

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное
учреждение высшего профессионального образования
«Московский физико-технический институт
(Государственный университет)»

Факультет управления и прикладной математики
Кафедра информатики МФТИ

**«УДАЛЕНИЕ ШУМА ИЗ ИЗОБРАЖЕНИЯ
С ИСПОЛЬЗОВАНИЕМ РАЗРЕЖЕННОГО
ПРЕДСТАВЛЕНИЯ»**

Выпускная квалификационная работа
(магистрская работа)

Направление подготовки: 03.04.01 Прикладные математика и физика

Выполнила:
студентка 6 курса 973 группы
Колдаева Мария Владимировна

Научный руководитель:
к.ф-м.н.
Кухаренко Борис Георгиевич

Москва, 2015

Содержание

1 Введение	2
2 Постановка задачи	3
3 Литературный обзор	4
3.1 Глобальный алгоритм с использованием порога	4
3.2 Локальный алгоритм с использованием жесткого порога	5
3.3 Локальный алгоритм с использованием нежесткого порога	6
4 Методы, основанные на обучении словаря	7
4.1 Алгоритм поиска разреженного представления (Matching pursuit-алгоритм)	8
4.2 Метод K-SVD для обучения словаря	9
4.3 Метод CSR для удаления шума из изображения	10
4.4 Результаты сравнения описываемых методов удаления шума	11
5 Статистическая интерпретация задачи разреженного представления	14
5.1 Пример: метод наименьших квадратов в статистической интерпретации	15
5.2 Предлагаемая модификация метода CSR	16
6 Выводы	18

1 Введение

Настоящая дипломная работа относится к практической области обработки изображений. В обработке изображений часто возникает задача удаления шума. Шум на изображениях можно определить как случайные колебания яркости, распространенные по всему изображению. Шум может возникать из-за несовершенства устройств, формирующих изображение, плохой освещенности, ошибок при передаче изображения. В данной дипломной работе рассматривается шум, образовавшийся в результате несовершенства камеры, а именно, камеры наномикроскопа.

Вследствие явления шума на изображениях, становится сложным различать детали изображения, проводить автоматическую обработку и распознавание изображений, а также портится общий вид изображений (что существенно, например, на фотографиях). Задача удаления шума может возникать не только для изображений, но также и для произвольных сигналов (радиосигналы, звуковые сигналы, и т.д.), и проблемы из-за наличия шума аналогичны в этих областях.

Существует множество методов удаления шума, от простых до самых сложных. Для улучшения общего вида изображения, например на пользовательских фотографиях, достаточно стандартных методов, основанных на фильтрации изображения. Эти методы широко распространены на практике: такие алгоритмы встраиваются в цифровые камеры, камеры смартфонов или встречаются в отдельных приложениях. В данной работе рассматриваются наноизображения, где, зачастую, из-за шума невозможно различить важные детали изображения, поэтому простейших методов для обработки таких изображений недостаточно. В связи с этим, в области обработки изображений, стоит актуальный вопрос улучшения качества методов удаления шума, а также адаптации данных методов под конкретные особенности изображений.

В работе рассматриваются наиболее сложные и современные методы удаления шума из изображения, основанные на разреженном представлении изображения. Данные методы берут свое начало с идеи Фурье-преобразования – разложения сигнала по элементарным функциям со всеми убывающими амплитудными коэффициентами. Однако методы Фурье-преобразования не являются наиболее подходящими для работы с изображениями, а именно, не отражают его особенностей, конечных размеров, в связи с чем, дальнейшим этапом исследований становятся специальные элементарные функции – вейвлеты, курвлеты, и т.д. Но и эти методы не представляют достаточного качества удаления шума. В связи с этим, последние разработки в этой области состоят в развитии методов удаления шума, основанных на обучении словаря элементарных функций для разложения изображения (обучение производится для адаптации словаря к особенностям конкретного изображения); а также методов, основанных на кластеризации похожих частей изображения.

Содержание диплома отражено в публикациях автора [1], [2]. На секции “информатика” конференции МФТИ №57 (2014 г.) выступление отмечено грамотой.

2 Постановка задачи

Целью работы является проверка возможности применения аппарата статистического машинного обучения к одной из областей анализа изображения, а именно к задаче разреженного представления. В связи с этим, необходимо разработать статистическую(байесовскую) интерпретацию задачи разреженного представления, что, по видимому, является еще недостаточно распространенным подходом в данной области.

Задача дипломной работы заключается в анализе, тестировании и сравнении существующих методов удаления шума, основанных на разреженном представлении и их адаптация к работе с фотографиями, полученными с трансмиссионного электронного микроскопа. Такое исследование необходимо для того, чтобы выяснить, какие из методов наиболее перспективны, и понять, какой подход принято применять для формулировки задач разреженного представления.

Оценка качества алгоритма производится визуально на целевых изображениях, или с использованием известного критерия «пиковое отношение сигнал-шум (PSNR)» на тестовых изображениях с добавленным синтетическим шумом. Предлагается исследовать именно алгоритмы, основанные на разреженном представлении, так как это является новой, перспективной и еще недостаточно изученной тематикой в области анализа изображений, и такие методы показывают лучшее качество удаления шума по сравнению с традиционными методами, основанными на фильтрации изображения.

Для удаления шума используется обработка изображений в формате YCbCr, а алгоритм действует в серошкольном канале изображения Y. Изображение представляется в виде матрицы значений его Y-канала. Разреженным представлением изображения называются векторы разложения изображения по некоторому базису, такие, что эти векторы минимальны по некоторой установленной норме (норма устанавливается исследователем в зависимости от его критериев).

Математической моделью шума является гауссова добавка к изображению:

$$y = y_0 + v, \quad (1)$$

где y - зашумленное изображение, y_0 - требуемое незашумленное изображение, а шум v представляется его плотностью вероятности:

$$p(v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(v-\mu)^2}{2\sigma^2}} \quad (2)$$

Задача удаления шума с использованием разреженного представления является задачей оптимизации:

$$(P^0) \min_{\mathbf{x}} \|\mathbf{x}\|_0 : \|\mathbf{Dx} - \mathbf{y}\|_2^2 = 0, \quad (3)$$

где \mathbf{D} - некоторый базис, \mathbf{x} - разреженное представление изображения в этом базисе, а \mathbf{y} - исходное изображение.

