

Министерство образования и науки Российской Федерации

УДК
ГРНТИ
Инв. №

УТВЕРЖДЕНО:

Исполнитель:

Государственное учебно-научное учреждение
Факультет вычислительной математики и
кибернетики Московского государственного
университета имени М.В.Ломоносова

От имени Руководителя организации

_____ / Моисеев Е.И. /
М.П.

НАУЧНО-ТЕХНИЧЕСКИЙ ОТЧЕТ

о выполнении 3 этапа Государственного контракта
№ 14.740.11.0996 от 23 мая 2011 г.

Исполнитель: Государственное учебно-научное учреждение Факультет
вычислительной математики и кибернетики Московского государственного
университета имени М.В.Ломоносова

Программа (мероприятие): Федеральная целевая программа «Научные и научно-
педагогические кадры инновационной России» на 2009-2013 гг., в рамках реализации
мероприятия № 1.2.2 Проведение научных исследований научными группами под
руководством кандидатов наук.

Проект: Математические методы анализа и обработки стохастических изображений.

Руководитель проекта:

_____ /Шестаков Олег Владимирович
(подпись)

Москва
2012 г.

СПИСОК ОСНОВНЫХ ИСПОЛНИТЕЛЕЙ
по Государственному контракту 14.740.11.0996 от 23 мая 2011 на выполнение
поисковых научно-исследовательских работ для государственных нужд

Организация-Исполнитель: Государственное учебно-научное учреждение Факультет вычислительной математики и кибернетики Московского государственного университета имени М.В.Ломоносова

Руководитель темы:

кандидат физико-математических наук, доцент _____ Шестаков О. В.
подпись, дата

Исполнители темы:

кандидат физико-математических наук, без ученого звания _____ Кудрявцев А. А.
подпись, дата

кандидат физико-математических наук, без ученого звания _____ Мизотин М. М.
подпись, дата

без ученой степени, без ученого звания _____ Насонова А. А.
подпись, дата

кандидат физико-математических наук, без ученого звания _____ Насонов А. В.
подпись, дата

без ученой степени, без ученого звания _____ Семашко А. С.
подпись, дата

без ученой степени, без ученого звания _____ Сторожилова М. В.
подпись, дата

без ученой степени, без ученого звания _____ Сергеев В. В.
подпись, дата

без ученой степени, без
ученого звания

_____ Садовой И. А.
подпись, дата

без ученой степени, без
ученого звания

_____ Кузнецова М. Г.
подпись, дата

Реферат

Отчет 65 с., 1 ч., 14 рис., 2 табл., 36 источн., 0 прил.

стохастические изображения, томографические изображения, локальные дескрипторы, ортогональные многочлены, разреженные представления

В отчете представлены результаты исследований, выполненных по 3 этапу Государственного контракта № 14.740.11.0996 "Математические методы анализа и обработки стохастических изображений." (шифр "2011-1.2.2-111-001") от 23 мая 2011 по направлению "Проведение научных исследований научными группами под руководством кандидатов наук в следующих областях:- математика;- механика" в рамках мероприятия 1.2.2 "Проведение научных исследований научными группами под руководством кандидатов наук.", мероприятия 1.2 "Проведение научных исследований научными группами под руководством докторов наук и кандидатов наук", направления 1 "Стимулирование закрепления молодежи в сфере науки, образования и высоких технологий." федеральной целевой программы "Научные и научно-педагогические кадры инновационной России" на 2009-2013 годы.

Цель работы - разработка методов обработки и анализа стохастических и медицинских изображений. Основные направления работ - создание и исследование метода нелинейной пороговой обработки, направленного на подавление шума, а также метода индексации базы данных томографических изображений на основе локальных особенностей изображений, выделяемых с помощью базисов ортогональных многочленов. Применение аппарата разреженных представлений для создания методов обработки и анализа стохастических изображений.

При выполнении работ по третьему этапу государственного контракта для решения поставленных задач были применены методы вейвлет анализа, пороговой обработки, аппарат разложения функций по различным базисам, в том числе по ортогональным многочленам и словарю разреженных представлений, а также различные методы математической статистики для построения оценок параметров.

Для проведения вычислительных экспериментов на втором была задействована вычислительная и программная база кафедры математической статистики и лаборатории математических методов обработки изображений факультета ВМК МГУ имени М.В. Ломоносова.

На третьем этапе работ получены следующие результаты:

1. Методы нелинейной пороговой обработки сигналов и изображений,

направленных на подавление шума.

2. Метод обработки и анализа изображений на основе разреженных представлений.

3. Метод индексации и поиска изображений на основе систем ортогональных многочленов.

СОДЕРЖАНИЕ

Введение.....	7
Проведение III этапа исследований по проблеме «Математические методы анализа и обработки стохастических изображений»	9
1 Разработка методов нелинейной пороговой обработки сигналов и изображений, направленных на подавление шума.....	9
2 Разработка методов обработки и анализа изображений с помощью разреженных представлений.....	23
3 Создание методов обработки и анализа изображений на основе систем ортогональных многочленов.....	35
Заключение	60
Список использованных источников	62

Введение

Вейвлет-разложение применяется для обработки сигналов и изображений в самых разнообразных областях, включая геофизику, физику плазмы, вычислительную томографию, компьютерную графику и т.д. Одной из основных задач, для решения которой используется вейвлет-разложение, является удаление шума. При этом наиболее популярным методом является пороговая обработка вейвлет-коэффициентов, которая обнуляет коэффициенты, не превышающие заданного порога. Порог можно выбирать различными способами, исходя из постановки задачи и целей обработки. Наличие шума неизбежно приводит к погрешностям в оцениваемом сигнале, что особенно сказывается на качестве изображений, восстанавливаемых по сигналам, представляющим собой некоторые интегральные характеристики исходного изображения, например, в томографических экспериментах.

Математический аппарат разреженных представлений динамично развивается и все больше и больше применяется для решения задач обработки изображений, видео и др. сигналов. В то время как в некоторых типах задач (например, шумоподавлении, сжатии, восстановлении потерянных данных) методы разреженных представлений применяются широко и успешно, представляют интерес и другие способы их использования – например, для классификации изображений или их частей. В работе исследуется применение разреженных представлений для обнаружения хориоретинальной атрофии с помощью поблочной классификации.

Среди различных применений томографических методов особенно важным является их применение в медицине. Магнитно-резонансная томография (МРТ) является незаменимым современным диагностическим средством и уже получила широкое распространение. Число медицинских изображений быстро растет в связи с появлением новых методов медицинских исследований, а также широкого распространения оборудования для проведения обследований, и ручной анализ такого

огромного объема данных очень трудоемок. В литературе можно найти результаты исследований, которые показали, что эффективность работы докторов увеличивается при использовании поисковых систем по изображениям. В связи с этим, задача поиска МРТ является актуальной.

В разделе 3 рассматривается применение дескрипторов на основе ортогональных многочленов к поиску медицинских изображений, в частности для поиска МРТ изображений мозга и обнаружения болезни Альцгеймера (БА). Наличие у радиологов и неврологов «под рукой» МРТ изображений похожих случаев позволяет вынести более обоснованное решения о состоянии здоровья и диагнозе.

**Проведение III этапа исследований по проблеме
«Математические методы анализа и обработки стохастических
изображений»**

1 Разработка методов нелинейной пороговой обработки сигналов и изображений, направленных на подавление шума.

1.1 Метод реконструкции изображений

В данном разделе на примере интегрального преобразования, используемого в томографических задачах, разрабатывается метод нелинейной пороговой обработки, который также применим к широкому кругу других приложений, связанных с подавлением шума в регистрируемых данных. Томографические методы реконструкции изображений получили широкое распространение в медицине, биологии, физике плазмы, газовой динамике, геофизике, астрономии и радиолокации (см., например, [1–3]). Во многих видах томографических экспериментов основную роль при реконструкции изображений играет преобразование (или оператор) Радона:

$$\begin{aligned} Rf(\varphi, s) &= \int_{L_{\varphi, s}} f(x, y) dl = \\ &= \int_{-\infty}^{\infty} f(s \cos \varphi - t \sin \varphi, s \sin \varphi + t \cos \varphi) dt, \quad s \in \mathbb{P}, \varphi \in [0, 2\pi), \end{aligned} \tag{1.1}$$

где $f(x, y) \in \mathbf{L}^2(\mathbb{P}^2)$ - функция с компактным носителем, описывающая искомое изображение, а $L_{\varphi, s}$ - прямая на плоскости, по которой интегрируется функция $f(x, y)$ (здесь s - расстояние от начала координат до прямой, а φ - угол, образованный с осью x перпендикуляром, опущенным из начала координат на эту прямую). При фиксированном значении φ функцию $Rf(\varphi, s)$ также называют проекцией функции $f(x, y)$ под углом φ . Для

преобразования Радона справедлива так называемая теорема о центральном сечении, играющая ключевую роль во многих методах реконструкции изображений по проекциям:

$$\hat{R}f(\varphi, \omega) = \sqrt{2\pi} \hat{f}(\omega \cos \varphi, \omega \sin \varphi), \quad (1.2)$$

где $\hat{R}f(\varphi, \omega)$ - одномерное преобразование Фурье функции $Rf(\varphi, s)$ по второй переменной, а $\hat{f}(\omega_1, \omega_2)$ - двумерное преобразование Фурье функции $f(x, y)$.

На практике проекционные данные регистрируются с некоторой погрешностью (шумом), возникающей из-за несовершенства оборудования и различных случайных помех. И поскольку задача обращения преобразования Радона относится к классу так называемых некорректно поставленных задач, при реконструкции изображений приходится использовать методы регуляризации. В работах [4–7] рассматривается метод регуляризации, заключающийся в разложении функции изображения по двумерному ортонормированному вейвлет-базису с последующей пороговой обработкой коэффициентов разложения. При этом функция Rf раскладывается по так называемому вейглет-базису, вообще говоря не являющемуся ортонормированным. Однако в силу «почти ортогональности» вейглетов среднеквадратичный риск метода оценивается при дополнительных предположениях, эквивалентных условию ортонормированности вейглет-базиса.

Методы анализа среднеквадратичного риска и его оценки в задаче реконструкции томографических изображений во многом опираются на тот факт, что преобразование Радона обладает многими свойствами линейных однородных операторов, которые рассматривались в работах [8] и [9]. В указанных работах рассматривается метод пороговой обработки коэффициентов вейглет-вейвлет разложения сигнала, и показывается, что при некоторых условиях регулярности оценка среднеквадратичного риска этого метода асимптотически нормальна. Далее мы представим метод

обращения преобразования Радона, альтернативный методу вейвлет-вейглет разложения и исследуем свойства оценки среднеквадратичного риска этого метода.

Разложим функцию $Rf(\varphi, s)$ в ряд Фурье по переменной φ на $[0, 2\pi]$. Поскольку функция $Rf(\varphi, s)$ вещественна, мы будем использовать вещественный базис Фурье

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(n\varphi)}{\sqrt{\pi}}, \frac{\sin(n\varphi)}{\sqrt{\pi}} \right\}_{n=1}^{\infty}.$$

Для сокращения записи обозначим элементы этого базиса через $\{Z_n\}_{n=0}^{\infty}$, перенумеровав их соответствующим образом. Имеем

$$Rf(\varphi, s) = \sum_{n=0}^{\infty} Rf_n(s) Z_n(\varphi),$$

где

$$Rf_n(s) = \int_0^{2\pi} Rf(\varphi, s) Z_n(\varphi) d\varphi.$$

Рассмотрим вейвлет-базис $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$, где $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$, а $\psi(x)$ - некоторая вейвлет-функция. Индекс j называется масштабом, а индекс k - сдвигом. Семейство $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ образует ортонормированный базис в $L^2(\mathbb{P})$ (см. [10]). Далее разложим функции $Rf_n(s)$ по базису $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$:

$$Rf(\varphi, s) = \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \langle Rf_n, \psi_{j,k} \rangle \psi_{j,k}(s) Z_n(\varphi). \quad (1.3)$$

После применения к разложению (1.3) обратного преобразования R^{-1} , получаем представление для функции f , которое является аналогом вейглет-вейвлет разложения, рассмотренного в работах [8], [9] и [11]:

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \beta_{n,j,k} \langle Rf_n, \psi_{j,k} \rangle u_{n,j,k}(x, y), \quad (1.4)$$

где

$$u_{n,j,k}(x,y) = \frac{R^{-1}[Z_n \psi_{j,k}](x,y)}{\beta_{n,j,k}}, \quad \beta_{n,j,k} = \|R^{-1}[Z_n \psi_{j,k}]\|.$$

Последовательность $\{u_{n,j,k}\}$ не образует ортонормированную систему, однако если вейвлет-функция ψ удовлетворяет некоторым условиям гладкости, то последовательность $\{u_{n,j,k}\}$ образует устойчивый базис, т.е. существуют такие константы $0 < A \leq B < \infty$, что

$$A \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} c_{n,j,k}^2 \leq \left\| \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} c_{n,j,k} u_{n,j,k} \right\|^2 \leq B \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} c_{n,j,k}^2 \quad (1.5)$$

для всех квадратично суммируемых последовательностей $\{c_{n,j,k}\}$.

Иногда свойство (1.5) называют «почти ортогональностью» (см. [7]). Функции $u_{n,j,k}$ по аналогии с терминологией работ [5] и [11] назовем «вейглетами». Справедливо следующее утверждение, являющееся полным аналогом леммы 1.1 из работы [8].

Лемма 1.1 Пусть существуют такие константы $A > 0$, $a > 1/2$ и $b > 3/2$, что

$$|\hat{\psi}(\omega)| \leq A |\omega|^a (1 + |\omega|^2)^{-(b+a)/2} \quad (1.6)$$

для всех $\omega \in \mathbb{P}$, тогда последовательность $\{u_{n,j,k}\}$ образует устойчивый базис в $\mathbf{L}^2(\mathbb{P})$.

Доказательство. Требуется показать, что существуют константы $0 < \tilde{A} \leq \tilde{B} < \infty$ такие, что для произвольной функции $f \in \mathbf{L}^2(\mathbb{P})$ выполнено

$$\tilde{A} \|f\|^2 \leq \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} |\langle f, u_{n,j,k} \rangle|^2 \leq \tilde{B} \|f\|^2 \quad (1.7)$$

(это соотношение эквивалентно условию (1.5)).

