

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧЕРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»

КАФЕДРА АСТРОФИЗИКИ

Костенков Александр Евгеньевич

АНАЛИЗ ДВУМЕРНЫХ АСТРОНОМИЧЕСКИХ
ИЗОБРАЖЕНИЙ МЕТОДАМИ ФУРЬЕ И ВЕЙВЛЕТ АНАЛИЗА

ДИПЛОМНАЯ РАБОТА

Допущен к защите.
Заведующий кафедрой астрофизики:
д.ф.-м.н., профессор Гаген-Торн В.А.

Научный руководитель:
д.ф.-м.н., профессор Холтыгин А.Ф.

Научный консультант:
д.ф.-м.н., профессор С.Н. Фабрика

Рецензент:
д.ф.-м.н., профессор М.А. Погодин

Санкт-Петербург, 2016

SAINT-PETERSBURG
STATE UNIVERSITY

DEPARTMENT OF ASTROPHYSICS

Kostenkov Alexander

**ANALYSIS OF TWO-DIMENSIONAL ASTRONOMICAL IMAGES
BY METHODS OF FOURIER AND WAVELET ANALYSIS**

GRADUATION THESIS

Admitted for defence.

Head of Department:

Doctor of Sciences, Professor V.A. Hagen-Thorn

Scientific supervisor:

Doctor of Sciences, Professor A.F. Kholtygin

Scientific supervisor:

Doctor of Sciences, Professor S.N. Fabrika

Reviewer:

Doctor of Sciences, Professor M.A. Pogodin

Saint-Petersburg, 2016

Оглавление

| | | |
|----------|---|-----------|
| 1 | Основные преобразования и соотношения | 7 |
| 1.1 | Пространственная фильтрация | 8 |
| 1.1.1 | Сглаживающие пространственные фильтры | 9 |
| 2 | Частотная фильтрация | 11 |
| 2.1 | Дискретное преобразование Фурье | 11 |
| 2.2 | Вейвлет преобразование | 15 |
| 2.2.1 | Общие понятия вейвлет анализа | 16 |
| 2.2.2 | Кратномасштабный анализ | 16 |
| 2.2.3 | Дискретное вейвлет преобразование | 19 |
| 2.2.4 | Двумерный КМА | 21 |
| 2.2.5 | Альтернативные алгоритмы вейвлет преобразований | 22 |
| 3 | Trous-вейвлет преобразование | 23 |
| 3.1 | Trous-вейвлет алгоритм | 23 |
| 3.2 | Одномерные многоуровневые преобразования | 27 |
| 3.3 | Двумерные многоуровневые преобразования | 31 |
| 4 | Многоуровневая обработка астрономических снимков | 35 |
| 4.1 | Модель сигнал-шум | 35 |
| 4.1.1 | Soft and hard thresholding | 36 |
| 4.1.2 | 3σ -clipping и винеровская фильтрация | 37 |

| | | |
|----------|---|-----------|
| 5 | Многоуровневое представление данных | 41 |
| 5.1 | Описание многоуровневой модели | 41 |
| 5.2 | Поиск объектов и структур | 43 |
| 5.3 | Алгоритм поиска объектов | 44 |
| 5.4 | SExtractor | 45 |
| 5.4.1 | Поиск рассеянных структур и скоплений | 48 |
| 5.4.2 | Проблемы выделения звезд и скоплений | 49 |
| 6 | Поиск звездных скоплений в М33 | 52 |
| 6.1 | Luminous blue variable | 52 |
| 6.2 | Выбор исходных данных | 54 |
| 6.3 | Определение параметров групп звезд | 55 |
| 6.4 | Идентификация объектов | 57 |
| | Приложение А | 70 |
| | Приложение Б | 72 |
| | Приложение В | 74 |
| | Приложение Г | 76 |
| | Приложение Д | 77 |

Абстракт

В дипломной работе представлены способы обработки двумерных астрономических изображений, основанные на различных методах матричной свертки, фильтрации, а так же двумерном вейвлет анализе. Представлен подход к декомпозиции изображений адаптированный для обработки астрономических снимков. Отладка алгоритмов производилась с использованием модельных решений.

Первая часть диплома - описание математического аппарата и существующих способов, изложение теории многоуровневого представления в астрономии. Второй этап работы является демонстрацией использования различных методов анализа изображений на модельных решениях, сравнение результатов, полученных разными способами. Третья часть - применение вейвлет анализа в идентификации скоплений, адаптивный поиск объектов на снимках.

Используя этот подход, были выделены скопления в галактике М33, которые могут быть связаны с LBV-звездами. Относительно скоплений было составлено распределение расстояний до звезд. Такой же подход можно применить и для других галактик Местной группы.

Введение

Обработка и анализ данных становятся все более значимыми в астрономии. Это можно объяснить эволюцией приемников излучения - в наше время размеры ПЗС матриц могут достигать 16000×16000 пикселей и более. Следствием этого является огромное количество получаемой информации, которую невозможно обрабатывать вручную. Поэтому нужны методы и алгоритмы, позволяющие автоматически обрабатывать полученные данные.

Большую часть данных получают из снимков, о них и пойдет речь в этой работе. Чаще всего изображения сильно зашумлены, объекты в низком разрешении на них сложно идентифицировать, присутствует эффект проекции на снимке, особенно в переполненных звездами полях, что осложняет классификацию. Для того, чтобы извлечь информацию о каждом объекте, применяют методы математической обработки снимков, начиная от простых преобразований, например, матричной свертки, до сложных декомпозиций изображений при помощи вейвлет преобразований. Это позволяет анализировать снимок на разных масштабах, составлять многоуровневое представление объекта - выделить самые мелкие и более крупные образования. Этот подход является очень сильным инструментом в анализе данных.

В работе ставилась цель выделить характеристики скоплений, связанных с LBV, путем обработки снимков, идентифицировать аналогичные объекты на всем изображении. Решение этой задачи можно разделить на три этапа:

- Анализ и сбор данных - первичная редукция
- Обработка снимков - математическая обработка, например, очистка от шума, выделение исследуемых объектов, фильтрация изображения
- Выделение необходимых критериев для идентификации скоплений, классификация объектов на снимке

В данный момент существует большое количество алгоритмов и моделей, которые используются в обработке снимков. Применяют как простые преобразования изображений с помощью фильтров низких и высоких частот (НЧ и ВЧ), так и обработка в частотной области при помощи Фурье и вейвлет анализа. Широко используется многоуровневое представление данных. Отличным примером реализации такого подхода является SExtractor (Source Extractor, см. [3]). Это программное обеспечение, которое выделяет и классифицирует астрономические объекты на снимках. В ней заложен интересный подход к анализу данных: поиск объектов и оценка фона осуществляется после свертки изображения с выбранным фильтром, вариации фильтров позволяют искать объекты разной формы и размеров (например, Point Spread Function (PSF) для поиска одиночных звезд, для протяженных объектов широкий НЧ фильтр). Иногда специфика задачи не предполагает точных критериев - хорошим примером являются поиск молодых скоплений, которые могут занимать довольно обширную площадь, а применение широких фильтров приводит к размыванию и захвату лишних звезд в переполненных звездами областях ("crowded fields").

В данной работе предлагается использовать многоуровневый подход в задаче - декомпозиция изображения на структуры разного уровня, поиск интересных объектов с помощью набора фильтров на каждом уровне. Так можно найти как крупные, так и мелкие структуры.

Проблема поиска скоплений в переполненных областях все еще остается открытой, нашу задачу упрощает тот факт, что производился поиск

молодых скоплений рядом с которыми находится эмиссионная туманность (область III с высоким излучением в линии H_{α} , возраст $< 10^8$ - 10^9 лет).

Глава 1

Основные преобразования и соотношения

Информация, содержащаяся в снимках, может быть представлена разными способами. В данной работе будет рассматриваться пространственная интерпретация. Растровые изображения представляются пространственным распределением освещенности на плоскости. Ее можно описать как функцию двух переменных:

$$f = (x, y) \quad (1.1)$$

где f - освещенность клетки с координатами x, y . Таким образом изображение A может быть представлено в виде двумерных массивов:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (1.2)$$

где a_{mn} - освещенность пикселя с заданными координатами.

В дипломной работе рассматриваются полутоновые и бинарные изображения. Полутоновое изображение кодируется с помощью битовой сетки,

то есть каждый пиксель описывается разным количеством бит, что и определяет количество полутонов. Например, 2 бита - 4 полутона, 3 хранят 8 и т.д. Бинарное изображение - вырожденный случай полутонового (1 бит на 1 пиксель), пиксель передает лишь 2 полутона - черный или белый (в цифровом представлении матрица из 0 и 1).

Далее под обработкой снимка будет подразумеваться преобразование матрицы изображения для какой-то конкретной задачи. Большинство таких методов делится на две большие категории: пространственная обработка (например, фильтрация по маске) и частотная обработка (преобразование Фурье и т.д.). Ниже будет рассмотрен математический аппарат тех способов, которые были использованы в работе (описание взято из [33], [31]).

1.1 Пространственная фильтрация

Некоторые локальные преобразования оперируют одновременно как со значениями пикселей в окрестности, так и с соответствующими им значениями некоторой матрицы, имеющей те же размеры, что и окрестность. Соседние клетки определяются стандартно: все комбинации $(x \pm 1, y \pm 1)$, где (x, y) - координаты выбранного пикселя в матрице.

Схема пространственной фильтрации основана на простом перемещении маски фильтра от точки к точке изображения. Обработка изображения f , имеющего размеры $M \times N$, с помощью фильтра размерами $m \times n$ задается выражением общего вида:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (1.3)$$

где, как следует из предыдущего абзаца, $a = (m - 1)/2$ и $b = (n - 1)/2$. При вычислении итоговой интенсивности всего изображения данная формула должны быть вычислена для всех сочетаний $x = 0, 1, 2 \dots M - 1$ и

$y = 0, 1, 2 \dots N-1$. Это означает что все элементы изображения будут обработаны по заданной маске. Процедура линейной фильтрации, задаваемая этим уравнением в частотной области аналогична процедуре свертки в Фурье пространстве. Поэтому эту процедуру называют "свертку маски(ядра) с изображением". Важным вопросом при применении формулы 1.3 остается ситуация, когда центр маски находится вблизи границы изображения. Можно ограничить перемещение фильтра по границе изображения либо достраивать изображение с помощью нулей, а потом убрать их.

В зависимости от выбора ядра можно сгладить изображение или, наоборот, выделить границы.

1.1.1 Сглаживающие пространственные фильтры

Сглаживающие пространственные фильтры применяются для расфокусировки изображения и подавления шума. Расфокусировка может применяться как предварительный шаг для обработки изображений, например для удаления мелких деталей для обнаружения более крупных объектов или для устранения разрывов в линиях или деталях.

Примером линейного низкочастотного фильтра является размытие по Гауссу. Производится свертка матрицы изображения с ядром вида:

$$g(s, t) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (1.4)$$

где (x, y) - координаты в маске, σ - среднеквадратичное отклонение.

Медианный фильтр - один из видов нелинейных цифровых фильтров, широко используемый в обработке сигналов изображений. Медианная фильтрация в меньшей степени сглаживает границы изображения, чем любая линейная свертка. Значения отсчетов внутри окна фильтра сортируются в порядке возрастания (убывания) и значение, находящееся в середине упорядоченного списка, поступает на выход фильтра. В случае четного чис-

ла отсчетов в окне выходное значение фильтра равно среднему значению двух отсчетов в середине упорядоченного списка. Окно перемещается вдоль фильтруемого сигнала и вычисления повторяются.

Низкочастотные фильтры используются в астрономии для выделения более крупных структур или объектов на снимках (например, галактики или скопления). Они очень просты в использовании, но существуют более тонкие алгоритмы по выделению структур в частотной области (об этом в следующей главе).

Глава 2

Частотная фильтрация

В данной главе представлена математическая интерпретация методов, основанных на обработке в частотной области - Фурье и вейвлет преобразованиях. Они представляют собой очень гибкие инструменты для фильтрации и анализа данных. Для визуализации примеров в этой главе использовались их реализации среды OCTAVE. Теория изложена в [32], [33].

2.1 Дискретное преобразование Фурье

Дискретное преобразование Фурье переводит конечную последовательность вещественных чисел в конечную последовательность коэффициентов Фурье. Пусть $\{x_i\}, i = 0, \dots, N-1$ - последовательность вещественных чисел - например, отчеты яркости пикселей по строке изображения. Эту последовательность можно представить в виде комбинации конечных сумм вида

$$x_i = a_0 + \sum_{n=1}^{N/2} a_n \cos\left(\frac{2\pi ni}{N}\right) + \sum_{n=1}^{N/2} b_n \sin\left(\frac{2\pi ni}{N}\right) \quad (2.1)$$

где

$$a_0 = \frac{1}{N} \sum_{i=0}^{N-1} x_i, \quad a_{N/2} = \frac{1}{N} \sum_{i=0}^{N-1} x_i (-1)^i, \quad a_k = \frac{2}{N} \sum_{i=0}^{N-1} x_i \cos\left(\frac{2\pi ik}{N}\right) \quad (2.2)$$

$$b_k = \frac{2}{N} \sum_{i=0}^{N-1} x_i \sin \left(\frac{2\pi i k}{N} \right), \quad 1 \leq k < N/2. \quad (2.3)$$

Часто используют комплексную запись, основанную на формуле Эйлера:

$$G_m = \frac{1}{N} \sum_{n=1}^{N-1} x_n e^{-\frac{2\pi i m n}{N}} \quad (2.4)$$

Дискретное преобразование Фурье для двумерного массива чисел размера $M \times N$ определяется следующим образом:

$$G_{uw} = \frac{1}{NM} \sum_{n=1}^{N-1} \sum_{m=1}^{M-1} x_{mn} e^{-2\pi i \left[\frac{mu}{M} + \frac{nw}{N} \right]} \quad (2.5)$$

Двумерное преобразование Фурье может быть представлено в виде композиции двух одномерных преобразований:

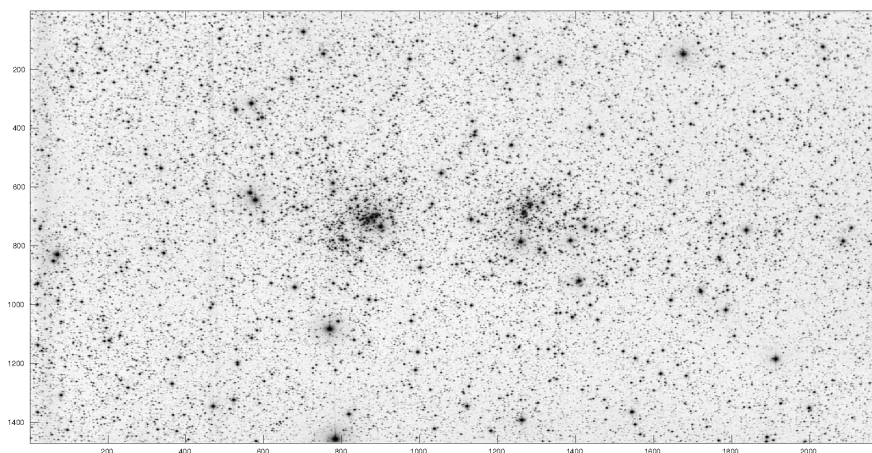
$$G_{uw} = \frac{1}{N} \sum_{n=1}^{N-1} \left[\frac{1}{M} \sum_{m=0}^{M-1} x_{mn} e^{-\frac{2\pi i m w}{M}} \right] e^{-\frac{2\pi i n u}{N}} \quad (2.6)$$

Здесь выражение в квадратных скобках есть одномерное преобразование строки матрицы данных. Таким образом, для получения двумерного преобразования Фурье нужно сначала вычислить одномерные преобразования строк, записать результаты в исходную матрицу и вычислить одномерные преобразования для столбцов полученной матрицы. Алгоритмы фильтрации в частотной области основываются на свертке. В двумерном случае преобразование свертки выглядит следующим образом:

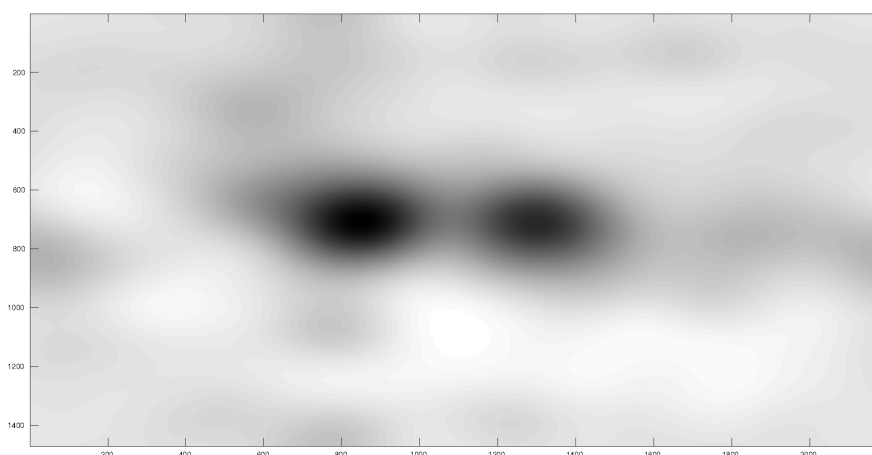
$$G(u, v) = H(u, v) F(u, v) \quad (2.7)$$

где G - Фурье-образ результата свертки, H - Фурье-образ фильтра, а F - Фурье-образ исходного изображения.

На Рис. 2.1 приведен пример обработки снимка в частотной области при помощи НЧ фильтра Гаусса ($\sigma = 3$). На изображении представлены два скопления - NGC 869 и NGC 884 (более известные как χ и h Персея, любительский снимок - SBIG ST-8300M, Astro-Tech AT90DT, f/6.7 603 мм, выдержка 285 минут, Dardenne Prairie, Missouri). Используя широкий фильтр можно убрать высокочастотные шумы для того, чтобы идентифицировать скопления. Можно выбрать маску на границе обработанного изображения и бинаризовать по порогу, чтобы определить точное положение. Чем уже профиль фильтра в частотной области (чем больше σ), тем он шире в пространственной.



а)



б)

Рис. 2.1: а) Исходное изображение б) Результат свертки с НЧ фильтром

С помощью аналогичных преобразований можно выделить ВЧ (например, фон, см. Рис. 2.2)

Комбинируя разные маски $H(u, v)$ можно выделить почти любую часть изображения. Легко подобрать фильтр индивидуально, т.к. в частотном пространстве двумерная свертка заменяется поэлементным перемножением образов исходного изображения и соответствующего ядра.

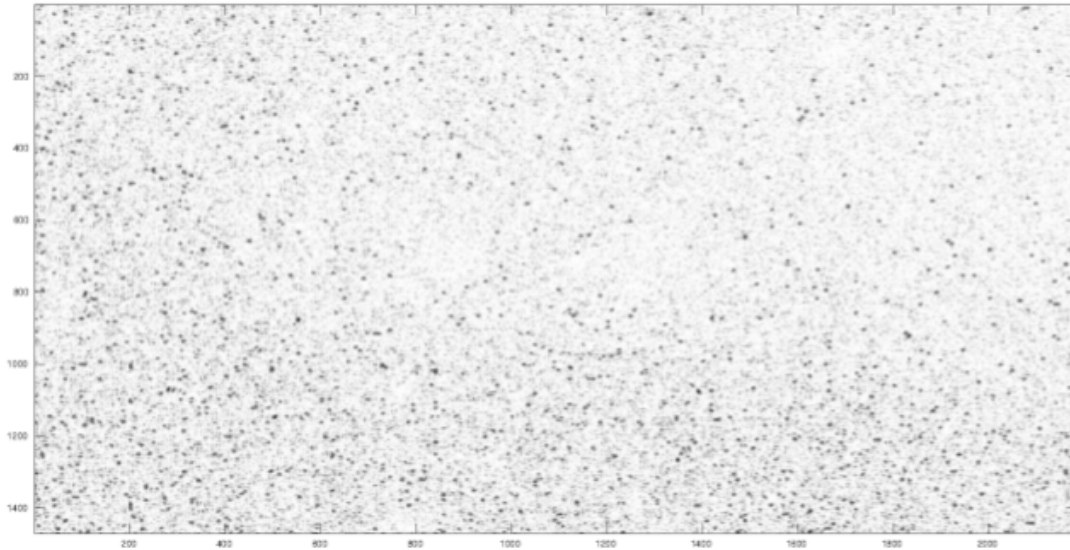


Рис. 2.2: Вклад ВЧ в исходное изображение

С позиции точного представления произвольных сигналов преобразование Фурье имеет ряд ограничений и недостатков. Оно не учитывает колебания сигнала во времени. Локальные особенности сигнала (разрывы, ступеньки, пики и т.п.) дают едва заметные составляющие спектра, по которым обнаружить эти особенности, и тем более их место и характер, практически невозможно. Эти ограничения можно избежать, используя оконное преобразование Фурье, однако появляется ограничение по времени и частоте для всех точек, которое не может быть адаптировано к локальным свойствам сигнала.

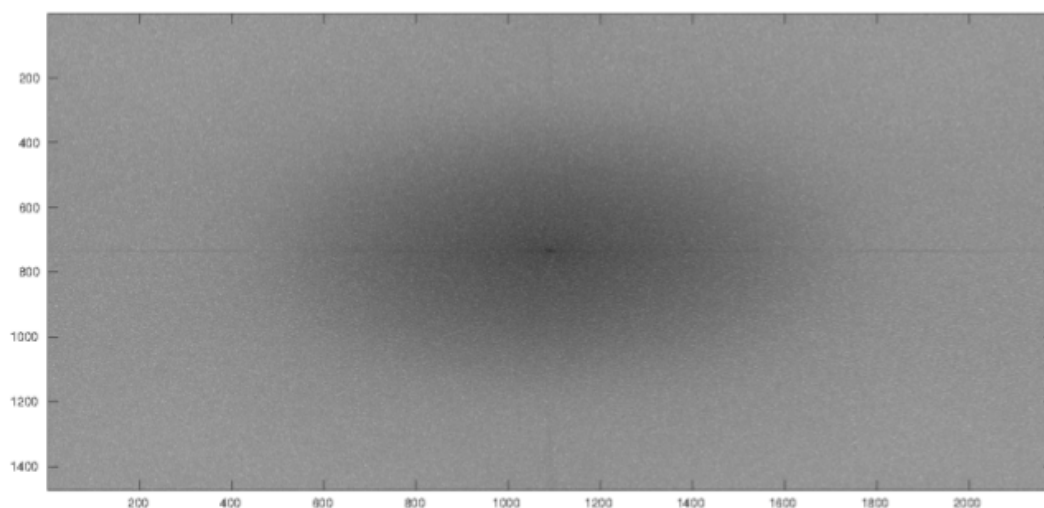


Рис. 2.3: Мощность спектра исходного изображения, логарифмическая шкала

Такие методы, как преобразование Фурье, дискретное косинус преобразование и т.д., позволяют восстановить исходный сигнал по определенной частотной полосе, большая часть информации о нем содержится в нескольких коэффициентах (спектр мощности 2.3). Это существенный недостаток. Многоуровневые вейвлет-преобразования позволяют фильтровать изображения по всем масштабам, независимо друг от друга.