Нахождение разреженного представления тесно связано с задачей выпуклой оптимизации с ограничениями, и от изменений алгоритма (подбора словаря, нормы разреженного представления), меняется и итоговая задача оптимизации. В связи с этим, в работе необходимо формулировать оптимационные задачи, и использовать наиболее эффективные методы их решения.

Также для решения задачи эффективного удаления шума используются методы обучения словаря – набора базисных векторов для разложения по ним изображения. Такое обучение нацелено на адаптацию этих векторов под конкретное изображение, в связи с чем удается получить наиболее разреженное его представление, а следствием, наилучшее качество удаленного шума. В работе исследуются наиболее современные методы, комбинирующие подход обучения с подходом кластеризации изображения: анализируются идентичные участки изображения, и вследствие этого накладываются дополнительные ограничения на задачу оптимизации.

Научная новизна дипломной работы состоит в следующем:

- Разрабатываемый метод дает лучшее качество удаления шума по сравнению с традиционными
- Сохраняется максимальная резкость изображения, детали различимы
- Применимость изучаемых методов к другим задачам улучшения изображений

3 Литературный обзор

В связи с задачей удаления шума существует множество исследований, изучающих данную тематику. Среди них удаление шума посредством фильтраций, интерполяций, порядковых статистик и прочих методов аппроксимации. В дипломе не рассматриваются описанные подходы; изучаются только методы, основанные на разреженном представлении.

Методами удаления шума из изображения с использованием разреженного представления (в дальнейшем просто методы) называют различные вариации задачи (3). Данная задача является невыпуклой и дискретной, что усложняет ее решение. Однако, заменив нормы, можно свести данную задачу к выпуклой задаче оптимизации [3], что значительно упрощает ее решение, что и будет сделано в дальнейшем. Побольше про задачу оптимизации.

Методы удаления шума подразделяются на две основных категории: глобальные – использующие изображение целиком, и локальные – основанные на анализе изображения по частям для удаления шума с изображения целиком.

3.1 Глобальный алгоритм с использованием порога

В качестве примера глобального метода приводится алгоритм с использованием порога (Thresholding algorithm) [3]. Исходная задача представляется выпуклой задачей

оптимизации:

$$(P_\epsilon^0) \min_{\mathbf{x}} \|\mathbf{x}\|_0 : \|\mathbf{Dx} - \mathbf{y}\|_2^2 \leq \epsilon. \quad (4)$$

Обозначения \mathbf{D} , \mathbf{x} и \mathbf{y} аналогичны (3) (и останутся таковыми в дальнейшем), а ϵ - это величина порога, зависящая от величины шума (обычно $\epsilon = cN\sigma^2$, где $0.5 \leq c \leq 1.5$) [3]. Тогда, если $\hat{\mathbf{x}}$ - решение задачи (4), то

$$\hat{\mathbf{y}} = \mathbf{D}\hat{\mathbf{x}} \quad (5)$$

- итоговое изображение с удаленным шумом. Выбор словаря \mathbf{D} осуществляется исследователем. Опыт показывает, что вейвлеты не подходят в качестве словаря [3], поэтому применяются видоизмененные вейвлеты: курвлеты, бэндлеты и т.д. [4], а также матрица Хаара [3]. Здесь и далее, предполагается, что словарь \mathbf{D} нормирован на l_2 - норму его векторов. Из формулы (5) следует, что итоговое изображение:

$$\hat{\mathbf{y}} = \mathbf{D}\mathcal{S}_T(\mathbf{D}^T\mathbf{y}), \quad (6)$$

где \mathcal{S}_T – т.н. оператор порога: $\mathcal{S}_T(z) = 0$, при $|z| \leq \epsilon$ и $\mathcal{S}_T(z) = z$, иначе.

Таким образом, отсекаются малые значения координат разреженного представления, которые предположительно представляют шум, и остаются те, которые представляют незашумленное изображение.

Такой подход и его улучшения [6] показывают удовлетворительные результаты и хорошую производительность (время работы на конкретной картинке). Минусами данных подходов является необходимость подбора словаря, так, чтобы получить разреженное представление для всего изображения, а также слишком сложная структура итоговых алгоритмов, при их дальнейшем улучшении.

В связи с указанными недостатками глобального подхода, наибольшую перспективу составляют локальные методы, а также локальные методы с обучением словаря (его адаптацией к конкретному изображению). При таком подходе, изображение обрабатывается по элементам. Эти элементы должны перекрываться, с тем, чтобы не возникало артефактов на границах их соприкосновения, а затем полученные элементы усредняются. Так как дипломная работа представляет собой разбор и улучшение локального метода, приводится краткий обзор основных локальных методов.

3.2 Локальный алгоритм с использованием жесткого порога

Данный алгоритм предложен в статье [7]:

1. Выделение областей изображения размером $\sqrt{n} \times \sqrt{n}$, где n – заданное число пикселей в области (например, 64). Каждый пиксель изображения является центром такой области. Область обозначается, как \mathbf{p}_{ij} .
2. Для каждой области находится ее разреженное представление: $\tilde{\mathbf{q}}_{ij} = \mathbf{D}^T \mathbf{p}_{ij}$. Затем к нему применяется пороговый алгоритм (6): получается итоговое представление $\hat{\mathbf{q}}_{ij}$. Тогда итоговая область с удаленным шумом: $\hat{\mathbf{p}}_{ij} = \mathbf{D}\hat{\mathbf{q}}_{ij}$.

3. Итоговые области $\hat{\mathbf{p}}_{ij}$ усредняются так, чтобы получить гладкое изображение.

В качестве словаря в данном алгоритме могут быть выбраны векторы дискретного косинусного преобразования:

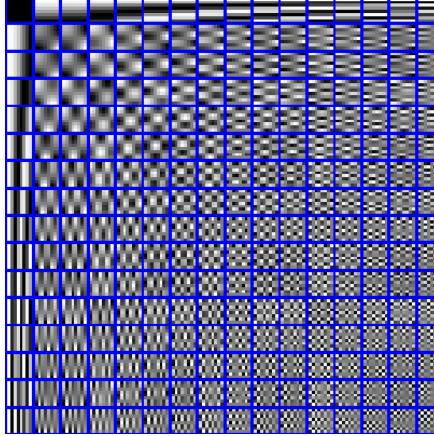


Рис. 1: Словарь DCT

3.3 Локальный алгоритм с использованием нежесткого порога

Вышеописанный метод дает неплохие результаты, однако не учитывает особенностей изображения, в связи с чем, итоговое изображение теряет мелкие детали и текстуры. В связи с этим недостатком предлагается [8] адаптировать пороговый оператор, обучаясь на части незашумленного изображения (в дальнейшем будет описано, как действуют, если нет незашумленного изображения).

