства признаков с целью поиска и распознавания лиц. Представлены сравнительные вычислительные сложности ряда методов.*

Ключевые слова: *разреженное представление, система распознавания, вычислительная сложность, преобразование, поиск лиц, распознавание лиц.*

Abstract. *The article presents the analysis of methods for sparse representation and converting the digital image, extraction and reduce the original feature space search and recognition. Considered comparative computational complexities of these methods: DCT, FFT, DWT, PCA, NNDA.*

Keywords: *sparse representation, recognition system, computational complexity, transformation, facial search, face recognition.*

В настоящее время такое направление в области распознавания изображений, как поиск и идентификация лиц, является популярной и хорошо изученной темой.

Системы распознавания, основанные на физических признаках цифрового изображения, подразумевают работу с очень большим пространством признаков. По этой причине возникает задача уменьшения или представления исходного пространства признаков разреженным.

В системах распознавания данную работу выполняет блок экстракции признаков, поэтому проблема представления изображения существенно меньшего размера, чем размер исходного изображения, для уменьшения количества вычислительных операций, является достаточно актуальной для поиска и распознавания лиц [1].

Основной задачей разреженного представления данных является минимизация входного сигнала для последующей обработки с возможностью реконструирования. Разреженные представления используются для хранения сравнительно небольшого объема данных, которые располагаются в большой области данных.

Методы разреженного представления имеют широкую популярность в обработке не только изображений, но и сигналов в целом.

Методам разреженного представления изображений с целью распознавания лиц посвящено немалое количество работ, большинство из которых основано на следующих методах и их объединениях [2]:

- метод главных компонент (Principal Component Analysis, PCA);
- дискретное косинус-преобразование (Discrete Cosine Transform);
- анализ ближайших соседних точек (Nearest Neighbor Discriminant Analysis, NNDA);
- дискретное вейвлет-преобразование (Discrete Wavelet Transform);
- быстрое преобразование Фурье (FT).

Рассмотрим подробнее эти методы и алгоритмы, в которых они применяются.

Дискретное косинус-преобразование (ДКП) — одно из ортогональных преобразований, имеющее очень важное значение в области обработки сигналов, начиная от сжатия аудиосигналов и заканчивая сжатием изображений, а также для спектрального представления информации.

Основная идея подхода состоит в представлении данных изображения коэффициентами их дискретного преобразования (трансформантами). Дискретное косинусное преобразование очень тесно взаимосвязано с дискретным преобразованием Фурье, но в отличие от преобразования Фурье использует только вещественные числа.

Пикселы изображения имеют корреляцию по двум направлениям, а не только по одному. Поэтому методы сжатия изображений используют двумерное ДКП, которое задается формулой

$$G_{ij} = \frac{1}{\sqrt{2n}} C_i C_j \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} P_{xy} \cos\left(\frac{(2y+1)j\pi}{2n}\right) \cos\left(\frac{(2x+1)j\pi}{2n}\right).$$

При $0 \le i, j \le n-1$ изображение разбивается на блоки пикселов p_{xy} размера $n \times n$ и уравнения используются для нахождения коэффициентов G_{ij} для каждого блока пикселов. Если допускается частичная потеря информации, то коэффициенты квантуются [2].

Прямое применение этой формулы потребует операций $O(N^2)$, но можно вычислить с помощью сложности $O(N \cdot \log(N))$, факторизуя вычисление аналогично быстрому преобразованию Фурье.

Дискретное косинусное преобразование тесно связано с дискретным преобразованием Фурье и является гомоморфизмом его векторного пространства [5].

Быстрое преобразование Фурье (БПФ) — алгоритм быстрого вычисления дискретного преобразования Фурье (ДПФ), т. е. алгоритм вычисления за количество действий, меньшее чем $O(N^2)$, требуемых для прямого вычисления ДПФ. Наиболее распространенным алгоритмом БПФ является алгоритм Кули-Тьюки, при котором ДПФ от $N = N_1 N_2$ выражается как сумма ДПФ более малых размерностей N_1 и N_2 рекурсивно для того, чтобы достичь сложность $O(N \cdot log(N))$ [3].