Благодаря выявлению локальных особенностей сигнала, принципиально отсутствующему у преобразования Фурье, вейвлет преобразование нашло широкое применение для анализа тонкой структуры сигналов и изображений, для их сжатия и очистки от шума, что важно и полезно в радиотехнике, электронике, гидроакустике, геофизике, медицине и других областях науки и техники.

2.2 Вейвлет преобразование

В этом разделе будет изложен математический аппарат, который в дальнейшем понадобится для описание многоуровневого представления (Multiscale Vision Model) данных. А так же основы вейвлет-анализа и основные формулы для дискретного преобразования.

2.2.1 Общие понятия вейвлет анализа

Вейвлет преобразованием сигнала будем называть его представление в виде обобщенного ряда по системе базисных функций вида:

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (2.8)$$

сконструированных из материнского (исходного) вейвлета $\psi(t)$, обладающего определенными свойствами за счет операций сдвига во времени (b) и изменения временного масштаба (a). Множитель $1/\sqrt{a}$ обеспечивает независимость нормы этих функций от масштабирующего числа.

При непрерывном изменении параметров a и b для расчета вейвлет-спектра необходимы большие вычислительные затраты. Множество функций $\psi_{ab}(t)$ избыточно. Необходима дискретизация этих параметров при сохранении возможности восстановления сигнала из его преобразования. Дискретизация, как правило, осуществляется через степень двойки:

$$\psi_{mk} = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) = \frac{1}{\sqrt{2^m}}\psi(2^{-m}n - k) \quad (2.9)$$

где $a = 2^m$, $b = k \cdot 2^m$, m и k - целые числа. В этом случае плоскость a, b превращается в соответствующую сетку m, k . Параметр m называется параметром масштаба, n - это сдвиг.

2.2.2 Кратномасштабный анализ

При анализе изображения зачастую удобно представить его в виде совокупности последовательных приближений (Описание взято из [29]).

Кратномасштабный анализ (КМА) обладает целым рядом полезных свойств, главным из которых является возможность выделения из исходного сигнала его деталей различных масштабов. Это позволяет адаптировать изображение к исходной задаче.

Кратномасштабным анализом называется описание пространства $L^2(R)$ через иерархически вложенные подпространства V_m , которые не пересекаются (ортогональность по условию задания), и объединение которых дает в пределе всё $L^2(R)$, т.е.

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots, \quad \bigcap_{m \in \mathbb{Z}} V_m = \{0\}, \quad \bigcup_{m \in \mathbb{Z}} V_m = L^2(R) \quad (2.10)$$

Далее, эти пространства имеют следующее свойство: для любой функции $f(x) \in V_m$, ее сжатая версия будет принадлежать V_{m-1} ,

$$f(x) \in V_m \longleftrightarrow f(2x) \in V_{m-1} \quad (2.11)$$

А так же существует такая функция $\phi(x) \in V_0$, что ее сдвиги $\phi_{0,n}(x) = \phi(x - n)$, $n \in \mathbb{Z}$ образуют ортонормированный базис пространства V_0 .

Из последнего свойства вытекает, что функции $\phi_{m,n}(x) = 2^{-m/2} \phi(2^{-m}x - n)$ образуют ортонормированный базис пространства V_m . Эти базисные функции называются масштабирующими.

Таким образом, любая функция $f(x) \in L^2(R)$ может быть представлена на множестве последовательных ее приближений $f_m(x) \in V_m$. Другими словами, функция $f(x)$ есть предел аппроксимаций $f_m(x) \in V_m$:

$$f(x) = \lim_{x \rightarrow \infty} f_m(x) \quad (2.12)$$

Из определения кратномасштабного анализа следует, что функция $f_m(x)$ есть ортогональная проекция $f(x)$ на подпространство V_m , то есть

$$f(x) = \sum_n \langle \phi_{m,n}, f(x) \rangle \phi_{m,n}(x) = \sum_n c_{m,n} \phi_{m,n}(x) \quad (2.13)$$

Так как $\phi(x) = \phi_{0,0}(x) \in V_0 \subset V_{-1}$, то

$$\phi_{0,0}(x) = \sqrt{2} \sum_n h_n \phi_{-1,n}(x) = 2 \sum_n h_n \phi(2x - n) \quad (2.14)$$

где h_n - некоторая последовательность. Это равенство принято называть масштабирующим уравнением.

Имеют место следующие соотношения:

$$\sum_n h_n = 1 \quad (2.15)$$

$$\delta_{0,k} = \langle \phi_{0,0}(x) \phi_{0,k}(x) \rangle = 2 \sum_n h_n h_{n+2k} \quad (2.16)$$

и в спектральной области $|H(\omega)|^2 + |H(\omega + \pi)|^2 = 1$, что приводит к условиям $H(0) = 1$ и $H(\pi) = 0$.

Введем в рассмотрение еще один объект - пространства W_m :

$$V_{m-1} = V_m \oplus W_m, \quad \bigcap_{m \in \mathbb{Z}} W_m = \{0\}, \quad \overline{\bigcup_{m \in \mathbb{Z}} W_m} = L^2(\mathbb{R}) \quad (2.17)$$

Пусть $\psi(x) = \psi_{0,0}(x)$ - базисная функция W_0 , тогда

$$\psi_{0,0}(x) = \sqrt{2} \sum_n g_n \phi_{-1,n}(x) \quad (2.18)$$

для некоторой последовательности g_n .

Определим семейство вейвлет-функций $\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m}x - n)$. Имеет место соотношение: $0 = \langle \phi_{0,0}, \psi_{0,0} \rangle = 2 \sum_n h_n g_n$. Легко видеть, что выбор $g_n = (-1)^n = h_{-n+2t+1}$ корректен для всех целых t .

Определения вейвлет-функций позволяют записать любую функцию $f(x) \in L^2(\mathbb{R})$ в виде суммы ее проекций на W_j , $j \in \mathbb{Z}$: $f(x) = \sum_{j=-\infty}^{\infty} e_j(x)$

Если осуществить анализ до некоторого масштаба m , то $f_m(x)$ будет

представлена суммой ее грубой аппроксимации $f_m(x) \in V_m$ и множества деталей $e_j \in W_j$:

$$f(x) = f_0(x) + \sum_{j=-\infty}^m e_j(x) = \sum_n c_{m,n} \phi_{m,n}(x) + \sum_{j=-\infty}^m d_{j,k} \psi_{j,k}(x) \quad (2.19)$$

2.2.3 Дискретное вейвлет преобразование

Пусть имеется некоторый дискретный сигнал c_n . Интерпретируем его как коэффициенты разложения некоторой функции $f_0(x) \in V_0$ по базису масштабирующих функций пространства V_0 :

$$f_0(x) = \sum_n c_{0,n} \phi_{0,n}(x) \quad (2.20)$$

где $c_{0,n} = c_n$. Тогда мы можем вычислить аппроксимации этой функции, принадлежащие пространствам V_1, V_2, \dots . Согласно идее кратномасштабного анализа функция $f_0(x)$ раскладывается на сумму:

$$f_0(x) = f_1(x) + e_1(x) = \sum_k c_{1,k} \phi_{1,k}(x) + \sum_k d_{1,k} \psi_{1,k}(x) \quad (2.21)$$

Таким образом мы получили две новые последовательности коэффициентов $c_{1,n}$ и $d_{1,n}$. Этот процесс может быть продолжен разложением $f_1(x), f_2(x), \dots$, и функция $f_0(x)$ (а следовательно и исходный дискретный сигнал c_n) будет представлена совокупностью коэффициентов $d_{m,n}, m \in Z^+, n \in Z$. Так определяются дискретные ряды вейвлетов.

Учитывая, что масштабирующие функции образуют базисы соответствующих пространств, можно показать, что:

$$c_{j,k} = \sqrt{2} \sum_n c_{j-1,n} h_{n+2k} \text{ и } d_{j,k} = \sqrt{2} \sum_n c_{j-1,n} g_{n+2k} \quad (2.22)$$

А обратный преобразование:

$$c_{j-1,n} = \sum_k c_{j,k} h_{n+2k} + \sum_k d_{j,k} g_{n+2k} \quad (2.23)$$

Последовательности h_n и g_n называются фильтрами. Они связаны с базисными функциями $\psi(x)$ и $\phi(x)$:

$$\phi(x) = 2 \sum_k h_k \phi(2x - k) \quad (2.24)$$

$$\psi(x) = 2 \sum_k (-1)^k h_{1-k} \phi(2x - k) = 2 \sum_k g_k \phi(2x - k) \quad (2.25)$$

$$g_k = (-1)^k h_{2n-1-k} \quad (2.26)$$

где $k = 0, 1, 2 \dots 2n-1$, n - порядок вейвлета. Например, для вейвлетов Добеши (подробнее в [27]), решая уравнение 2.25, можно получить следующий фильтр (для 4 точек):

$$h = \left(\frac{1 + \sqrt{3}}{4\sqrt{2}}, \frac{3 + \sqrt{3}}{4\sqrt{2}}, \frac{3 - \sqrt{3}}{4\sqrt{2}}, \frac{1 - \sqrt{3}}{4\sqrt{2}} \right) \quad (2.27)$$

Для ВЧ фильтра можно получить соответственно $g_0 = h_3$, $g_1 = -h_2$, $g_2 = h_1$, $g_3 = -h_0$.

Так как фильтры пропускают только половину всех частотных компонентов сигнала, то не попавшие в полосу прозрачности составляющие могут быть удалены. Поэтому выполняется децимация $\downarrow 2$, т.е. прореживание в два раза.

На стадии реконструкции сигнала осуществляется операция интерполяции $\uparrow 2$, обратная децимации $\downarrow 2$, путем увеличения в два раза числа составляющих добавлением нулевых компонентов вперемежку с имеющимися. При сложении сигналов, полученных на выходе фильтров, будем иметь сигнал, близкий к исходному, т.е. произойдет его реконструкция на

начальном уровне.

2.2.4 Двумерный КМА

На практике двумерный КМА строят как тензорное произведение одномерных КМА. При таком подходе отцовский и материнский вейвлеты будут сформированы следующим образом:

$$\phi(x, y) = \phi(x)\phi(y) \quad (2.28)$$

$$\psi_{LH}(x, y) = \phi(x)\psi(x) \text{ и } \psi_{HL}(x, y) = \psi(x)\phi(y) \quad (2.29)$$

$$\psi_{HH}(x, y) = \psi(x)\psi(y) \quad (2.30)$$

где индексы H и L означают реализацию фильтров ВЧ и НЧ составляющих.

Предположим, что мы имеем изображение размером $N \times N$. первоначально каждая строчка фильтруется на НЧ и ВЧ половины. В результате получаются два изображения размером $N \times N/2$. Далее каждый столбец делится таким же образом. В итоге получаются четыре изображения размером $N/2 \times N/2$. НЧ по горизонтали и вертикали (НЧНЧ₁), ВЧ по горизонтали и вертикали (ВЧВЧ₁), НЧ по горизонтали и ВЧ по вертикали (НЧВЧ₁) и ВЧ по горизонтали и НЧ по вертикали (ВЧНЧ₁). Первое из указанных выше изображений делится аналогичным образом на следующем шаге (уровне) преобразования. Это называется пирамидальным алгоритмом Маала. На Рис. 2.4 показан пример разложения изображения скоплений NGC 869 и NGC 884 на коэффициенты 1 уровня с помощью вейвлета Хаара (для удобства вейвлет коэффициенты нормированы по яркости, выполнено в среде OCTAVE, [35]).

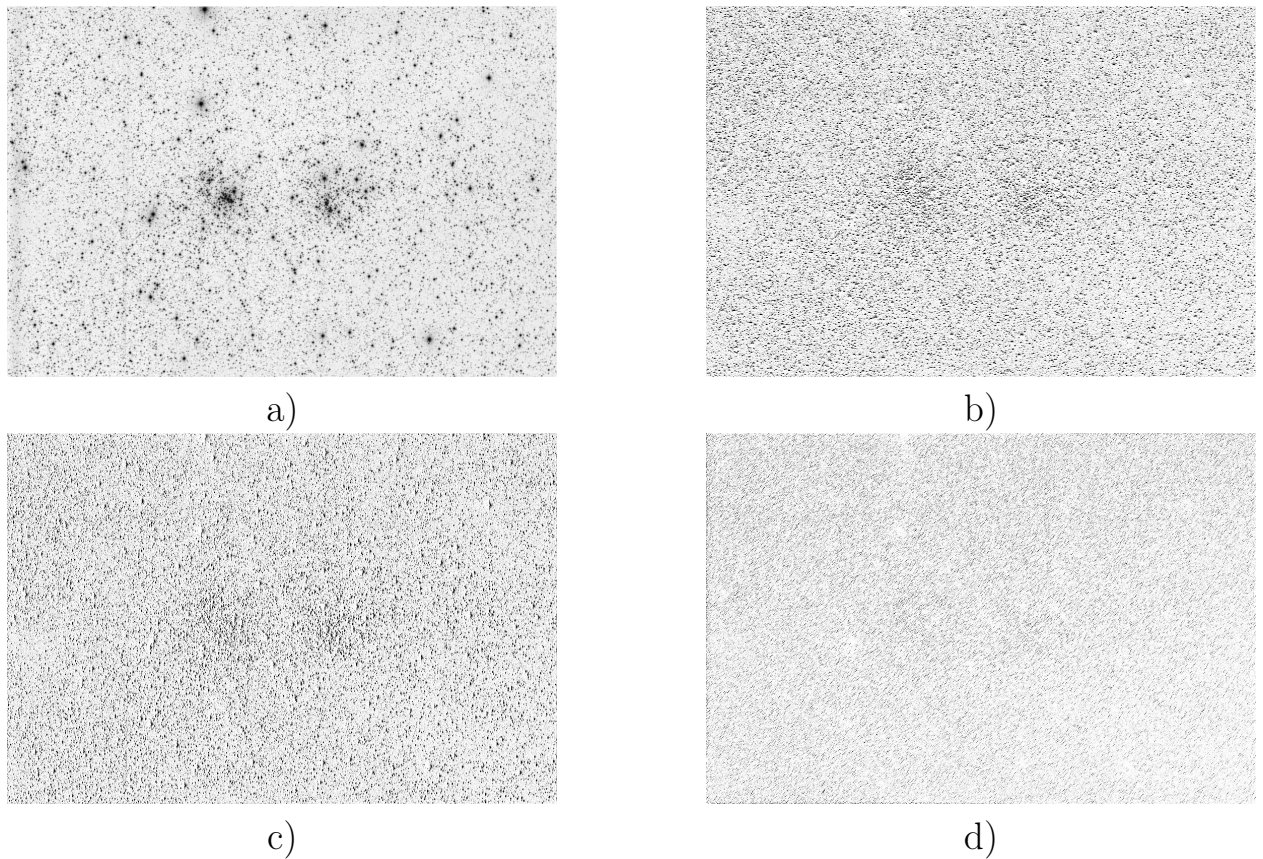


Рис. 2.4: Коэффициенты пирамидального алгоритма Маала: а) Аппроксимация исходного изображения 1-ого уровня (HCH_1) б) HCH_1 в) VCH_1 д) VCH_1

2.2.5 Альтернативные алгоритмы вейвлет преобразований

С практической точки зрения существуют более адаптированные для астрономических задач методы, чем алгоритм Маала. Например, trous-вейвлет преобразование (название связано с особым способом свертки). Оно является многоуровневым: на выходе N изображений такого же размера как и исходное (N задано заранее), причем каждое из них отображает уровень исходного изображения - от высокочастотных составляющих до самых низкочастотных, суммируя все уровни получаем исходный снимок.

Это позволяет фильтровать данные на всех уровнях, а затем просто складывать все уровни преобразования. Подробнее о реализации и тестировании будет рассказано в следующей главе.

Глава 3

Trous-вейвлет преобразование

Для реализации алгоритмов, представленных в этой главе, была выбрана среда OCTAVE. Однако, использованы только стандартные алгебраические библиотеки, поэтому в случае необходимости методы можно легко перенести на другой язык (например, Python+Numpy).

3.1 Trous-вейвлет алгоритм

Основной идеей trous-вейвлет преобразование является отсутствие децимации, то есть пропуская сигнал через НЧ и ВЧ фильтры, мы не прореживаем его, а заполняем "дырки" (отсюда и название с французского "algorithme a trous"). В работе был использован алгоритм, изложенный в [11].

На выходе у данного метода J изображений одинакового размера (того же как исходное), где J задано заранее. Каждое изображение с номером j будет называться j уровнем и обозначаться w_j . Точку уровня j с номером l обозначим как $w_{j,l}$. Уровни упорядочены так, что первый отвечает за самую высокочастотную составляющую изображения, а последний за низкочастотную. Введем исходный сигнал $\{c_{0,l}\}$ и масштабирующую функцию

$\phi(x)$, которой соответствует НЧ фильтр.

Вейвлет-функция удовлетворяет масштабирующему уравнению 2.14:

$$\frac{1}{2}\psi\left(\frac{x}{2}\right) = \sum_k g(k)\phi(x-k) \quad (3.1)$$

Коэффициенты дискретного вейвлет преобразование вычисляются как:

$$w_{j+1,l} = \sum_k g(k)c_{j,l+2^j k} \quad (3.2)$$

В итоге каждый уровень вейвлет преобразования представляет собой разность сигналов:

$$w_{j+1} = c_{j,l} - c_{j+1,l} \quad (3.3)$$

Разность сигналов $\{c_{0,l}\} - \{c_{1,l}\}$ содержит информацию о разности масштабов и этот ряд связан с вейвлет преобразованием, которое соответствует $\phi(x)$. Соответствующая вейвлет-функция определяется, как

$$\frac{1}{2}\psi\left(\frac{x}{2}\right) = \phi(x) - \frac{1}{2}\phi\left(\frac{x}{2}\right) \quad (3.4)$$

В итоге сигнал вычисляется, как

$$c_{j,l} = \sum_k h(k)c_{j-1,l+2^j k} \quad (3.5)$$

Таким образом происходит свертка предыдущего уровня с фильтром, при этом расстояние между пикселями в свертке увеличивает как степень двойки.

Коэффициенты $\{h(k)\}$ определяются через масштабирующую функцию $\phi(x)$:

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_l h(l)\phi(x-l) \quad (3.6)$$

Полный алгоритм trous-вейвлет преобразования выглядит так:

1. j инициализируется нулем, а $c_{0,l}$ начальным сигналом.
2. Вычисляем свертку уровня j с фильтром h . Расстояние между пикселями в свертке равно 2^j .
3. Вычисляем коэффициенты дискретного вейвлет-преобразования, как разность между соседними уровнями:

$$w_{j+1,l} = c_{j,l} - c_{j+1,l} \quad (3.7)$$

4. Если j не достигло заданного значения, увеличиваем j на единицу и переходим к шагу 2..
5. Множество $W = w_1, w_2, \dots, w_J, c_J$ является вейвлет преобразованием исходного сигнала.

Можно легко осуществить обратное преобразование:

$$c_{0,l} = c_{J,l} + \sum_j^J w_{j,l} \quad (3.8)$$

В работе в качестве масштабирующей функции использовался кубический $B_3(x)$ -сплайн (рис. 3.1):

$$\phi(x) = B_3(x) = \frac{1}{12}(|x - 2|^3 - 4|x - 1|^3 + 6|x|^3 - 4|x + 1|^3 + |x + 2|^3) \quad (3.9)$$

В этом случае коэффициенты матрицы свертки равны в одномерном слу-

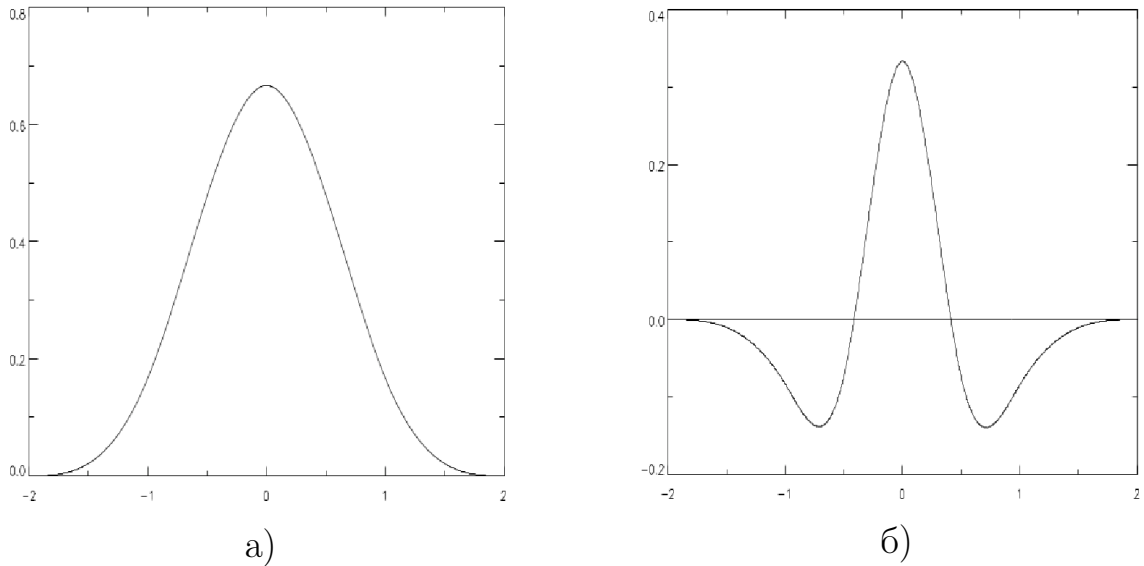


Рис. 3.1: а) Масштабирующая функция $\phi(x)$ (кубический сплайн) б) Соответствующая вейвлет функция $\psi(x)$

чае $(\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16})$, а в двумерном:

$$(1/16, 1/4, 3/8, 1/4, 1/16) \otimes \begin{pmatrix} \frac{1}{16} \\ \frac{1}{4} \\ \frac{3}{8} \\ \frac{1}{4} \\ \frac{1}{16} \end{pmatrix} \quad (3.10)$$

У данного алгоритма есть недостатки - отрицательные значения коэффициентов и отсутствие устойчивости. Среднее значение вейвлет коэффициентов - ноль. Все уровни засоряются шумами и при фильтрации можно отсечь значимые коэффициенты, что приводит при реконструкции к изменению яркости объектов. Поэтому для выполнения работы так же рассматривались многоуровневые преобразования, основанные не на вейвлет-анализе. Например, многоуровневое медианное преобразование, основанное на соответствующем нелинейном НЧ фильтре.

Обозначим исходное изображение как f , S - заданное количество уровней разложения, тогда точный алгоритм выглядит так:

1. Обозначим $c_j = f$, где $j = 1$. Пусть $s = 1$.
2. Следующий уровень c_{j+1} определяется как результат медианной фильтрации c_j с окном $2s + 1$.
3. Многоуровневые коэффициенты w_{j+1} определяются разницей уровней:

$$w_{j+1} = c_j - c_{j+1}.$$
4. Перейдем на следующий уровень $j = j + 1$, увеличим окно в два раза $s = 2s$. Возвращаемся ко второму шагу, если $j < S$.