Такой метод называется методом с использованием нежесткого порогового оператора. Описанная адаптация предназначается для того, чтобы низкочастотные компоненты, связанные с деталями изображения, «обрезались» меньше, а высокочастотные – предполагаемые шумовые – «обрезались» сильнее. Эта интерпретация следует из теории анализа сигналов и из экспериментального опыта обработки изображений.

Метод удаления шума с использованием нежесткого порога не приводится в дипломе полностью, однако необходимо описать идею данного алгоритма, как пример метода, способного обучаться на изображении.

Математическое представление описываемого метода аналогично задаче (3.2), но пороговый оператор \mathcal{S}_T предлагается представить в виде вектора различных порогов (пороги различны для различных частот):

$$\mathcal{S}\{\mathbf{u}\} = \begin{bmatrix} \mathcal{S}_1\{u[1]\} \\ \mathcal{S}_2\{u[2]\} \\ \vdots \\ \mathcal{S}_m\{u[m]\} \end{bmatrix}, \quad (7)$$

где $\mathbf{u}_k = \mathbf{D}^T \mathbf{p}_k$, а \mathcal{S}_i могут быть различны. Тогда штрафная функция для обучения:

$$F_{local}(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_m) = \sum_{i=1}^m \sum_{k=1}^M (u_k^0[i] - \mathcal{S}_i\{u_k[i]\})^2. \quad (8)$$

Представляя каждый из m пороговых операторов полиномом: $\mathcal{S}_i\{u\} = \sum_{j=0}^J c_i[j]u^j$ и подставив в (8), получим:

$$F_{local}(\mathbf{c}) = \sum_{k=1}^M \mathbf{U}_k^T \mathbf{D}^T (\mathbf{D} \mathbf{U}_k \mathbf{c} - \mathbf{p}_k^0), \quad (9)$$

где \mathbf{U}_k - это блочно-диагональная матрица векторов $\mathbf{b}_k = [u^k[i]^0, u^k[i]^1, \dots, u^k[i]^J]$, и $i = 1, \dots, m$. Тогда, из условия $\frac{\partial F_{local}(\mathbf{c})}{\partial \mathbf{c}} = 0$ (минимум штрафной функции), получим оптимальные коэффициенты для нахождения порогового оператора:

$$\mathbf{c}_{opt} = \left[\sum_{k=1}^M \mathbf{U}_k^T \mathbf{D}^T \mathbf{D} \mathbf{U}_k \right]^{-1} \sum_{k=1}^M \mathbf{U}_k^T \mathbf{D}^T \mathbf{p}_k^0. \quad (10)$$

Таким образом, было получено однозначное значение мягкого порогового оператора в зависимости от известных величин: разреженного представления зашумленного изображения и векторов обучающей выборки: незашумленного изображения. Подход обучения на изображении используется повсеместно в данной тематике. В дальнейшем можно использовать полученный пороговый оператор (множество кривых усадки, shrinkage curves) и для других изображений с той же величиной шума. Чем больше целевые изображения похожи на тестовое изображение, тем лучшим будет качество описываемого метода для удаления шума.

4 Методы, основанные на обучении словаря

Следующие описываемые методы наиболее качественны и перспективны. Они подробно изучены и описаны в дипломной работе, для того, чтобы предложить дальнейший метод, сочетающий в себе подходы описываемых методов.

Методы удаления шума, основанные на обучении словаря, являются наиболее эффективными для решения поставленной задачи, и поэтому представляют наибольший интерес для современных исследований. Основной идеей данных методов является построение словаря с использованием обучающей выборки. В общем виде такие методы можно описать основными этапами:

- Стадия инициализации словаря. Словарь \mathbf{D} задается случайными векторами или известным базисом (вейвлеты, курвлеты). Также словарь можно задать обученным ранее другим методом словарем.

- Стадия задания обучающей выборки. Задается обучающая выборка \mathbf{y}_i . Выборка составляется из разбиения небольшой части изображения (64×64 пикселя) на более мелкие части (8×8 пикселей), либо набором таких мелких частей из различных изображений. При этом необходимо выбирать обучающую выборку так, чтобы она максимально была похожа на контрольное изображение (текстуры, детали). Основой для обучающей выборки также может служить незашумленная частью заданного изображения (невыполнимо для реального случая); часть исходного изображения, очищенная ранее другим методом (например локальным пороговым алгоритмом).
- Стадия нахождения разреженного представления. По словарю на данной стадии находится разреженное представление для синтетически-зашумленной обучающей выборки. Применяются методы для решения задачи (4): метод поиска совпадения (matching pursuit, MP) [9], ортогональный метод поиска совпадений (OMP) [10], фокальный метод решения неопределенной системы уравнений (focal undetermined system solver, FOCUSS) [11].
- Стадия обновления словаря - ключевая стадия алгоритма. Векторы словаря обновляются по некоторому правилу (например максимум правдоподобия - maximum likelihood; максимум апостериорной вероятности - maximum a posteriori, MAP), которое формулируется, основываясь на сравнении обучающего набора частей изображения (патчей) и тех же частей с удаленным шумом (рассчитанным по формуле (6)). Правило обновления словаря определяет алгоритм удаления шума, основанного на обучении словаря.

4.1 Алгоритм поиска разреженного представления (Matching pursuit-алгоритм)

Так как нахождение разреженного представления является неотъемлемой частью задачи удаления шума из изображения с обучением словаря, необходимо упомянуть о методах, которые позволяют найти необходимое разложение по избыточному словарю. Данная задача является NP-трудной [12], поэтому нахождение ее точного решения представляется слишком сложным. В связи с этим фактом, предлагается [12] искать приближенное решение задачи разреженного представления (4).