Определим оператор обратного проецирования, действующий на функцию $g(\varphi, s)$:

$$R^* g(x, y) = \int_0^{2\pi} g(\varphi, x \cos \varphi + y \sin \varphi) d\varphi. \quad (1.8)$$

Оператор R^* является сопряженным к оператору R (см. [13]).
Обозначим

$$u_{n,j,k}^*(x, y) = \beta_{n,j,k} R^*[Z_n \psi_{j,k}](x, y).$$

Таким же образом, как в доказательстве леммы 1.1 из работы [8], можно убедиться, что (1.7) эквивалентно существованию таких положительных констант C_1 и C_2 , что

$$\sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} |\langle f, u_{n,j,k} \rangle|^2 \leq C_1 \|f\|^2 \quad (1.9)$$

и

$$\sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} |\langle f, u_{n,j,k}^* \rangle|^2 \leq C_2 \|f\|^2. \quad (1.10)$$

Найдем константы $\beta_{n,j,k}$. Используя теорему о центральном сечении (1.2), имеем

$$\begin{aligned} \beta_{n,j,k}^2 &= \left\| R^{-1}[Z_n \psi_{j,k}] \right\|^2 = \\ &= \int_{-\infty-\infty}^{\infty} \int_{-\infty}^{\infty} \left| R^{-1}[Z_n \psi_{j,k}](x, y) \right|^2 dx dy = \int_{-\infty-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \hat{R}^{-1}[Z_n \psi_{j,k}](\omega_1, \omega_2) \right|^2 d\omega_1 d\omega_2 = \\ &= \frac{1}{2} \int_0^{2\pi} \int_{-\infty}^{\infty} \left| \hat{R}^{-1}[Z_n \psi_{j,k}](\omega \cos \varphi, \omega \sin \varphi) \right|^2 |\omega| d\omega d\varphi = \frac{1}{4\pi} \int_{-\infty}^{\infty} |\hat{\psi}_{j,k}(\omega)|^2 |\omega| d\omega = \\ &= \frac{2^j}{4\pi} \int_{-\infty}^{\infty} |\hat{\psi}(\omega)|^2 |\omega| d\omega = 2^j \|\xi\|^2, \end{aligned}$$

где через ξ обозначена функция, для которой

$$\hat{\xi}(\omega) = \frac{\hat{\psi}(\omega) |\omega|^{1/2}}{2\sqrt{\pi}}. \quad (1.11)$$

Таким образом, константы $\beta_{n,j,k}$ не зависят от n и k и пропорциональны $2^{j/2}$. Докажем теперь (1.9). Имеем

$$\begin{aligned}
U(f) &= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} |\langle f, u_{n,j,k} \rangle|^2 = \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \overline{u_{n,j,k}(x,y)} dx dy \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\omega_1, \omega_2) \overline{\hat{u}_{n,j,k}(\omega_1, \omega_2)} d\omega_1 d\omega_2 \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \int_0^{2\pi} \int_0^{2\pi} \hat{f}(\omega \cos \varphi, \omega \sin \varphi) \overline{\hat{u}_{n,j,k}(\omega \cos \varphi, \omega \sin \varphi)} \omega d\omega d\varphi \right|^2 = \\
&= \frac{1}{(2\pi)^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \frac{1}{2} \int_0^{2\pi} \int_{-\infty}^{\infty} \hat{R}f(\varphi, \omega) \overline{\hat{R}u_{n,j,k}(\varphi, \omega)} |\omega| d\omega d\varphi \right|^2 = \\
&= \frac{1}{(2\pi)^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \frac{1}{4\beta_{n,j,k}^2} \left| \int_0^{2\pi} \int_{-\infty}^{\infty} \hat{R}f(\varphi, \omega) Z_n(\varphi) \overline{\hat{\psi}_{j,k}(\omega)} |\omega| d\omega d\varphi \right|^2 = \\
&= \frac{1}{(2\pi)^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \frac{1}{4\beta_{n,j,k}^2} \left| \int_{-\infty}^{\infty} \hat{R}f_n(\omega) \overline{\hat{\psi}_{j,k}(\omega)} |\omega| d\omega \right|^2 = \\
&= \frac{1}{(2\pi)^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \frac{1}{4\beta_{n,j,k}^2} \left| \int_{-\infty}^{\infty} \hat{R}f_n(\omega) 2^{-j/2} \overline{\hat{\psi}(2^{-j}\omega)} e^{i\omega k 2^{-j}} |\omega| d\omega \right|^2 = \\
&= \frac{1}{2\|\xi\|^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \frac{1}{\pi 2^{j+1}} \left| \int_{-\infty}^{\infty} \hat{R}f_n(\omega) \overline{\hat{\xi}(2^{-j}\omega)} |\omega|^{1/2} e^{i\omega k 2^{-j}} d\omega \right|^2 = \\
&= \frac{1}{2\|\xi\|^2} \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \frac{1}{\pi 2^{j+1}} \left| \int_0^{\pi 2^{j+1}} e^{i\omega k 2^{-j}} \sum_{m \in \mathbb{Z}} \hat{R}f_n(\omega + \pi 2^{j+1} m) |\omega + \pi 2^{j+1} m|^{1/2} \overline{\hat{\xi}(2^{-j}\omega + 2\pi m)} d\omega \right|^2.
\end{aligned}$$

Функция

$$\sum_{m \in \mathbb{Z}} \hat{R}f_n(\omega + \pi 2^{j+1} m) |\omega + \pi 2^{j+1} m|^{1/2} \overline{\hat{\xi}(2^{-j}\omega + 2\pi m)}$$

периодична с периодом $\pi 2^{j+1}$. Следовательно,

$$\begin{aligned}
U(f) &= \frac{1}{2\|\xi\|^2} \sum_{n=0}^{\infty} \sum_{j \in \mathbb{Z}} \int_0^{\pi 2^{j+1}} \left| \sum_{m \in \mathbb{Z}} \hat{R}f_n(\omega + \pi 2^{j+1} m) |\omega + \pi 2^{j+1} m|^{1/2} \overline{\hat{\xi}(2^{-j}\omega + 2\pi m)} \right|^2 d\omega = \\
&= \frac{1}{2\|\xi\|^2} \sum_{n=0}^{\infty} \sum_{j \in \mathbb{Z}} \int_0^{\pi 2^{j+1}} \sum_{m \in \mathbb{Z}} \hat{R}f_n(\omega + \pi 2^{j+1} m) |\omega + \pi 2^{j+1} m|^{1/2} \overline{\hat{\xi}(2^{-j}\omega + 2\pi m)} \times
\end{aligned}$$

$$\begin{aligned}
& \times \sum_{l \in \mathbb{Z}} \overline{\hat{R}f_n(\omega + \pi 2^{j+1} l)} \left| \omega + \pi 2^{j+1} l \right|^{1/2} \hat{\xi}(2^{-j} \omega + 2\pi l) d\omega = \\
& = \frac{1}{2 \|\xi\|^2} \sum_{n=0}^{\infty} \sum_{j, m \in \mathbb{Z}} \int \hat{R}f_n(\omega) |\omega|^{1/2} \overline{\hat{R}f_n(\omega + \pi 2^{j+1} m)} \left| \omega + \pi 2^{j+1} m \right|^{1/2} \overline{\hat{\xi}(2^{-j} \omega)} \hat{\xi}(2^{-j} \omega + 2\pi m) d\omega.
\end{aligned}$$

Далее, поступая, как в доказательстве леммы 1.1 из работы [8], получаем, что при выполнении условия (1.6) для некоторой положительной константы \tilde{C}_1 выполнено

$$U(f) \leq \sum_{n=0}^{\infty} \tilde{C}_1 \|g_n\|^2,$$

где $g_n(\omega) = \hat{R}f_n(\omega) |\omega|^{1/2}$. Обозначим $g(\varphi, \omega) = \hat{R}f(\varphi, \omega) |\omega|^{1/2}$. По теореме о центральном сечении

$$\begin{aligned}
\|g\|^2 &= \int_0^{2\pi} \int_{-\infty}^{\infty} |g(\varphi, \omega)|^2 d\omega d\varphi = \int_0^{2\pi} \int_{-\infty}^{\infty} |\hat{R}f(\varphi, \omega)|^2 |\omega| d\omega d\varphi = \\
&= 4\pi \int_0^{2\pi} \int_0^{\infty} |\hat{f}(\omega \cos \varphi, \omega \sin \varphi)|^2 \omega d\omega d\varphi = 4\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy = 4\pi \|\hat{f}\|^2 = 4\pi \|f\|^2.
\end{aligned}$$

Кроме того,

$$\begin{aligned}
\|g\|^2 &= \int_0^{2\pi} \int_{-\infty}^{\infty} |\omega| \sum_{l=0}^{\infty} \hat{R}f_l(\omega) Z_l(\varphi) \overline{\sum_{n=0}^{\infty} \hat{R}f_n(\omega) Z_n(\varphi)} d\omega d\varphi = \\
&= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} |\hat{R}f_n(\omega)|^2 |\omega| d\omega = \sum_{n=0}^{\infty} \|g_n\|^2.
\end{aligned}$$

Следовательно, (1.9) выполнено. Докажем теперь (1.10).

$$\begin{aligned}
U^*(f) &= \sum_{n=0}^{\infty} \sum_{j, k \in \mathbb{Z}} |\langle f, u_{n, j, k}^* \rangle|^2 = \sum_{n=0}^{\infty} \sum_{j, k} |\langle f, \beta_{n, j, k} R^*[Z_n \psi_{j, k}] \rangle|^2 = \sum_{n=0}^{\infty} \sum_{j, k \in \mathbb{Z}} |\langle Rf, \beta_{n, j, k} Z_n \psi_{j, k} \rangle|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j, k \in \mathbb{Z}} \left| \|\xi\|^2 2^{j/2} \int_0^{2\pi} \int_{-\infty}^{\infty} Rf(\varphi, s) \overline{Z_n(\varphi) \psi_{j, k}(s)} ds d\varphi \right|^2 =
\end{aligned}$$

$$\begin{aligned}
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \|\xi\| 2^{j/2} \int_0^{2\pi} \int_{-\infty}^{\infty} \hat{R}f(\varphi, \omega) Z_n(\varphi) \overline{\hat{\psi}_{j,k}(\omega)} d\omega d\varphi \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \|\xi\| 2^{j/2} \int_0^{2\pi} \int_{-\infty}^{\infty} \hat{R}f(\varphi, \omega) |\omega|^{1/2} Z_n(\varphi) \overline{\hat{\psi}_{j,k}(\omega)} |\omega|^{-1/2} d\omega d\varphi \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \|\xi\| 2^{j/2} \int_{-\infty}^{\infty} \hat{R}f_n(\omega) |\omega|^{1/2} \overline{\hat{\psi}_{j,k}(\omega)} |\omega|^{-1/2} d\omega \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \|\xi\| 2^{j/2} \int_{-\infty}^{\infty} \hat{R}f_n(\omega) |\omega|^{1/2} 2^{-j/2} \overline{\hat{\psi}(2^{-j}\omega)} e^{i\omega k 2^{-j}} |\omega|^{-1/2} d\omega \right|^2 = \\
&= \sum_{n=0}^{\infty} \sum_{j,k \in \mathbb{Z}} \left| \|\xi\| 2^{-j/2} \int_{-\infty}^{\infty} \hat{R}f_n(\omega) \overline{\hat{\eta}(2^{-j}\omega)} |\omega|^{1/2} e^{i\omega k 2^{-j}} d\omega \right|^2.
\end{aligned}$$

Здесь через η обозначена функция, для которой $\hat{\eta}(\omega) = \hat{\psi}(\omega) |\omega|^{-1/2}$. Далее, поступая, как при доказательстве (1.9), получаем, что найдется положительная константа C_2 такая, что

$$U^*(f) \leq C_2 \|f\|^2.$$

Лемма доказана.

1.2 Пороговая обработка и статистическая оценка среднеквадратичного риска

Далее будем предполагать, что функция f , описывающая изображения, имеет компактный носитель и равномерно регулярна по Липшицу с некоторым показателем $\gamma > 0$. Оказывается, что при этом функция Rf равномерно регулярна по Липшицу с показателем $\gamma + 1/2$ (см. [13] и [14]). Также будем считать, что вейвлет-функция ψ удовлетворяет условиям теоремы 6.3 из [10] и леммы 1.1. Таким образом, найдется такая константа $A > 0$, что

$$\langle Rf, Z_n \psi_{j,k} \rangle \leq \frac{A}{2^{j(\gamma+1)}} \quad (1.12)$$

при всех неотрицательных $j \in \mathbb{Z}$ и всех $k \in \mathbb{Z}$.

При практической реализации разложения (1.3) дискретные коэффициенты получаются умножением матрицы значений функции Rf на ортогональную матрицу F и ортогональную матрицу W , определяемые базисом Фурье $\{Z_n\}$ и вейвлет-базисом $\{\psi_{j,k}\}$. Причем, поскольку в проекциях присутствует шум, рассматривается следующая модель измеряемых данных:

$$X_{u,v} = (Rf)_{u,v} + \varepsilon_{u,v}, \quad u = 1, \dots, 2^J, \quad v = 1, \dots, 2^J, \quad (1.13)$$

где J - некоторое положительное число (“разрешение” изображения), $X_{u,v}$ - наблюдаемые данные, R - преобразование Радона, f - истинная (незашумленная) функция изображения, а $\varepsilon_{u,v}$ - независимые случайные погрешности, которые имеют одинаковое нормальное распределение с нулевым средним и дисперсией σ^2 . Таким образом, в силу ортогональности матриц преобразования, дискретные коэффициенты разложения наблюдаемых данных (1.13), которые мы обозначим через $Y_{n,j,k}$, описываются следующей моделью:

$$Y_{n,j,k} = \mu_{n,j,k} + \varepsilon_{n,j,k}^{FW}, \quad n = 0, \dots, 2^J - 1, \quad j = 0, \dots, J - 1, \quad k = 0, \dots, 2^j - 1, \quad (1.14)$$

где $\mu_{n,j,k} = 2^j \langle Rf, Z_n \psi_{j,k} \rangle$, а случайные погрешности $\varepsilon_{n,j,k}^{FW}$ также независимы и имеют такое же распределение, как и $\varepsilon_{u,v}$.

Поскольку в измерениях присутствует шум, необходимо использовать некоторые процедуры для его удаления. Мы рассмотрим процедуру пороговой обработки (см. [10]). Смысл пороговой обработки вейвлет-коэффициентов измеряемого сигнала заключается в удалении достаточно маленьких коэффициентов, которые считаются шумом. Будем использовать так называемую мягкую пороговую обработку с порогом T_j , зависящим от масштаба j . Для подавления шума к каждому коэффициенту $Y_{n,j,k}$ применяется функция $\rho_{T_j}(x)$, определяемая соотношением

$$\rho_{T_j}(x) = \text{sgn}(x)(|x| - T_j).$$

При такой пороговой обработке коэффициенты, которые по модулю меньше порога T_j , обнуляются, а абсолютные величины остальных коэффициентов уменьшаются на величину порога. В результате функция сигнала f в разложении (1.4) (точнее, ее нормированная версия) оценивается следующим образом:

$$\overset{\circ}{f} = \sum_{n=0}^{2^J-1} \sum_{j=0}^{J-1} \sum_{k=0}^{2^{j-1}} \beta_{n,j,k} \rho_{T_j}(Y_{n,j,k}) u_{n,j,k}.$$

Среднеквадратичная ошибка (риск) оценки $\overset{\circ}{f}$ определяется как

$$r_\sigma = E \left\| f - \overset{\circ}{f} \right\|^2.$$

Так как система функций $\{u_{n,j,k}\}$ не является ортонормированной, риск нельзя представить в виде суммы вкладов погрешностей отдельных коэффициентов, однако в силу устойчивости базиса $\{u_{n,j,k}\}$ найдутся такие константы A и B , зависящие от ψ , что $Ar_J \leq r_\sigma \leq Br_J$, где

$$r_J = r_J(f, \sigma, \mathbf{T}) = \sum_{n=0}^{2^J-1} \sum_{j=0}^{J-1} \sum_{k=0}^{2^{j-1}} \beta_{n,j,k}^2 E(\mu_{n,j,k} - \rho_{T_j}(Y_{n,j,k}))^2, \quad (1.15)$$

(здесь $\mathbf{T} = (T_0, \dots, T_{J-1})$ - вектор порогов). Таким образом, с точностью до множителя величина r_σ асимптотически ведет себя так же, как r_J , поэтому в дальнейшем мы будем рассматривать асимптотическое поведение r_J .

В выражении (1.15) присутствуют неизвестные величины $\mu_{n,j,k}$, поэтому вычислить значение r_J нельзя. Однако его можно оценить. В каждом слагаемом если $|Y_{n,j,k}| > T_j$, то вклад этого слагаемого в риск составляет $\beta_{n,j,k}^2 (\sigma^2 + T_j^2)$, а если $|Y_{n,j,k}| \leq T_j$, то вклад составляет $\beta_{n,j,k}^2 \mu_{n,j,k}^2$.

Поскольку $EY_{n,j,k}^2 = \sigma^2 + \mu_{n,j,k}^2$, значение $\mu_{n,j,k}^2$ можно оценить разностью $Y_{n,j,k}^2 - \sigma^2$.

Таким образом, в качестве оценки риска можно использовать следующую величину:

$$\hat{r}_J(f, \sigma, \mathbf{T}) = \sum_{n=0}^{2^J-1} \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \beta_{n,j,k}^2 F[Y_{n,j,k}^2, \sigma, T_j],$$

$$F[x, \sigma, T_j] = (x - \sigma^2) \mathbb{1}_{\{|x| \leq T_j^2\}} + (\sigma^2 + T_j^2) \mathbb{1}_{\{|x| > T_j^2\}}, \quad (1.16)$$

причем эта оценка является несмещенной (см. [10]).

1.3 Предельные теоремы для оценки риска

По аналогии с работами [6], [8] и [9] мы рассмотрим ситуацию, когда для каждого масштаба j выбирается “универсальный” порог. Такой порог определяется соотношением $T_{U,j} = \sigma \sqrt{2 \ln 2^{j+J}}$, поскольку на каждом масштабе количество коэффициентов равно 2^{j+J} , и получил название «универсальный», так как он не зависит от наблюдаемых данных и при выборе этого порога риск r_j близок к минимальному (см. [12]).

Для оценки риска (1.16) справедлива следующая теорема, являющаяся аналогом теоремы 1.1 из [8].

Теорема 1.1 Пусть вейвлет-функция ψ удовлетворяет условиям теоремы 6.3 из [10] и леммы 1.1, а функция $f \in \mathbf{L}^2(\mathbb{P})$ имеет компактный носитель и равномерно регулярна по Липшицу с показателем $\gamma > 0$, тогда

$$\mathbb{P} \left(\frac{\hat{r}_J(f, \sigma, \mathbf{T}_U) - r_J(f, \sigma, \mathbf{T}_U)}{D_J} < x \right) \Rightarrow \Phi(x) \text{ при } J \rightarrow \infty, \quad (1.17)$$

где $D_J = \sigma^2 \|\xi\|^2 \sqrt{2/7} 2^{2J}$, функция ξ задана соотношением (1.11), а $\mathbf{T}_U = (T_{U,0}, \dots, T_{U,J-1})$.