Исходный сигнал восстанавливается аналогично trous-вейвлет преобразованию:

$$f = c_p + \sum_j w_j \quad (3.11)$$

Здесь коэффициенты w_j не близки к нулю, поэтому яркость восстановленного изображения будет соответствовать исходной.

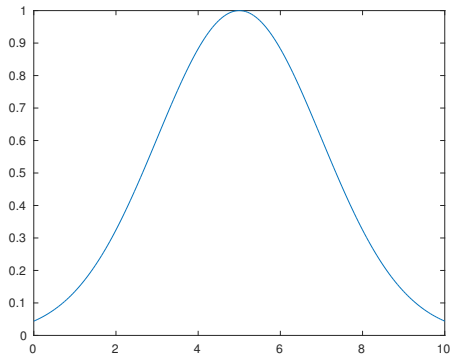
Представленные в данной главе методы декомпозиции являются многоуровневыми, то есть они переводят одно изображение в набор из $J + 1$ изображений, каждое из которых отвечает за свой масштаб исходного снимка. Рассмотрим наглядно примеры таких преобразований.

3.2 Одномерные многоуровневые преобразования

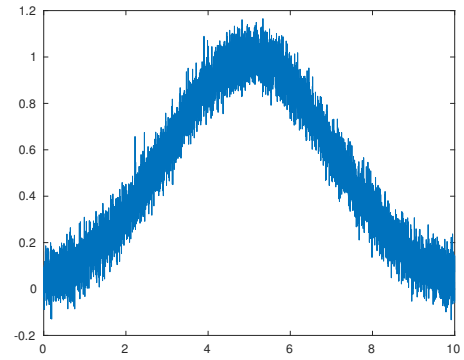
Начнем с одномерного случая. Смоделируем нормальное распределение вида:

$$f(x, \sigma, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.12)$$

Возьмем $N = 10000$ точек из промежутка $[0, 10]$. Пусть $\sigma = 2$, $\mu = 5$. Добавим аддитивный белый гауссовский шум к получившемуся сигналу (Рис. 3.2).



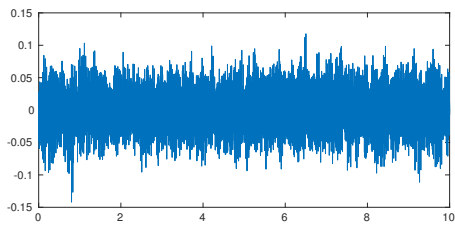
а)



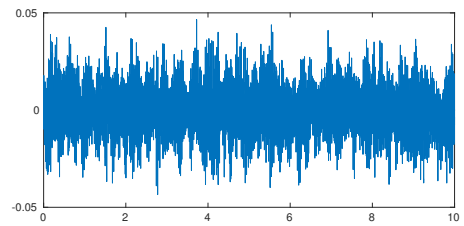
б)

Рис. 3.2: а) Исходное распределение б) Распределение с добавлением белого шума

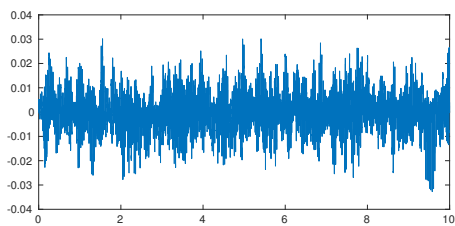
Результатом многоуровневого преобразования исходного распределения будут 6 графиков, каждый из которых отвечает за свой масштаб изображения (Рис. 3.3, 6 масштабов).



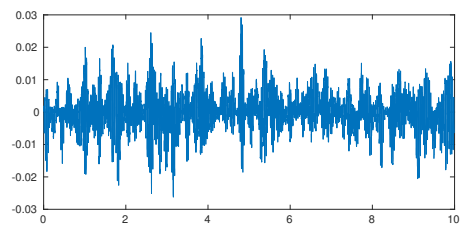
а)



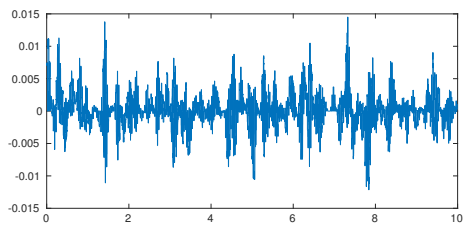
б)



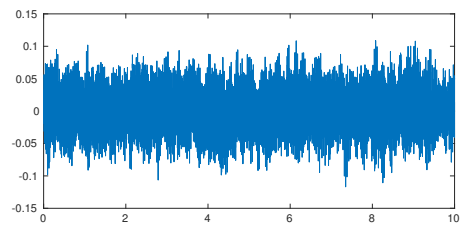
в)



г)



д)



е)

Рис. 3.3: а) - д) - масштабы распределения, начиная от самых высоких частот, е) - сумма всех шумовых компонент

На Рис. 3.4 изображен 6 уровень преобразования (самая НЧ компонента).

Сумма всех уровней в точности представляет собой исходное распределение. Интересно что, частоты уровней белого шума отличаются, это может быть связано с работой генератора случайных чисел при моделировании.

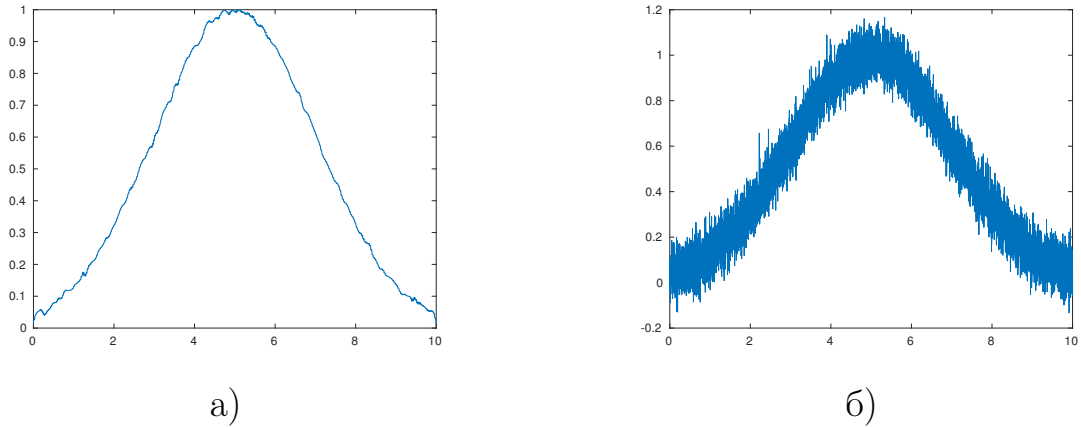


Рис. 3.4: а) 6 гармоника преобразования б) Сумма всех уровней

Как видно из 3.3 и 3.4, используя trous-вейвлет алгоритм и подобные ему методы, можно рассматривать каждый масштаб отдельно в пространственных координатах. Это удобнее с практической точки зрения. Для того чтобы получить исходную гладкую кривую, можно сгладить НЧ масштаб, предварительно убрав все остальные гармоники.

Рассмотрим пример экстракции линий в спектре ультраяркого рентгеновского источника (NGC 4395, ULX-1, Рис. 3.5)

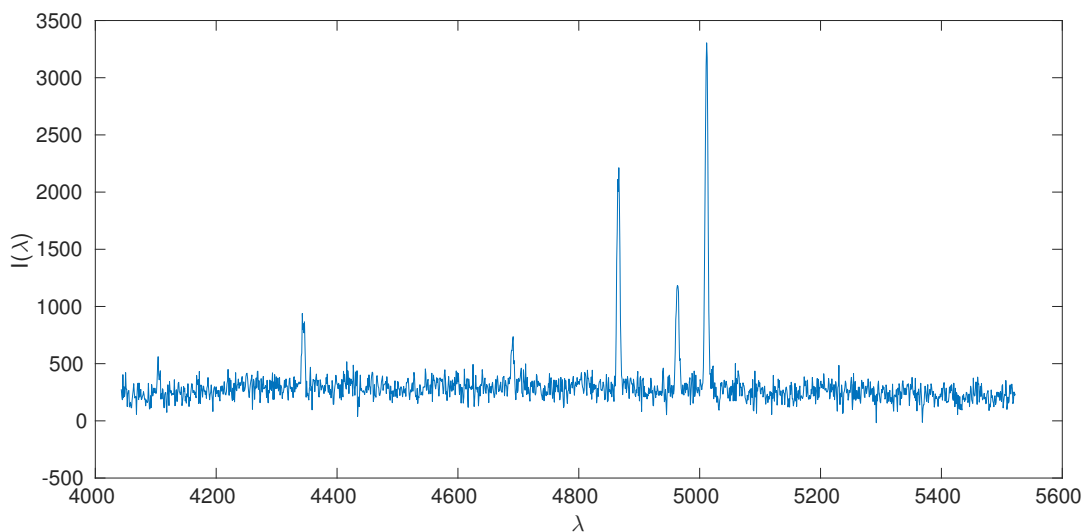


Рис. 3.5: Спектр ULX-1, NGC 4395

Аналогично модельному решению, разделим спектр на 6 гармоник. Сразу отбросим самые высокочастотные и низкочастотные части (1 и 6).

Пусть $w_{j,l}$ - коэффициенты J уровня ($J = 2 \dots 5$). Для каждого J выберем порог, отвечающие пикам в спектрах и ниже которого обнулим все вейвлет-коэффициенты уровня. После этого сложим обратно все J уровни преобразования (Рис. 3.6). Так как мы намеренно не суммировали самую ВЧ гармонику (Рис. 3.7) в итоге $I(\lambda)$ изменилась примерно на 7 – 8%.

Примеры, представленные в этом разделе были вычислены с помощью многоуровневого медианного преобразования. Существенный недостаток его с по сравнению с trous-вейвлет алгоритмом заключается во времени работы, как правило это $O(n^2 \log(n))$ в отличии от $O(n \log(n))$ у вейвлет преобразования. Если в одномерном случае для $n = 10^4 - 10^5$ это не так критично, то для изображений, где число элементов больше 10^8 это существенная разница (об этом в следующем разделе).

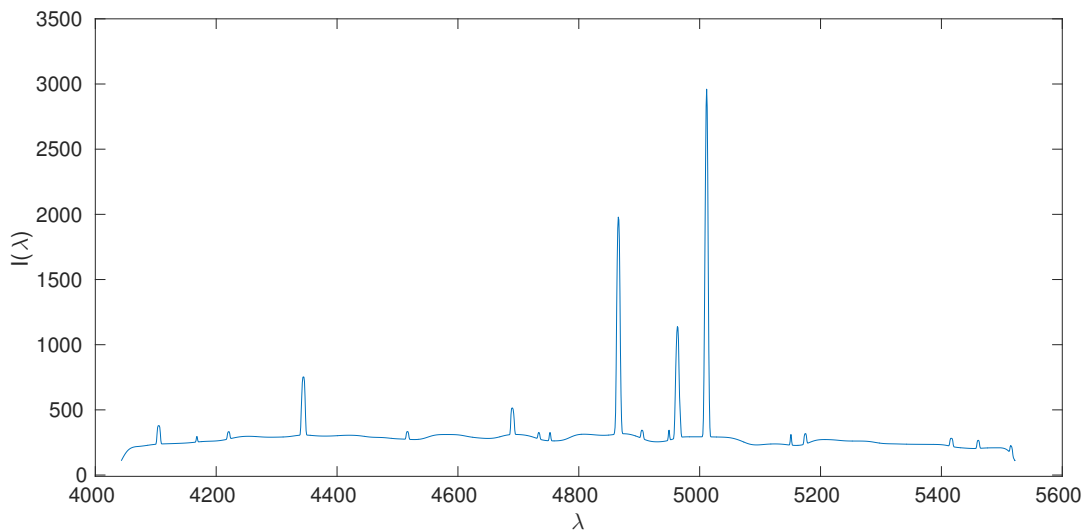


Рис. 3.6: Отфильтрованный спектр ULX-1, NGC 4395

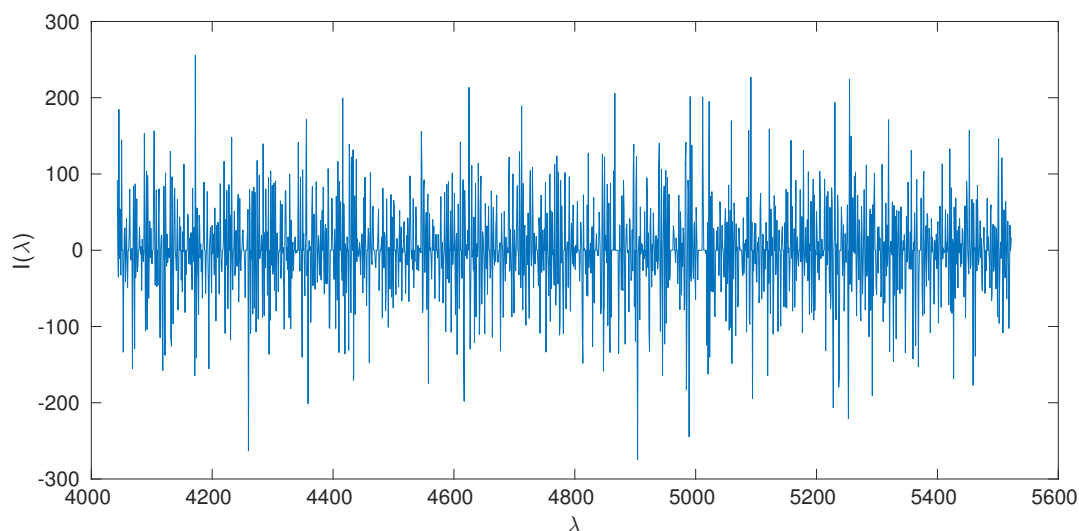


Рис. 3.7: ВЧ компонента спектра

3.3 Двумерные многоуровневые преобразования

Аналогично смоделируем гауссово распределение для двумерного случая и заполним поле белым шумом (Рис. 3.8, размер 500×500 пикселей). Для того чтобы выделить скопления можно воспользоваться бинаризацией по маске (выбрать на НЧ гармонике маску, далее бинаризовать по этому значению, все что ниже этого значения будет приравнено к 0). На рис. 3.8 представлено trous-вейвлет разложение (с использованием B_3 -сплайна в качестве масштабирующей функции). Многоуровневое медианное преобразование показывает аналогичные результаты, средние значения коэффициентов в этих алгоритмах почти равны, это связано с отсутствием сильной ВЧ зашумленности при моделировании.

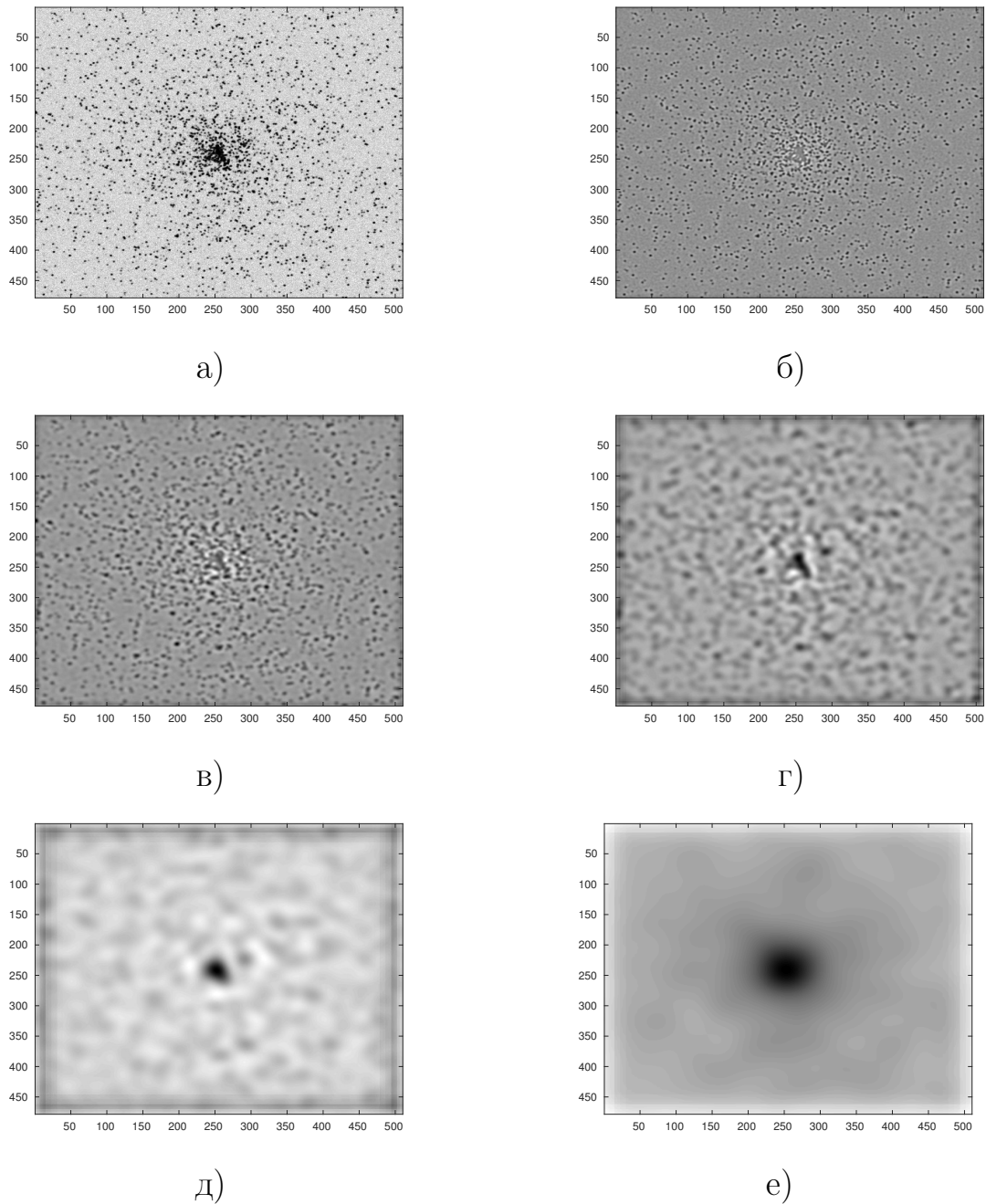


Рис. 3.8: а) Исходная модель, б) - е) Гармоники, начиная с НЧ, масштабированные по яркости, в сумме дают первоначальное распределение

Пусть $W_1 \dots W_k$ - набор вейвлет коэффициентов каждого уровня, f - исходное изображение. Сравним преобразования по нескольким критериям:

1. Средние значения вейвлет коэффициентов на каждом уровне J : $\langle w_{j,k} \rangle$. Для trous-вейвлет преобразования среднее значение коэффициентов 0. Это может приводить к неустойчивости метода. Например, нарушению интенсивности объектов на итоговом изображении.

2. Разница между суммой вейвлет-шкал и интенсивностью исходного изображения: $\langle \sum_k W_k - f \rangle$. Это проверка первого критерия.
3. Скорость работы в секундах. Важный параметр, особенно на больших N ($> 10^8$).

В качестве примеров снимков объектов будем использовать NGC 2997 (DFOSC, Danish 1.54m telescope, 27/02/98, размер 2036×2042 пикселей) и NGC 5194 (Kitt Peak National Observatory, 2.1m, 2001/03/28, 2013×2822). Исходные изображения на Рис. 3.9.

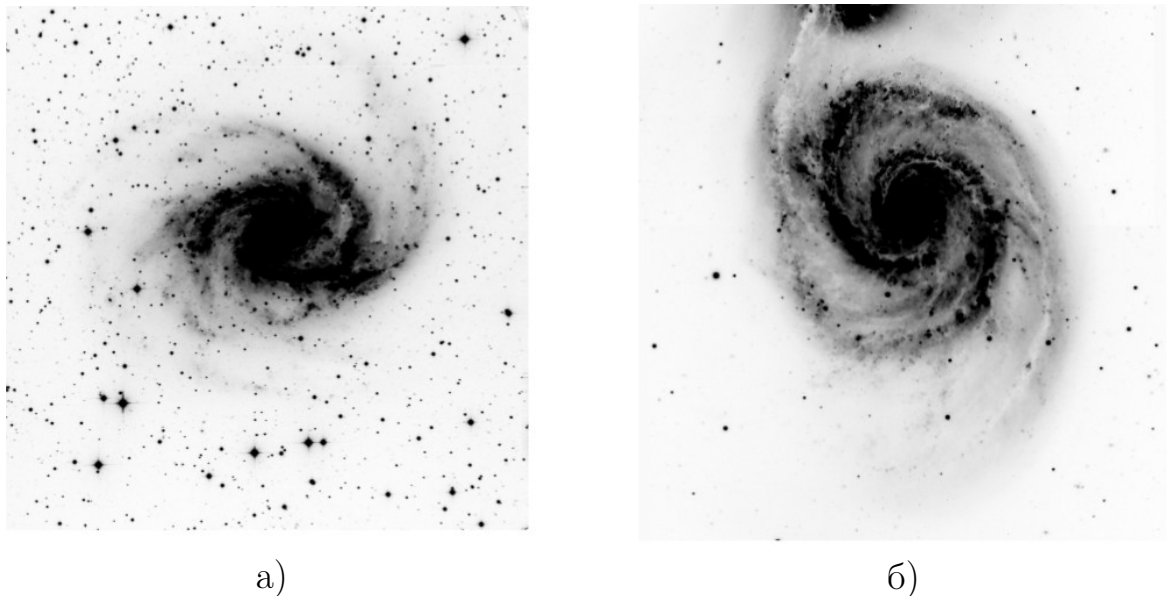


Рис. 3.9: а) NGC 2997 б) NGC 5194 (M51, Галактика Водоворот)

Выполним декомпозицию этих снимков на $J = 6$ уровней. Все уровни разложения представлены в Приложении А для вейвлет-преобразования и в Приложении Б для медианного метода.

Заметим, что в случае trous-вейвлет преобразования (Рис. А.2 и А.1) в местах ярких объектов на масштабах $J = 1 \dots 5$ появляются "провалы" (нулевая интенсивность в окрестностях), которые заполняются при суммировании с НЧ гармоникой. Для изображения NGC 2997 метод является устойчивым: все $\langle w_{j,k} \rangle$ больше 0.2 для $j = 1$ (чем больше уровень, тем больше среднее значение коэффициентов), суммирование гармоник да-

ет в точности исходное изображение: $\langle \sum_k W_k - f \rangle = 0$ для всех элементов W_k . В случае NGC 5194 среднее значение коэффициентов равно $\langle w_{1,k} \rangle = 1.4 \cdot 10^{-6}$, однако $\langle \sum_k W_k - f \rangle = 10^{-6}$, то есть получаем практически абсолютные точные интенсивности для всех элементов. В среде MATLAB вейвлет преобразование занимает 1.113 секунд для NGC 2997 и 1.141 секунд для NGC 5194 (core i3-2330M, 2.2GHz, 4 GB RAM). Скорость работы одно из главных преимуществ этого подхода.

Многоуровневый медианный алгоритм (Рис. Б.2 и Б.1) устойчив, т.к. коэффициенты на всех уровнях больше нуля и для всех элементов: $\langle \sum_k W_k - f \rangle = 0$. Преобразование выполняется за 163.294 секунды для NGC 2997 и 247.211 секунд для NGC 5194. Время выполнения можно сократить, используя быстрые реализации медианного фильтра (в работе использовался базовый алгоритм).