Для краткости, в дипломной работе приводится только простейший жадный алгоритм поиска совпадения (matching pursuit, далее MP) [9]. Остальные алгоритмы подобны ему и могут быть найдены в указанной ранее литературе.

Инициализация алгоритма MP:

- Входом алгоритма является словарь \mathbf{D} и патч (часть) изображения \mathbf{p}_{ij} для которого необходимо найти разреженное представление \mathbf{q}_{ij} . Результатом данного алгоритма являются коэффициенты итогового вектора разреженного представления.
- Начальный вектор для поиска полагается $\mathbf{r}_1 = \mathbf{p}_{ij}$, номер шага алгоритма $k = 1$.

Шаг алгоритма:

- Найдем вектор словаря, дающий максимальное скалярное произведение с исходным патчем: $\mathbf{d}_{\alpha_k} = \arg \max_{\mathbf{d}_\alpha} |\langle \mathbf{r}_k, \mathbf{d}_\alpha \rangle|$.
- Коэффициент разложения по найденному вектору словаря: $q_{ijk} = \langle \mathbf{r}_k, \mathbf{d}_\alpha \rangle$.
- Положим $\mathbf{r}_{k+1} = \mathbf{r}_k - q_{ijk} \mathbf{d}_{\alpha_k}$.
- Увеличим номер итерации алгоритма $k \leftarrow k + 1$.

Алгоритм завершает работу при $\|\mathbf{r}_k\|_2 < \epsilon$, где ϵ - это заданный порог.

Приведенный алгоритм применяется на практике для поиска разреженного представления в задачах удаления шума, связанных с обучением словаря.

4.2 Метод K-SVD для обучения словаря

Метод К-сингулярного разложения (K-SVD) для удаления шума основан на обучении разреженного словаря [13]. Задачу разреженного разложения изображения по словарю можно представить, как обобщение задачи кластеризации: в задаче кластеризации точки принадлежат одному кластеру, и не принадлежат остальным, то есть вектор представления точки пространством кластеров описывается, как n -мерный вектор $(0, 0, \dots, 1, \dots, 0)^T$, в котором одна единица и остальные нули (точка принадлежит только одному кластеру). В задаче разреженного представления точка (патч изображения) описывается подобным вектором, коэффициенты которого являются коэффициентами разложения по словарю. Основываясь на предложенной интерпретации был предложен метод, в котором обучение словаря происходит способом, подобным методу кластеризации К-средних (K-means) [14].

Как было описано ранее, алгоритм определяется стадией обновления словаря: Для каждого столбца $k = 1, 2, \dots, K$ из словаря $\mathbf{D}^{(J-1)}$:

- Определим множество обучающих векторов (частей изображения), использующих этот столбец словаря: $\omega_k = \{i | 1 \leq i \leq N, \mathbf{x}_T^k \neq \mathbf{0}\}$
- Рассчитаем матрицу ошибки: $\mathbf{E}_k = \mathbf{Y} - \sum_{j \neq k} \mathbf{d}_j \mathbf{x}_T^j$
- Ограничим \mathbf{E}_k , выбрав только столбцы, соответствующие ω_k , получим матрицу \mathbf{E}_k^R .
- Применим сингулярное (SVD) разложение: $\mathbf{E}_k^R = \mathbf{U} \Delta \mathbf{V}^T$.
- Тогда обновленный столбец словаря $\tilde{\mathbf{d}}_j$ - первый столбец \mathbf{U} , а новый вектор разложения - \mathbf{x}_R^k - это первый столбец \mathbf{V} , помноженный на $\Delta(\mathbf{1}, \mathbf{1})$

Ниже приводится наиболее продвинутый метод для удаления шума из изображения. В настоящей работе проводится статистический анализ данного метода и предлагается его улучшение с точки зрения статистической интерпретации.

4.3 Метод CSR для удаления шума из изображения

Метод разреженного представления, основанный на кластеризации изображения (Clustering-based Sparse Representation, CSR) был предложен в статье [15]. Метод основан на наблюдении свойства изображения иметь идентичные участки (похожие патчи). В связи с чем предлагается наложить дополнительное ограничение на задачу разреженного представления (4), которое будет учитывать факт похожести частей изображения, а следственно, их разреженных представлений.

Напомним исходную задачу разреженного представления для части(патча) изображения:

$$(P_\epsilon^0) \min_{\vec{\alpha}_i} \|\vec{\alpha}_i\|_0 : \|\Phi \vec{\alpha}_i - \mathbf{x}_i\|_2^2 \leq \epsilon, \quad (11)$$

где $\mathbf{x}_i \in \mathbb{R}^n$ - патч изображения, “вытянутый” в вектор, $\vec{\alpha}_i \in \mathbb{R}^m$ - разреженное представление патча, $\Phi \in \mathbb{R}^{n \times m}, m \geq n$ - словарь, векторы которого совпадают с размерностью патча. Вектор патча представляется вектором полного изображения, как: $\mathbf{x}_i = \mathbf{R}_i \mathbf{X}$, где \mathbf{X} - это вектор полного изображения, а \mathbf{R}_i - т.н. оконный оператор, выделяющий часть изображения.

Полное изображение тогда составляется из патчей по методу наименьших квадратов: $\mathbf{X} = \left(\sum_i \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \left(\sum_i \mathbf{R}_i^T \mathbf{x}_i \right)$. Для полного изображения вводится вектор разреженных коэффициентов, полученный конкатенацией аналогичных векторов для каждого патча: $\vec{\alpha} = (\vec{\alpha}_1, \dots, \vec{\alpha}_N)^T$, где N - это общее число патчей в изображении. Вводится оператор \mathbf{D} восстановления полного изображения из его разреженного представления

$$\mathbf{X} = \mathbf{D} \vec{\alpha}, \text{ где} \quad (12)$$

$$\mathbf{D} = \left(\sum_i \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} (\mathbf{R}_1^T \Phi, \dots, \mathbf{R}_N^T \Phi). \quad (13)$$

Для формулировки метода CSR необходимо записать Лагранжиан для задачи (4):

$$\vec{\alpha} = \arg \min_{\vec{\alpha}} \frac{1}{2} \|\mathbf{Y} - \mathbf{D} \vec{\alpha}\|_2^2 + \lambda \|\vec{\alpha}\|_1, \quad (14)$$

где $\|\vec{\alpha}\|_0$ заменена на $\|\vec{\alpha}\|_1$ для обеспечения выпуклости задачи (4). $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ - зашумленное изображение с \mathbf{N} - гауссовым белым шумом.