Зачастую дисперсия шума σ^2 не известна и ее также необходимо оценивать с помощью некоторой оценки $\hat{\sigma}^2$. При этом выражение (??) принимает вид

$$\hat{r}_J(f, \hat{\sigma}, \mathbf{T}) = \sum_{n=0}^{2^J-1} \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \beta_{n,j,k}^2 F[Y_{n,j,k}^2, \hat{\sigma}, T_j], \quad (1.18)$$

и вместо порогов $T_{U,j}$ используются пороги $\hat{T}_{U,j} = \hat{\sigma} \sqrt{2 \ln 2^{j+J}}$ (обозначим вектор этих порогов через $\hat{\mathbf{T}}_U$).

Обычно дисперсия σ^2 оценивается по выборке сигнала (или изображения), однако ее можно оценить и по независимой выборке. Для этого следует произвести измерение пустого сигнала, тогда наблюдения будут представлять из себя чистый шум, по которому и оценивается σ^2 . Далее мы рассмотрим оба случая.

В случае, когда $\hat{\sigma}^2$ строится по независимой выборке, для оценки риска (1.18) справедливо следующее утверждение.

Теорема 1.2 Пусть вейвлет-функция ψ удовлетворяет условиям теоремы 6.3 из [10] и леммы 1.1, а функция $f \in \mathbf{L}^2(\mathbb{P})$ имеет компактный носитель и равномерно регулярна по Липшицу с показателем $\gamma > 0$. Пусть $\hat{\sigma}^2$ - не зависящая от $Y_{n,j,k}$ оценка дисперсии σ^2 , для которой выполнено

$$\mathbb{P}(2^J (\hat{\sigma}^2 - \sigma^2) < x) \Rightarrow \Phi_{\Sigma}(x) \text{ при } J \rightarrow \infty, \quad (1.19)$$

где $\Phi_{\Sigma}(x)$ - функция распределения нормального закона с нулевым средним и дисперсией Σ^2 . Тогда

$$\mathbb{P}\left(\frac{\hat{r}_J(f, \hat{\sigma}, \hat{\mathbf{T}}_U) - r_J(f, \sigma, \mathbf{T}_U)}{D_J} < x\right) \Rightarrow \Phi_{\Upsilon}(x) \text{ при } J \rightarrow \infty, \quad (1.20)$$

где $\Phi_{\Upsilon}(x)$ - функция распределения нормального закона с нулевым средним и дисперсией

$$\Upsilon^2 = 1 + \frac{7}{18} \frac{\Sigma^2}{\sigma^4}.$$

Если дисперсия оценивается по коэффициентам разложения $Y_{n,j,k}$ и функция f удовлетворяет требуемым условиям регулярности, то в силу (1.12) ее оценивают по половине всех коэффициентов на уровне $j = J - 1$, так как эти коэффициенты фактически содержат только шум. Мы рассмотрим популярные оценки неизвестной дисперсии: $\hat{\sigma}_S^2$ - выборочная дисперсия, $\hat{\sigma}_R$ - интерквартильный размах и $\hat{\sigma}_M$ - абсолютное медианное отклонение от медианы. Эти оценки определяются следующими соотношениями:

$$\hat{\sigma}_S^2 = \frac{1}{2^{2J-1}} \sum_{n=0}^{2^J-1} \sum_{k=0}^{2^{J-1}-1} Y_{n,J-1,k}^2 - \left(\frac{1}{2^{2J-1}} \sum_{n=0}^{2^J-1} \sum_{k=0}^{2^{J-1}-1} Y_{n,J-1,k} \right)^2, \quad (1.21)$$

$$\hat{\sigma}_R = \frac{Y_{2^{J-1},(3/4)} - Y_{2^{J-1},(1/4)}}{2\zeta_{3/4}}, \quad (1.22)$$

$$\hat{\sigma}_M = \frac{\frac{\text{med}_{0 \leq n \leq 2^J-1, 0 \leq k \leq 2^{J-1}-1} |Y_{n,J-1,k} - \text{med}_{0 \leq n \leq 2^J-1, 0 \leq l \leq 2^{J-1}-1} Y_{n,J-1,l}|}{\zeta_{3/4}}}{\zeta_{3/4}}, \quad (1.23)$$

где $Y_{2^{J-1},(1/4)}$ и $Y_{2^{J-1},(3/4)}$ - выборочные квантили порядка 1/4 и 3/4, построенные по набору коэффициентов $Y_{n,J-1,k}$, $0 \leq n \leq 2^J - 1$, $0 \leq k \leq 2^{J-1} - 1$, $\zeta_{3/4}$ - теоретическая квантиль порядка 3/4 стандартного нормального распределения ($\zeta_{3/4} \approx 0.6745$), а «*med*» обозначает выборочную медиану.

Теорема 1.3 Пусть вейвлет-функция ψ удовлетворяет условиям теоремы 6.3 из [10] и леммы 1.1, а функция $f \in \mathbf{L}^2(\mathbb{P})$ имеет компактный носитель и равномерно регулярна по Липшицу с показателем $\gamma > 1/2$. Пусть оценка дисперсии $\hat{\sigma}^2$ определяется выражением (1.21). Тогда

$$\mathbb{P} \left(\frac{\hat{r}_J(f, \hat{\sigma}, \hat{\mathbf{T}}_U) - r_J(f, \sigma, \mathbf{T}_U)}{D_J} < x \right) \Rightarrow \Phi_\Upsilon(x) \text{ при } J \rightarrow \infty, \quad (1.24)$$

где $\Phi_\Upsilon(x)$ - функция распределения нормального закона с нулевым средним и дисперсией $\Upsilon^2 = 2/9$.

Теорема 1.4 Пусть вейвлет-функция ψ удовлетворяет условиям теоремы 6.3 из [10] и леммы 1.1, а функция $f \in \mathbf{L}^2(\mathbf{P})$ имеет компактный носитель и равномерно регулярна по Липшицу с показателем $\gamma > 1$. Пусть оценка дисперсии $\hat{\sigma}^2$ определяется выражением (1.22) или (1.23). Тогда

$$\mathbf{P}\left(\frac{\hat{r}_J(f, \hat{\sigma}, \hat{\mathbf{T}}_U) - r_J(f, \sigma, \mathbf{T}_U)}{D_J} < x\right) \Rightarrow \Phi_\Upsilon(x) \text{ при } J \rightarrow \infty, \quad (1.25)$$

где $\Phi_\Upsilon(x)$ - функция распределения нормального закона с нулевым средним и дисперсией

$$\Upsilon^2 = \frac{7}{36\zeta_{3/4}^2 (\phi(\zeta_{3/4}))^2} - \frac{4}{3}.$$

Доказательство теорем 1-4 полностью аналогично доказательству соответствующих теорем из работ [8] и [9].

Замечание 1.2 Вектор порогов \mathbf{T}_U фактически используется для построения оценки функции Rf . В работе [11] вместо вектора порогов \mathbf{T}_U рассматривается единый для всех масштабов порог (в данном случае он равен $T_R = \sigma\sqrt{4\ln 2^{2^J}}$), который также назван «универсальным», и показывается, что при построении оценки функции f этот порог является асимптотически близким к оптимальному в минимаксном смысле. При рассмотрении порога T_R все представленные результаты останутся неизменными.

2 Разработка методов обработки и анализа изображений с помощью разреженных представлений

Методы разреженных представлений применимы во многих задачах обработки изображений – сжатии, подавлении шума, ретушировании и др. При этом изображение, закодированное в виде вектор-столбца \mathbf{y} из n элементов, представляется в виде линейной комбинации нескольких изображений, называемых атомами, из определенного набора изображений – словаря $\mathbf{A} \in \mathbb{R}^{n \times m}$, $n < m$: $\mathbf{y} \approx \mathbf{A}\mathbf{x}$. Представление можно считать разреженным, если число ненулевых элементов вектора \mathbf{x} гораздо меньше числа атомов (столбцов матрицы \mathbf{A}); при этом необходимо учитывать величину ошибки представления $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$.

Будем обозначать $\|\mathbf{x}\|_0$ число ненулевых элементов в \mathbf{x} , а $\text{supp } \mathbf{x}$ – множество индексов ненулевых элементов \mathbf{x} (носитель вектора \mathbf{x}), таким образом, $\|\mathbf{x}\|_0 = |\text{supp } \mathbf{x}|$.

2.1 Построение разреженных представлений

Математическая постановка задачи построения разреженного представления данного изображения \mathbf{y} с данным словарем \mathbf{A} возможна в двух вариантах:

а) с ограничением количества k_0 атомов, используемых для построения представления:

$$(P_0): \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2, \quad \|\mathbf{x}\|_0 = k_0;$$

б) с ограничением ошибки представления ε :

$$(P_0): \min_{\mathbf{x}} \|\mathbf{x}\|_0, \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \varepsilon.$$

Поиск точного решения данных задач на практике невозможен для типичных размерностей данных, встречающихся в реальных задачах обработки изображений, из-за крайне большой вычислительной сложности – при фиксированном значении мощности носителя k_0 сложность поиска

оптимального (в смысле ошибки представления) решения экспоненциально зависит от k_0 .

Существует множество разнообразных методов построения приближенно оптимального разреженного представления, среди которых – так называемые жадные методы. Базовый алгоритм OMP (orthogonal matching pursuit), использующийся в этих методах, основан на итеративном расширении носителя решения. На каждой итерации выбирается элемент, добавление которого в носитель позволяет максимально уменьшить ошибку представления.

Введем обозначения: k – номер итерации, \mathbf{x}^k – решение, \mathbf{r}^k – невязка, $S^k = \text{supp } \mathbf{x}^k$ – носитель решения, \mathbf{a}_i – i -й столбец матрицы \mathbf{A} . Тогда алгоритм OMP можно записать следующим образом:

1) положим $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{r}^0 = \mathbf{y} - \mathbf{A}\mathbf{x}^0 = \mathbf{y}$, $S^0 = \text{supp } \mathbf{x}^0 = \emptyset$;

2) на каждой итерации увеличиваем k на 1 и выполняем следующие шаги:

$\forall i \notin S^{k-1}$ вычисляем минимальную норму невязки, которую можно достичь, добавляя в носитель решения i -й элемент:

$$\varepsilon(i) = \min_{z_i} \|\mathbf{a}_i z_i - \mathbf{r}^{k-1}\|_2^2 = \|\mathbf{r}^{k-1}\|_2^2 - \frac{(\mathbf{a}_i^T \mathbf{r}^{k-1})^2}{\|\mathbf{a}_i\|_2^2} \quad \left(z_i = \frac{\mathbf{a}_i^T \mathbf{r}^{k-1}}{\|\mathbf{a}_i\|_2^2} \right)$$

добавляем элемент, позволяющий достичь минимальную норму невязки:

$$j: \quad \forall i \notin S^{k-1} \quad \varepsilon(j) \leq \varepsilon(i);$$

$$S^k = S^{k-1} \cup \{j\}$$

вычисляем новое решение:

$$\begin{aligned} \mathbf{x}^k &= \arg \min_{\mathbf{x}, \text{supp } \mathbf{x} = S^k} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2; \\ \mathbf{A}_{S^k}^T (\mathbf{A}_{S^k} \mathbf{x}_{S^k} - \mathbf{y}) &= 0 \quad (= -\mathbf{A}_{S^k}^T \mathbf{r}^k); \\ \mathbf{x}_{S^k} &= (\mathbf{A}_{S^k}^T \mathbf{A}_{S^k})^{-1} \mathbf{A}_{S^k}^T \mathbf{y}, \end{aligned}$$

где A_S обозначает матрицу, состоящую из столбцов матрицы A с номерами, входящими в множество S , а x_S обозначает вектор из элементов вектора x с номерами из этого множества.

проверяем критерии останова:

$k = k_0$, если задана максимальная мощность носителя

$\|y - Ax^k\|_2 \leq \varepsilon$, если задана максимальная ошибка представления.

В данной работе применялся вариант вышеописанного алгоритма, позволяющий в некоторых случаях получить более точное решение. Его отличие от базового алгоритма в том, что для выбора наилучшего элемента для добавления при вычислении остаточных ошибок $\varepsilon(i)$ представления варьируются значения всех элементов решения, входящих в носитель.

$$\hat{S}_i^k = S^{k-1} \cup \{i\},$$

$$\varepsilon(i) = \min_{x_{\hat{S}_i^k}, z_i} \left\| \begin{bmatrix} \mathbf{A}_{\hat{S}_i^k} & \mathbf{a}_i \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\hat{S}_i^k} \\ z_i \end{bmatrix} - \mathbf{b} \right\|_2^2.$$

2.2 Построение словаря

Для применения разреженных представлений крайне важно выбрать способ построения словарей, используемых в решении задачи. Можно выделить два подхода к выбору словарей:

- использование априорно заданных словарей, соответствующих таким преобразованиям, как вейвлетное, дискретное косинусное, и др.;
- построение словаря с помощью машинного обучения.

В первом случае часто возможно получение разреженного представления более быстрыми, чем стандартные, алгоритмами, но в большинстве случаев для реальных изображений представления, получаемые с такими словарями, не могут быть сильно разреженными (т.е. мощность носителя представления сравнима с размерностью изображения). Используемый же нами второй подход применим только, если

обрабатываемые изображения принадлежат к достаточно узкому классу. Требуется наличие достаточного количества изображений данного класса – обучающая выборка. При этом словарь, построенный с помощью машинного обучения, позволяет получать довольно точные представления, включающие относительно малое количество атомов.

Для построения словарей в данной работе применялся алгоритм K-SVD, позволяющий построить словарь заданного размера, наилучшим образом описывающий заданный набор изображений в расчете на заданную мощность носителя представлений для изображений из набора.

Обозначим изображения обучающей выборки $(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_d) = \mathbf{Y} \in \mathbb{R}^{n \times d}$. Требуется построить матрицу $\mathbf{A} \in \mathbb{R}^{n \times m}$ – словарь, и представления изображений обучающей выборки $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d) = \mathbf{X} \in \mathbb{R}^{m \times d}$.

Формальные требования можно записать следующим образом:

$$\min_{\mathbf{A}, \mathbf{X}} \sum_{i=1}^d \|\mathbf{y}_i - \mathbf{A}\mathbf{x}_i\|_2, \quad \|\mathbf{x}_i\|_0 \leq k_0,$$

если ограничена мощность носителей представлений \mathbf{y}_i , или

$$\min_{\mathbf{A}, \mathbf{X}} \sum_{i=1}^d \|\mathbf{x}_i\|_0, \quad \|\mathbf{y}_i - \mathbf{A}\mathbf{x}_i\|_2 \leq \varepsilon,$$

если ограничена допустимая ошибка представления для \mathbf{y}_i . Как и в случае с задачей нахождения разреженного представления по заданному словарю, поставленные выше задачи на практике нерешаемы из-за слишком большой вычислительной сложности при характерных размерностях данных.