Оба алгоритма показывают себя хорошо, однако у каждого есть свои недостатки: возможная неустойчивость и провалы на ярких объектах у вейвлет-преобразования и медленная работа медианного метода. Несмотря на это, они вполне применимы на практике.

Глава 4

Многоуровневая обработка астрономических снимков

В этой главе рассматривается подход к фильтрации изображений, используя многоуровневые разложения, построенные с помощью trous-вейвлет преобразования или медианного алгоритма.

4.1 Модель сигнал-шум

Пусть $I(x, y)$ — исходное изображение. Тогда его можно представить в виде структуры:

$$I(x, y) = \sum_{i=1}^n O_i(x, y) + F(x, y) + B(x, y) \quad (4.1)$$

где $O_1 \dots O_n$ — набор из n объектов снимка, $F(x, y)$ — фон, $B(x, y)$ — шумовая компонента. С другой стороны изображение можно представить как сумму вейвлет шкал:

$$I(x, y) = \sum_{j=1}^J W_j \quad (4.2)$$

Здесь W_j — коэффициенты j масштаба, J - количество уровней разложения. Все наборы коэффициентов $\{W_1 \dots W_J\}$ определены в пространственных координатах.

Как уже упоминалось ранее, большинство снимков сильно зашумлены, это приводит к переполнению всех уровней коэффициентами, которые не являются значимыми (т.е. не несут информацию об объектах изображения). Основная идея подхода к фильтрации снимка заключается в последовательной очистке каждого уровня от элементов, которые относятся к шумовой компоненте. Затем изображение восстанавливается при сложении всех масштабов.

Многоуровневые разложения являются мощным инструментом в обработке астрономических изображений, они позволяют легко анализировать любые масштабы снимка. Ниже будут предложены несколько вариантов применения такого подхода.

4.1.1 Soft and hard thresholding

Hard thresholding — алгоритм фильтрации изображения по пороговому значению. Все вейвлет коэффициенты ниже определенного значения T_j приравниваются к 0 (считаются незначимыми):

$$\widetilde{w}_{j,k} = \begin{cases} w_{j,k} & \text{если } |w_{j,k}| \geq T_j \\ 0 & \text{если } |w_{j,k}| < T_j \end{cases} \quad (4.3)$$

где $w_{j,k}$ - вейвлет коэффициент масштаба j в позиции k .

Soft thresholding предполагает замену каждого коэффициента на значение $\widehat{w}_{j,k}$:

$$\widehat{w}_{j,k} = \begin{cases} \text{sign}(w_{j,k})(w_{j,k} - T_j) & \text{если } |w_{j,k}| \geq T_j \\ 0 & \text{если } |w_{j,k}| < T_j \end{cases} \quad (4.4)$$

В случае Гауссовского шума $T_j = K\sigma_j$, j - уровень масштаба, σ_j — отклонение от шума на шкале j . Значение K чаще всего принимается за 3.

Рассмотрим снимок Юпитера (HST, 2007-04-27, 6730Å). Заполним изображение равномерным Гауссовским шумом. Разложим снимок на 6 гармоник при помощи медианного преобразования и применим фильтрацию по порогу с 4σ к каждому масштабу. На Рис. 4.1 видно, что алгоритм показывает хорошие результаты: изображение немного потеряло в детализации, но границы сохранены (большое преимущество в сравнении с сглаживающими фильтрами типа Гаусса).

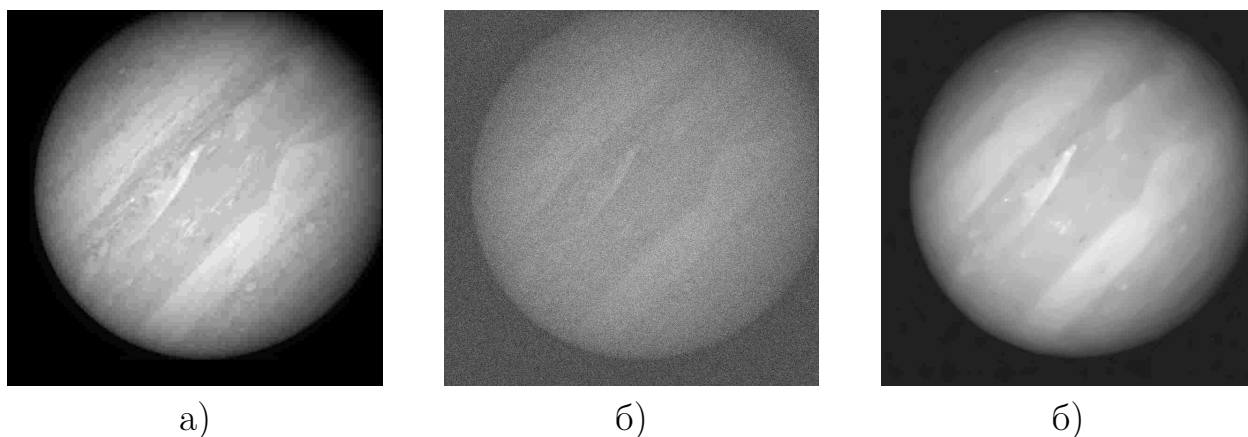


Рис. 4.1: а) Оригинальный снимок б) Зашумленное изображение в) Изображение после применения пороговой обработки

4.1.2 3σ -clipping и винеровская фильтрация

Одним из самых эффективных методов по экстракции фона изображения является 3σ -clipping, этот подход используется в программном обеспечении DAOPHOT (Stetson's DAOPHOT program, [6]). Снимок итеративно делится на части, пока элементы находящиеся в $\pm 3\sigma$ от среднего значения (Median) не меняются больше, чем на 20% за проход. Тогда можно сказать, что поле не является переполненным звездами ("crowded fields") и вычислить среднее значение фона (Mean). Отфильтрованное изображение представляется в виде:

$$\text{Image} = 2.5 \times \text{Median} - 1.5 \times \text{Mean} \quad (4.5)$$

На практике используют формулу вида:

$$\text{Image} = 3 \times \text{Median} - 2 \times \text{Mean} \quad (4.6)$$

Аналогичный подход использован для выделения фона в SExtractor (Bertin E., S. Arnouts, [3]).

Так же хорошие результаты показывает алгоритм адаптивной винеровской фильтрации для подавления аддитивного гауссова белого шума. Данный метод основан на статистических оценках фрагментов изображения в пределах скользящего окна размера $m \times n$ пикселей (подробнее в [31], [32]).

Для всех положений скользящего окна с центральным пикселем и координатами (x, y) среднее значение яркости равно:

$$\mu = \frac{1}{mn} \sum_{(x,y) \in \eta} I(x, y) \quad (4.7)$$

где η — окрестность $m \times n$ вокруг каждого пикселя изображения I . Дисперсия вычисляется как:

$$\sigma^2 = \frac{1}{mn} \sum_{(x,y) \in \eta} I^2(x, y) - \mu^2 \quad (4.8)$$

Данная формула применяется для всех положений окна.

Отфильтрованное изображение записывается в виде:

$$F(x, y) = \mu + \frac{\sigma^2 - \nu^2}{\sigma^2} (I(x, y) - \mu) \quad (4.9)$$

Здесь ν^2 — мощность белого шума. Можно оценить ν как среднее по всем σ на снимке.

Приведем пример оценки самых мелкомасштабных колебаний интенсивности фона снимка. Рассмотрим изображение M33 (Рис. 4.2, KPNO 4.0m telescope, 2000-10-04, Massey, фильтр V 5386.76Å, размер снимка 8278 × 8384)

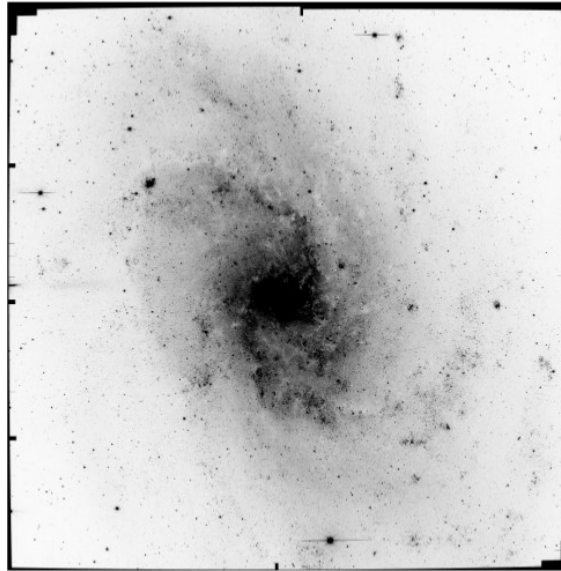


Рис. 4.2: M33, 5386.76Å, Massey, ds9 scale 92.5%

В Приложении В приведены разложения снимка по 6 гармоникам с помощью изложенных выше алгоритмов. Рассмотрим самый высокочастотный масштаб: обнулим очевидные шумы ($< 2 \times \text{Mean}$) и применим винеровскую фильтрацию с окном 5×5 , оценивая мощность шума, как среднее отклонение по участку. В итоге на Рис. 4.3 видны ВЧ гармоники и фон изображения. На изображении стало заметно больше деталей - наблюдается спиральная структура вблизи ядра, ярко выраженные рукава галактики. Интересная особенность - после фильтрации видно, что снимок разделен на 8 частей (в шапке FITS-файла указано: Number of CCDs = 8).

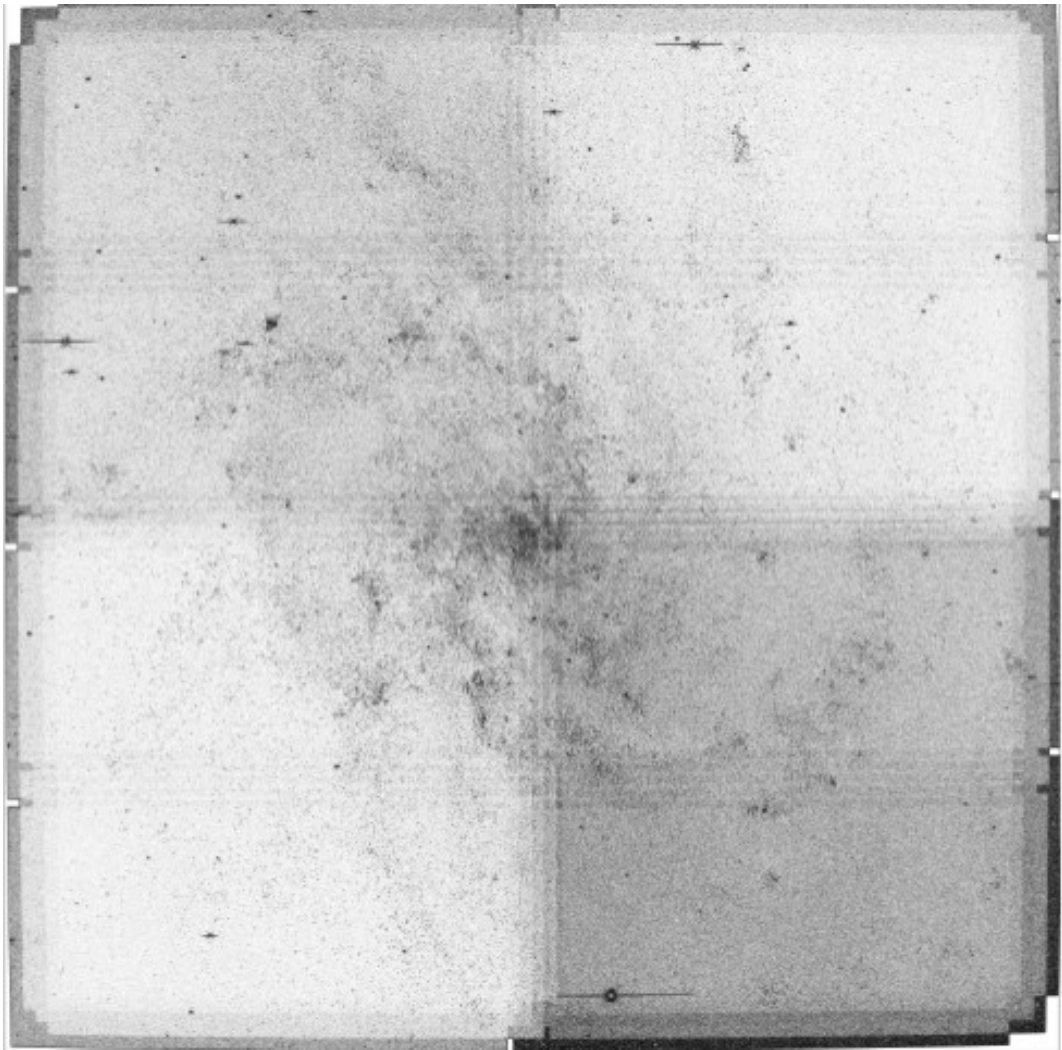


Рис. 4.3: МЗЗ, ВЧ гармоника, ds9 scale 92.5%

Глава 5

Многоуровневое представление данных

Одно из основных применений математических алгоритмов попиксельной обработки в астрономии - это поиск и отождествление объектов на снимках с целью составления каталогов с основными параметрами (видимая звездная величина, размеры, координаты) и их дальнейшей классификации. Ниже будет представлен подход к идентификации объектов на изображении при помощи многоуровневого представления данных.

5.1 Описание многоуровневой модели

Пусть f — исходное изображение. Построим его бинарную модель из многоуровневого представления: если на уровне j содержится информация в пикселе с координатами (i, k) , то его значение $M_{i,k,j} = 1$, остальные элементы приравниваются к 0. Таким образом модель строится следующим образом:

1. Преобразование исходного изображения f с помощью многоуровневого алгоритма в структуру вида $\{w_1, \dots, w_j\}$, где w_k — набор коэффициентов для масштаба k .

2. Определение значимых элементов на каждом уровне.
3. Бинаризация всех уровней по определенному порогу значимости.

В итоге многоуровневая модель данных (Multiscale Vision Model — MVM) представляет собой следующий набор структур:

- Наборы значимых вейвлет коэффициентов $\{W_1, \dots, W_j\}$, определяемых при помощи алгоритмов, приведенных в Главе 4.
- Структуры $S_{j,k}$, которые представляют собой наборы значимых вейвлет коэффициентов на одной шкале j :

$$S_{j,k} = \{w_j[k_1, l_1], w_j[k_2, l_2], \dots, w_j[k_p, l_p]\} \quad (5.1)$$

где p — число значимых коэффициентов, входящих в структуру $S_{j,k}$,
 w_{j,x_i,y_i} — вейвлет коэффициент шкалы j в позиции (x_i, y_i) .

- Объекты — наборы структур $S_{j,k}$:

$$O_l = \{S_{j_1,k_1}, \dots, S_{j_n,k_n}\} \quad (5.2)$$

Так же определим оператор \mathcal{L} , отображающий принадлежность структуры данному объекту: $\mathcal{L}(S_{i,j}) = l$, если $S_{i,k} \in O_l$ и $\mathcal{L}(S_{i,j}) = 0$ в остальных случаях.

- Уровень масштаба объекта, который определяется уровнем вейвлет коэффициента с максимальным значением.
- Набор связей между структурами: набор правил по которым можно отнести несколько структур к одному объекту.
- Подобъекты, которые имеют свои и максимумы в вейвлет пространстве, но сами являются частью более крупного объекта.

5.2 Поиск объектов и структур

После применения многоуровневого представления снимка мы получим несколько бинарных изображений. Как уже было оговорено ранее, если вейвлет коэффициент в позиции (x, y) является значимым, то соответствующее ему значение пикселя равно 1, в остальных случаях 0. Последовательности ненулевых пикселей i образуют структуры $S_{i,j}$ на масштабе j .

Любой объект представляет собой набор иерархических структур. Правила, которые определяют какие структуры входят в объект определяются набором связей между структурами (см. Раздел 5.1). Если пиксель с максимальным значением структуры $S_{i,j}$ входит в структуру $S_{i,j+1}$, тогда эти структуры связаны. Полные наборы связанных структур являются объектами. Данная система наглядно представлена на Рис. 5.1. Много объектов и связанных структур появляется из-за зашумленности данных, поэтому важно применять эффективную фильтрацию к каждому уровню. На Рис. 5.2 изображена структура суммы с 1 по 5 уровня разложения снимка МЗЗ из Приложения В и отдельно пятая гармоника. Можно видеть как объекты по отдельности входят в структуры.

В вейвлет пространстве решения о разделении и слиянии структур, определяется по локальным максимум на различных масштабах. Структура определяет объект на уровне j , если выполняется условие: $w_j^m > w_{j-1}^m$ и $w_j^m > w_{j+1}^m$, где:

- w_j^m — максимальный вейвлет коэффициент структуры j
- $w_{j-1}^m = 0$, если коэффициент не связан ни с какой структурой масштаба $j - 1$ или максимальное значение $S_{j-1,l}$, где $\mathcal{L}(S_{j-1,k'}) = \mathcal{L}(S_{j,k})$.
- $w_{j+1}^m = \max \{w_{j+1,x_1,y_1}, \dots, w_{j+1,x_n,y_n}\}$, где $w_{j,x,y}$ — коэффициенты $S_{j,k}$ (все $w_{j,x,y} \in S_{j,k}$).

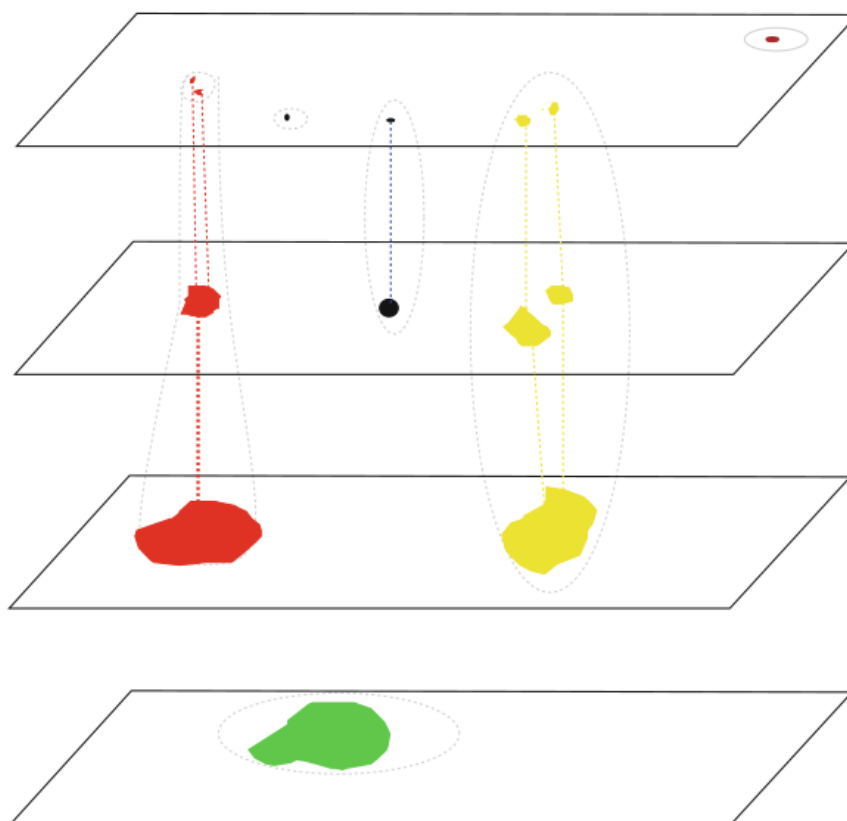
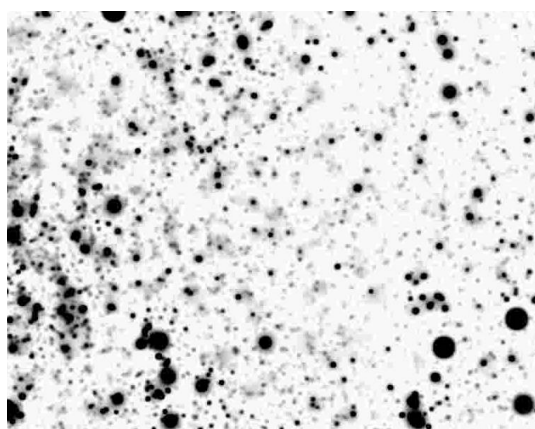
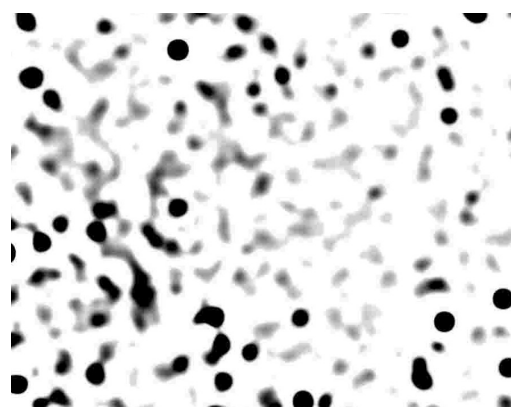


Рис. 5.1: Схематическое представление связности структур. Пример взят из работы [14]



а)



б)

Рис. 5.2: а) сумма пяти гармоник разложения М33 б) Пятый уровень преобразования

5.3 Алгоритм поиска объектов

Метод поиска объектов на снимках можно разделить на несколько этапов:

1. Выполнить декомпозицию изображения, получив набор коэффициентов $\{w_1, \dots, w_j\}$ для каждого уровня.
2. Определить отклонение шума на каждом масштабе j .
3. Определить пороговые значения для шума на каждом уровне.
4. Отдельно выполнить фильтрацию для каждого масштаба.
5. Определить межуровневые отношения по бинаризованным уровням.
6. Найти все максимальные значения коэффициентов для каждой структуры уровня.
7. Вычислить все связанные структуры и разделить их.

5.4 SExtractor

SExtractor (Source-Extractor, Bertin E., S. Arnouts, [3]) — программное обеспечение для идентификации объектов на астрономических снимках и составления каталогов. Алгоритм экстракции объектов, использованный авторами, представляет собой некоторое подобие многоуровневого подхода (схематично метод изображен на Рис. 5.3):

- Large Image — исходный снимок в FITS-формате.
- Background Algorithm — использован 3σ -clipping алгоритм для экстракции фона изображения (описан в разделе 4.1.2, использована формула 4.6)
- Thresholding включает в себя свертку с фильтром (SExtractor хорошо оптимизирован для свертки матриц с маленьким ядром) и определение порогового значения для объектов (подробнее об этом ниже)

- Deblending — разделение структур на источники, используя граф связности и априорные представления об искомым объектах. Размеры объектов вычисляются по бинарной карте.
- Photometry — фотометрия выделенных объектов.