В методе CSR предлагаются разбивать патчи изображения на K кластеров, принимая во внимание их схожесть по $\|\cdot\|_2$ и использовать эмпирическое свойство разреженных представлений быть похожими для похожих патчей изображения. Для того, чтобы учесть описанное свойство коэффициентов $\vec{\alpha}_i$ для патчей, необходимо добавить дополнительный член в функцию стоимости (14):

$$(\vec{\alpha}, \vec{\mu}) = \arg \min_{\vec{\alpha}} \frac{1}{2} \|\mathbf{Y} - \mathbf{D} \vec{\alpha}\|_2^2 + \lambda_1 \|\vec{\alpha}\|_1 + \lambda_2 \sum_{k=1}^K \sum_{i \in C_k} \|\Phi \vec{\alpha}_i - \vec{\mu}_k\|_2^2, \quad (15)$$

где K - это число кластеров похожих частей изображения, C_k обозначает k -й кластер, а $\vec{\mu}_k$ обозначает центр C_k .

Далее задача (15) оптимизируется для упрощения ее понимания и дальнейшего решения путем введения разреженного представления центра кластера $\vec{\beta}$ по словарю, применявшемуся для разложения патчей изображения: $\vec{\mu} = \Phi\vec{\beta}$. Тогда искомая оптимизация запишется в виде:

$$(\vec{\alpha}, \vec{\beta}) = \arg \min_{\vec{\alpha}} \frac{1}{2} \|\mathbf{Y} - \mathbf{D}\vec{\alpha}\|_2^2 + \lambda_1 \|\vec{\alpha}\|_1 + \lambda_2 \sum_{k=1}^K \sum_{i \in C_k} \|\Phi\vec{\alpha}_i - \Phi\vec{\beta}_k\|_2^2. \quad (16)$$

Принимая во внимание, что словарь нормирован, и заменив в последнем члене $\|\cdot\|_2^2$ на $\|\cdot\|_1$ для наилучшего решения (эмпирическое наблюдение [16]), получается итоговая задача:

$$(\vec{\alpha}, \vec{\beta}) = \arg \min_{\vec{\alpha}} \frac{1}{2} \|\mathbf{Y} - \mathbf{D}\vec{\alpha}\|_2^2 + \lambda_1 \|\vec{\alpha}\|_1 + \lambda_2 \sum_{k=1}^K \sum_{i \in C_k} \|\vec{\alpha}_i - \vec{\beta}_k\|_1. \quad (17)$$

Решение полученной задачи оптимизации не приводится в дипломе, и может быть найдено в статьях [15], [17].

4.4 Результаты сравнения описываемых методов удаления шума

Приводятся результаты сравнения вышеизложенных методов. Методы были адаптированы и апробированы на серии изображений с синтетическим шумом; на изображениях, полученных с электронного микроскопа, а также на реальных фотографиях низкого качества:



Рис. 2: Исходное изображение

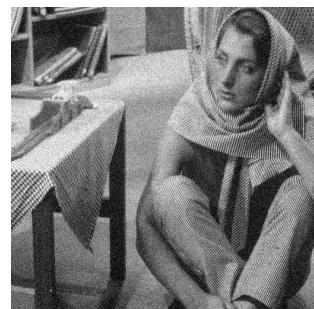


Рис. 3: Зашумленное изображение, $\sigma = 20$



Рис. 4: Изображение с удаленным шумом,
KSVD



Рис. 5: Изображение с удаленным шумом,
CSR

Значение пиковое отношение сигнал-шум для зашумленного изображения: $PSNR = 22.12$. Для изображения с удаленным методом KSVD шумом это значение составляет $PSNR_{KSVD} = 31.63$. Для метода CSR $PSNR_{CSR} = 31.78$.



Рис. 6: Исходное изображение



Рис. 7: Зашумленное изображение, $\sigma = 20$



Рис. 8: Изображение с удаленным шумом,
KSVD





Рис. 10: Исходное изображение



Рис. 11: Зашумленное изображение, $\sigma = 30$



Рис. 12: Изображение с удаленным шумом,
KSVD



Рис. 13: Изображение с удаленным шумом,
CSR

Для естественных изображений оценка с помощью PSNR не применима, поэтому
качество оценивается визуально.



Рис. 14: Исходное изображе-
ние



Рис. 15: Изображение с уда-
ленным шумом, KSVD



Рис. 16: Изображение с уда-
ленным шумом, CSR

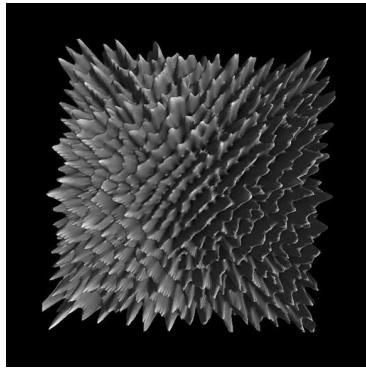


Рис. 17: Исходное изображение

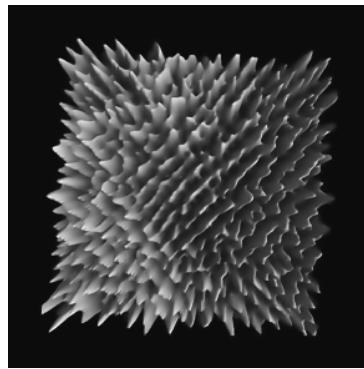


Рис. 18: Изображение с удаленным шумом, KSVD

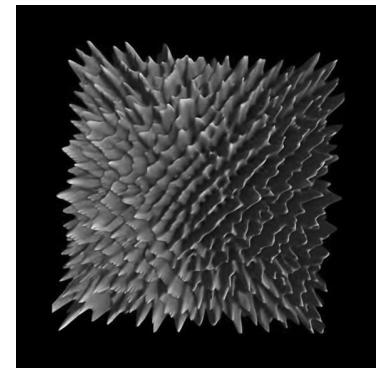


Рис. 19: Изображение с удаленным шумом, CSR



Рис. 20: Исходное изображение



Рис. 21: Изображение с удаленным шумом, KSVD



Рис. 22: Изображение с удаленным шумом, CSR

Как можно видеть, метод CSR дает лучшие результаты по сравнению с KSVD, поэтому считается наиболее перспективным.