K-SVD – эвристический алгоритм, основанный на итеративном улучшении словаря. Изначально словарь задается произвольным образом (например, в качестве атомов берется часть изображений из обучающей выборки), и по нему каким-либо алгоритмом вычисляются начальные представления \mathbf{x}_i для \mathbf{y}_i . Под улучшением словаря понимается минимизация суммы квадратов ошибок представлений изображений обучающей выборки:

$$f = \sum_{i=1}^d \|\mathbf{y}_i - \mathbf{A}\mathbf{x}_i\|_2^2 = \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2.$$

Для этого на каждой итерации последовательно выбирается оптимальное значение для каждого из атомов словаря одновременно с соответствующими элементами представлений, использующих этот атом. Рассмотрим, как происходит обновление значений для j_0 -го атома. Перепишем минимизируемую функцию в виде

$$f = \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 = \left\| \mathbf{Y} - \sum_{j=1}^m \mathbf{a}_j \mathbf{x}_j \right\|_F^2 = \left\| \left(\mathbf{Y} - \sum_{j \neq j_0} \mathbf{a}_j \mathbf{x}_j \right) - \mathbf{a}_{j_0} \mathbf{x}_{j_0} \right\|_F^2.$$

Здесь и далее \mathbf{x}_{j_0} обозначает не разреженное представление для \mathbf{y}_{j_0} , а строку из j_0 -ых элементов разреженных представлений всех изображений обучающей выборки. Задачу минимизации можно теперь рассматривать, как задачу поиска наилучшего приближения для матрицы $\mathbf{E}_{j_0} = \mathbf{Y} - \sum_{j \neq j_0} \mathbf{a}_j \mathbf{x}_j$ матрицей ранга 1. При этом, варьируя \mathbf{x}_{j_0} , необходимо сохранять не только количество нулей в нем, но и их расположение. Для этого зададим матрицу проекции \mathbf{P}_{j_0} , такую, что вектор-строка $\mathbf{x}_{j_0} \mathbf{P}_{j_0}$ содержит только ненулевые элементы \mathbf{x}_{j_0} , и обозначим $\mathbf{x}_{j_0}^R = \mathbf{x}_{j_0} \mathbf{P}_{j_0}$, $\mathbf{E}_{j_0}^R = \mathbf{E}_{j_0} \mathbf{P}_{j_0}$. Наилучшее приближение матрицей ранга 1 можно найти, вычислив сингулярное разложение:

$$\mathbf{E}_{j_0}^R = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

и использовав наибольшее сингулярное число:

$$\mathbf{a}_{j_0} = \mathbf{u}_1, \quad \mathbf{x}_{j_0} = \sigma_1 \mathbf{v}_1 \mathbf{P}_{j_0}^T.$$

Вместо ресурсоемкого вычисления сингулярного разложения иногда этот шаг заменяют на несколько итераций поочередной оптимизации по $\mathbf{x}_{j_0}^R$ и \mathbf{a}_{j_0} с помощью метода наименьших квадратов:

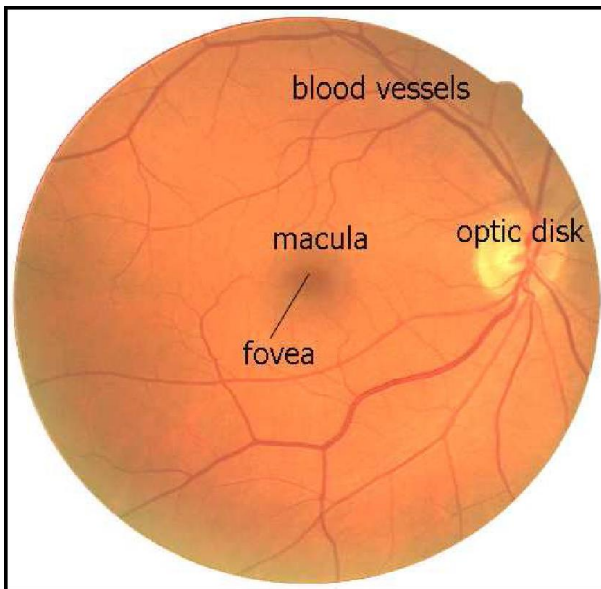
$$\min_{\mathbf{x}_{j_0}^R} \left\| \mathbf{E}_{j_0}^R - \mathbf{a}_{j_0} \mathbf{x}_{j_0}^R \right\|_F^2 \Rightarrow \mathbf{x}_{j_0}^R = \frac{\mathbf{a}_{j_0}^T \mathbf{E}_{j_0}^R}{\left\| \mathbf{a}_{j_0} \right\|_2},$$

$$\min_{\mathbf{a}_{j_0}} \left\| \mathbf{E}_{j_0}^R - \mathbf{a}_{j_0} \mathbf{x}_{j_0}^R \right\|_F^2 \Rightarrow \mathbf{a}_{j_0} = \frac{\mathbf{E}_{j_0}^R (\mathbf{x}_{j_0}^R)^T}{\left\| \mathbf{x}_{j_0}^R \right\|_2}.$$

Критерий остановки алгоритма K-SVD – изменение минимизируемой функции становится меньше заданного порога.

2.3 Применение разреженных представлений для задачи обнаружения хориоретинальной атрофии

В данной работе методы разреженного представления применялись для решения задачи выявления областей хориоретинальной атрофии (ХРА) сетчатки на изображениях человеческого глазного дна. При автоматическом анализе фотографий сетчатки информация об областях ХРА позволяет точнее определять границу диска зрительного нерва. Это важно, так как характеристики диска зрительного нерва (ДЗН) составляют основные критерии, по которым диагностируется глаукома, и, кроме того, размеры ДЗН часто служат опорой при определении размеров других важных составляющих сетчатки. Также, выраженность ХРА или ее отсутствие позволяют делать более точные предположения о наличии у пациента специфических болезней, таких как диабетическая ретинопатия и глаукома.



а) Цветное изображение здорового глазного дна.
 Optic disk – диск зрительного нерва, также называемый слепым пятном. Macula (macula lutea) – желтое пятно. Fovea – центральная ямка макулы, содержащая только фоторецепторы.



б) Пример атрофированного участка сетчатки. Зеленым показана граница области атрофии, черным – граница диска зрительного нерва.

Рисунок 2.1 – Пример изображения глазного дна и атрофированной области сетчатки.

Наиболее распространен в литературе подход для детектирования областей атрофии, основанный на классификации по текстурным признакам. Изображение разбивается на блоки небольшого размера, для каждого из которых вычисляется некоторое количество признаков (чаще всего эти признаки - характеристики матрицы смежности), после чего к полученным данным применяется какой-либо классификатор. Можно либо говорить на основании результатов о наличии или отсутствии в целом заметной патологии, либо использовать результаты классификации отдельных блоков в дальнейшем анализе изображения. В первом случае, хотя и достижима точность классификации порядка 80%, но практическая ценность такого результата крайне мала. Во втором же случае достаточно надежный алгоритм позволял бы получить упомянутые выше преимущества от детектирования

ХРА, но точность классификации отдельных блоков слишком мала (порядка 60%). При такой низкой точности использование результатов детектирования в других этапах анализа изображения может привести к ухудшению их результатов, поэтому на практике в системах автоматического анализа изображений глазного дна детектирование областей ХРА обычно не производится.

Данная работа нацелена на создание более точного метода детектирования областей ХРА с помощью составного классификатора, опирающегося не только на значения текстурных признаков, и не только на характеристики разреженных представлений анализируемых блоков. Такой подход может дать значительно более высокую точность классификации, чем у каждого из двух классификаторов по отдельности.

2.3.1 Классификатор на основе разреженных представлений

Основная идея классификации изображений с помощью разреженных представлений заключается в построении представлений изображения по словарям, соответствующим различным классам. Можно считать, что изображение принадлежит к данному классу, если при использовании соответствующего этому классу словаря ошибка представления получается заметно меньше, чем для других классов. Величины ошибки представления изображения по словарям разных классов можно интерпретировать, как показатели вероятности принадлежности этого изображения к тому или иному классу следующим образом:

$$p_i = \left(1 - \frac{e_i}{\sum e_k} \right) / (n - 1),$$

где n - количество классов, e_i - ошибка представления изображения по словарю i -го класса, p_i - вероятность того, что изображение принадлежит к i -му классу.

Для обучения классификаторов была использована база из 96 изображений, на которых специалист-офтальмолог вручную отметил области

ХРА, для тестирования - база из 130 изображений. Все изображения были разбиты на блоки 8x8 пикселей со всеми возможными перекрытиями. Были отобраны блоки, более чем на половину принадлежащие областям атрофии, для построения словаря класса А. Для построения словаря класса Б были отобраны блоки, целиком лежащие внутри области видимости камеры и не имеющие пересечений с областями атрофии. Для усиления значимости величины ошибки представления как показателя принадлежности к классу были вычислены усредненные изображения всех блоков, принадлежащих к каждому из классов. Эти усредненные изображения были вычтены из всех блоков, входящих в обучающие выборки для соответствующих классов. Соответственно, в дальнейшем перед нахождением разреженного представления для какого-либо блока из него вычитался соответствующий усредненный блок.

Для каждого класса методом K-SVD были построены словари из 100, 200 и 400 атомов со значениями параметра k_0 3, 4, 6 и 8. Для полученных 12 пар словарей была оценена точность классификации с использованием разреженных представлений. Результаты показали, что оптимальные параметры составляют: размер словаря – 200 атомов, $k_0 = 4$; точность классификации при этом составила 59,3%.

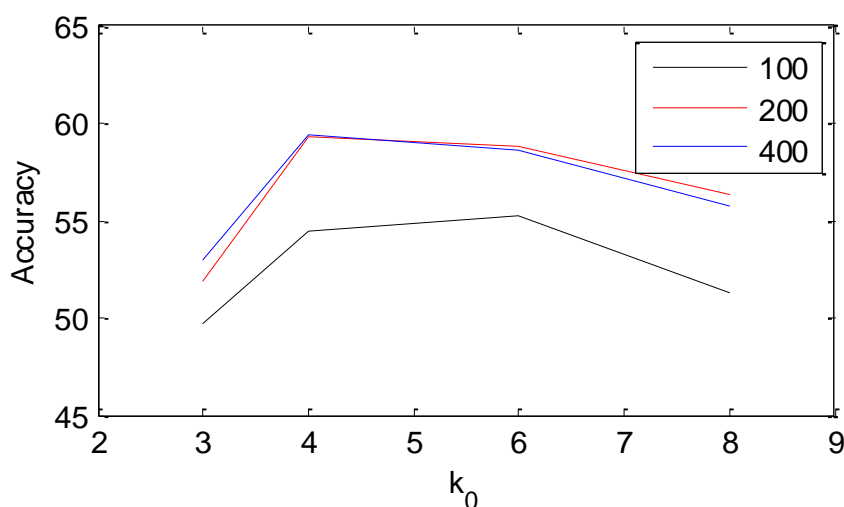


Рисунок 2.2 – Результаты оценки точности классификации для словарей размером в 100, 200 и 400 атомов в зависимости от параметра k_0 .

Увеличение размера словаря практически не дает прироста точности, а увеличение мощности носителя представлений может привести к уменьшению точности из-за того, что классы А и Б не слишком сильно отличаются (при достаточно большом k_0 для блоков обучающей выборки класса А возможно достаточно точное представление по словарю класса Б, и наоборот).

2.3.2 Классификатор на основе текстурных признаков

Текстурные признаки также вычислялись для всех блоков 8x8 пикселей со всеми перекрытиями; критерии принадлежности блоков к классам при обучении были те же, что описаны в предыдущем разделе.

Для каждого блока вычислялись матрицы смежности уровней яркости $C_{1,1}$, $C_{1,-1}$, $C_{-1,1}$, $C_{-1,-1}$ для синего канала I изображения с квантованием на 32 уровня по формуле:

$$C_{dx,dy}(i, j) = |\{(x, y) : I(x, y) = i, I(x + dx, y + dy) = j\}|.$$

Для каждой из четырех матриц вычислялись следующие значения: контраст, коэффициент корреляции, дисперсия, обратный момент разности, энтропия и неоднородность. Далее для коэффициента корреляции, дисперсии, обратного момента разности и неоднородности брались максимальные значения по четырем матрицам смежности, а для контраста и энтропии – минимальные. Таким образом, каждому блоку приписывались шесть признаков, по которым блоки классифицировались с помощью линейного дискриминантного анализа.

Такой набор признаков был выбран с помощью поэтапного исключения малозначащих и нерелевантных признаков. Изначально характеристики для красного и зеленого каналов тоже учитывались, и, кроме шести перечисленных свойств матриц смежности, рассматривался также второй угловой момент. В первую очередь была проверена гипотеза, что уменьшение числа анализируемых каналов не сильно ухудшит качество классификации, так как при такой редукции сильнее всего уменьшается

объем вычислений. Проверка производилась с использованием наибольших, наименьших и усредненных значений каждого из признаков по четырем матрицам. При сравнении четырех групп признаков: RGB (63 признака), R, G и B (по 21 признаку для каждого из каналов по отдельности) были получены следующие значения точности классификации: 65,6% (RGB), 63,2% (R), 65,1% (G), 65,4% (B). Поэтому решено было использовать признаки только синего канала изображения. Дальнейший переход к шести наиболее значащим признакам при увеличении скорости работы в 2,4 раза привел к падению точности классификации всего на 0,5%. Таким образом, точность классификатора на основе текстурных признаков составила 64,9%.

2.3.3 Фильтрация результатов классификации

Описанные в предыдущих разделах классификаторы оперируют перекрывающимися блоками размером 8x8 пикселей. Чтобы получить область, классифицированную как атрофическую, необходимо

- для каждого пикселя учесть результаты классификации всех блоков, включающих данный пиксель (их может быть меньше, чем 64, для пикселей, находящихся рядом с границей поля зрения камеры),
- и совместить результаты классификации для соседних пикселей (атрофированный регион должен быть сплошным и с гладкими границами; он не может, например, состоять из пикселей, расположенных в шахматном порядке).

Таким образом, для каждого пикселя имеется несколько ответов классификатора, обозначающих правдоподобность того, что соответствующие блоки изображения принадлежат области атрофии. Среднее значение этих ответов приписывалось каждому пикселю, затем полученная «карта вероятности атрофии» сглаживалась гауссовским фильтром с $\sigma = 5$. Результат бинаризовался с порогом, равным 0,5, и подвергался морфологическому закрытию со структурным элементом в виде

круга, радиус которого равен $1/50$ диаметра поля зрения камеры. Полученное бинарное изображение и считалось маской области атрофии – итоговым результатом работы классификатора. Именно эти результаты использовались для оценки качества работы классификаторов; несколько таких результатов приведено ниже в качестве примеров.

2.3.4 Составной классификатор

Для совместного использования двух базовых классификаторов была использована несимметричная схема, в которой результат работы классификатора, основанного на разреженных представлениях – ошибки представления блока по двум словарям – использовался в качестве пары дополнительных параметров для классификатора, основанного на текстурных признаках. Таким образом, для блоков, имеющих значительные ошибки представления по обоим словарям, соотношение этих ошибок автоматически учитывалось с малым весом, и наоборот, при ненадежности ответа, полученного по текстурным признакам, результат классификации по ошибкам разреженных представлений имел больший вес. Такой подход позволил достичь точности классификации в 69,7%. Примеры результатов, полученных с помощью описанного алгоритма, приведены на рисунке 2.3.

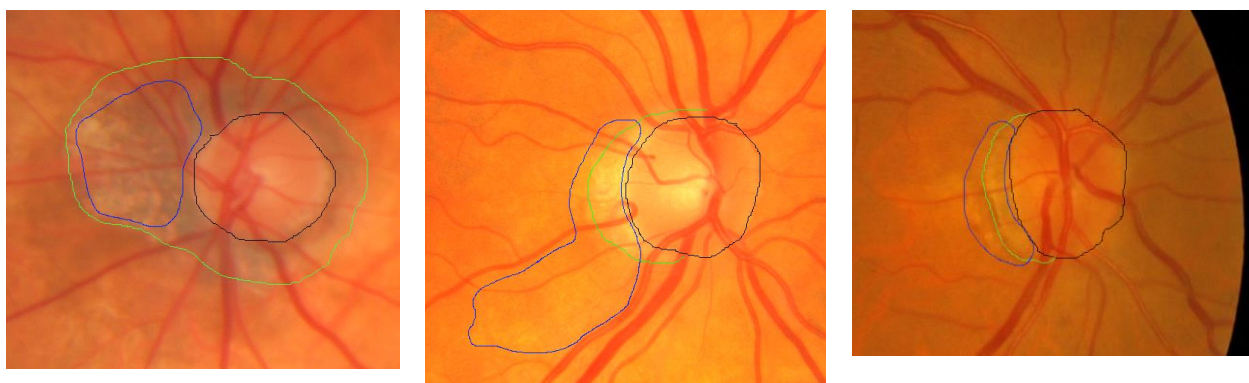


Рисунок 2.3 – Примеры результатов классификации. Синим цветом обозначены границы детектированных областей атрофии, зеленым – эталонные границы областей атрофии, черным – границы диска зрительного нерва. Перед фильтрацией результатов классификации исключались пиксели, принадлежащие диску зрительного нерва.

3 Создание методов обработки и анализа изображений на основе систем ортогональных многочленов.

3.1 Анализ изображений с помощью детектирования локальных особенностей

Извлечение ключевых точек и их параметризация - одна из основных низкоуровневых задач обработки изображения. Например, это - начальный шаг в задачах сопоставления стереоизображений [15], распознавания объектов [16], индексация видео и изображений [17], склейки панорам и других. Существует большое количество подходов к задаче выделения ключевых точек, например, детектор уголков Харриса [18], детектор пятен на основе разности гауссиан (DoG), предложенный Lowe [19], подход, основанный на круговых гармонических функциях [20, 21], и другие. Задача построения дескриптора ключевой точки также широко представлена в литературе [19, 20, 22]. В построения дескриптора основным требованием является инвариантность к классу определенных проективных и фотометрических преобразований. Это требование крайне важно для того чтобы получить высокий процент правильных соответствий ключевых точек на изображениях, полученных с разных ракурсов и при различных условиях освещения.