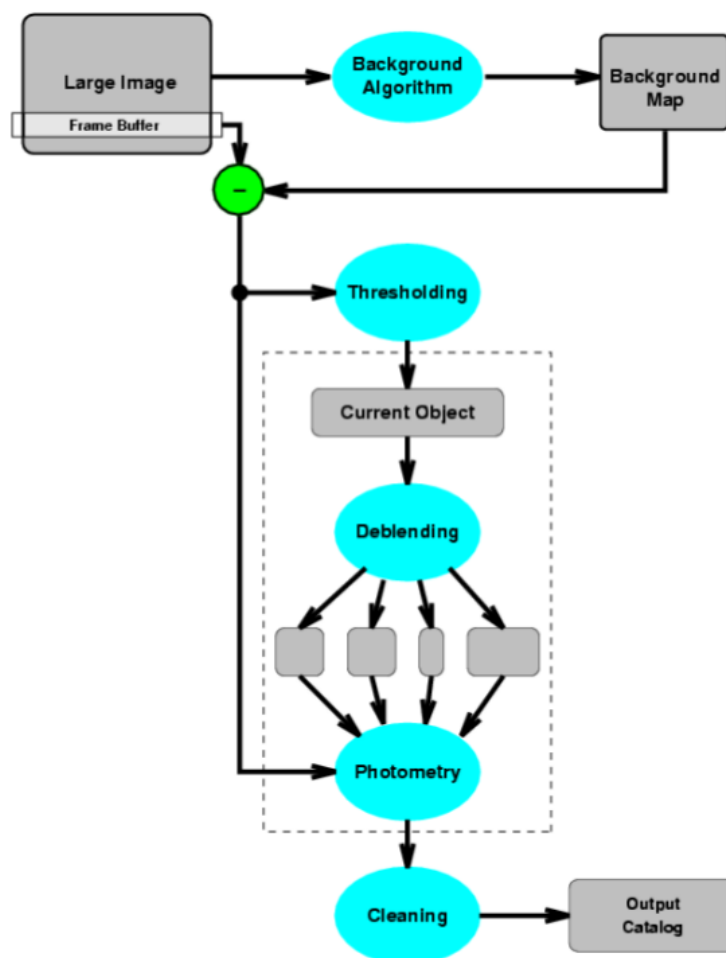


Рис. 5.3: Алгоритм работы Source-Extractor, Bertin E., S. Arnouts, [3]

Перед определением пороговых значений выполняется свертка изображения с одним из следующих фильтров:

- Набор сглаживающих фильтров Гаусса (3×3 , 5×5 и 9×9). Применяются для идентификации объектов с $1.5 < \text{FWHM} < 5$.
- Top-hat — морфологическая фильтрация изображения, вычитает замкнутое изображение (морфологическое замыкание) из исходного. Используется для поиска слабых протяженных объектов ($\text{FWHM} < 5$).

- Mex-hat — преобразования при помощи набора вейвлет функций типа "мексиканская шляпа". Применяется в переполненных звездами полях.
- Small pyramidal function - быстрая свертка (по умолчанию).

В зависимости от типа искомых объектов, может быть выбран узкий фильтр (например, для звезд это PSF), либо широкий (для небольших скоплений матрица Гаусса 9×9 , для туманностей хорошие результаты показывает Top-hat). Таким образом, комбинацией разных фильтров можно выделить все структуры снимка, в этом заключается некое подобие многоуровневого подхода, примененного в SExtractor.

Настройка экстракции производится через конфигурационный файл, который содержит в себе такие основные параметры, как:

- BACK_SIZE — размер элемента сетки при выделении фона.
- DETECT_THRESH — пороговое значение для определения значимости пикселей.
- DETECT_MINAREA и DETECT_MAXAREA — границы размеров определяемых объектов.
- FILTER_NAME и FILTER — выбор типа фильтра для свертки.
- DEBLEND_MINCONT и DEBLEND_NTHRESH — минимальная разница в яркости для разделения объекта на структуры и максимальное количество разделений.

Весь список параметров представлен в [7] и [3]. Особое внимание стоит уделить выбору порога и фильтра, так как эти характеристики очень сильно влияют на качество выборки.

5.4.1 Поиск рассеянных структур и скоплений

В Приложении Г приведен пример работы SExtractor в поиске галактик в Hubble Ultra Deep Field 2009 (F105W). Изначально это программное обеспечение создавалось для выделения галактик и подобных структур на снимках, поэтому на данном поле не возникает проблем с поиском и используются стандартные параметры.

Пример выделения рассеянных структур и небольших скоплений (H_α туманностей M33) приведен на Рис. 5.4.

Часто SExtractor используют для поиска небольших скоплений на больших картах. Например, такая работа представлена в [8]. Выделение скоплений производилось по нескольким признакам:

- Параметр DETECT_THRESH для определение порога идентификации.
- Размеры выделяемого объекта: FWHM, Area.
- Вытянутость — отношение малой полуоси к большой.
- Видимая звездная величина объекта.

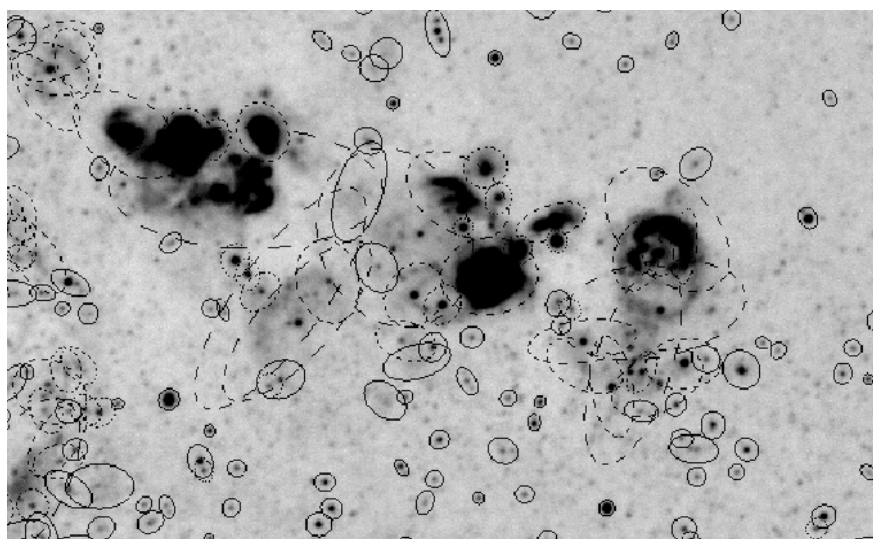


Рис. 5.4: SExtractor, пример выделения объектов и туманностей M33, H_α , Massey, KPNO

5.4.2 Проблемы выделения звезд и скоплений

Существует проблема выделения объектов в переполненных звездами полях, особенно если стоит задача выделения отдельных звезд или небольших групп. Причем понижение порога для алгоритма 4.6 или комбинирование фильтров (уменьшение размера сглаживающей матрицы Гаусса) не очень помогают в этой ситуации. Пример работы программы с параметром `-OBJECTS` для разных фильтров представлен на Рис. 5.5 (в качестве примера использована центральная область галактики M33, Massey, KPNO). Первые три изображения на Рис. 5.5 — результат поиска с применением малых сглаживающих фильтров. Как можно видеть, они показывают хорошие идентичные результаты, но отождествить отдельные звезды в таких заполненных полях проблематично. Авторы программы Bertin E. и Arnouts S рекомендуют в [3] использовать Mex-hat преобразование для скученных полей, если стоит цель отождествления малых объектов и структур (изображение г), Рис. 5.5). Метод позволяет идентифицировать некоторые звезды по отдельности, но большая часть информации потеряна. Уменьшение порога $< 1\sigma$ не улучшает результат.

Попробуем использовать для этого многоуровневое вейвлет преобразование, представленное в разделе 3.3, и виенеровскую фильтрацию 4.6 и пороговую обработку 4.1.2 на разных масштабах исходного изображения (Рис. 5.7). Комбинацией параметров фильтров можно выделять или скрывать источники. Чтобы выделить группы звезд, можно применить Гауссовское сглаживание, описанное в 1.1.1.

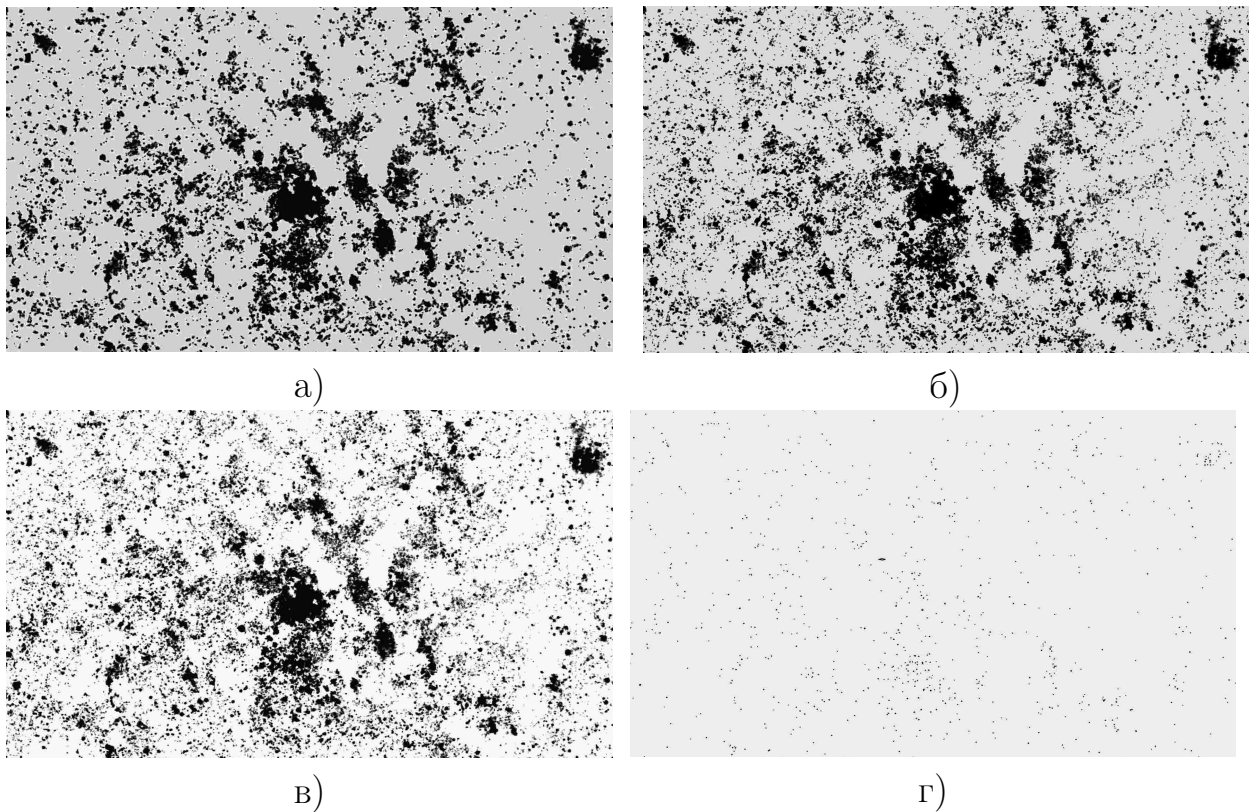


Рис. 5.5: Вывод промежуточных изображений с помощью SExtractor (-OBJECTS) с использованием разных фильтров: а) Матрица Гаусса 3x3 б) Top-hat в) Без использования фильтра г) Mex-hat преобразование

Как видно на Рис. 5.6 trous-вейвлет преобразование с применением вивенеровской фильтрации показывает хорошие результаты в выделении точечных источников, лучше чем предложенное в SExtractor Mex-hat преобразование с экстракцией фона при помощи 3σ -clipping алгоритма. Преимущество второго подхода заключается в высокой скорости работы, большей автоматизированности и универсальности.

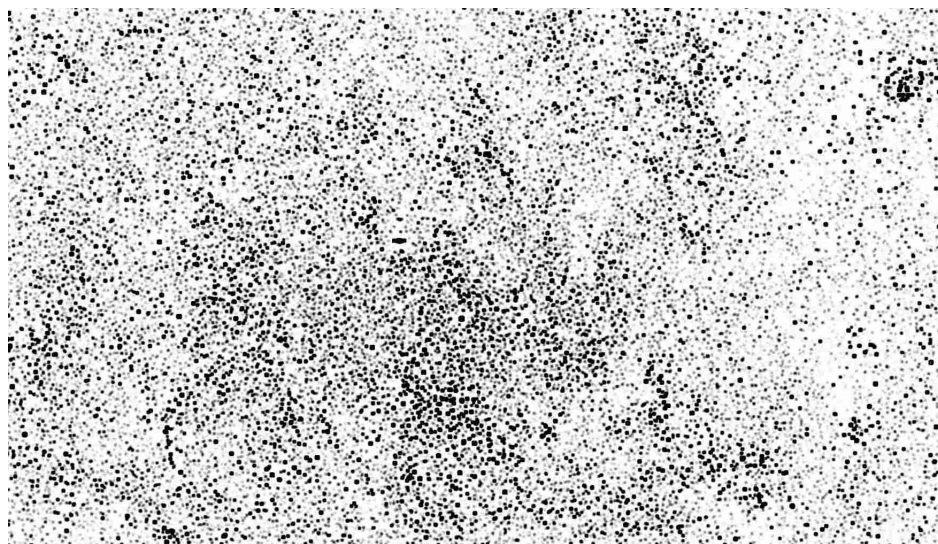
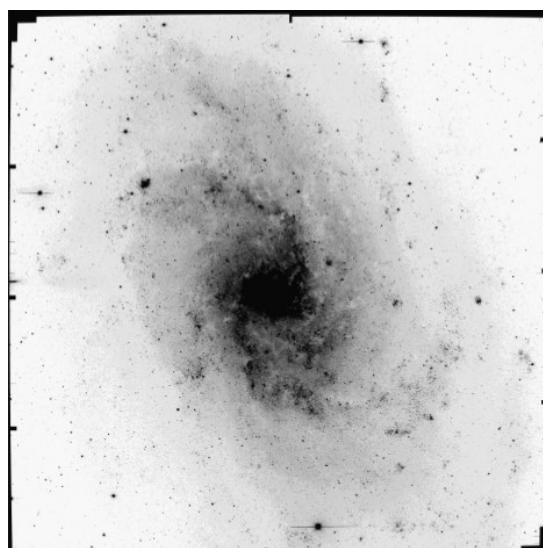
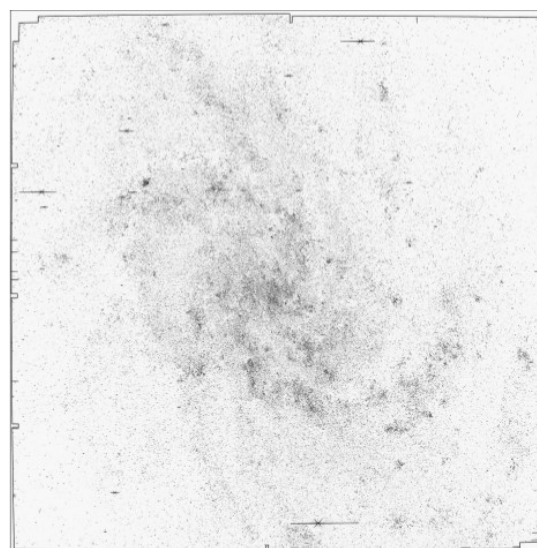


Рис. 5.6: Центр снимка М33, trous-вейвлет разложение (Рис. 5.7), виенеровская фильтрация, пороговая обработка



а)



б)

Рис. 5.7: а) М33, Massey, оригинал б) Отфильтрованный снимок

Глава 6

Поиск звездных скоплений в М33

В данной главе будет рассмотрен пример поиска скоплений в галактике М33. Основной задачей стояло выделение молодых групп звезд, которые могут быть связаны с яркими голубыми переменными (LBV — luminous blue variable).

6.1 Luminous blue variable

Яркие голубые переменные (LBVs) — одна из наименее понятных стадий в эволюции массивных звезд (до $150 M_{\odot}$). Трудности в понимании этих объектов определяются как малочисленностью известных LBV-звезд, их разнообразием, сильной спектральной и фотометрической переменностью [19, 20], так и неточным знанием их фундаментальных параметров. Очевидно, что массивная звезда может проявить себя как LBV после стадии O-звезды Главной Последовательности и перед стадией поздней WR-звезды азотной последовательности (WNL). Однако, более точное положение LBV-звезд в переходе $OV \rightarrow WNL$ пока неясно [15]. Весьма вероятно, что стадия LBV и соответствующие этой стадии неустойчивости вет-

ров определяется содержанием водорода в атмосфере звезды. Однако, как определение содержания водорода, так и точное определение светимости и температуры звезды задача нелегкая. Определение фундаментальных параметров массивных звезд в близких галактиках, например в богатой массивными звездами галактике М 33, может быть более уверенным чем в нашей Галактике, так как расстояния до галактик измерены с достаточной точностью.

У LBV-звезд наблюдаются сильные изменения блеска и спектра при примерно постоянной болометрической светимости. Характерные времена такой глобальной переменности составляют месяцы–годы. При увеличении визуального блеска LBV-звезды температура фотосферы понижается до 9000-10000 К, размер фотосферы звезды увеличивается, это высокое/холодное состояние. Здесь и далее под фотосферой мы понимаем "псевдофотосферу", то есть то место в ветре звезды, где формируется наблюдаемое континуальное излучение. При понижении визуального блеска температура заметно возрастает до 35000 К или даже выше [16, 17, 19], размер звезды уменьшается, это низкое/горячее состояние. В горячем состоянии спектры LBV весьма похожи на спектры WNL-звезд, однако, истинные звезды типа WN7/8/9 отличаются от LBV в низком/горячем состоянии рядом спектральных особенностей [15, 21] и низким содержанием водорода. Так называемые переходные (in transition или "slash") звезды типа Ofpe/WN9, выделенные в отдельный класс звезд Walborn [22], похожие на звезды LBV в низком/горячем состоянии. Ofpe/WN9-звезды характеризуются смешанным спектром с эмиссионными линиями Of и WNL звезд. На примере двух галактических звезд типа LBV AG Car и He 3–519, которые в низком/горячем состоянии соответствовали типу Ofpe/WN9, Smith, Crowther & Prinja [21] предложили продление WNL-классификации до WN10 и WN11.

Тем не менее, до сих пор остается открытым вопрос, отличаются ли ис-

тинные звезды LBV в низком/горячем состоянии от самых поздних звезд WN9/10/11. Другими словами, возможно ли, что все звезды самых поздних типов WNL являются звездами LBV, которые за время их наблюдений (всего несколько десятков лет) не проявили сильной переменности блеска. Такие звезды, которые можно подозревать в принадлежности к типу LBV, принято называть спящими (dormant) LBV. Заметим, что спящие LBV-звезды могут быть не только среди звезд типа WNL. Две классические звезды LBV, η Car и P Cyg, сейчас имеют спектр OB- гипергигантов, переменность их блеска мала. Однако, несколько сот лет назад у них наблюдались гигантские вспышки (giant eruptions) [19]. Если бы не эти гигантские поярчания η Car и P Cyg и не знаменитая туманность вокруг звезды η Car, сейчас мы не смогли бы заподозрить их в принадлежности к классу LBV.

N-body simulation предлагают следующую модель: на стадии образования скопления, плотный центр выбрасывает массивную звезду (LBV), и она начинает удаляться от скопления. Для того, чтобы проверить эту теорию, предлагается идентифицировать все скопления, которые могут связаны с LBV и статистически оценить вероятность такого распределения.

LBV — массивные и молодые звезды с возрастом $< 10^7$ лет. Поэтому в работе велся поиск групп звезд, рядом с которыми находятся области III с высоким излучением в H_α . Если звезды формировались в этих областях, то их возраст не больше $10^7 - 10^8$ лет.

6.2 Выбор исходных данных

В качестве исходных были использованы снимки МЗЗ, выполненные группой Massey (KPNO, 4.0m) в В, V, H_α фильтрах. Изображение в диапазоне V уже было использовано в предыдущих разделах (например, Рис.4.2). Снимки в фильтрах В (4381\AA) и H_α (6563\AA), рассмотренные в работе, представлены на Рис. 6.1.

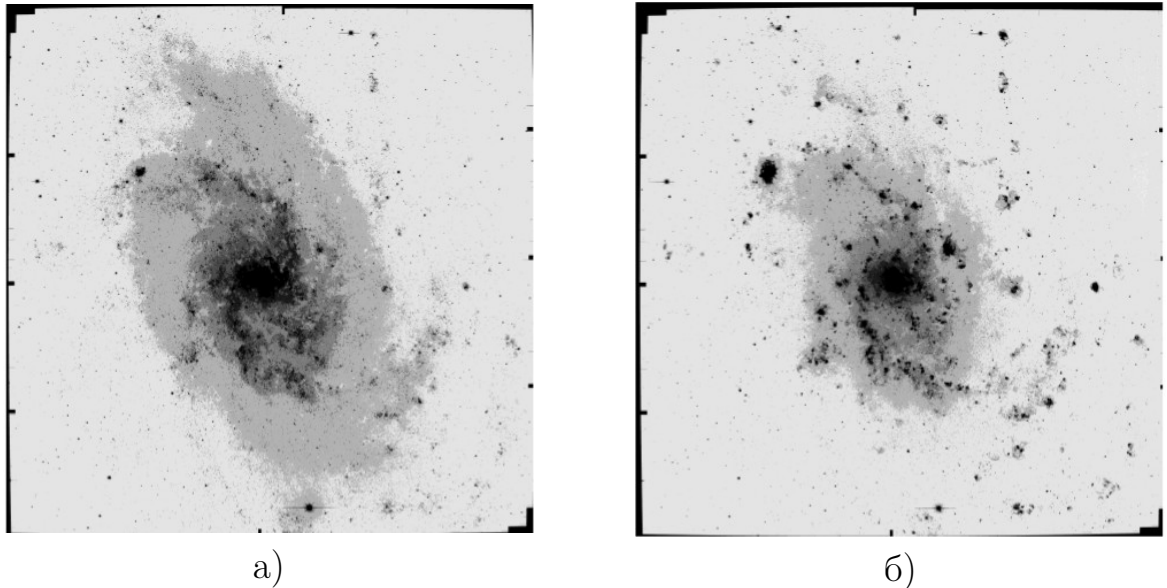
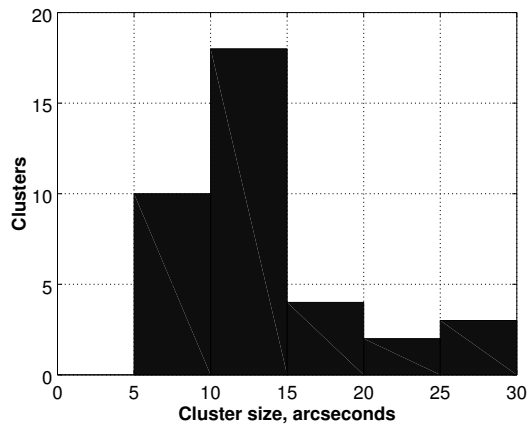


Рис. 6.1: а) V (4381\AA) б) H_{α} (6563\AA) M33, Massey, ds9 scale 92.5%

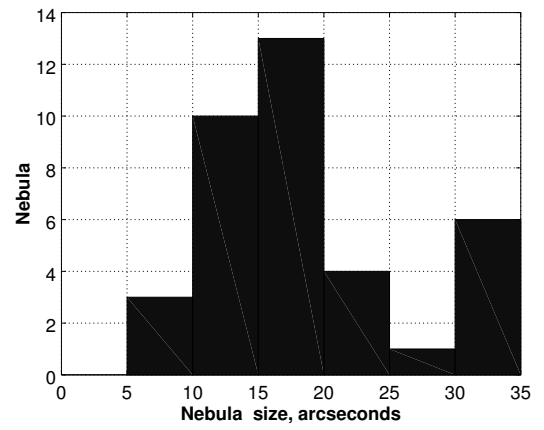
6.3 Определение параметров групп звезд

Для работы были выбраны два списка звезда: подтвержденные LBV (8 звезд) и 29 кандидатов в LBV-звезды в M33 (данные были взяты из [24, 25])

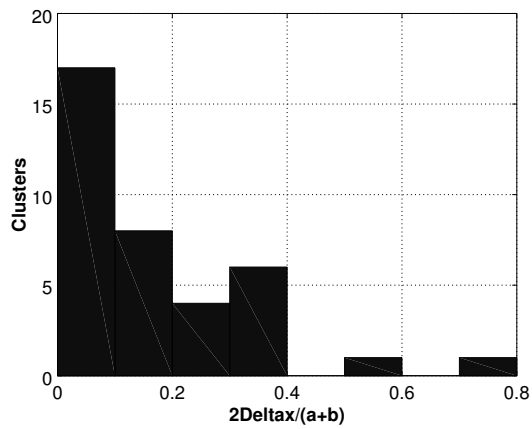
В начале было необходимо вручную (визуально оценить размеры по снимкам B, V, H_{α} в RGB цветах) замерить характерные параметры скоплений, которые могут связаны с LBV-звездами. Все характеристики представлены в Приложении Д (Табл. Д.1, Д.2 для кандидатов, Д.3, Д.4 для подтвержденных LBV). Гистограммы распределение параметров представлены на Рис.6.2 для кандидатов и Рис.6.3 для подтвержденных LBV. Эти диаграммы далее использовались для подбора критериев поиска объектов на снимках.