5 Статистическая интерпретация задачи разреженного представления

Вышеизложенные методы для удаления шума из изображения основаны на различных подходах к решению оптимизационной задачи (4), а также модификаций этой задачи. Из обзора литературы становится ясно, что качество удаления шума зависит от формулировки задачи (4) и от качества найденного разреженного представления в данной формулировке. Методы поиска наилучшего разреженного представления при фиксированной формулировке ясны и описаны в литературе.

В настоящее время для постановки задач разреженного представления используются эмпирические наблюдения, касающиеся изображения (например то, что части изображения похожи, или что необходимо адаптировать словарь к конкретному изображению). Эмпирический подход ограничивает возможности по созданию новых методов удаления шума, а математическая интерпретация и происхождение некоторых

существующих методов неясны. В связи с указанными недостатками эмпирического подхода, в дипломной работе предлагается статистический (байесовский) подход к анализу изображений. Приводятся байесовские интерпретации и анализ существующих методов, выносятся предложения по улучшению этих методов, а также предлагается улучшение метода CSR для удаления шума.

Статистическим подходом называется оценка математических величин с точки зрения математической статистики. Изображение и его разреженное представление в таком случае представляются некоторыми случайными величинами и их распределениями. Накладывая такие априорные представления о характеристиках изображения, можно представить задачи оптимизации для удаления шума с использованием разреженного представления, как максимум правдоподобия, либо максимум апостериорной вероятности.

Например, подобный подход ранее был использован для метода наименьших квадратов.

5.1 Пример: метод наименьших квадратов в статистической интерпретации

Изначально, метод наименьших квадратов вводился без использования статистических предположений и задача формулировалась, как задача минимизации квадратичной невязки. В дипломе приводится случай системы линейных уравнений:

$$X\vec{w} = \vec{y}, \quad (18)$$

где $\vec{w} \in \mathbb{R}^m$, $\vec{y} \in \mathbb{R}^n$ и $X \in \mathbb{R}^{n \times m}$, и $n \geq m$, т.е. система является избыточной, и, в общем случае, не имеет решений. Тогда подобная система решается минимизацией “расстояния” между векторами $X\vec{w}$ и \vec{y} . Для этого применяется евклидова норма разности векторов:

$$(X\vec{w} - \vec{y})^T (X\vec{w} - \vec{y}) \rightarrow \min. \quad (19)$$

Решением задачи (19) является формула, которая и называется методом наименьших квадратов:

$$\vec{w} = (X^T X)^{-1} X^T \vec{y}. \quad (20)$$

В статистической интерпретации предлагается следующая модель:

$$X\vec{w} + \vec{\epsilon} = \vec{y}, \quad (21)$$

где $\vec{\epsilon}$ - центрированный нормально распределенный случайный вектор с нулевым матожиданием и попарно независимыми компонентами, т.е.: $\vec{\epsilon} \sim \mathcal{N}(\vec{0}, \sigma^2 I)$, или:

$p(\vec{\epsilon}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\sigma^2 I)}} \exp\left(-\frac{1}{2} (\vec{\epsilon} - \vec{m})^T (\sigma^2 I)^{-1} (\vec{\epsilon} - \vec{m})\right)$. Эта модель в литературе называется моделью линейной регрессии [19].

Будем считать \vec{y} случайным вектором, а X и \vec{w} фиксированными, тогда можно записать условие максимального правдоподобия:

$$p(\vec{y} | X, \vec{w}) \rightarrow \max_{\vec{w}} \quad (22)$$

Считая \vec{y} переменной, запишем (22) подробнее:

$$\begin{aligned} p(\vec{y} | X, \vec{w}) &= p(X\vec{w} + \vec{\epsilon} = \vec{y} | X, \vec{w}) = p(\vec{\epsilon} = \vec{y} - X\vec{w} | X, \vec{w}) = p(\vec{\epsilon}) = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \exp\left(-\frac{1}{2\sigma^2} \vec{\epsilon}^T \vec{\epsilon}\right) \rightarrow \max_{\vec{w}}. \end{aligned} \quad (23)$$

Применив логарифмирование, отбросив константу и подставив значение $\vec{\epsilon}$ из (21), получим:

$$-\frac{1}{2\sigma^2} (\vec{y} - X\vec{w})^T (\vec{y} - X\vec{w}) \rightarrow \max_{\vec{w}}, \quad (24)$$

что, при смене знака совпадает с эмпирической формулировкой (21).

Такая статистическая интерпретация в литературе [19] называется 0-м уровнем байесовского вывода. Дальнейшим развитием метода наименьших квадратов могут послужить априорное предположение о распределении случайного вектора \vec{w} (1-й уровень байесовского вывода), либо, далее, априорные предположения о скрытых параметрах указанного распределения (2-й уровень байесовского вывода). Возможно и более высокие уровни, однако такие предположения сложны и плохо интерпретируются.

Таким образом, приведен пример того, как можно статистически интерпретировать задачу, и сделано предложение по возможным вариациям статистической интерпретации. Подобное углубление накладывает дополнительные штрафы в итоговом методе (21), а их вид зависит от вышеописанного апостерионого предположения. Подобная схема байесовского вывода была предложена для метода удаления шума из изображения CSR.

5.2 Предлагаемая модификация метода CSR

Отметим, что итоговая задача оптимизации в методе CSR (17) получена с помощью замены квадратичного штрафа на штраф l_1 для улучшения качества результата, а целесообразность замены подтверждена эмпирически. При таком подходе дальнейшее изменение модели, а также учет априорных знаний затруднен в силу отсутствия статистической интерпретации. Предложим далее модификацию алгоритма CSR, построив соответствующую статистическую модель генерации изображения. Отметим, что предложенная модификация обладает преимуществом интерпретируемости, а также может быть легко изменена для учета имеющейся априорной информации.