В этом разделе рассматривается многомасштабный подход к выделению ключевых точек и построению дескрипторов на основе круговых гармонических функций Гаусса-Лагерра (КГФ ГЛ) [20] и применяется к проблеме поиска МРТ изображений мозга. Дескрипторы на основе КГФ ГЛ имеют несколько преимуществ по сравнению с другими дескрипторами, особенно для рассматриваемой задачи. Например, они не содержат основанную на гистограмме часть, как, например, в дескрипторе SIFT, которая делает дескриптор инвариантным к небольшим локальным сдвигам пикселей на описываемой дескриптором окрестности (патче). В определенных прикладных задачах, такая инвариантность нежелательна,

поскольку известно, что геометрические изменения внутри патчей практически отсутствуют. Как показано в работах [23, 24], дескрипторы без применения гистограммы в некоторых случаях превосходят SIFT, который на данный момент является де-факто текущим стандартом качества.

Поскольку обнаружение ключевых точек и построение дескрипторов зачастую являются основными частями более сложных систем, их вычислительная эффективность также очень важна. Поэтому в мы также рассматриваем ускорение расчета дескрипторов Гаусса-Лагерра с помощью 2-мерных функций Эрмита.

Другие преимущества и свойства дескрипторов Гаусса-Лагерра рассматриваются в разделе 3.2. Применение дескрипторов и постановка основной задачи приводится в разделе 3.3. В разделе 3.4 описан предложенный алгоритм для поиска МРТ мозга. В разделе 3.5 приводится алгоритм предварительной классификации, призванный улучшить общую эффективность работы системы. Наконец, результаты представлены в разделе **Ошибка! Источник ссылки не найден.**

3.2 Ключевые точки и дескрипторы на основе круговых гармонических функций Гаусса-Лагерра (КГФ ГЛ)

3.2.1 Обнаружение ключевых точек

Рассмотрим семейство ортонормированных и сепарабельных в полярных координатах функций:

$$\Psi_n^\alpha(r, \gamma; \sigma) = \psi_n^{|\alpha|}(r^2 / \sigma) e^{i\alpha\gamma}.$$

Их радиальные профили - функции Лагерра:

$$\psi_n^\alpha(x) = \frac{1}{\sqrt{n! \Gamma(n + \alpha + 1)}} x^{\alpha/2} e^{-x/2} L_n^\alpha(x),$$

где $n = 0, 1, \dots; \alpha = 0, \pm 1, \pm 2, \dots$ и $L_n^\alpha(x)$ полиномы Лагерра:

$$L_n^\alpha(x) = (-1)^n x^{-\alpha} e^x \frac{d}{dx^n} (x^{n+\alpha} e^{-x}).$$

Функции Лагерра $\psi_n^\alpha(x)$ может быть вычислены, используя следующие рекуррентные соотношения:

$$\begin{aligned} \psi_{n+1}^\alpha(x) &= \frac{(x - \alpha - 2n - 1)}{\sqrt{(n+1)(n+\alpha+1)}} \psi_n^\alpha(x) - \\ &- \sqrt{\frac{n(n+\alpha)}{(n+1)(n+\alpha+1)}} \psi_{n-1}^\alpha(x), \quad n = 0, 1, \dots, \\ \psi_0^\alpha(x) &= \frac{1}{\sqrt{\Gamma(\alpha+1)}} x^{\alpha/2} e^{-x/2}, \quad \psi_{-1}^\alpha(x) \equiv 0. \end{aligned}$$

Применение рекуррентных формул позволяет избежать проблем потери точности при умножении экспонент на степенную функцию.

Эти функции Ψ_n^α , называются круговыми гармоническими функциями Гаусса-Лагерра (КГФ ГЛ) и имеют 2 индекса, определяющих радиальный порядок n и угловой порядок α . Действительные части Ψ_n^α ($n = 0, 1, \dots, 4$; $\alpha = 1, 2, \dots, 5$) приведены на рис. 3.1.

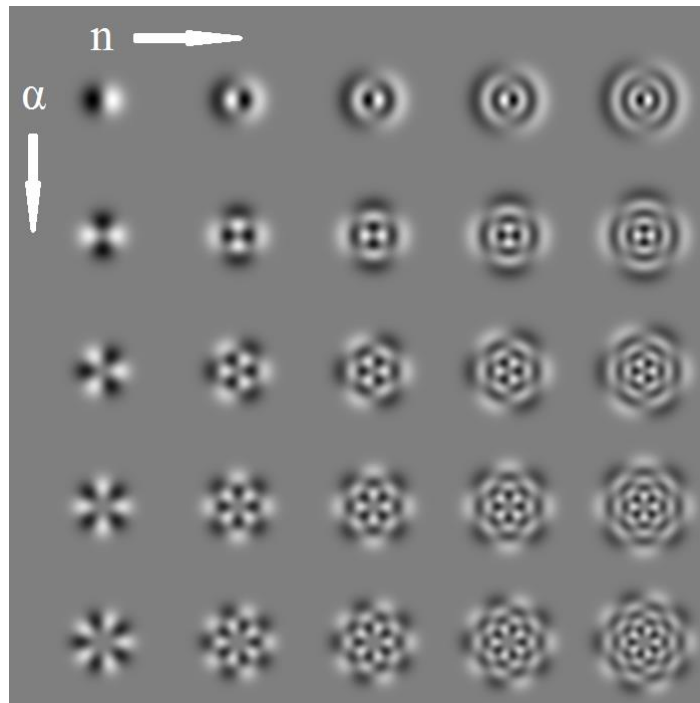


Рисунок 3.1 – Действительная часть Ψ_n^α ($n = 0, 1, \dots, 4$; $\alpha = 1, 2, \dots, 5$).

Функции КГФ ГЛ являются самоуправляемыми, то есть они могут быть повернуты на угол θ путем умножения на $e^{i\alpha\theta}$. Они также сохраняют свою

форму при применении преобразования Фурье. Также, они подходят для многомасштабного и многокомпонентного анализа изображения [20, 25].

Рассмотрим данное изображение $I(x, y)$ определенное на действительной плоскости R^2 . Вследствие ортогональности системы функций Ψ_n^α , изображение $I(x, y)$ может быть представлено в окрестности точки x_0, y_0 для любого фиксированного масштаба σ в декартовой системе координат следующим образом:

$$I(x_0, y_0) = \sum_{\alpha=-\infty}^{\infty} \sum_{n=0}^{\infty} g_{\alpha,n}(x_0, y_0; \sigma) \Psi_n^\alpha(\rho, \omega; \sigma),$$

где

$$g_{\alpha,n}(x_0, y_0; \sigma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) \overline{\Psi_n^\alpha(\rho, \omega; \sigma)} dx dy,$$

и

$$\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

$$\omega = \arctg\left(\frac{y - y_0}{x - x_0}\right).$$

Рассмотрим кратко алгоритм обнаружения ключевых точек, предложенный в [20]. Пусть σ будет коэффициентом масштаба и $\sigma \in [2^{-s_{\max}}, 2^{s_{\max}}]$ дискретизовано в $(2s_{\max} + 1)$ октавы, каждая октава в свою очередь содержит N_s масштабов с равным шагом. Таким образом, набор весов определяется как $\{\sigma_j\}$, где $j = 0, 1, \dots, 2N_s(2s_{\max} + 1) - 1$. Принимая во внимание свойства функций КГФ ГЛ, оказывается, что при значениях параметра $n = 0, \alpha = 3, 4$ они являются детекторами особенностей типа «развилка» и «крест». Набор $2N_s(2s_{\max} + 1)$ откликов детектора определяется как:

$$S(x, y; \sigma) = |g_{3,0}(x, y; \sigma)|^2 + |g_{4,0}(x, y; \sigma)|^2, \quad \sigma \in \{\sigma_j\}$$

и называется скалограммой. Скалограмма просматривается скользящим 3D окном (5x5x3 пикселей). Кандидаты ключевых точек $\bar{K} = (\bar{x}, \bar{y}, \bar{\sigma})$

определяются как локальные максимумы скалограммы в пределах окна. Здесь (\bar{x}, \bar{y}) координата ключевой точки и $\bar{\sigma}$ соответствующий масштаб ключевой точки. Таким образом, получается набор ключевых точек изображения $\{\bar{K}\}$. Этот набор прореживается, выбрасывая те ключевые точки \bar{K} у которых одни и те же координаты (\bar{x}, \bar{y}) , но на разных (два и более) масштабах. И, наконец, отбрасываются ключевые точки \bar{K} , в которых отклик детектора $S(\bar{x}, \bar{y}; \bar{\sigma})$ по величине меньше, чем выбранный порог:

$$S(\bar{x}, \bar{y}; \bar{\sigma}) < T \cdot \max_{x,y} (S(x, y; \bar{\sigma}))$$

3.2.2 Дескрипторы на основе КГФ ГЛ

Здесь мы кратко рассматриваем алгоритм построения дескрипторов ключевых точек, предложенный в [6]. Любая ключевая точка $\bar{K} = (\bar{x}, \bar{y}, \bar{\sigma})$ связана с локальным дескриптором $\bar{\chi} = \{\bar{\chi}(n, \alpha, j)\}$. В данном случае дескриптор – это комплекснозначный вектор, состоявший из локальной проекции изображения на систему функций КГФ ГЛ Ψ_n^α на $2j_{\max}$ масштабах, соседних с масштабом $\bar{\sigma}$ ключевой точки \bar{K} . Элементы вектора $\bar{\chi}$ определяются как:

$$\bar{\chi}(n, \alpha, j) = A_{norm} \cdot g_{\alpha,n}(x, y; \sigma_j) \cdot e^{-i\alpha\theta_j},$$

$$n = 0, \dots, n_{\max}, \quad \alpha = 1, \dots, \alpha_{\max}, \quad j = -j_{\max}, \dots, j_{\max}.$$

где σ_j – j -й масштаб последующий за $\bar{\sigma}$ если $j > 0$, или предшествующий $\bar{\sigma}$ если $j < 0$. A_{norm} – коэффициент нормализации, который делает дескриптор инвариантным к изменению освещенности изображения. Сдвиг фазы $e^{-i\alpha\theta_j}$ используется для того, чтобы сделать дескриптор инвариантным к ориентации ключевой точки, где

$$\theta_j = \arg(g_{1,0}(\bar{x}, \bar{y}; \sigma_j)).$$

Эффективность сопоставления ключевых точек данным методом была продемонстрирована в [20] и сравнена с алгоритмом SIFT. Результаты сопоставления методом КГФ ГЛ превосходят результаты SIFT, особенно в

случае вращения, масштабирования и сдвига изображений. Однако, при этом вычислительная сложность алгоритма КГФ ГЛ выше.

Оценим число операций, требуемых для вычисления дескриптора для одной ключевой точки. Предположим, что ширина и высота области изображения, соответствующей ключевой точке $\bar{K} = (\bar{x}, \bar{y}, \sigma_j)$ в справочном масштабе σ_j равна r_j . Тогда, порядок количества операций для вычисления элементов $\bar{\chi}$ для одной ключевой точки равен $r_j^2 j_{\max} n_{\max} \alpha_{\max}$.

Чтобы уменьшить сложность вычисления, предлагается новый метод вычисления дескрипторов ключевых точек, основанный на разложении по 2-мерным функциям Эрмита.

3.2.3 Вычисление дескрипторов с помощью разложения по 2D функциям Эрмита

2-мерные функции Эрмита $\Phi_{m,n}(x, y; \sigma)$ образуют полную ортонормированную систему в пространстве L_2 , и может быть представлена в виде:

$$\Phi_{m,n}(x, y; \sigma) = \frac{1}{\sigma} \varphi_m\left(\frac{x}{\sigma}\right) \varphi_n\left(\frac{y}{\sigma}\right),$$

$$\varphi_n(x) = \frac{(-1)^n e^{-\frac{x^2}{2}}}{\sqrt{2^n n! \sqrt{\pi}}} \cdot H_n(x).$$

где $n = 0, 1, 2, \dots$ и $H_n(x)$ многочлены Эрмита:

$$H_n(x) = (-1)^n e^{x^2} \frac{d}{dx^n} (e^{-x^2}).$$

Функции Эрмита $\varphi_n(x)$ могут быть вычислены с помощью следующих рекуррентных соотношений:

$$\varphi_n(x) = x \sqrt{\frac{2}{n}} \varphi_{n-1}(x) - \sqrt{\frac{n-1}{n}} \varphi_{n-2}(x), \quad n = 2, 3, \dots,$$

$$\varphi_0(x) = \frac{1}{\sqrt[4]{\pi}} e^{-x^2/2}, \quad \varphi_1(x) = \frac{\sqrt{2}x}{\sqrt[4]{\pi}} e^{-x^2/2}.$$

2-мерное разложение Эрмита изображения $I(x, y)$ в окрестности точки x_0, y_0 для любого фиксированного σ может быть определено как:

$$I(x_0, y_0) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} h_{m,n}(x_0, y_0; \sigma) \Phi_{m,n}(x, y; \sigma)$$

Где

$$h_{m,n}(x_0, y_0; \sigma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x_0 + x, y_0 + y) \Phi_{m,n}(x, y; \sigma) dx dy.$$

Поскольку из определения 2-мерных функций Эрмита $\Phi_{m,n}(x, y; \sigma)$ можно видеть, что они сепарабельны в декартовой системе координат x, y — независимый, таким образом, вычисление $h_{m,n}$ может быть выполнено следующим образом:

$$\bar{h}_{m,n}(x_0, y; \sigma) = \int_{-\infty}^{\infty} I(x_0 + x, y) \varphi_m(x/\sigma) dx,$$

для каждого фиксированного y и после этого

$$h_{m,n}(x_0, y_0; \sigma) = \frac{1}{\sigma} \int_{-\infty}^{\infty} \bar{h}_{m,n}(x_0, y_0 + y) \varphi_n(y/\sigma) dy$$

Коэффициенты 2-мерных разложений в ряд по функциям Эрмита и КГФ ГЛ поочередно связаны между собой линейно [26, 27]. Для заданного $m \geq 0$ функции КГФ ГЛ линейно связаны с 2-мерными функциями Эрмита как:

$$\Psi_{m-l,l}(x, y; \sigma) = \sum_{h=0}^m T_{l,h}^{(m)} \Phi_{m-h,h}(x, y; \sigma),$$

для $l \geq m-l$, где

$$T_{l,h}^{(m)} = i^h \sqrt{\frac{(m-h)!h!}{2^m(m-l)!}} \sum_{q=\max(0,h-l)}^{\min(h,m-l)} (-1)^q C_{m-l-q}^{m-l} C_{l-h+q}^l.$$

Для $l < m-l$ уравнения получены, принимая во внимание, что

$$\Psi_{m-l,l}(x, y; \sigma) = \overline{\Psi_{m-l,l}(x, y; \sigma)}.$$

$m-l, l$ индексы функций КГФ ГЛ $\Psi_{m-l,l}(x, y; \sigma)$ связаны с радиальным порядком n и угловым порядком α следующим образом:

$$m-l = \max(n, n-\alpha), \quad l = \max(n, n+\alpha),$$

$$n = \min(m-l, l), \quad \alpha = l - (m-l).$$

Как пример, соотношение между функциями КГФ ГЛ и 2-мерными функциями Эрмита для $m=3$ иллюстрировано на рис. 3.2.

Принимая во внимание эти формулы, любой коэффициент разложения по функциям КГФ ГЛ $g_{\alpha,n}(x, y; \sigma)$ может быть вычислен как линейная комбинация коэффициентов разложения в ряд по функциям Эрмита $h_{m,n}(x, y; \sigma)$.

$\Phi_{3,0}$	$\Phi_{2,1}$	$\Phi_{1,2}$	$\Phi_{0,3}$	Re	Im	
						$\Psi_{0,-3}$
$\frac{1}{2\sqrt{2}}$	$-\frac{3i}{2\sqrt{2}}$	$-\frac{3}{2\sqrt{2}}$	$\frac{i}{2\sqrt{2}}$			$\Psi_{1,-1}$
$\frac{3}{2\sqrt{2}}$	$-\frac{i}{2\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$	$-\frac{3}{2\sqrt{2}}$			$\Psi_{1,1}$
$\frac{3}{2\sqrt{2}}$	$\frac{i}{2\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$	$\frac{3}{2\sqrt{2}}$			$\Psi_{0,3}$
$\frac{1}{2\sqrt{2}}$	$\frac{3i}{2\sqrt{2}}$	$-\frac{3}{2\sqrt{2}}$	$\frac{i}{2\sqrt{2}}$			

Рисунок 3.2 – Пример отношения между КГФ ГЛ и 2-мерными функциями Эрмита.

Вследствие сепарабельности 2-мерных функций Эрмита $\Phi_{m,n}(x, y; \sigma)$ число операций, требуемых для вычисления дескриптора ключевой точки $\bar{\chi}$ уменьшается до порядка $r_j^2 j_{\max} (2n_{\max} + \alpha_{\max})$.