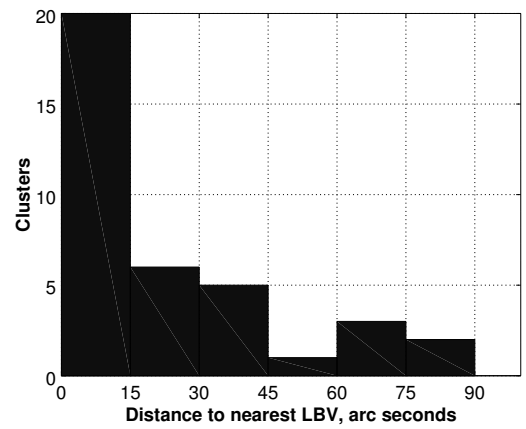
а)



б)



в)



г)

Рис. 6.2: а), б) — распределение размеров ($(a + b)/2$ из Табл. Д.3,Д.4), в) — гистограмма для $2\Delta x/(a + b)$, г) — распределение расстояний для ближайшей LBV

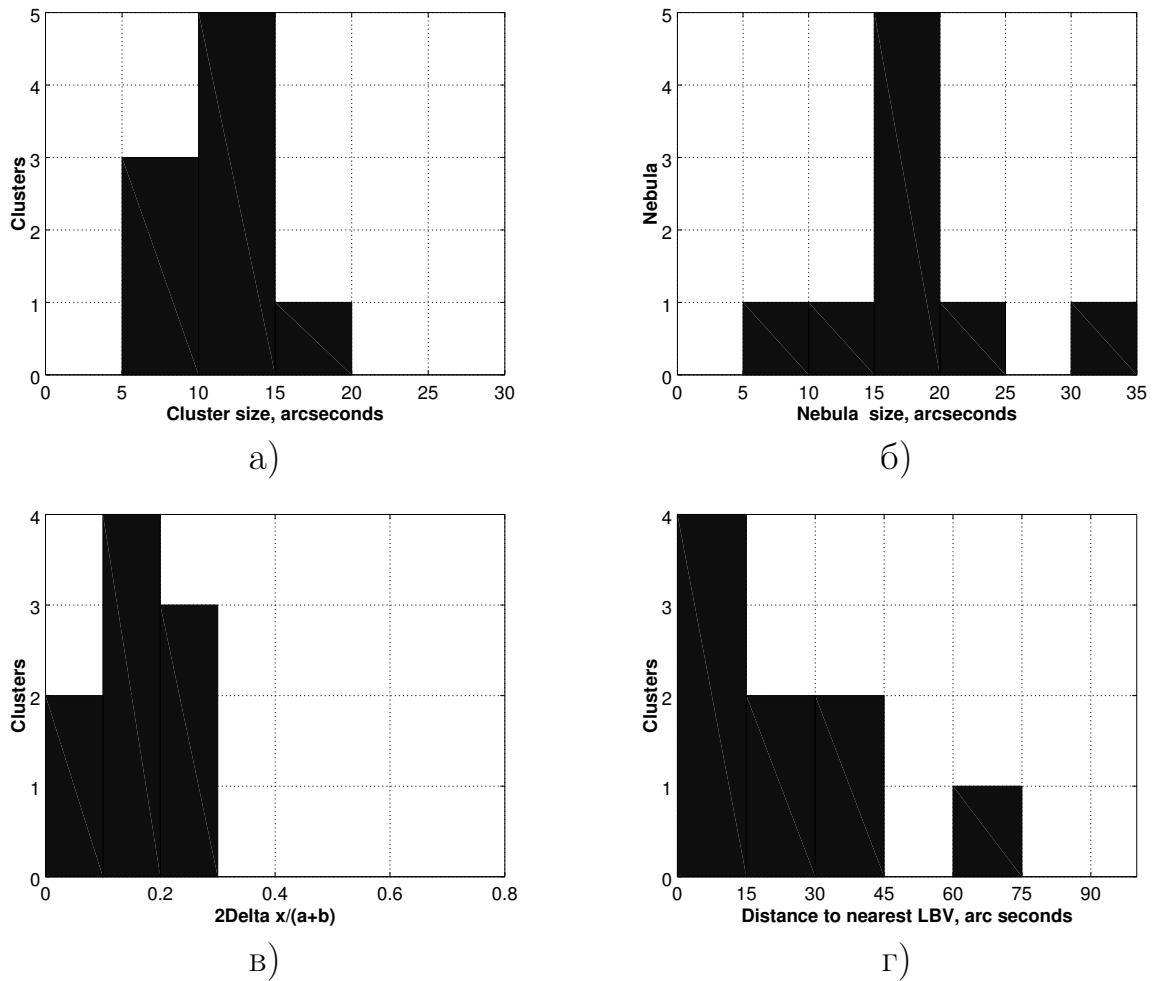


Рис. 6.3: а), б) — распределение размеров ($(a + b)/2$ из Табл. Д.1,Д.2), в) — гистограмма для $2\Delta x/(a + b)$, г) — распределение расстояний для ближайшей LBV

6.4 Идентификация объектов

Эта часть диплома является продолжением работы по выделению скоплений МЗЗ, начатой в прошлом году.

В начале предполагалось из каталога звезд МЗЗ (около 150000 звезд, Massey) выделить скопления, которые могут быть связаны с LBV-звездами. Модель заключалась в том, что каждая звезда представляла собой точку с заданными координатами (α, δ) .

Первым был реализован самый очевидный способ — метод звездных подсчетов. Он заключается в том, чтобы поделить карту на участки и по-

считать количество звезд в каждом их них, а затем вычислить среднеквадратическое отклонение этой величины от значений окружающих ее участков (например, в данном случае брали участок 3×3 с центром в определяющей клетке). Таким образом, скопления идентифицировались как участки с локальным повышением плотности звезд. Схематично это представлено на Рис.6.4. Выбирались секторы с отклонением $> 5\sigma$, центры которых наносились на карту вместе с LBV-кандидатами. В этой работе вся область была поделена по участкам размером $10''$, это оптимально, чтобы в каждой клетке было не более одного скопления.

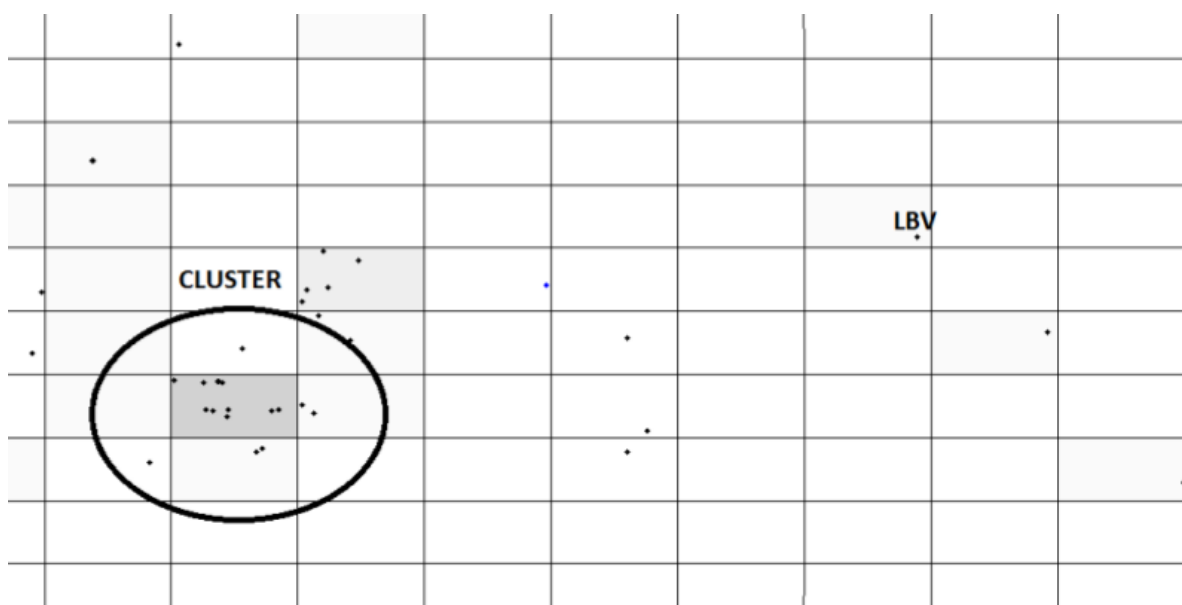


Рис. 6.4: Метод звездных подсчетов, схематичное представление

Однако этот алгоритм плохо работал в областях с высокой плотностью звезд, метод удалось улучшить при помощи отсеечения гармоник по секторам в переполненных звездами зонах, однако одним из его главных недостатков являлось то, что он не давал никакой информации о скоплении кроме координат сектора, в котором оно находится.

После этого был использован альтернативный подход. Применялась аналогичная предыдущему методу модель: на карту наносились звезды, как точки с координатами (α, δ) , были выбраны голубые звезды с $B - V < 0.5$ и $m < 22^m$ (Рис. 6.5). К полученной карте применялось вейвлет-преобразование

и подбирался оптимальный масштаб для выделения скоплений. Далее применялся *hard-thresholding*, маска выбиралась экспериментально по границе предполагаемых скоплений (Например, Рис.6.6).

Для каждой обособленной группы пикселей можно определить центр масс (это будет центра скопления) и площадь, которую она занимает (некий критерий отбора).

Подход показывает хорошие результаты в областях с высокой плотностью (Рис. 6.6), однако большинство предполагаемых скоплений выделить не удалось — если взять недостаточно крупный масштаб, то после пороговой обработки останутся все точечные источники. При увеличении масштаба сглаживания сильно теряется качество выборки, в таком случае невозможно выбрать скопления по формальным критериям.

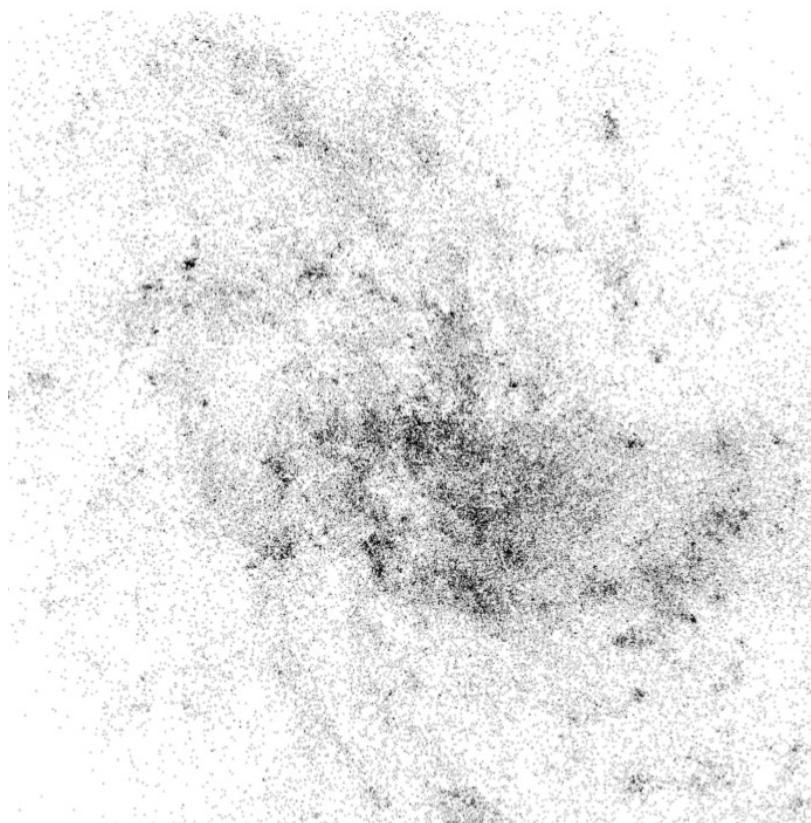


Рис. 6.5: Карта звезд по каталогу Massey

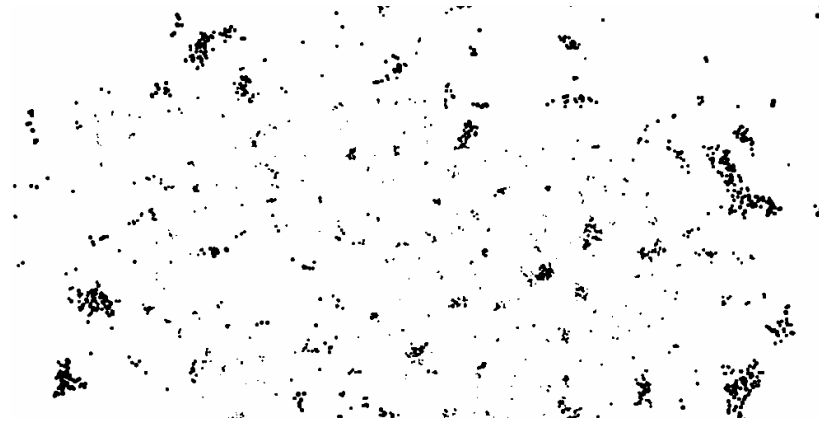


Рис. 6.6: Центральная часть М33, пороговая обработка

Можно идентифицировать изображения по снимкам М33. Как видно из Рис.6.2,6.3, искомые объекты имеют сравнительно крупные размеры. Используем SExtractor для экстракции объектов по сглаженным и отфильтрованным изображениям (надо подобрать максимально широкий фильтр, не теряя информации о размерах объектов в центральных областях).

Основные критерии поиска скоплений — размеры и близость к H_α туманности. В начале были выделены все возможные объекты на снимке в В и H_α фильтрах, а затем выбраны туманности по критериям из Рис.6.2,6.3(размеры от 8'' до 23''). После этого составлено распределение (Рис.6.7) $2\Delta x/(a + b)$ по каждому скоплению, где Δx — расстояние от скопления до ближайшей туманности, a, b — размеры туманности. Будем считать, что если этот параметр меньше 0.7, то скопление находится рядом с туманностью.

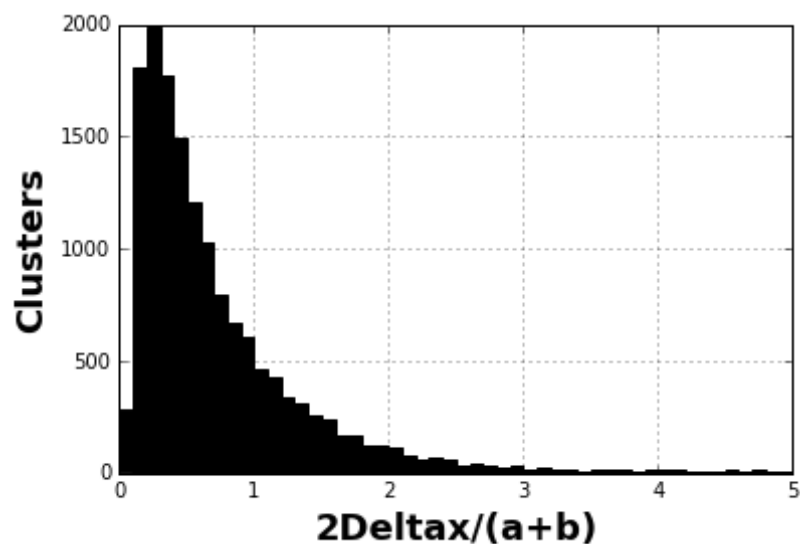


Рис. 6.7: Распределение $2\Delta x/(a+b)$ по всем выделенным объектам

Далее были выделены предполагаемые скопления по размеру (a, b взяты из Рис.6.2,6.3, $9'' < (a+b)/2 < 14''$). После получено распределение по расстояниям до ближайшей LBV для каждого скопления (Рис.6.8). Всего было выделено 109 объектов.

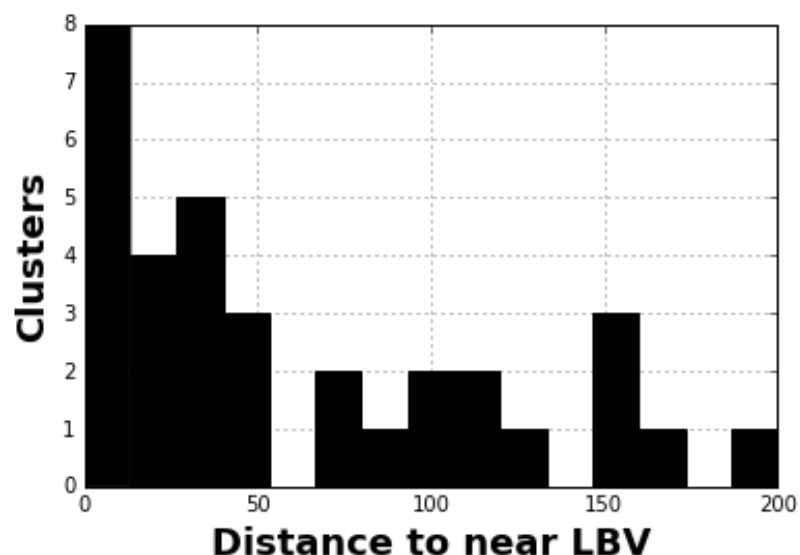


Рис. 6.8: Распределение расстояний до ближайших LBV по всем выделенным объектам

Для того, чтобы рассчитать вероятность такого распределения, используем метод моделирования Монте-Карло. На Рис.6.9 представлен снимок в фильтре H_α с нанесенными скоплениями, LBV-звездами и кандидатами.

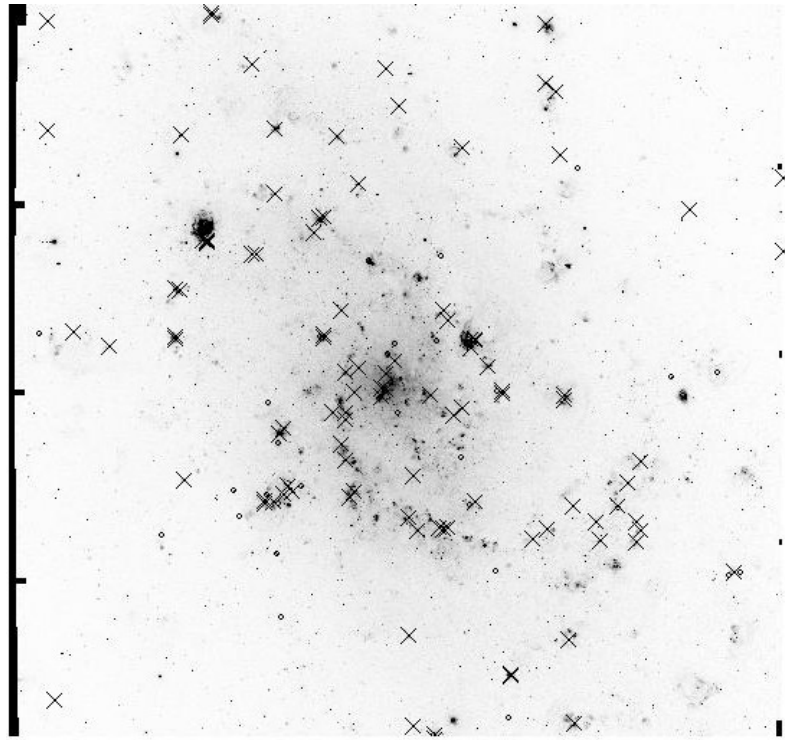


Рис. 6.9: М33, H_{α} , Massey. \times — смоделированные скопления, \circ — LBV и кандидаты.

Нанесем на карту $N = 1000$ случайных точек и составим гистограмму относительно расстояний до ближайшей LBV для каждой точки. Попробуем аппроксимировать получившиеся графики разными распределениями (Рис.6.10). Как видно, распределение растет быстро до $100''$, а затем медленно убывает, поэтому точно аппроксимировать его нормальным распределением не получается (Рис.6.10).

На Рис.6.11 представлены диаграммы из Рис.6.2 с соответствующими им гамма и бета распределения. Из графиков видно, что они существенно отличаются от смоделированных методом Монте-Карло. Это можно строго доказать.

Классическим примером критерия принадлежности выбранных данных к какому-либо закону распределения является критерий χ^2 (критерий согласия Пирсона). Для сравнения данных полученных практическим путем и

теоретического закона распределения используется формула

$$\chi^2 = N \sum_{i=1}^N ((p_i^t - p_i^e)^2 / p_i^t) \quad (6.1)$$

где p_i^t и p_i^e — вероятности попадания величин в выбранный интервал i .

Главным недостатком критерия χ^2 является необходимость строгой привязки начальных данных к интервалам и объединении интервалов при малом числе данных. Поэтому в работе для доказательства неслучайного распределения скоплений использовался критерий Колмогорова-Смирнова (KS test).