Предлагается следующая статистическая модель генерации изображения \mathbf{X}

$$\begin{cases} \mathbf{X} = \mathbf{D}\vec{\alpha} + \vec{\epsilon} \\ \vec{\alpha}_i \sim Laplace\left(\vec{\beta}_k, \tau_1^{-1} I\right) \\ \vec{\beta}_k \sim Laplace\left(\vec{0}, \tau_2^{-1} I\right) \end{cases}, \quad (25)$$

где $\vec{\alpha} = [\alpha_1, \dots, \alpha_N]^T$ есть объединенный вектор разреженных представлений, \mathbf{D} есть введенный ранее (13) оператор осреднения, который восстанавливает изображение по патчам, $Laplace(\vec{m}, \tau^{-1} I)$ - это распределение Лапласа: $p(\vec{w}) = \prod_{i=1}^n \tau/2 \exp(-\tau|w_i - m_i|)$ (рассматривается вектор с независимыми в совокупности

компонентами), а $\vec{\epsilon} \sim \mathcal{N}(\vec{0}, \sigma^2 I)$. Приведенная статистическая модель (25) имеет следующую интерпретацию. Сначала генерируются разреженные представления центров кластеров $\vec{\beta}_k$. Разреженность здесь индуцируется априорным распределением Лапласа на компоненты вектора представления. Распределение Лапласа является поощряющим разреженность и применяется, например, как априорное для вектора параметров \mathbf{w} в линейной регрессии для отбора признаков [18]. Затем по сгенерированным центрам кластеров генерируются его представители – коэффициенты разреженного представления патчей соответствующего кластера на изображении. Как и в случае с центрами кластеров разреженность индуцируется с помощью распределения Лапласа. По полученным разреженным представлениям патчей вычисляется истинное изображение, как среднее по патчам \mathbf{x}_i . Итоговое изображение есть смесь истинного с независимым нормальным шумом.

Описанная схема отражает предположение авторами [15] о центрах кластеров разреженных представлений $\vec{\beta}_k$, о том, что разреженные представления патчей $\vec{\alpha}_k$ разбиваются на кластеры (из (25) матожиданием является $\vec{\beta}_k$), а итоговое изображение представляет собой зашумленное независимым нормальным шумом истинное изображение, восстановливаемое по своему разреженному представлению.

Коэффициенты разреженных представлений центров кластеров и патчей получим из принципа максимума апостериорной вероятности.

$$p(\vec{\alpha}_1, \dots, \vec{\alpha}_N, \vec{\beta}_1, \dots, \vec{\beta}_K | \Phi, \mathbf{X}) \rightarrow \max_{\vec{\alpha}_1, \dots, \vec{\alpha}_N, \vec{\beta}_1, \dots, \vec{\beta}_K}. \quad (26)$$

Обозначим совокупность векторов $\vec{\alpha}_1, \dots, \vec{\alpha}_N$ $\tilde{\alpha}$, а совокупность векторов $\vec{\beta}_1, \dots, \vec{\beta}_K$ $\tilde{\beta}$ для сокращения записи. Используя формулу Байеса, получим:

$$p(\tilde{\alpha}, \tilde{\beta} | \Phi, \mathbf{X}) = \frac{p(\mathbf{X} | \Phi, \tilde{\alpha}, \tilde{\beta}) p(\tilde{\alpha}, \tilde{\beta})}{\int p(\mathbf{X} | \Phi, \tilde{\alpha}', \tilde{\beta}') p(\tilde{\alpha}', \tilde{\beta}') d\tilde{\alpha}' d\tilde{\beta}'} \propto p(\tilde{\alpha}, \tilde{\beta}) p(\mathbf{X} | \Phi, \tilde{\alpha}, \tilde{\beta}), \quad (27)$$

где $p(\tilde{\alpha}, \tilde{\beta}) = p(\tilde{\alpha} | \tilde{\beta}) p(\tilde{\beta})$.

Подставим значения для вероятностей из (25) в (27):

$$\begin{aligned} p(\tilde{\alpha}, \tilde{\beta} | \Phi, \mathbf{X}) &\propto p(\mathbf{X} | \Phi, \tilde{\alpha}, \tilde{\beta}) p(\tilde{\alpha}, \tilde{\beta}) = \\ &= p(\vec{\epsilon} = \mathbf{X} - \mathbf{D}\tilde{\alpha} | \mathbf{D}, \tilde{\alpha}, \tilde{\beta}) p(\tilde{\alpha} | \tilde{\beta}) p(\tilde{\beta}) = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \exp\left(-\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{D}\tilde{\alpha})^T (\mathbf{X} - \mathbf{D}\tilde{\alpha})\right) \prod_{i \in C_k} \frac{\tau_1}{2} \exp(-\tau_1 |\alpha_i - \beta_i|) \cdot \\ &\quad \cdot \prod_{k=1}^K \frac{\tau_2}{2} \exp(-\tau_2 |\beta_k|). \end{aligned} \quad (28)$$

Введем обозначение $\vec{\beta} = [\vec{\beta}_1, \dots, \vec{\beta}_K]^T$. Прологарифмировав выражение (28),

получим следующую задачу оптимизации:

$$\left(\vec{\alpha}, \vec{\beta}\right) = \arg \min_{\vec{\alpha}, \vec{\beta}} \frac{1}{2} \|\mathbf{X} - \mathbf{D}\vec{\alpha}\|_2^2 + \lambda_1 \|\vec{\beta}\|_1 + \lambda_2 \sum_{k=1}^K \sum_{i \in C_k} \|\vec{\alpha}_i - \vec{\beta}_k\|_1. \quad (29)$$

Заметим, что задача оптимизации, полученная из предложенной вероятностной модели, отличается от задачи оптимизации в методе CSR (17) тем, что использует штраф за l_1 норму $\vec{\beta}$, а не $\vec{\alpha}$. Эмпирическая задача оптимизации, предложенная в методе CSR фактически означает, что средний вектор кластера генерируется по его представителям, а не наоборот. Представляется более естественным обратный порядок генерации, который и используется в предложенной вероятностной модели.

Сравним далее предложенный метод с методом CSR и методом K-SVD на паре изображений со стандартным отклонением шума $\sigma = 20$ и $\sigma = 30$ (см. рис. 23). Значения соотношения PSNR для всех трех методов приведено в табл. 1.