В общем виде формула дискрентного преобразования Фурье выглядит следующим образом:

$$X_k = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{-\frac{2xi}{N}kn}.$$

При использовании алгоритма преобразования Кули-Тьюки, взяв за основание 2 и выразив ДПФ как сумму двух частей: сумму четных индексов $m = 2n$ и сумму нечетных индексов $m = 2n + 1$, в результате упрощений, обозначив ДПФ четных индексов x_{2m} через E_k (англ., even — четный) и ДПФ нечетных индексов x_{2m+1} через O_k (англ., *odd* — нечетный), для $0 \le n \le N/2$

$$X_m = E_m + e^{-\frac{2mi}{N}m} O_m,$$

$$X_{m+\frac{N}{2}} = E_m - e^{-\frac{mi}{N}m} O_m.$$

Данная запись является базой алгоритма Кули-Тьюки с основанием 2 для вычисления БПФ. При рекурсивном делении ДПФ от N входных значений на сумму 2 ДПФ по $N/2$ входных значений сложность алгоритма становится равной $O(N \log(N))$ [3].

Методы ДКТ и БПФ могут иметь серьезную информационную избыточность в связи с тем, что изображение в данных методах разбивается на блоки, между которыми возникает корреляция. Подходом, позволяющим снизить межблочную избыточность и фрагментарность преобразования, является рекурсивное блочное кодирование. Приобретают популярность также различные виды *дискретного вейвлет-преобразования (ДВП)*.

Алгоритм ДВП является близким видом преобразования к ДКП, но его преимущество в том, что нестационарные сигналы локализуются в малом числе вейвлет-коэффициентов. Это приводит к возможности лучшего восстановления нестационарного сигнала по неполным данным и решает проблемы, встречающиеся в методах БПФ и ДКП.

Простое ДВП для входного сигнала, представленного массивом $2n$ чисел, группирует элементы по 2 и образует из них суммы и разности. Группировка сумм проводится рекурсивно для образования следующего уровня разложения. В итоге получается разность и одна общая сумма.

Вейвлеты обладают общими полезными свойствами: во-первых, это преобразование можно выполнить за $O(d \cdot \log(d))$ операций. Во-вторых, оно не только раскладывает сигнал на некоторое подобие частотных полос, но и представляет временную область, т. е. моменты возникновения тех или иных частот в сигнале [8].

Преобразование Карунена — Лоэва (KLT) имеет также большую сравнительную близость к косинус-преобразованию. В преобразовании KLT вычисляется самый оптимальный базис для нескольких векторов в отличие от всех остальных видов преобразований, которые являются преобразованиями с постоянным базисом.

Базисные векторы для KLT вычисляются с помощью пикселов исходного изображения, т. е. они зависят от исходных данных. В конкретном методе сжатия эти векторы следует записывать в сжатый файл для использования декодером. Кроме того, не известен быстрый метод вычисления этих векторов. Все эти факты делают метод KLT сугубо теоретическим без реальных приложений [7].

Вычислительная сложность в алгоритме KLT так же, как и в ДКП, равна $O(d \cdot \log(d))$.

Анализ ближайших соседних точек (Nearest Neighbor Discriminant Analysis, NNDA) заключается в отыскании среди множества элементов, расположенных в метрическом пространстве, элементов, близких к заданному, согласно некоторой заданной функции близости, определяющей это метрическое пространство. В зависимости от используемого алгоритма в прикладной задаче сложности алгоритма могут быть разными. Наиболее распространенными алгоритмами по разбиению пространства являются диаграммы Вороного и различные алгоритмы, основанные на древовидной структуре.

Диаграмма Вороного конечного множества точек S на плоскости представляет такое разбиение плоскости, при котором каждая область этого разбиения образует множество точек, более близких к одному из элементов множества S, чем к любому другому элементу множества [6]. Алгоритм может быть реализован с вычислительной сложностью $O(n^2\log(n))$.

Сравним вычислительную сложность рассмотренных алгоритмов. В таблице 1 показана сравнительная вычислительная сложность методов ДКП, БПФ, ДВП, KLT, NNDA. N — число изображений, d — количество пикселей (MN-размер), d' — разрешение низкочастотного поддиапазона ($d' < d$).