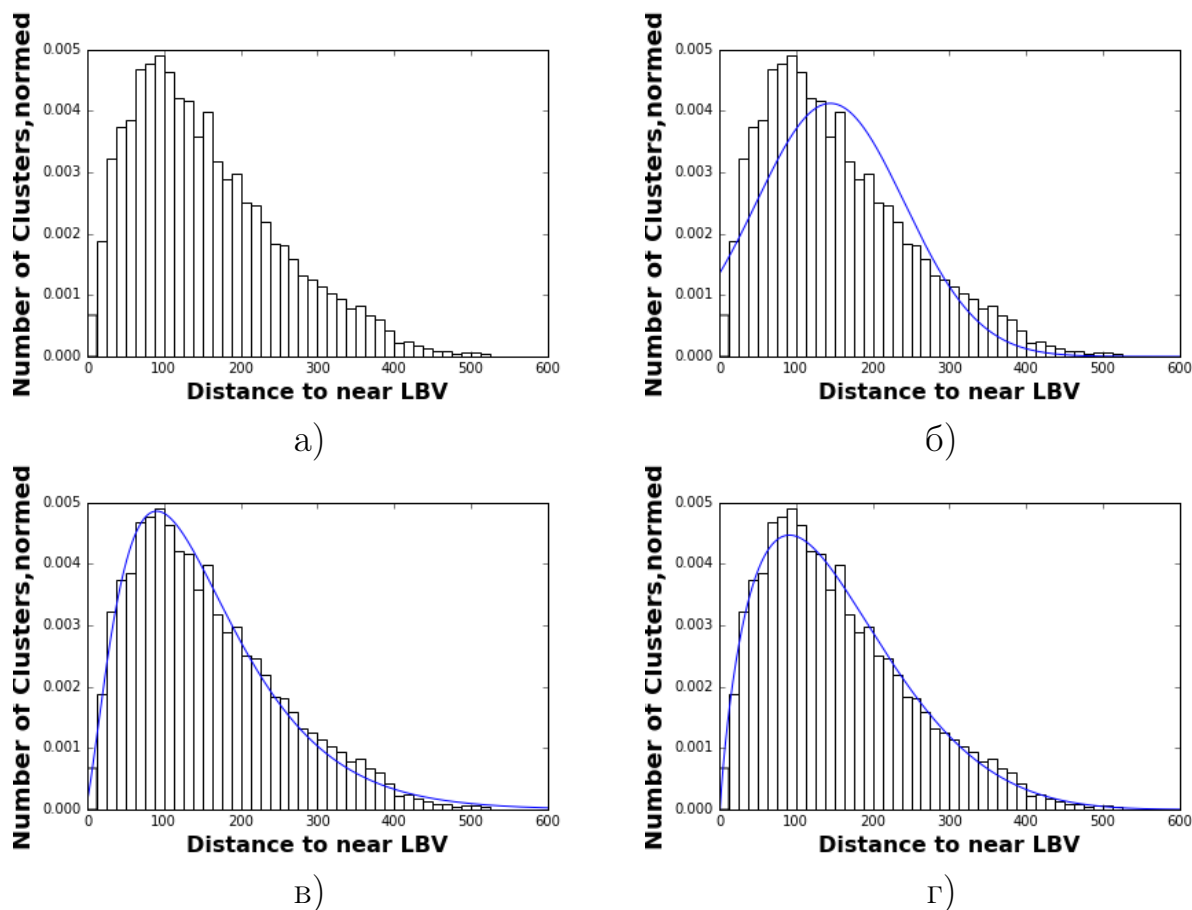


Рис. 6.10: а) — распределение для моделирования Монте-Карло $N = 1000$ б) — аппроксимация нормальным распределением в) — гамма-распределение г) — бета-распределение

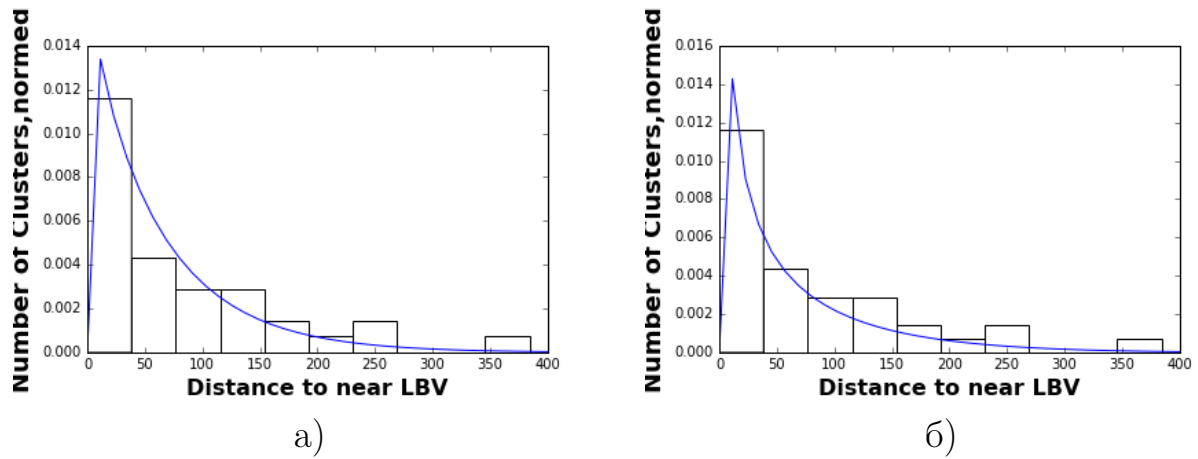


Рис. 6.11: а),б) — распределение для LBV и кандидатов из Рис.6.2, аппроксимированные гамма и бета распределениями

Формула для вычисления критерия Колмогорова-Смирнова представляет собой:

$$D_{m,n} = \sup |F_m(x) - G_m(x)| \quad (6.2)$$

где $D_{m,n}$ — расхождение между распределениями, $F_m(x), G_m(x)$ — распределение частоты встречаемости значений.

Для KS-теста использовался язык статистической обработки данных R(см. [36]).

Статистический анализ двух распределений проводился при помощи p-value (уровень значимости). Проведем численный эксперимент: 100 раз смоделируем Монте-Карло точки на карте ($N = 1000$) и составим распределение p-value. Для каждой итерации значение p-value не превышает 10^{-5} . Это значит, что с вероятностью $> 99.99\%$ эти распределения не соответствуют друг другу.

На этом основании можно сделать вывод, что распределение выделенных объектов не случайно относительно LBV-звезд. Это с высокой долей вероятности означает, что они вылетели из молодых скоплений на стадии их формирования.

Заключение

В работе были изложены основные методы обработки астрономических снимков при помощи пространственных и адаптивных фильтров, пороговой обработки, Фурье и вейвлет-анализа.

В первой части диплома на примерах показано преимущество многоуровневых преобразований над стандартными методами фильтрации изображений (например, Фурье). Далее представлен алгоритм trous-вейвлет преобразования, алгоритм и его отладка на модельных решениях — одномерных, двумерных (гауссов шум) и реальных данных (спектр NGC 4395, снимки NGC 2997, NGC 5194).

Во второй части работы изложены алгоритмы, используемые в современном астрономическом программном обеспечении, например 3σ -clipping или пороговая обработка (hard-thresholding), а так же их адаптация в рамках многоуровневой модели. Такой подход к фильтрации показал очень хорошие результаты, что показано в Гл.4.

Далее в рамках многоуровневой модели был представлен алгоритм поиска объектов на снимках, а так же примеры использования ПО (SExtractor) с аналогичным подходом, который был улучшен в некоторых случаях с помощью представленных преобразований (например, при поиске объектов в переполненных звездами областях).

В последней главе представленные инструменты были использованы для поиска молодых звездных образований в галактике M33. Статистически доказана связь LBV-звезд с найденными скоплениями.

Литература

- [1] Albert Bijaoui, A multiscale vision model adapted to the astronomical images, *Signal Processing* 46 (1995) 345-362
- [2] Albert Bijaoui, Christophe Benoistand, Eric Slezak, Analysis of multiband astronomical images using multiscale tools, Casiopee laboratory, B.P. 4229, 06304 Nice Cedex 4, France
- [3] Bertin E., S. Arnouts SExtractor: Software for source extraction — *Astronomy & Astrophysics Supplement*, 1996
- [4] Bertin, E. and Arnouts, S.: 1996, *Astronomy and Astrophysics, Supplement Series* 117, 393
- [5] Eric H. Neilsen, Jr., Notes on the Essentials of Astronomy Data, 2014-06-16
- [6] DAOPHOT: A COMPUTER PROGRAM FOR CROWDED-FIELD STELLAR PHOTOMETRY, PETER B. STETSON, *Publications of the astronomical society of the pacific*, March 1987
- [7] Dr. Benne, W. Holwerda, Source extractor for dummies, Baltimore MD 21218
- [8] Izaskun San Roman and Ata Sarajedini, Photometric Properties of the M33 Star Cluster System, Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611-2055

- [9] Mallat S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989. — Vol. 11. — P. 674–693
- [10] Nathan Smith, and Ryan Tombleson, Luminous blue variables are antisocial: their isolation implies that they are kicked mass gainers in binary evolution, MNRAS 447, 598–617 (2015)
- [11] J.L. Starck, F. Murtagh Handbook of Astronomical Data Analysis - 2002
- [12] J.L. Starck, F. Murtagh, Albert Bijaoui, Image processing and data analysis: The multiscale approach
- [13] J.L. Starck, Astronomical image representation by the curvelet transform, A&A 398, 785–800 (2003)
- [14] Jean-Luc Starck, Fionn Murtagh, Mario Bertero, Starlet Transform in Astronomical Data Processing, Springer Science+Business Media LLC 2011
- [15] N. Smith and P. S. Conti, *Astrophys. J.* 679, 1467 (2008).
- [16] T. Szeifert, *LIACo* 33, 459 (1996).
- [17] J. S. Clark, V. M. Larionov, and A. Arkharov, *Astronom. and Astrophys.* 435, 239 (2005).
- [18] Roberta M. Humphreys, Kerstin Weis, Kris Davidson and Michael S. Gordon, On the Social Traits of LBVs, Minnesota Institute for Astrophysics, University of Minnesota, Astronomical Institute, Ruhr-Universitaet Bochum, Germany, Minnesota Institute for Astrophysics, University of Minnesota, 2016
- [19] R. M. Humphreys and K. Davidson, *Publ. Astronom. Soc. Pacific* 106, 1025 (1994).

- [20] A. M. van Genderen, *Astronom. and Astrophys.* 366, 508 (2001).
- [21] L. J. Smith, P. A. Crowther, and R. K. Prinja, *Astronom. and Astrophys.* 281, 833 (1994).
- [22] N. R. Walborn, *Astrophys. J.* 215, 53 (1977).
- [23] Shashikiran Ganesh, Physical Research Laboratory, *Wavelets in the Astronomical Context*, 2003
- [24] Humphreys, Roberta M.; Weis, Kerstin; Davidson, Kris; Bomans, D. J.; Burggraf, Birgitta, *Luminous and Variable Stars in M31 and M33. II. Luminous Blue Variables, Candidate LBVs, Fe II Emission Line Stars, and Other Supergiants*, *The Astrophysical Journal*, 790: 48 (21pp), 2014 July
- [25] Humphreys, Roberta M.; Martin, John C.; Gordon, Michael S., *A New Luminous Blue Variable in M31*, *PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC*, 127: 347–350, 2015 April
- [26] А.Н. Яковлев, *Введение в вейвлет-преобразования*, Новосибирск: Изд-во НГТУ, 2003
- [27] Добеши И., *Десять лекций по вейвлетам*, 2001, 464 стр.
- [28] И.М. Дремин, О.В. Иванов, В.А. Нечитайло, *Вейвлеты и их использование*, *Успехи физических наук*, май 2001, том 171, № 5
- [29] М. А. Иванов, *Применение вейвлет-преобразований в кодировании изображений*, работа по гранту № 01-01-794
- [30] Н.Ф. Астафьева, *Вейвлет-анализ: основы теории и примеры применения*, *Успехи физических наук*, ноябрь 1996, том 166, № 11

- [31] Прэтт У Цифровая обработка изображений В 2 т М Мир, 198
- [32] Писаревский А. Н. и др. Системы технического зрения (Принципиальные основы, аппаратное и математическое обеспечение) Л: Машиностроение, 1988.
- [33] Р. Гонсалес, Р. Вудс, Цифровая обработка изображений, ТЕХНОСФЕРА, Москва, 2005
- [34] О.Н. Шолухова, С.Н. Фабрика, А.В. Жарова, А.Ф. Валеев, В.П. Горанский, Спектральная переменность LBV-звезды V 532 (ROMANO'S STAR), Астрофизический Бюллетень, 2011, том 66, № 2, с. 135–157
- [35] <https://www.gnu.org/software/octave/>
- [36] <https://www.r-project.org/>

Приложение А

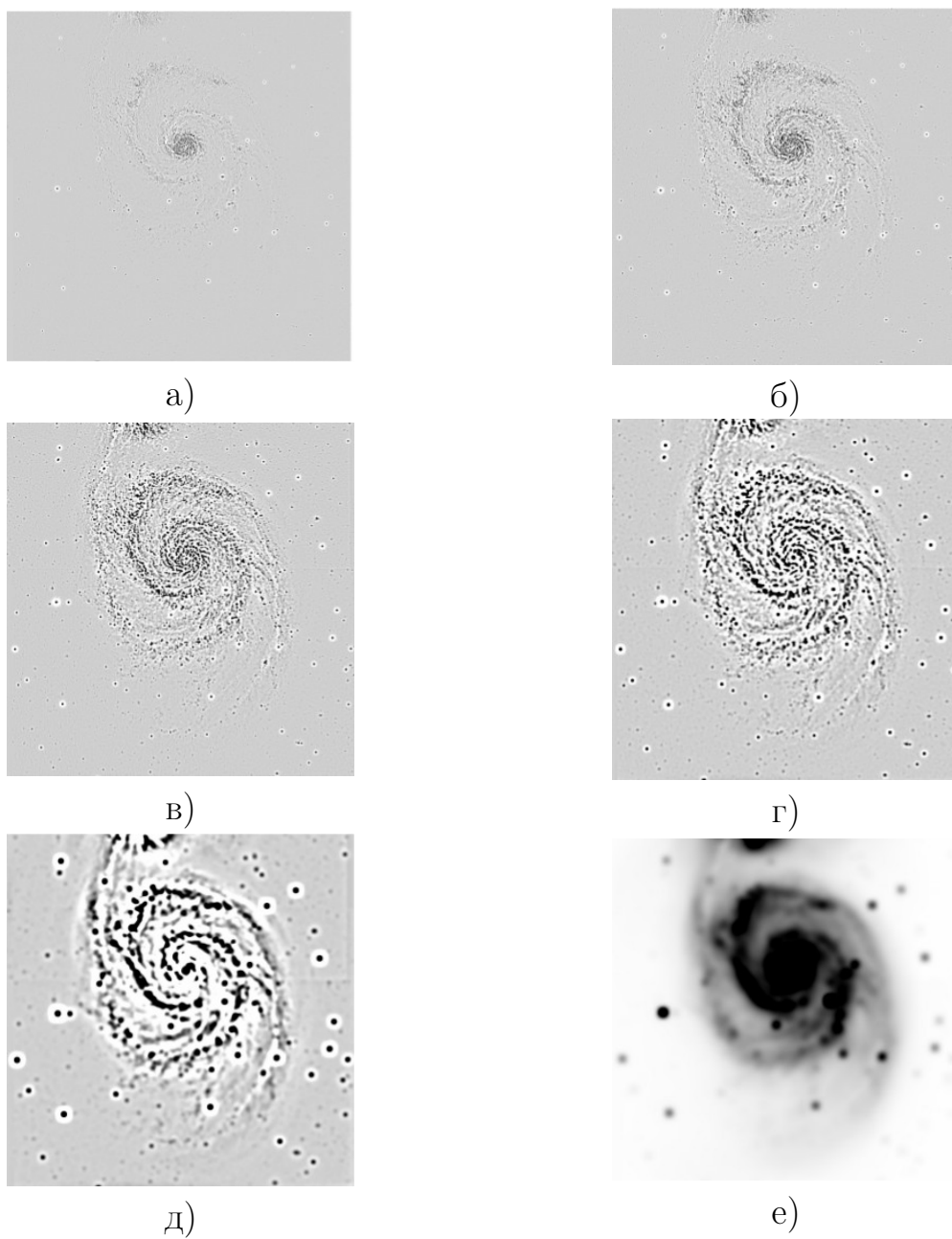
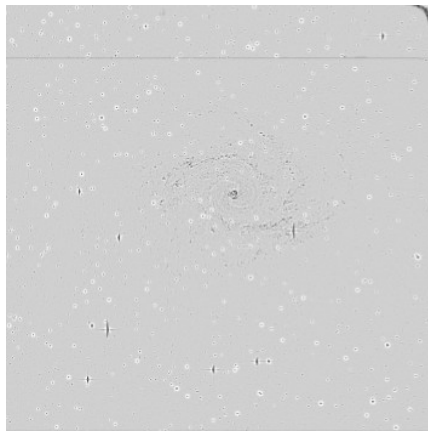
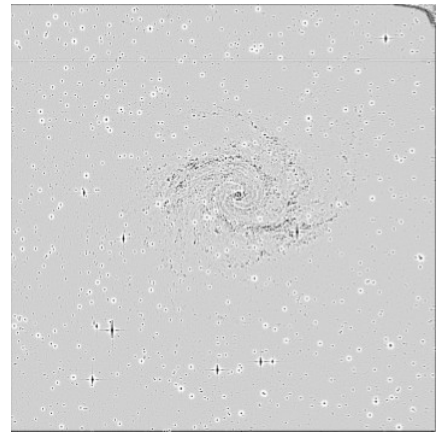


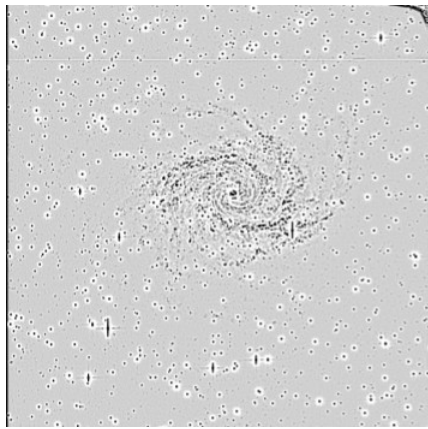
Рис. А.1: а) - е) - разложение снимка NGC 5194 (trous-вейвлет преобразование)



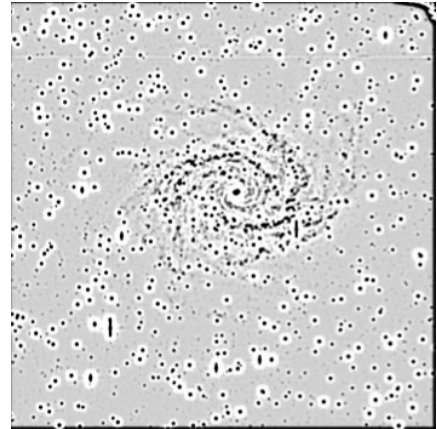
а)



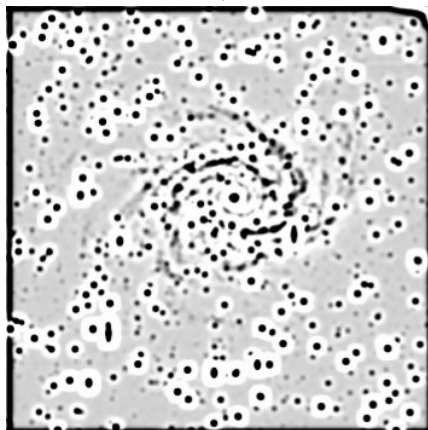
б)



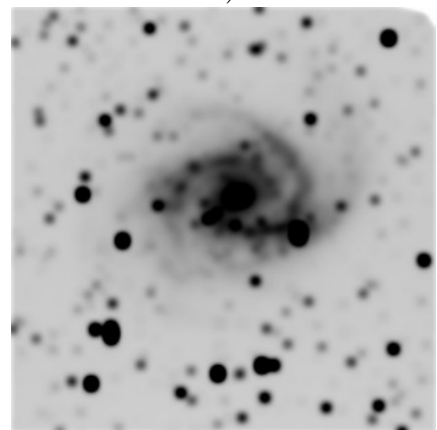
в)



г)



д)



е)

Рис. А.2: а) - е) - разложение снимка NGC 2997 (trous-вейвлет преобразование)

Приложение Б

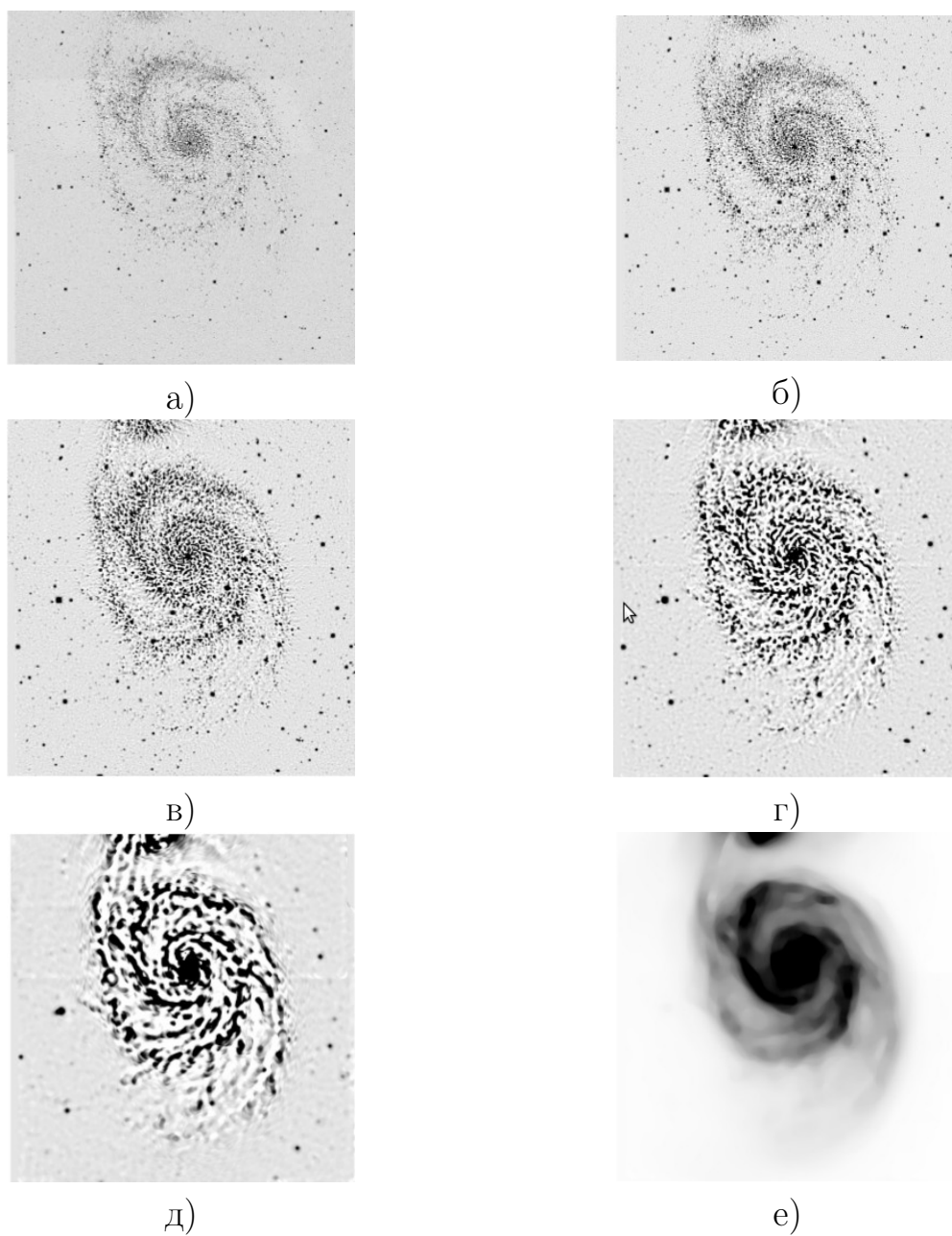
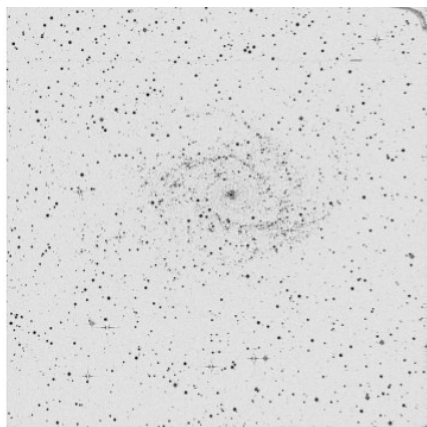
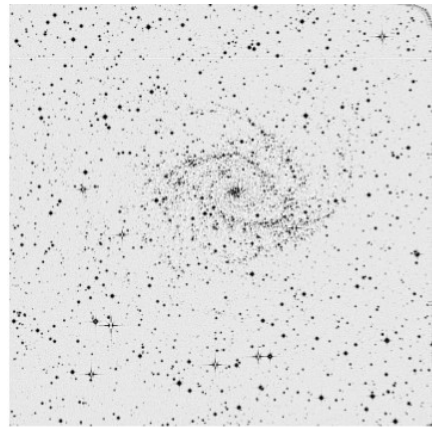