Таблица 1: Сравнение по PSNR предложенного метода удаления шума с методами CSR и K-SVD

	K-SVD	CSR	Предложенный метод
$\sigma = 20$	31.63	31.78	31.79
$\sigma = 30$	29.79	29.73	29.75

Из табл. 1 заключаем, что предложенный метод имеет сходное качество удаления шума по отношению к методу CSR, а потому представляется перспективным для решения задачи удаления шума. В дальнейшем в рамках предложенной статистической модели планируется изучить другие априорные распределения, которые имеют более сильные свойства поощрения разреженности, например, распределение Дирихле [19, 21]. Отметим, однако, что использование распределения Дирихле может сделать оптимизационную задачу невыпуклой, что затрудняет поиск ее решения [20].

6 Выводы

В настоящей дипломной работе были исследованы и апробированы методы удаления шума из изображения с использованием разреженного представления. Была показана эволюция данных методов от глобального порогового алгоритма (4) до современных методов, основанных на обучении словаря (4.2) и кластеризации похожих патчей изображения (17). Методы были апробированы на изображениях с добавленным синтетически шумом; сравнивалось их качество. Также, была проведена серия экспериментов на реальных изображениях, полученных с трансмиссионного электронного микроскопа, и на обычных нечетких фотографиях. Приведенные исследования являются довольно полными в данной тематике, и могут быть использованы для дальнейшего улучшения качества удаления шума с использованием разреженного представления.

Стандартно, поиск разреженного представления основывается на решении задачи оптимизации с эмпирическими, либо эвристическими ограничениями. Постановка



(а) Изображение с удаленным шумом, KSVD



(б) Изображение с удаленным шумом, CSR



(в) Изображение с удаленным шумом, предложенный метод

Рис. 23: Сравнение предложенного метода с методами K-SVD и CSR на изображении с $\sigma = 20$.

задачи различается от метода к методу по используемым при оптимизации ограничениям, и поэтому она основана на наблюдаемых свойствах изображений. Такой подход является естественным, однако он плохо обоснован и было необходимо выяснить предпосылки к использованию этих эмпирических ограничений оптимизаций.

В магистерской работе была построена байесовская схема, которая приводит к формулировке задач оптимизации через максимум апостериорной вероятности. Указанная схема позволяет дать интерпретацию известных методов с точки зрения статистических моделей и предлагать новые формулировки оптимационных задач, основываясь на ясных математических предположениях о распределениях случайных величин. Описанная схема была применена для наиболее продвинутого метода, позволила обосновать некоторые его неточности, и даже предложить новую формулировку задачи оптимизации для этого метода.

Предложенная схема может быть использована в дальнейшей работе с разреженными представлениями, и может быть применена не только для удаления шума, но и в других областях, связанных с анализом изображений: восстановление пикселей, повышение разрешения и скалирование.

Список литературы

- [1] Колдаева М.В Удаление шума из изображения с использованием разреженного представления. //Сборник тезисов 56-й научной конференции МФТИ, секция информатика - 2013.
- [2] Колдаева М.В Удаление шума из изображения с использованием разреженного представления. Методы, основанные на обучении словаря. //Сборник тезисов 57-й научной конференции МФТИ, секция информатика - 2014.
- [3] Elad M. Sparse and redundant representations: from theory to applications in signal and image processing. – Springer, 2010.
- [4] Starck J. L., Candes E. J., Donoho D. L. The curvelet transform for image denoising //Image Processing, IEEE Transactions on. – 2002. – Т. 11. – №. 6. – С. 670-684.
- [5] Aharon M., Elad M. Image denoising via sparse and redundant representations over learned dictionaries //Image Processing, IEEE Transactions on. – 2006. – Т. 15. – №. 12. – С. 3736-3745
- [6] Chang S. G., Yu B., Vetterli M. Adaptive wavelet thresholding for image denoising and compression //Image Processing, IEEE Transactions on. – 2000. – Т. 9. – №. 9. – С. 1532-1546.
- [7] Guleryuz O. G. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: theory //Image Processing, IEEE Transactions on. – 2006. – Т. 15. – №. 3. – С. 539-554.
- [8] Hel-Or Y., Shaked D. A discriminative approach for wavelet denoising //Image Processing, IEEE Transactions on. – 2008. – Т. 17. – №. 4. – С. 443-457.
- [9] Mallat S. G., Zhang Z. Matching pursuits with time-frequency dictionaries //Signal Processing, IEEE Transactions on. – 1993. – Т. 41. – №. 12. – С. 3397-3415.
- [10] Chen S., Billings S. A., Luo W. Orthogonal least squares methods and their application to non-linear system identification //International Journal of control. – 1989. – Т. 50. – №. 5. – С. 1873-1896.
- [11] Gorodnitsky I. F., Rao B. D. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm //Signal Processing, IEEE Transactions on. – 1997. – Т. 45. – №. 3. – С. 600-616.
- [12] Davis G., Mallat S., Avellaneda M. Adaptive greedy approximations //Constructive approximation. – 1997. – Т. 13. – №. 1. – С. 57-98.
- [13] Aharon M., Elad M., Bruckstein A. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation //Signal Processing, IEEE Transactions on. – 2006. – Т. 54. – №. 11. – С. 4311-4322.
- [14] Gersho A., Gray R. M. Vector quantization and signal compression. – Springer Science and Business Media, 1992.

- [15] Dong W. et al. Sparsity-based image denoising via dictionary learning and structural clustering //Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. – IEEE, 2011. – C. 457-464.
- [16] Candes E. J., Romberg J., Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information // Information Theory, IEEE Transactions on. – 2006. – T. 52. – №. 2. – C. 489-509.
- [17] Daubechies I., Defrise M., De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint // Communications on pure and applied mathematics. – 2004. – T. 57. – №. 11. – C. 1413-1457.
- [18] Williams P. M. Bayesian regularization and pruning using a Laplace prior // Neural computation. – 1995. – T. 7. – №. 1. – C. 117-143.
- [19] Bishop C. M. et al. Pattern recognition and machine learning. – New York : Springer, 2006. – Vol. 4. – No. 4. – P. 12.
- [20] Boyd S., Vandenberghe L. Convex optimization. – Cambridge university press, 2004.
- [21] Marlin B. M., Murphy K. P. Sparse Gaussian graphical models with unknown block structure // Proceedings of the 26th Annual International Conference on Machine Learning. – ACM, 2009. – C. 705-712.