Таким образом, нетрудно заметить, что в описанных выше методах сложности алгоритмов практически одинаковы, во многих случаях идентичны, это связано с тем, что большинство методов, а именно: ДКП, БПФ и ДВП — берут за основу один и тот же подход — ортогональное преобразование данных. К примеру, ДКП по сути является гомоморфиз-

Таблица 1

Сравнение сложности методов

Название метода	Сложность
Дискретное косинус-преобразование (ДКП)	$O\ (d \cdot \log(d))$
Быстрое преобразование Фурье (БПФ)	$O\ (d \cdot \log\ (d'))$
Дискретное вейвлет-преобразование (ДВП)	$O\ (d \cdot \log\ (d))$
Преобразование Карунена – Лоэва (KLT)	$O\ (d \cdot \log\ (d))$
Анализ ближайших соседних точек (NNDA)	$O\ ((d^2 \cdot \log(d))$

мом векторного пространства DFT, и их отличие в том, что преобразование Фурье работает с более общими комплексными числами, а косинусное преобразование — с действительными числами. Алгортм ДВП, как уже отмечалось выше, является очень близким видом преобразования к ДКП. Конечно, здесь необходимо отметить, что разбиение на блоки с последующим ДКП-сжатием приводит к снижению качества восстановленного изображения при больших коэффициентах сжатия. Следовательно, сжатие изображений посредством дискретного вейвлет-преобразования заслуживает определенного внимания, поскольку оно может быть эффективно применено к целому изображению и при нем не будут присутствовать артефакты блочного разбиения. Также следует отметить большую сравнительную близость косинус-преобразования преобразованию Карунена – Лоэва.

Исключением является метод анализа ближайших соседних точек (NNDA), который отличается от всех остальных, хотя и у него вычислительная сложность оказалось идентичной, но стоит заметить, что это только для алгоритма с модификацией. При больших объемах данных решающий алгоритм должен удовлетворять заданным требованиям к допустимой погрешности (точности) поиска и вычислительной сложности (быстродействию). Как правило, эти характеристики находятся в обратной зависимости: с увеличением быстродействия уменьшается точность (растет погрешность), и наоборот. Поэтому использование NNDA в этом плане имеет дополнительные сложности.

Исходя из всего вышесказанного, можно сделать вывод, что высокая схожесть алгоритмов сложности методов является логичной, каждый из них имеет одинаковую природу решения задачи, каждый из подходов является оптимальным для конкретных задач, которые необходимо решить, к примеру, если задача восстановления исходного изображения не является приоритетной, то нет существенной разницы, какой метод использовать: ДКП или ДВП. Если в рамках задачи необходимо восстановление изображения, то в таком случае ДВП будет более оптимальным методом.

При создании систем распознавания в большинстве случаев применяются комбинации данных методов для получения более качественных результатов.

Список литературы

1. *Аимбетова Д. Т. Распознавание изображений лиц для идентификации личности / Д. Т. Аимбетова, Б. Ж. Жарлыкасов, А. З. Муслимова // Актуальные научные исследования в современном мире: 32-я Международная научная конференция, Переслав-Хмельницкий, 26–27 дек. 2017 г. Вып. 12(32). С. 164.*
2. *Кухарев Г. А. Поиск изображений лиц в больших базах данных / Г. А. Кухарев // Мир измерений. 2009. № 4.*
3. *Методы обработки и распознавания изображений лиц в задачах биометрии / Г. А. Кухарев [и др.]. Санкт-Петербург: Политехника, 2013. 388 с.: ил.*
4. *Ремнева Е. А. Исторические аспекты биометрии / Е. А. Ремнева // Мир измерений. 2009. № 3.*
5. *Щеголева Н. Л. Простой алгоритм классификации линейно неразделимых данных / Н. Л. Щеголева, Г. А. Кухарев // Естественные и технические науки. 2012. № 1. С. 358–364.*
6. *Chao-Hsing Hsu. Comparison of Image Approximation Methods: Fourier Transform, Cosine Transform, Wavelets Packet and Karhunen-Loeve Transform / Chao-Hsing Hsu, Zhen Guo, Kang Yen. Department of Electrical Engineering Florida International University 10555 W. Flagler St.Miami Fl 33174.*
7. *Comparative Analysis of Face Recognition using DCT, DWT and PCA for Rotated faces / D. Rana [et al.] // International Journal of Scientific Research Engineering & Technology (IJSRET), ISSN 2278–0882. 2014. Vol. 3, Issue 5, August.*
8. *Tyagi S. K. Face Recognition Using Discrete Cosine Transform and Nearest Neighbor Discriminant Analysis / S. K. Tyagi, P. Khanna // International Journal of Engineering and Technology. 2012. Vol. 4, № 3. P. 311–314.*