3.3 Применение дескрипторов в задаче поиска МРТ изображений

В этом разделе мы рассмотрим применение дескрипторов КГФ к поиску медицинских изображений, в частности на поиске МРТ изображений мозга для обнаружения болезни Альцгеймера (БА). Актуальность этой задачи упоминается в работе [28]. Число медицинских изображений быстро растет в связи с появлением новых методов медицинских исследований, а также широкого распространения оборудования для проведения обследований, и ручной анализ такого огромного объема данных очень трудоемок. С другой стороны, как отмечено в [29], эффективность работы докторов увеличивается при использовании поисковых систем по изображениям, основанных на содержании. Таким образом, оба факта подталкивают к созданию автоматизированной системы, которая помогает доктору извлечь только релевантную информацию из базы данных изображений.

Типичная ситуация состоит в том, что радиологи и врачи, диагностирующие болезни мозга с помощью МРТ, заинтересованы в наличии под рукой МРТ изображений похожих случаев. Это позволяет вынести более обоснованное решения о состоянии здоровья и диагнозе.

В [28], авторы сосредотачиваются на объединении различных видов информации, включая текстовые данные, визуальные особенности, извлеченные из МРТ, а так же прямой ввод пользователя (доктора) о похожести изображений.

В этой работе мы рассматриваем поиск изображений, основанный только на визуальных особенностях. Эта задача очень сложна в общей постановке, но мы сосредотачиваемся только на изображениях пациентов с распространенным заболеванием, которое вызывает характерные изменения в мозговой структуре. В случае болезни Альцгеймера наиболее частое и ярко выраженное изменение в структуре мозга – сокращение объема мозговой ткани в области гиппокампа и в височных долях [30]. Таким образом, мы будем использовать визуальные особенности, построенные только в этих областях, чтобы подчеркнуть различие или подобие субъектов относительно

наличия болезни Альцгеймера. Примеры здорового и больного мозга с маской нанесенного гиппокампа приведены на рис. 3.3.

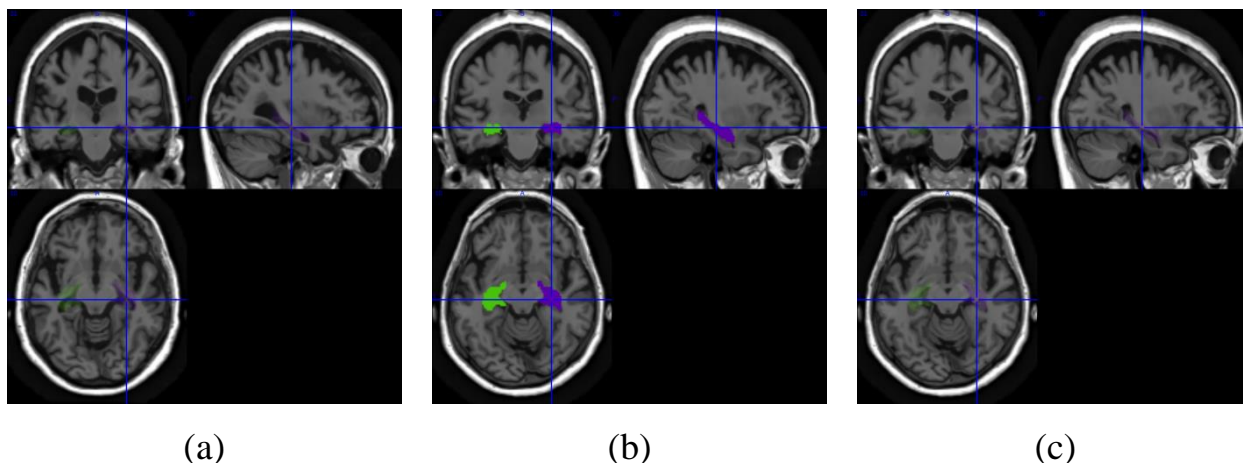


Рисунок 3.3 – (а) мозг с БА; (b) маска области гиппокампа нанесенная на изображение мозга; (с) пример здорового мозга.

Можно заметить, что в больном мозге количество спинномозговой жидкости (на изображении ей соответствует черный цвет) больше.

В то время как различие здорового и пациента с БА на данном примере могло выглядеть легким, необходимо помнить о большом естественном анатомическом различии в разных мозгах, а также необходимости обнаруживать так называемое Умеренное Когнитивное Нарушение (УКН), которое считается продромальной стадией БА.

3.4 Схема алгоритма

Опишем детально схему предложенного алгоритма. Он состоит из трех главных шагов: геометрическая нормализация МРТ изображения, извлечение визуальных особенностей и, наконец, генерирование сигнатуры изображения, на основе извлеченных особенностей.

3.4.1 Геометрическая нормализация изображения

Первая стадия визуального выделения признаков – выравнивание МРТ изображения и выделение объёма мозга на МРТ изображении. Этот шаг требуется, потому что различия в немозговой ткани влияют на визуальные особенности и ухудшают работу системы. Кроме того, заболевания мозга

часто проявляются лишь в определенных отделах головного мозга, таким образом, требуется приведение изображения к мозговому атласу, для извлечения особенностей областях, соответствующих болезни.

Необходимость процедуры приведения мозга в стандартную систему координат не вызывает сомнений, однако остается вопрос какой тип преобразования использовать, линейный или нелинейный. Линейное преобразование, такое как аффинное или евклидово, позволяет осуществить лишь грубое совмещение МРТ изображения, устраняя только сдвиг, поворот изображения и изменение масштаба. Небольшие анатомические структуры не будут точно приведены линейным преобразованием к шаблону мозга из-за естественных анатомических различий разных людей. Нелинейное (нежесткое) совмещение, изменяющее геометрию локально, позволяет произвести более точное совмещение мелких структур мозга, однако, при этом трудно гарантировать, что изображения не «притянуты» к шаблону сверх меры, то есть с нарушением анатомических особенностей. Эта проблема становится особенно остро в случае, если мы пытаемся совместить мозг с анатомическими отклонениями вследствие некоторой болезни к шаблону мозга, построенному по здоровым пациентам. Более подробный анализ этой проблемы представлен в [30]. Авторы отмечают несколько недостатков подхода Voxel-Based Morphometry (VBM), который основан на нелинейном совмещении.

В то время как цель статьи [31] отличается от рассматриваемой задачи, мы можем почерпнуть из нее некоторые полезные выводы. Во-первых, нежесткое совмещение нежелательно для подходов, основанных на выделении локальных особенностей, поскольку она искажает сами особенности, в то время как их общий вид необходимо сохранить. Во-вторых, для упрощения процесса мы можем анализировать МРТ изображения послойно, поскольку в результате линейного совмещения срезы мозга будут достаточно хорошо соответствовать друг другу. Подобная техника послойного анализа использовалась в работе [28].

Подводя итог по этапу выравнивания МРТ изображений мозга, в нашем алгоритме, применяется только аффинное совмещение всех МРТ к мозговому шаблону MNI152. Для выполнения данной операции используется набор инструментов VBM8, находящийся в свободном доступе [32] (шаблон MNI включен в комплект). Результатом данного этапа являются выровненные в одном пространстве МРТ изображения, в которых можно приблизительно выделить отделы мозга по атласу головного мозга.

3.4.2 Выделение признаков

Следующий шаг после предварительной обработки МРТ изображений мозга – вычисление локальных дескрипторов изображения. Как это было уже отмечено, мы должны рассчитать только те дескрипторы, которые содержат визуальную информацию, связанную с присутствием или отсутствием болезни Альцгеймера. Поэтому, сначала необходимо определить местонахождение областей интереса, то есть гиппокампа и височных долей в объеме мозга. Для этого был использован анатомический атлас отделов головного мозга, поставляемый в Anatomical Automatic Labeling toolbox (http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html) [33].

По мозговому атласу отмечаются маски интересующих областей МРТ изображения, и в них производится расчет дескрипторов. Поскольку было использовано линейное геометрическое совмещение с шаблоном мозга, которое не обеспечивает точное совмещение отделов мозга, дескрипторы вычисляются таким образом, чтобы покрыть маски областей. В этом случае соответствующие мозговые части будут захвачены даже в случае неточных масок.

Для выбора характерного размера области одного дескриптора размер и их местоположения возможно использование различных стратегий. Первая стратегия состоит в том, чтобы сначала применить детектором ключевых точек в интересующих областях и затем вычислить дескриптор для обнаруженных ключевых точек, используя их характерный масштаб и

положение. Например, возможно использование детектора и дескриптора SIFT [19], или детектора КГФ [20] и соответствующего дескриптора. Эта стратегия осуществления выборки дескрипторов проиллюстрирована на рис. 3.4а, б. Кругами обозначены характерные размеры и положение извлеченных ключевых точек (то есть область вычисления индивидуальных дескрипторов), черным обозначена маска области интереса. На рис. 3.4а, б представлена одна и та же область мозга, но для различных пациентов, с чем и связано различие в размерах и местоположении ключевых точек.

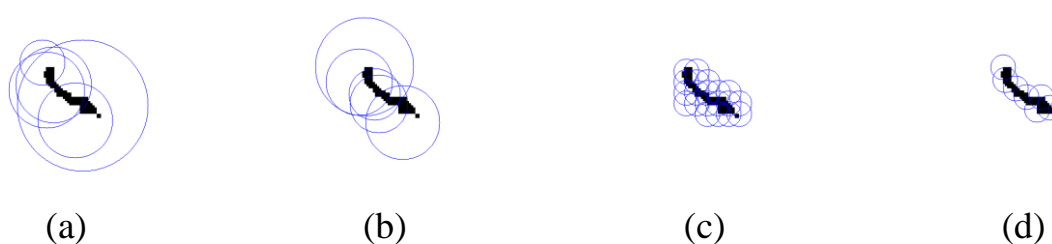


Рисунок 3.4. Примеры осуществления выборки стратегий дескрипторов, покрывающих область интереса. Области вычисления индивидуальных дескрипторов обозначены окружностями, область интереса отмечена черным. (а, б) выбор, используя детектор ключевых точек; (с) плотное поколение дескрипторами на регулярной сетке; (д) плотное покрытие дескрипторами на нерегулярной сетке.

Еще две стратегии плотного размещения дескрипторов представлены на рис. 3.4с, д. На рис. 3.4с сначала были сгенерированы центры областей на регулярной сетке и выбран фиксированный размер области построения дескриптора. Затем из всех областей отобраны только области, перекрывающие маску. На рис. 3.4д проиллюстрирована более «компактная» альтернатива. Области вычисления дескрипторов отобраны, путем построчного (попиксельного) сканирования изображения с маской области. Центр дескриптора помещается в первый попавшийся пиксель маски, и соответствующая часть маски, которая захвачена произведенной областью, исключается из дальнейшего рассмотрения. Процесс продолжается пока вся маска не будет покрыта.

В наших экспериментах, которые описаны в разделе 3.6, мы проверили все стратегии, но было обнаружено, что выбор конкретной стратегия осуществления выборки дескрипторов имеет небольшое влияние на качество поиска МРТ изображений. Этот факт согласуется с результатам, полученным в работе [34]. Однако, у описанных стратегий существуют незначительные преимущества и недостатки. Выделение ключевых точек занимает дополнительное вычислительное время и дает наименьшее количество дескрипторов (что не хорошо ввиду малого разрешения МРТ-изображений); стратегия выделения дескрипторов на рис. 3.4с, производит больше дескрипторов, чем стратегия на рис. 3.4d, также повышая вычислительную сложность, не давая при этом существенного улучшения качества поиска.

Таблица 3.1. Среднее число дескрипторов при различных стратегиях выборки.

	4a, b	4c	4d
Среднее число дескрипторов	157	560	213

3.4.3 Кластеризация дескрипторов и построение сигнатуры изображения

На этом шаге используется общая схема подхода «мешок визуальных слов» (Bag of Visual Words) [35]. Фактически, подход состоит из двух стадий: стадия обучения, на которой происходит кластеризация дескрипторов и стадия непосредственно поиска, на которой где изображение-запрос описывается относительными частотами появления дескрипторов в кластерах и производится сравнение таких сигнатур.

На стадии обучения для всех МРТ изображений $B_n(x, y, z), n = \overline{1, N}$ из базы вычисляются все дескрипторы $f_{n,i}, n = \overline{1, N}; i = \overline{1, I_N}$, после чего весь набор дескрипторов кластеризуется с помощью алгоритма k-средних [36]. Однако, не очевидно, можно ли сравнивать дескрипторы, полученные с МРТ мозга в различных проекциях (фронтальной, аксиальной, сагиттальной). Поэтому, процедура кластеризации выполняется 3 раза по отдельности для каждой проекции, например $f_{n,i}^{sg}, n = \overline{1, N}; i = \overline{1, I_n^{sg}}$ - дескрипторы сагиттальной

проекция кластеризуются и определяются центры кластеров $c_j^{sg}, j = \overline{1, K}$, где K – параметр алгоритма k -средних. То же самое проделывается для аксиальной $c_j^a, j = \overline{1, K}$ и фронтальной $c_j^c, j = \overline{1, K}$ проекции. В случае SIFT размерность дескрипторов и центров кластеров равняется 128, для КГФ размерность равна 150.

После построения центров кластеров можно построить изображений для всей базы данных. Каждый дескриптор $f_{n,i}$ помечается меткой $l_{n,i}$, обозначающей к какому кластеру он принадлежит. Метки назначаются по правилу ближайшего соседа $l_{n,i} = \arg \min_q d(f_{n,i}, c_q)$ отдельно для каждой проекции. Здесь расстояние $d(f_{n,i}, c_q)$ между дескрипторами, вообще говоря, зависит от используемого типа дескриптора, но для дескрипторов SIFT и КГФ хорошо подходит евклидово расстояние, так как все элементы дескриптора имеют одинаковый порядок.

Из меток $l_{n,i}^{sg}, l_{n,i}^a, l_{n,i}^c$ строятся гистограммы h_n^{sg}, h_n^a, h_n^c и нормируются. Наконец, сигнатура изображения h_n строится как конкатенация гистограмм всех проекций $h_n = [h_n^{sg}, h_n^a, h_n^c]$.

Для поиска МРТ в базе данных, схожих с изображением-запросом $B^*(x, y, z)$ сначала для изображения-запроса также рассчитываются дескрипторы $f_i^{*s}, n = \overline{1, N}; i = \overline{1, I_s^*}$, и строится сигнатура h^* по уже построенной кластеризации $l_i^* = \arg \min_q d(f_i^*, c_q)$. Схожесть изображений устанавливается путем сравнения сигнатур, меньшее расстояние означает большую схожесть. Для сравнения гистограмм использовалось расстояние L_1 :

$$d_n = D(h^*, h^n) = \sum_{i=1}^K |h_i^* - h_i^n|; \text{Подобие (к энному изображению)} = 1/d_n \quad (3.1)$$

Упрощенная схема соответствия

Шаги построения гистограммы и кластеризации дескрипторов могут быть опущены вследствие того, что МРТ головного мозга уже были грубо геометрически совмещены и могут быть сравнены послойно. Для удобства, обозначим дескрипторы изображения на срезе s как $f_{n,i}^s, n = \overline{1, N}; i = \overline{1, I_n^s}$. Число дескрипторов и их положение фактически, одинаковы для всех МРТ в случае стратегий плотного покрытия дескрипторами, поэтому возможно их непосредственное сравнение.

Непосредственная схема сравнения (только для плотных особенностей)

$$d_n = \sqrt{\sum_{s=1}^S \sum_{i=1}^{I_n^s} \|f_i^{*s} - f_{n,i}^s\|^2}; \text{ схожесть (к } n\text{-тому изображению)} = 1/d_n \quad (3.2)$$

Для особенностей, найденных детектором ключевых точек, их число может отличаться в каждом срезе, то есть они не могут быть сравнены непосредственно. Вместо этого можно использовать подход работы [19]. Для каждого дескриптора f_i^{*s} одного изображения определяется ближайший дескриптор $f_{n,i}^s$ другого изображения. Если f_i^{*s} также является самым близким дескриптором из всех на первом изображении для $f_{n,i}^s$ и расстояние между дескрипторами $\|f_i^{*s} - f_{n,i}^s\|$ ниже порога T , считается, что локальные особенности изображений «похожи». Число «похожих» особенностей может использоваться для определения степени схожести:

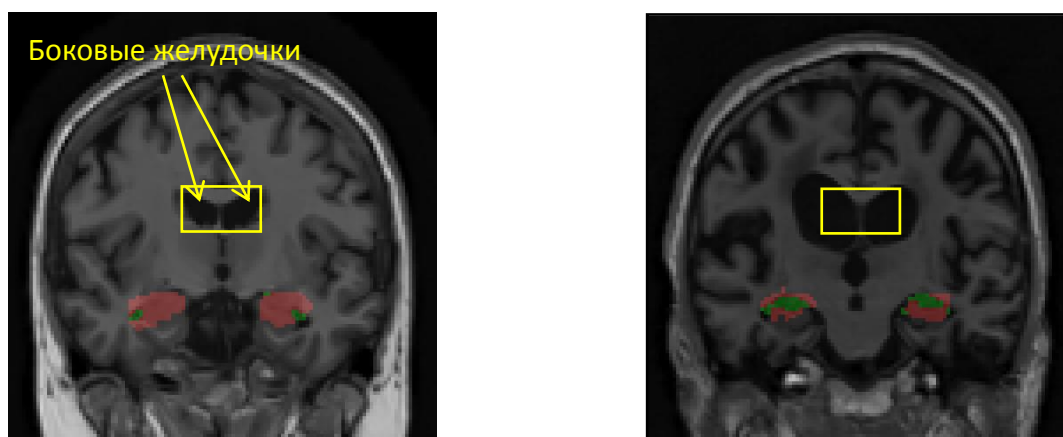
$$\text{Схожесть (к } n\text{-тому изображению)} = 1/d_n = \sum_{s=1}^S \sum_{i=1}^{I_s^*} \begin{cases} 1, & \min_{j=1 \dots I_n^s} \|f_i^{*s} - f_{n,j}^s\| < T \\ 0, & \text{otherwise} \end{cases} \quad (3.3)$$

Этот способ сравнения может использоваться и для стратегии плотного покрытия дескрипторами.