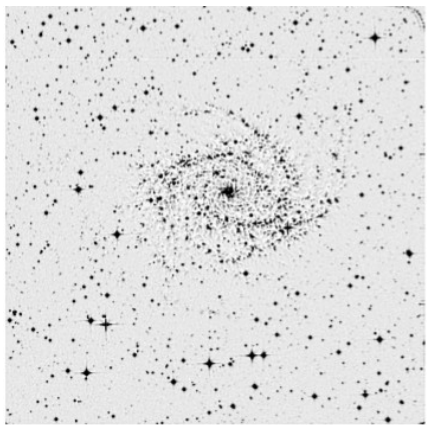
Рис. Б.1: а) - е) - разложение снимка NGC 5194 (медианный алгоритм)



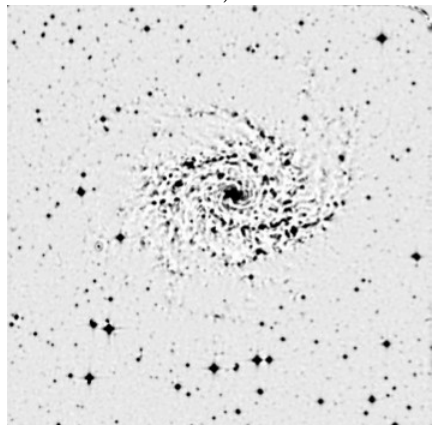
а)



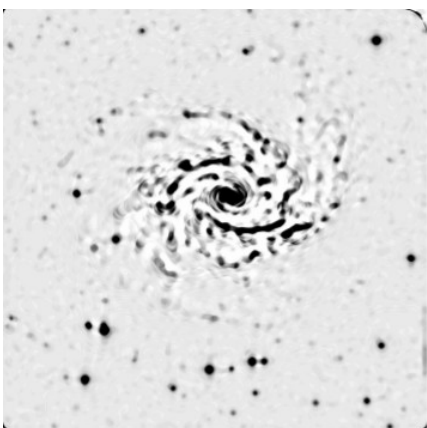
б)



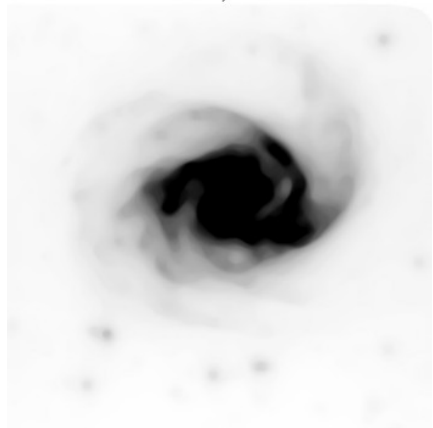
в)



г)



д)



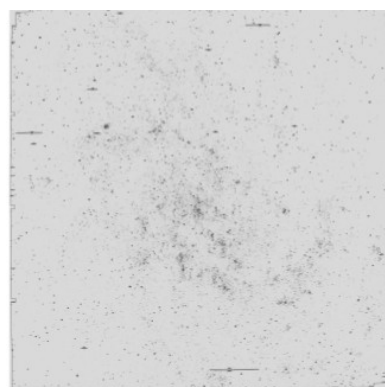
е)

Рис. Б.2: а) - е) - разложение снимка NGC 2997 (медианный алгоритм)

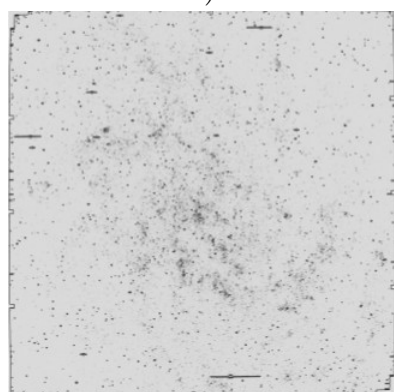
Приложение В



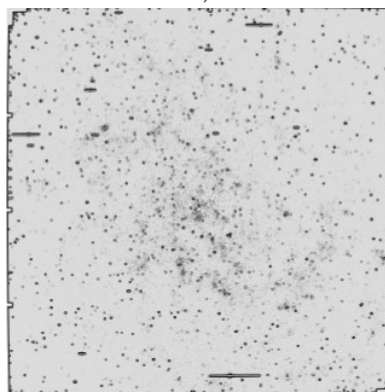
а)



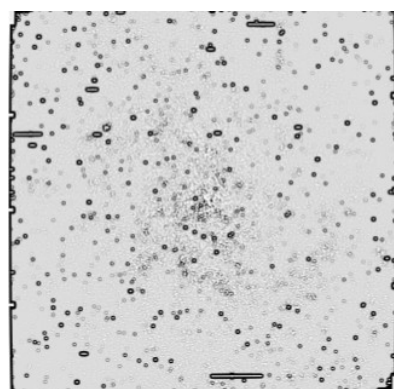
б)



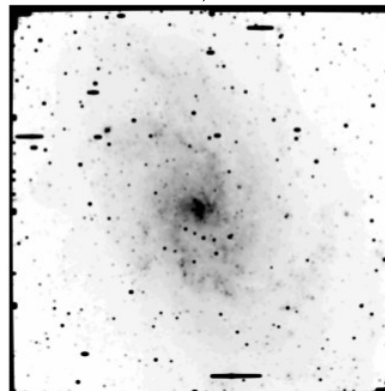
в)



г)



д)

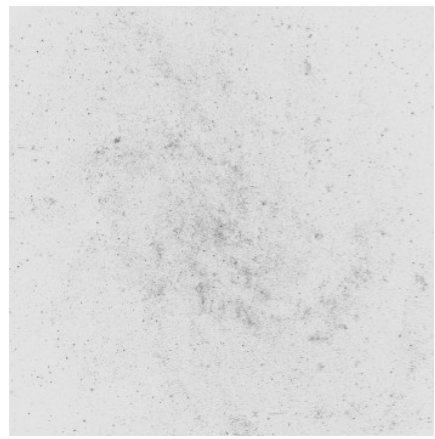


е)

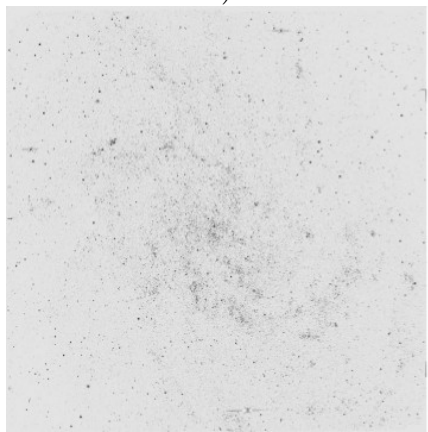
Рис. В.1: а) - е) - разложение снимка М33 (trous-вейвлет преобразование)



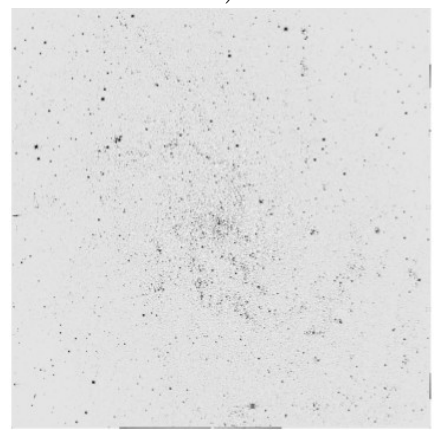
а)



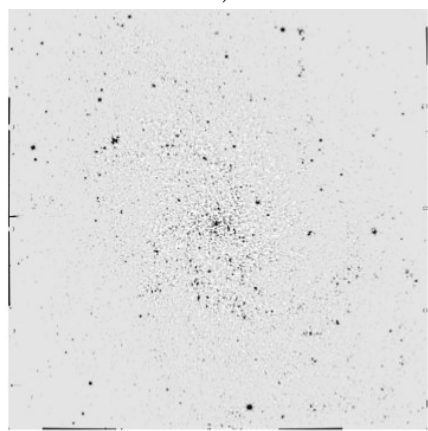
б)



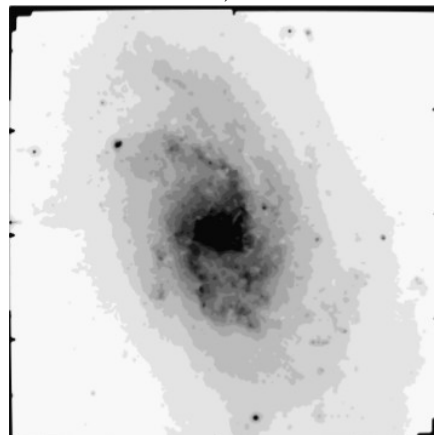
в)



г)



д)



е)

Рис. В.2: а) - е) - разложение снимка М33 (медианный алгоритм)

Приложение Г

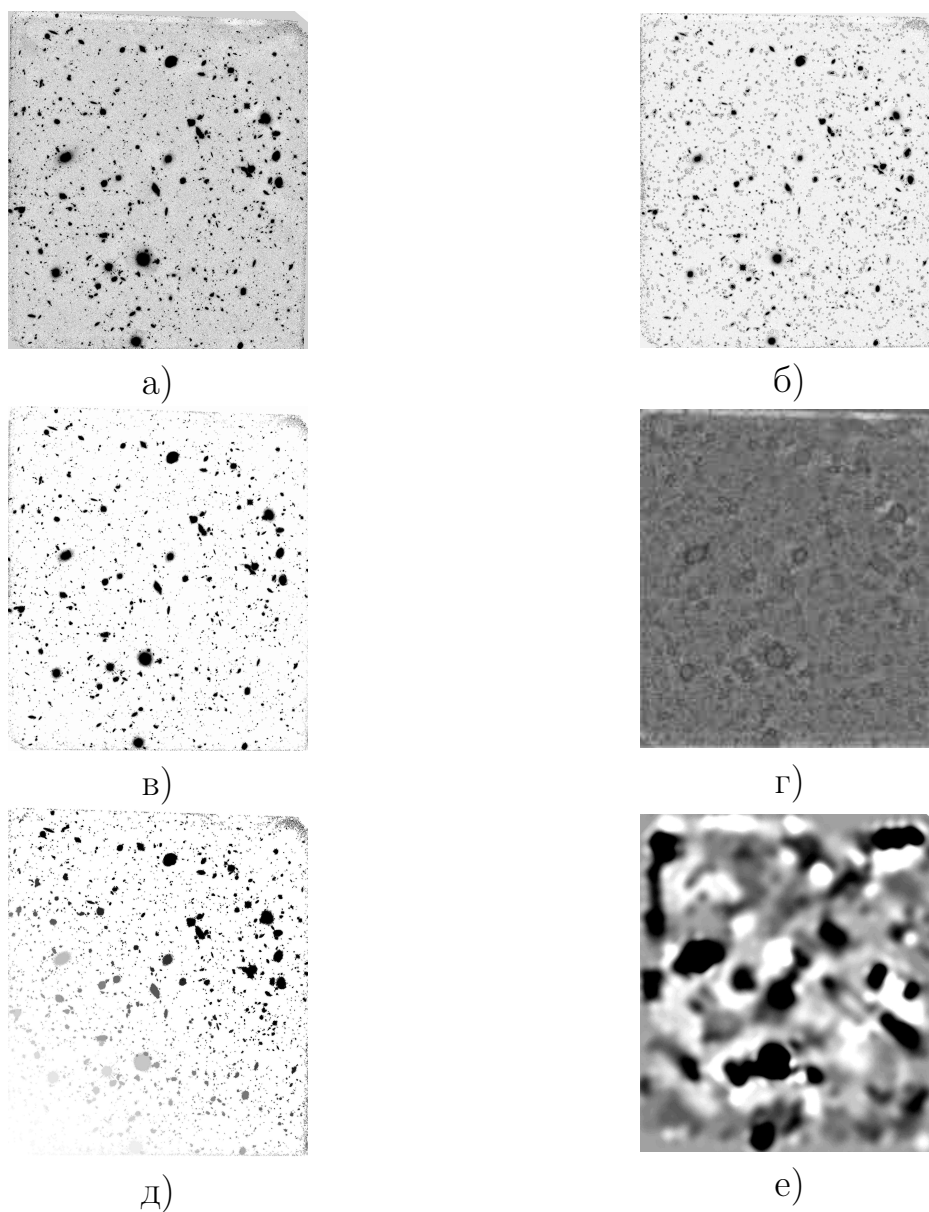


Рис. Г.1: Hubble Ultra Deep Field 2009 (F105W): а) Original б) Apertures (выделение границ) в) ОБЪЕКТЫ г) -ОБЪЕКТЫ (снимок без объектов) д) SEGMENTATION (бинаризованные уровни) е) BACKGROUND (фон)

Приложение Д

Таблица Д.1: Измеренные вручную размеры скоплений для подтвержденных LBV и кандидатов, α, δ — координаты в формате h:m:s, a, b — малая и большая полуось в угловых секундах

| α скопления | δ скопления | a скопления | b скопления |
|--------------------|--------------------|---------------|---------------|
| 1:32:35 | 30:30:26 | 8.26 | 11.6 |
| 1:32:35 | 30:30:26 | 8.26 | 11.6 |
| 1:33:26 | 30:24:07 | 11.86 | 15.0 |
| 1:33:41 | 30:22:37 | 12.0 | 15.8 |
| 1:33:59 | 30:23:20 | 12.0 | 19.0 |
| 1:34:16 | 30:28:26 | 12.0 | 14.0 |
| 1:34:18 | 30:31:37 | 13.0 | 17.5 |
| 1:34:42 | 30:31:40 | 5.5 | 13.0 |
| 1:35:07 | 30:41:46 | 12.5 | 25.0 |
| 1:35:00 | 30:37:18 | 5.0 | 8.5 |
| 1:32:37 | 30:40:07 | 15.0 | 16.0 |
| 1:32:45 | 30:38:58 | 24.0 | 36.0 |
| 1:32:45 | 30:38:58 | 24.0 | 36.0 |
| 1:33:11 | 30:48:57 | 12.8 | 30.0 |

| | | | |
|---------|----------|------|------|
| 1:33:28 | 30:31:31 | 12.0 | 15.0 |
| 1:32:59 | 30:33:33 | 10.5 | 17.5 |
| 1:34:22 | 30:33:13 | 11.5 | 15.5 |
| 1:34:20 | 30:33:57 | 8.0 | 14.0 |
| 1:34:15 | 30:33:41 | 12.0 | 15.0 |
| 1:34:18 | 30:34:10 | 9.5 | 10.0 |
| 1:34:10 | 30:34:38 | 9.5 | 15.0 |
| 1:34:16 | 30:36:42 | 14.0 | 16.0 |
| 1:34:18 | 30:38:36 | 7.0 | 12.5 |
| 1:34:32 | 30:47:10 | 25.0 | 29.0 |
| 1:33:40 | 30:45:55 | 12.0 | 16.0 |
| 1:33:26 | 30:39:50 | 12.0 | 17.0 |
| 1:33:33 | 30:41:34 | 24.0 | 25.0 |
| 1:33:40 | 30:41:38 | 10.0 | 11.0 |
| 1:33:35 | 30:36:28 | 9.0 | 12.0 |
| 1:33:55 | 30:45:26 | 8.0 | 14.0 |
| 1:34:06 | 30:41:44 | 12.0 | 14.0 |
| 1:33:51 | 30:40:56 | 7.0 | 11.0 |
| 1:33:33 | 30:33:44 | 7.5 | 14.0 |
| 1:33:50 | 30:41:30 | 9.0 | 14.0 |
| 1:33:47 | 30:38:39 | 8.0 | 10.0 |
| 1:33:53 | 30:38:53 | 4.8 | 7.5 |
| 1:33:52 | 30:39:15 | 5.0 | 6.0 |

Таблица Д.2: Измеренные вручную размеры туманностей для подтвержденных LBV и кандидатов, α, δ — координаты в формате h:m:s, a, b — малая и большая полуось в угловых секундах, Δx — расстояние между центрами скопления и туманности в угловых секундах

| α скопления | δ скопления | a скопления | b скопления | Δx |
|--------------------|--------------------|---------------|---------------|------------|
| 1:32:35 | 30:30:28 | 12.9 | 17.3 | 30.0 |
| 1:32:35 | 30:30:28 | 12.9 | 17.3 | 10.0 |
| 1:33:26 | 30:24:08 | 7.7 | 9.51 | 46.0 |
| 1:33:41 | 30:22:42 | 9.5 | 15.6 | 4.0 |
| 1:33:59 | 30:23:20 | 15.3 | 21.4 | 80.0 |
| 1:34:16 | 30:28:22 | 17.0 | 22.0 | 15.0 |
| 1:34:18 | 30:31:36 | 14.0 | 15.0 | 34.0 |
| 1:34:42 | 30:31:40 | 9.0 | 13.0 | 36.0 |
| 1:35:07 | 30:41:51 | 15.0 | 28.8 | 35.0 |
| 1:35:00 | 30:37:20 | 6.0 | 10.8 | 21.0 |
| 1:32:38 | 30:39:58 | 49.0 | 63.0 | 1.0 |
| 1:32:45 | 30:38:59 | 38.0 | 45.0 | 5.0 |
| 1:32:45 | 30:38:59 | 38.0 | 45.0 | 64.0 |
| 1:33:11 | 30:48:54 | 40.0 | 45.0 | 64.0 |
| 1:33:28 | 30:31:32 | 14.0 | 23.0 | 64.0 |
| 1:33:00 | 30:33:38 | 21.0 | 24.0 | 2.0 |
| 1:34:22 | 30:33:17 | 24.0 | 24.5 | 34.0 |
| 1:34:19 | 30:33:58 | 7.0 | 14.0 | 82.0 |
| 1:34:15 | 30:33:45 | 16.0 | 21.0 | 7.0 |
| 1:34:18 | 30:34:11 | 15.0 | 18.0 | 1.0 |
| 1:34:10 | 30:34:34 | 30.0 | 37.0 | 1.0 |

| | | | | |
|---------|----------|------|------|------|
| 1:34:16 | 30:36:41 | 14.0 | 22.0 | 4.5 |
| 1:34:18 | 30:38:36 | 9.0 | 14.0 | 4.0 |
| 1:34:32 | 30:47:10 | 65.0 | 95.0 | 30.0 |
| 1:33:40 | 30:45:58 | 11.5 | 16.5 | 19.0 |
| 1:33:26 | 30:39:06 | 14.5 | 22.5 | 9.5 |
| 1:33:33 | 30:41:34 | 28.0 | 32.0 | 16.0 |
| 1:33:40 | 30:41:37 | 14.0 | 15.0 | 3.0 |
| 1:33:35 | 30:36:28 | 9.0 | 12.0 | 30.0 |
| 1:33:55 | 30:45:26 | 16.0 | 27.0 | 8.0 |
| 1:34:06 | 30:41:45 | 15.0 | 20.0 | 4.0 |
| 1:33:51 | 30:40:56 | 11.0 | 16.0 | 7.0 |
| 1:33:33 | 30:33:46 | 16.0 | 8.0 | 1.0 |
| 1:33:50 | 30:41:30 | 14.0 | 20.0 | 3.0 |
| 1:33:47 | 30:38:41 | 14.0 | 21.0 | 35.0 |
| 1:33:53 | 30:38:52 | 16.0 | 20.0 | 4.0 |
| 1:33:52 | 30:39:15 | 7.0 | 7.5 | 7.0 |

Таблица Д.3: Измеренные вручную размеры скоплений для подтвержденных LBV, α, δ — координаты в формате h:m:s, a, b — малая и большая полуось в угловых секундах

| α скопления | δ скопления | a скопления | b скопления |
|--------------------|--------------------|---------------|---------------|
| 1:32:35 | 30:30:26 | 8.26 | 11.6 |
| 1:34:16 | 30:28:26 | 12.0 | 14.0 |
| 1:35:07 | 30:41:46 | 12.5 | 25.0 |
| 1:33:28 | 30:31:31 | 12.0 | 15.0 |
| 1:34:15 | 30:33:41 | 12.0 | 15.0 |
| 1:34:10 | 30:34:38 | 9.5 | 15.0 |

| | | | |
|---------|----------|-----|------|
| 1:33:35 | 30:36:28 | 9.0 | 12.0 |
| 1:33:47 | 30:38:39 | 8.0 | 10.0 |
| 1:33:52 | 30:39:15 | 5.0 | 6.0 |

Таблица Д.4: Измеренные вручную размеры туманностей для подтвержденных LBV, α, δ — координаты в формате h:m:s, a, b — малая и большая полуось в угловых секундах, Δx — расстояние между центрами скопления и туманности в угловых секундах

| α скопления | δ скопления | a скопления | b скопления | Δx |
|--------------------|--------------------|---------------|---------------|------------|
| 1:32:35 | 30:30:28 | 12.9 | 17.3 | 30.0 |
| 1:34:16 | 30:28:22 | 17.0 | 22.0 | 15.0 |
| 1:35:07 | 30:41:51 | 15.0 | 28.8 | 35.0 |
| 1:33:28 | 30:31:32 | 14.0 | 23.0 | 64.0 |
| 1:34:15 | 30:33:45 | 16.0 | 21.0 | 7.0 |
| 1:34:10 | 30:34:34 | 30.0 | 37.0 | 1.0 |
| 1:33:35 | 30:36:28 | 9.0 | 12.0 | 30.0 |
| 1:33:47 | 30:38:41 | 14.0 | 21.0 | 35.0 |
| 1:33:52 | 30:39:15 | 7.0 | 7.5 | 7.0 |