3.5 Предварительная классификация МРТ-изображений

В статье [15] было показано, что предварительная классификация изображений значительно улучшает качество поисковое. Поэтому, предлагается алгоритм предварительной классификации, основанный на

анализе уменьшения гиппокампа, поскольку это является одним из ранних и существенных маркеров болезни Альцгеймера. Чтобы проанализировать уменьшение, определяется объём спинномозговой жидкости (СМЖ) в области гиппокампа. На T1-взвешенных МРТ изображениях СМЖ представляет собой тёмные области, поэтому её относительно легко выделить пороговой обработкой. СМЖ: $B^*(x, y, z) < T_{dark}$. Но, из-за большой разницы в яркости и контрасте разных МРТ-изображений, сам порог должен быть выбран для каждого изображения индивидуально. Для выбора оптимального порога производится следующая процедура. Все воксели (объёмные пиксели) из области гиппокампа группируются вместе, и порог между темными и яркими вокселями устанавливается с помощью метода Оцу. Поскольку у здоровых пациентов мало СМЖ в области гиппокампа для того чтобы гарантировать правильное разделение пикселей на темные и светлые, мы добавляем к рассматриваемому набору пикселей дополнительную область, всегда заполненную СМЖ, например, боковые желудочки головного мозга. На рис. 3.5 добавленная область отмечена желтым прямоугольником.



(a) здоровый пациент

(b) Пациент БА

Рисунок 3.5 – Детектирование вокселей СМЖ в области гиппокампа. Область гиппокампа отмечена красным цветом, выбранные воксели СМЖ – зеленым.

Используя эту процедуру, определяется процент СМЖ в области гиппокампа для всех изображений в базе данных, и, используя этот параметр,

обучается классификатор Байеса. Средние значения и стандартные отклонения объёма СМЖ в вокселях приведены в таблице 3.2. Из таблицы видно, что разброс значений внутри классов велик, и класс здоровых пациентов более или менее отделен от класса УКН и БА, но разделение между УКН и БА плохое, в большой степени эти классы пересекаются, и между ними нет четкой границы. Однако и во врачебной практике наблюдается аналогичная ситуация когда болезнь пациента прогрессирует.

Таблица 3.2. Средние значения и стандартные отклонения объема СМЖ в вокселях для различных классов изображения.

Название класса	Среднее значение	Стандартное отклонение
Здоровый	1101.6	417.6
УКН	1481.7	559.3
БА	1693.4	599.4

В наших экспериментах только один этот параметр дает правильную классификацию в 74% если рассматривать 2 класса (БА и здоровый), но в случае 3 классов (добавляется класс УКН), точность классификации падает до 45%. Причина кроется в том, что класс УКН является переходным между здоровым и БА поэтому границы классов нечеткие, а также маска гиппокампа не достаточно точно соответствует каждому МРТ изображению.

Чтобы объединить предварительную классификацию с мерой схожести из раздела 4, метка класса не назначается изображению-запросу $V^*(x, y, z)$ сразу, вместо этого, вычисляются вероятности $[P_{AD}^*, P_{MCI}^*, P_N^*]$ принадлежности каждому классу. Объединенной мерой схожести тогда

$$\text{Похожесть (к } n\text{-тому изображению)} = 1 / (-\ln(p_{classof(n)}^* * d_n)), \quad (3.4)$$

где d_n любое расстояние из раздела 3.4.3. Логарифм вероятности взят, для устранения доминирования множителя вероятности над расстоянием

d_n , поскольку вероятность – это фактически, экспонента расстояния в параметре объёма СМЖ, и она «перевешивает» расстояние d_n .

3.6 Тестирование алгоритма на реальных данных и выбор оптимальной процедуры сравнения

3.6.1 Данные МРТ

Набор данных МРТ включал стандартные T1-взвешенные изображения, полученные различными типами сканера, используя последовательность MP RAGE. Использовались только изображения базового обследования на сканере 3T. Все изображения были предварительно обработаны как описано на веб-сайте ADNI (http://www.loni.ucla.edu/ADNI/Data/ADNI_Data.shtml), включая исправление дисторсии и исправление B1-неоднородности. Во всех тестах, если это не отмечено иначе, использовалось подмножество базы данных только с одним изображением каждого пациента.

3.6.2 Процедура проверки

Для тестирования алгоритма необходимо иметь референсные данные о похожести изображений. К сожалению, эта информация недоступна для известных баз данных МРТ свободного доступа, кроме того, похожесть с медицинской точки зрения отличается, если опросить различных экспертов. Таким образом, в нашей процедуре тестирования проверяется только правильность определенного класса. Действительно, изображения одного и того же класса должны быть более похожими, чем изображения других классов, поэтому точность поиска и можно оценить расчетом процента правильной классификации изображения в первых N наилучших найденных кандидатах.

Чтобы увеличить число экспериментов и точности статистики, выполняется полная перекрестная проверка, то есть тест повторяется для

каждого изображения из имеющихся данных, а в качестве базы данных, по которой ведется поиск выступают все остальные изображения.

3.6.3 Упрощенный подход

Вначале тестировался упрощенный подход. Используя этот тест, был проанализирован потенциал предложенного метода на основе дескрипторов. На рис. 3.6 показан процент правильной классификации в первых N наиболее релевантных изображений. Для этого теста использовалась только сагиттальная проекция. Результаты показали следующее. Во-первых, точность метода на основе дескрипторов значительно выше, чем произвольный выбор класса случайный шанс, но недостаточна для реальной работы. Во-вторых, качество поиска с дескрипторами КГФ несколько выше.

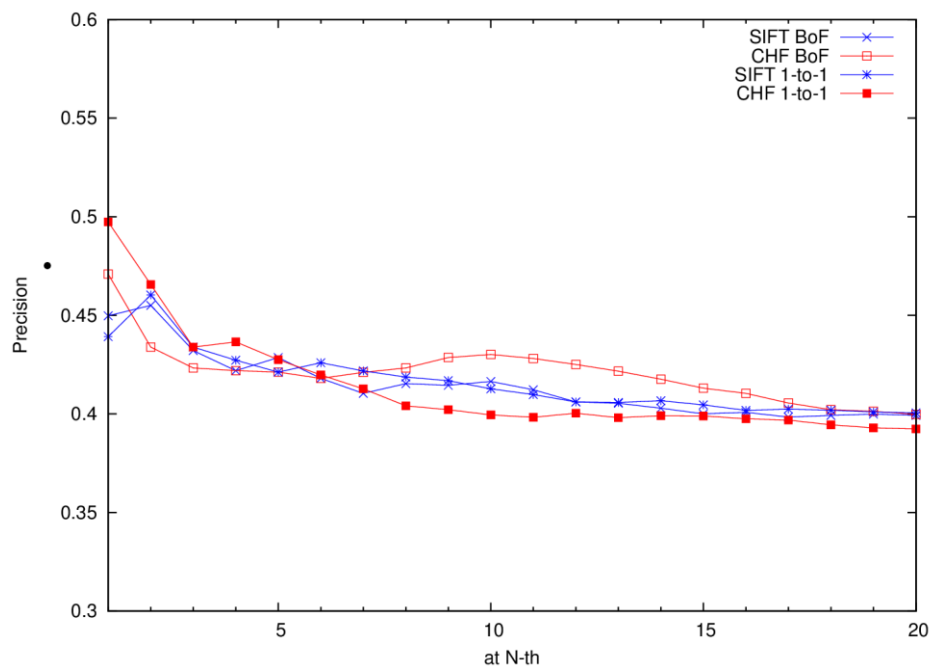


Рисунок 3.6 – Сравнение точности различных описателей и упрощенных схем соответствия.

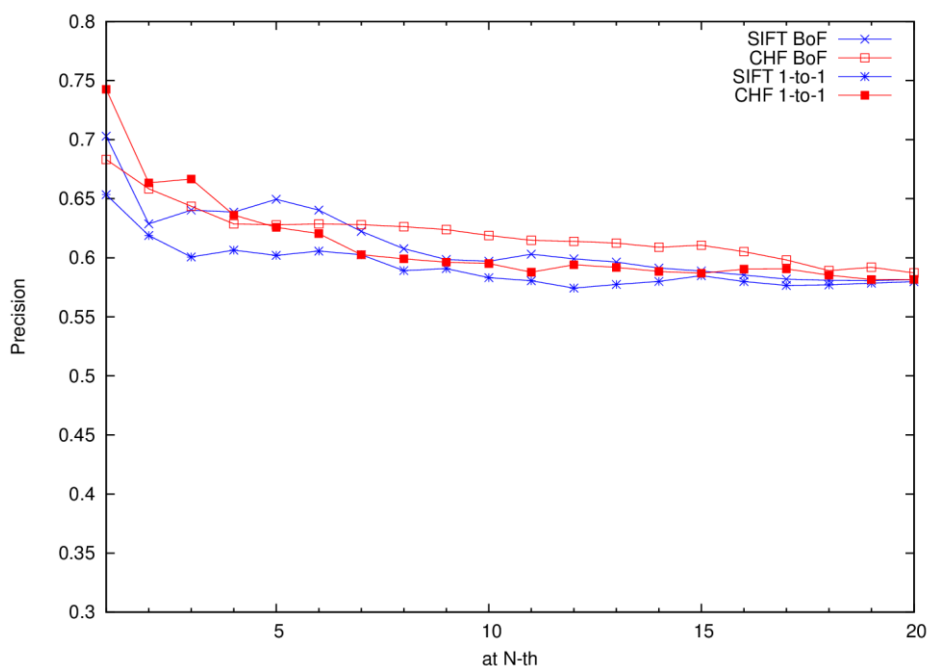


Рисунок 3.7 – Сравнение точности различных описателей и упрощенных схем соответствия 2 классов (Болезнь Альцгеймера и здоровый).

Для большего прояснения результатов был проведен дополнительный тест запросами к базе данных 10-ю различными МРТ пациентов уже присутствующих в базе данных, но сделанных 6 месяцев спустя. **Все** пациенты были правильно идентифицированы, кроме того, расстояния между самым близким (правильным) совпадением и вторым, было больше в 2-4 раза, в то время как для МРТ, не присутствующих в базе данных, это отношение близко к 1. Это приводит нас к выводу, что такое дескрипторное сравнение слишком сильно учитывает индивидуальные особенности пациента и не захватывает общие особенности классов изображения.

Такая же ситуация наблюдается в случае 2 классов (Болезнь Альцгеймера и здоровый), результаты приведены на рис. 3.7.

Для обобщения описания изображения была применена квантизация дескрипторов с помощью метода «мешок слов», как описано в секции 3.4.3.

3.6.4 Подход «мешок визуальных слов» (BOVW)

Для метода BOVW у нас есть только одна схема сравнения, описываемая формулой (3.1), но появляется дополнительный параметр –

количество кластеров. Поэтому, диаграммы точности были построены в зависимости от переменного числа классов. На рис. 8 приведены результаты для дескриптора SIFT, и на рис. 9 приведены результаты для КГФ.

Можно заметить из этих диаграмм, что зависимость от числа кластеров неоднородна, и не наблюдается никакой тенденции при увеличении числа кластеров. В лучшем случае, однако, точность этого подхода превосходит точность упрощенных схем.

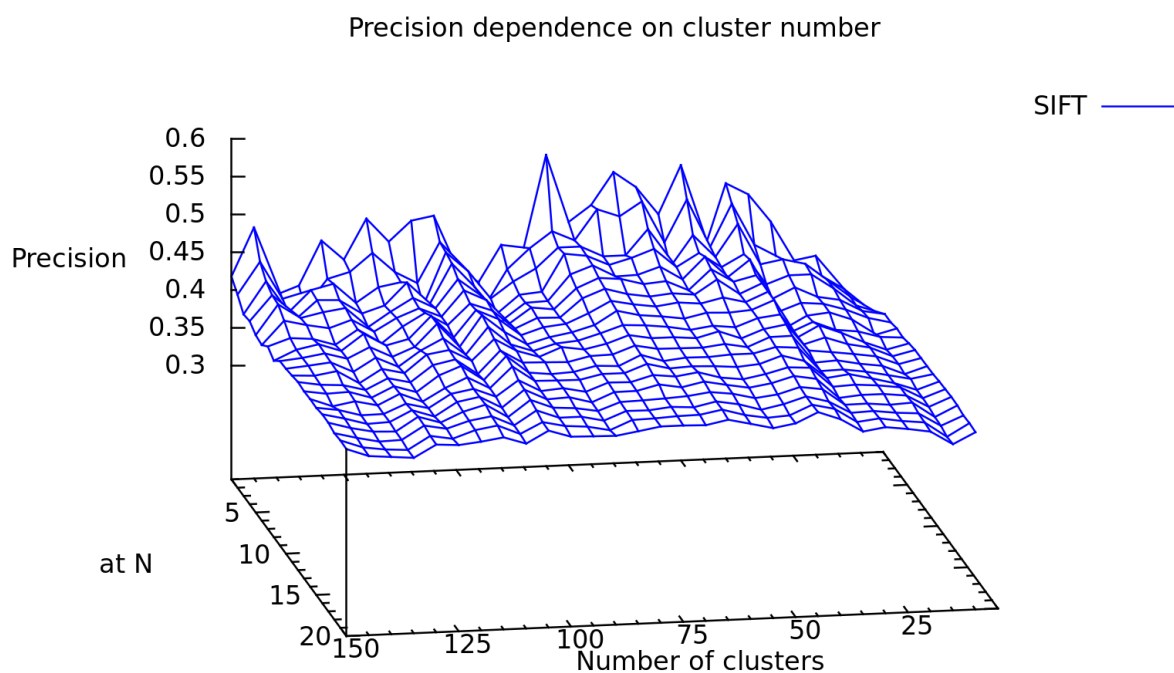


Рисунок 3.8 – Точность поиска, используя подход BOVW для дескриптора SIFT, в зависимости от количества кластеров.

3.6.5 Объединение с предварительной классификацией

Наконец, необходимо протестировать предварительную систему классификации, предложенную в разделе 3.5. Как это было уже упомянуто, одна только эта схема дает результаты, сопоставимые с упрощенным подходом. Объединенная схема, определенная формулой (3.4) дает приблизительно 5%-ое повышение в точности для всех типов расстояний (формулы (3.1)-(3.3)). Наилучшей комбинацией является предварительная

классификация, объединенная с расстоянием (3.3), то есть расстояние BOVW, её результаты представлены на рис. 3.10.

Увеличение точности для дескрипторов SIFT и КГФ приблизительно одинаковое.

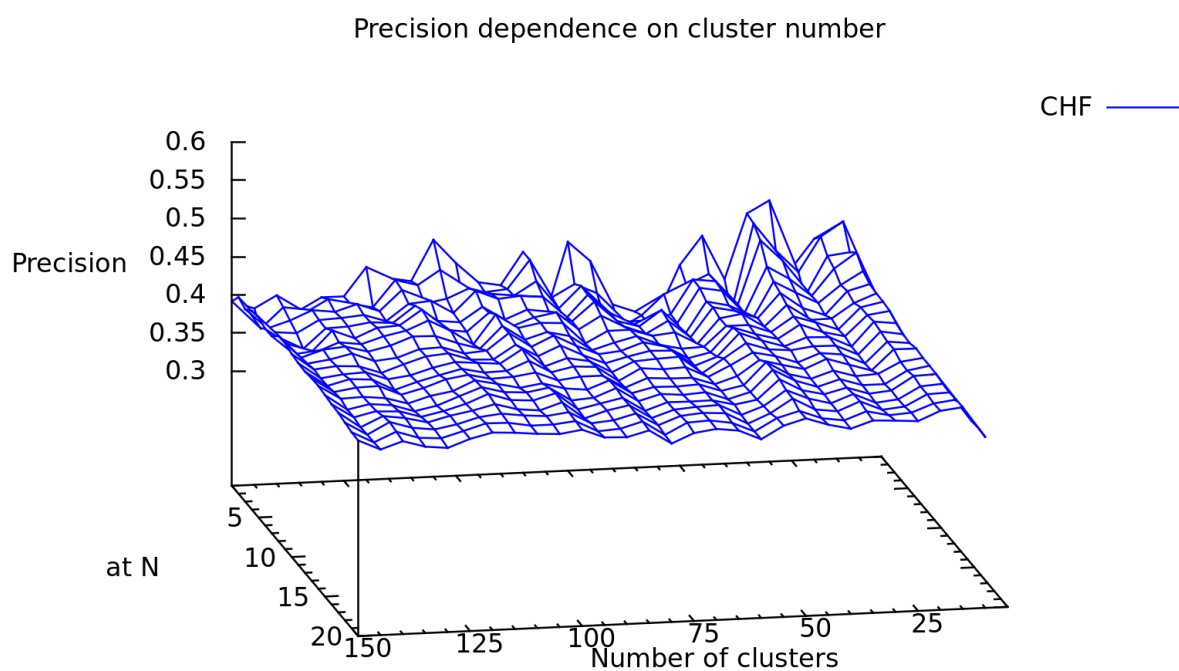


Рисунок 3.9 – Точность поиска, используя подход BOVW для дескриптора SIFT, в зависимости от количества кластеров.

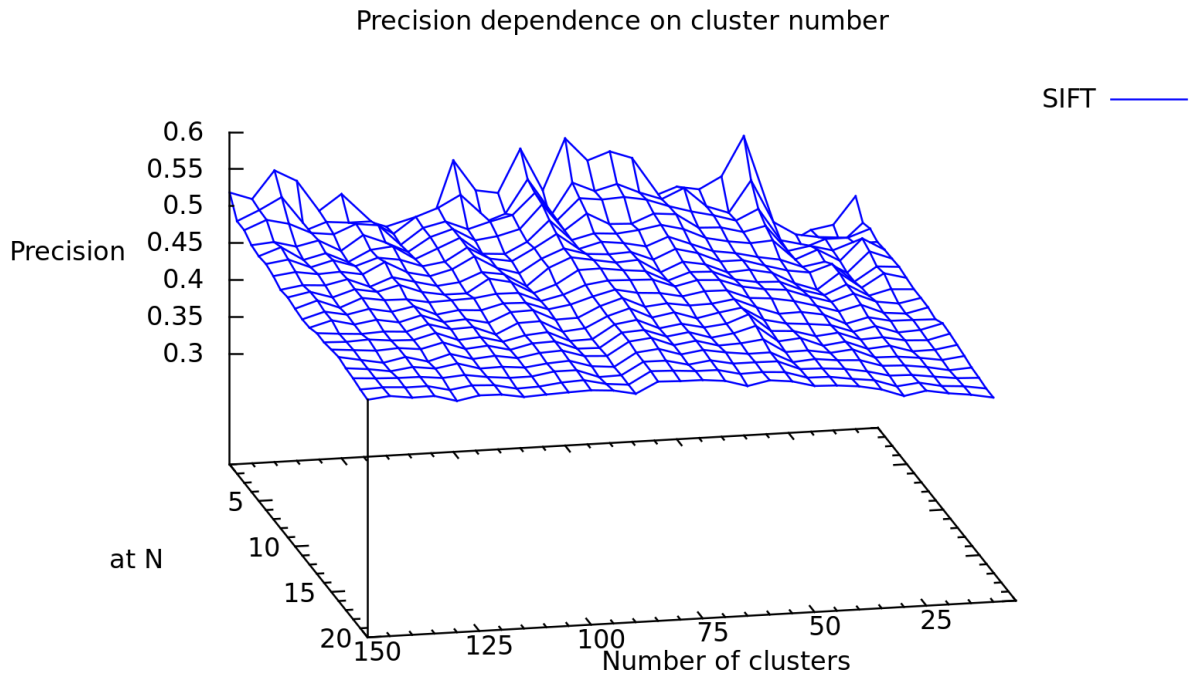


Рисунок 3.10 – Точность поиска, используя комбинацию BOVW с предварительной классификацией.

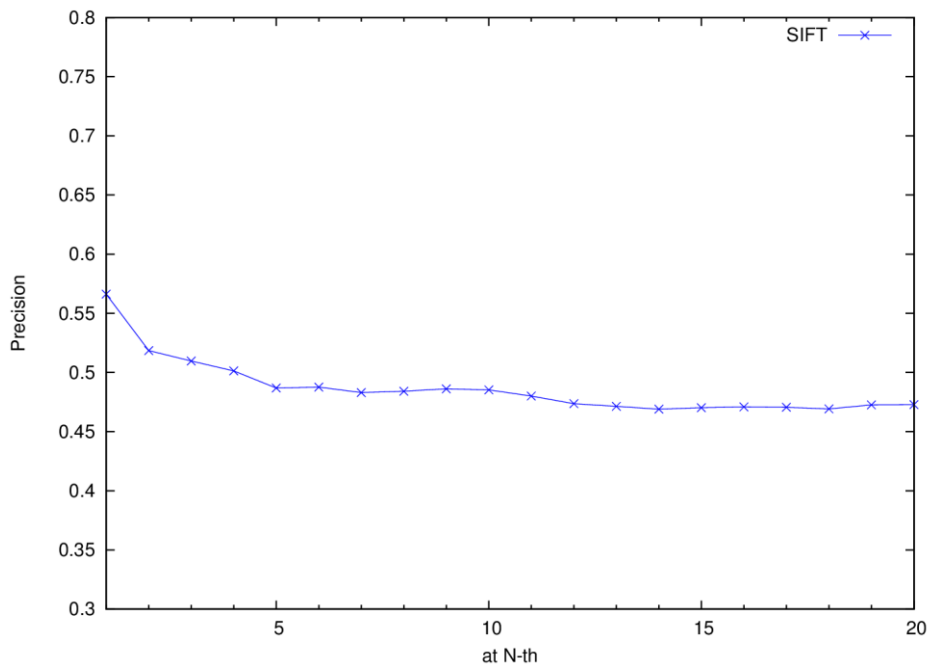


Рисунок 3.11 – Лучшее сечение рис. 3.10 (при 75 кластерах).

Подводя итог результатов экспериментов, можно прийти к заключению, что предложенный метод на основе дескрипторов обладает высоким потенциалом для поиска изображений по чисто визуальным

характеристикам по ряду причин. Во-первых, тест с запросами пациентов уже присутствующими в базе данных, показывает очень высокую точность поиска конкретных пациентов. Во-вторых, метод, основанный на дескрипторах, превосходит по точности простейший подход, базирующийся исключительно на нашем априорном знании об объеме гиппокампа с точки зрения правильной классификации. Здесь очень важен факт, что в схеме с дескрипторами нигде явно не определялся критерий, тем не менее наиболее похожее изображение чаще всего было из того же самого класса изображений.

Метод, однако, в существующей форме далек от идеала с точки зрения точности. Тестирование комбинированного метода показало, что, хотя основанный на BOVW метод в среднем классифицирует разумно, он не включает всю доступную информацию о классах пациентов, потому что даже добавления простого критерия процента СМЖ улучшило общую точность. Следовательно, необходимо дальнейшее исследование шага выделения признаков для «захвата» большего количества релевантной относительно рассматриваемой болезни информации и меньше информации об индивидуальной анатомии мозга.

Наконец, мы можем прийти к заключению, что надежность дескрипторов на основе КГФ Лагерра, немного лучше, чем у дескриптора SIFT, как видно из рис. 3.7.

Заключение

На примере интегрального преобразования, используемого в томографических задачах, разработан метод нелинейной пороговой обработки, который также применим к широкому кругу других приложений, связанных с подавлением шума в регистрируемых данных.

Разработан метод детектирования областей хориоретинальной атрофии на фотографиях глазного дна. Данный метод включает использование классификатора, основанного на разреженных представлениях вместе с традиционным алгоритмом, основанным на классификации по текстурным признакам, и позволяет получить значительно более точные результаты, чем традиционный метод.

Разработан метод поиска медицинских изображений на основе локальных дескрипторов изображения. Локальные дескрипторы строятся с помощью разложения окрестностей ключевых точек в частичный ряд Фурье по системе круговых гармонических функций Лагерра (КГФ ГЛ). В результате вычислительных экспериментов установлено, что разработанные дескрипторы показывают очень высокую точность поиска конкретных пациентов, то есть обладают высокой дискриминативностью. Установлено, что дескрипторы на основе КГФ ГЛ превосходят по точности дескриптор SIFT.

Исследованы три подхода к построению сигнатур МРТ изображения на основе локальных дескрипторов и соответствующие методы сравнения. Простейший подход заключается в генерировании плотной сетки ключевых точек и сравнении изображений по сумме евклидовых расстояний всех их дескрипторов. Второй способ является модификацией первого, но ключевые очки выделяются с помощью детектора ключевых точек, и допускается смещение точек на разных изображениях, для чего при сравнении выбираются пары дескрипторов с наименьшим расстоянием. Третий подход основывается на методе «мешок слов», для чего производится кластеризация

дескрипторов и построение гистограммы классов дескрипторов в качестве сигнатуры изображения. Сравнение подходов показало некоторое превосходство третьего подхода над остальными, при этом сам этап поиска в этом подходе является наиболее быстрым их трех.

Разработан метод предварительной классификации пациентов на здоровых и больных БА на основе анализа объема спинномозговой жидкости в области гиппокампа. Применение предварительной классификации позволяет улучшить общую точность поиска. Тем не менее, итоговая точность классификации БА говорит о том, что в текущей схеме всё еще слишком большая зависимость от индивидуальной анатомии человека, и для применения метода в системах диагностики необходимо совершенствование механизма выбора и кластеризации локальных дескрипторов.

Список использованных источников

1. Тихонов А.Н., Арсенин В.Я., Тимонов А.А. Математические задачи компьютерной томографии. - М.: Наука, 1987.
2. Троицкий И.Н. Статистическая теория томографии. - М.: Радио и связь, 1989.
3. Как А.С., Slaney M. Principles of Computerized Tomographic Imaging. - NY: IEEE Press, 1988.
4. Маркин А.В., Шестаков О.В. Асимптотики оценки риска при пороговой обработке вейвлет-вейглет коэффициентов в задаче томографии // Информатика и ее применения, 2010. Т. 4. № 2. С. 36-45.
5. Donoho D. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition // Applied and Computational Harmonic Analysis, 1995. Vol. 2. P. 101-126.
6. Kolaczyk E.D. A wavelet shrinkage approach to tomographic image reconstruction // J. Am. Statist. Assoc., 1996. Vol. 91. P. 1079-1090.
7. Lee N. Wavelet-vaguelette decompositions and homogenous equations: PhD dissertation. Purdue University, 1997.
8. Кудрявцев А.А., Шестаков О.В. Асимптотика оценки риска при вейглет-вейвлет разложении наблюдаемого сигнала // Т-Comm — Телекоммуникации и Транспорт, 2011. № 2. С. 54-57.
9. Кудрявцев А.А., Шестаков О.В. Асимптотическое распределение оценки риска пороговой обработки вейглет-коэффициентов сигнала при неизвестном уровне шума // Т-Comm — Телекоммуникации и Транспорт, 2011. № 5. С. 24-30.
10. Mallat S. A Wavelet Tour of Signal Processing. - NY: Academic Press, 1999.
11. Abramovich F., Silverman B.W. Wavelet Decomposition Approaches to Statistical Inverse Problems // Biometrika, 1998. Vol. 85. No. 1. P. 115-129.
12. Donoho D., Johnstone I.M. Ideal Spatial Adaptation via Wavelet Shrinkage // Biometrika., 1994. Vol. 81. No. 3. P. 425-455.

13. Natterer F. The Mathematics of Computerized Tomography. - NY: John Wiley & Sons. 1986.
14. Smith K., Solmon D., Wagner S. Practical and mathematical aspects of the problem of reconstructing objects from radiographs // Amer. Math. Soc., 1977. Vol. 83. No. 6. P. 1227-1270.
15. F. Schaffalitzky and A. Zisserman. Multi-view matching for unordered image sets // In Proc. ECCV2002. - 2002. - P. 414–431.
16. T. Tuytelaars, V. Ferrari and L. Van Gool. Simultaneous object recognition and segmentation from single or multiple model views // Int. J. of Computer Vision. - 2006. Vol. 67, No 2. - P. 159-188.
17. Cl. Morand, J. Benois-Pineau, J.-Ph. Domenger, J. Zepeda, E. Kijak, Ch. Guillemot. Scalable object-based video retrieval in HD video databases // Image Commun. - 2010. - Vol. 25, Issue 6. - P. 450-465.
18. C.G. Harris, M. Stephens. A combined corner and edge detector. // In Proc. 4th Alvey Vision Conf. Manchester. - 1998. - P. 147-151.
19. David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. International Journal of Computer Vision, Vol. 60, No. 2. (1 November 2004), pp. 91-110.
20. L. Sorigi, N. Cimminiello, and A. Neri, “Keypoints Selection in the Gauss Laguerre Transformed Domain,” in Proc. of BMVC06, 2006, pp. 539–547.
21. H. Hse, A.R. Newton. Sketched symbol recognition using Zernike moments. // In Proc. ICPR04. - 2004. P. 367-370.
22. K. Mikolajczyk, C. Schmid. A performance evaluation of local descriptors. // IEEE Trans. On PAMI. - 2004. - Vol. 27, No. 10. - P. 1615-1630.
23. Sorokin D. V., Mizotin M. M., Krylov A. S. Gauss-Laguerre Keypoints Extraction Using Fast Hermite Projection Method // Lecture Notes in Computer Science. 2011. Vol. 6753. Pp. 284–293.
24. Sorokin D. V., Krylov A. S. Gauss–Laguerre–Hermite Method of Keypoint Extraction // Pattern Recognition and Image Analysis. 2011. Vol. 21, no. 2. Pp. 332–334.

- 25.G. Jacovitti, A. Neri. Multiresolution circular harmonic decomposition. // IEEE Trans. Signal Process. - 2000. - Vol. 48, No. 11. - P. 3242-3247.
- 26.E. Zauderer. Complex argument Hermite-Gaussian and Laguerre-Gaussian beams. // J. Opt. Soc. Amer. A. - 1986. Vol. 3, No. 4. - P.465–469.
- 27.E. D. Di Claudio, G. Jacovitti, A. Laurenti. Maximum Likelihood Orientation Estimation of 1-D Patterns in Laguerre-Gauss Subspaces. // IEEE Tran. Image Processing. - Vol. 19, No. 5. - P. 1113-1125.
- 28.Agarwal, M.; Mostafa, J.; “Content-based image retrieval for Alzheimer's disease detection”, Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on, pp.13-18, 13-15 June 2011
- 29.A.M. Aisen, L.S. Broderick, H. Winer-Muram, C.E. Brodley, A.C. Kak, C. Pavlopoulou, J. Dy, C.-R. Shyu, and A. Marchiori. Automated Storage and Retrieval of Thin-section CT Images to Assist Diagnosis: System Description and Preliminary Assessment. Radiology, 228(3.1):265–270, 2003
- 30.B.H. Ridha, J. Barnes, L.A. van de Pol, J.M. Schott, R.G. Boyes, M.M. Siddique, M.N. Rossor, P. Scheltens, and N.C. Fox. Application of automated medial temporal lobe atrophy scale to Alzheimer disease. Archives of Neurology, 64(3.6):849–854, June 2007.
- 31.Toews M, Wells W 3rd, Collins DL, Arbel T. Feature-based morphometry: discovering group-related anatomical patterns. Neuroimage. 2010 Feb 1;49(3.3):2318-27. Epub 2009 Oct 21.
- 32.John Ashburner, Karl J. Friston, Unified segmentation. NeuroImage, Vol. 26, Issue 3, pp. 839-851, July 2005.
- 33.Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002 Jan; 15(3.1):273-89.
- 34.E. Nowak, F. Jurie, and B. Triggs, “Sampling strategies for bag-of-features image classification,” in Proc. ECCV, 2006

35. Chris Dance, Jutta Willamowski, Lixin Fan, Cedric Bray, and Gabriela Csurka. Visual categorization with bags of keypoints. In *ECCV International Workshop on Statistical Learning in Computer Vision*, 2004.
36. Arthur, D. and Vassilvitskii, S. (2007). "k-means++: the advantages of careful seeding". Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. pp. 1027–1035.