

УДК 519.7.007.52
КП
№ держреєстрації 0110U002636
Инв.№

Міністерство освіти і науки, молоді та спорту України
Харківський Національний університет радіоелектроніки
(ХНУРЕ)

61166, м. Харків, проспект Леніна, 14
тел. /факс: (057) 702 10 13

ЗАТВЕРДЖУЮ
проректор з наукової роботи,
д. ф.-м.н., проф.

_____ М.І.Сліпченко

З В І Т

ПРО НАУКОВО-ДОСЛІДНУ РОБОТУ

«Моделі і методи грануляції та інтерпретації багатовимірних даних»

(тема № 246)

(заключний)

Керівник НДР
доктор техн. наук,
професор

В.П. Машталір

2012

Рукопис завершено 23 грудня 2012 р.

Результати роботи розглянуті науково методичною радою ХНУРЕ,
протокол № 25 від 25 грудня 2012р.

СПИСОК АВТОРОВ

Руководитель темы докт. техн. наук, проф	Машталир В.П.
Ответственный исполнитель канд.техн.наук, доц.	Машталир С.В.
докт. техн. наук, проф	Пуятин Е.П.
докт. техн. наук, проф	Гороховатский В.А.
докт. техн. наук, проф	Шляхов В.В.
канд.техн. наук, доц.	Белоус Н.В.
канд.техн. наук, доц	Кобылин О.А.
канд.техн. наук	Киношенко Д.К.
канд.техн. наук	Гороховатский А.В.
канд.техн. наук	Передрий Е.А.
канд.техн. наук	Сакало Е.С.
канд.техн. наук	Щербинин К.С.
канд.техн. наук	Климов И.М.
с.н.с.	Ляшенко В.В.

м.н.с.	Белоус И.А.
м.н.с.	Горбач Т.В.
аспирант	Власенко Н.В.
аспирант	Красов А.И.
аспирант	Куликов Ю.В.
аспирант	Мантула Е.В.
аспирант	Полякова Т.В.
студент	Братчун А.С.
студент	Кузьменко Е.Ю.
студент	Удовиченко А.Г.

СОДЕРЖАНИЕ

Введение.....	8
1. Групповые структуры на фактор-множествах в задачах классификации.....	17
1.1 Мультиалгебраические системы с единым носителем в виде декартового куба.....	22
1.2 Аксиоматика мультигрупп.....	29
1.3 Условия существования мультигрупп.....	35
1.4 Примеры продуцирования мультигрупп.....	37
1.5 Анализ условий возникновения групповых преобразований на фактор-множествах при обработке видео.....	43
1.6 Сегментация видеопотоков во времени на основе сравнения фактор-множеств, индуцированных пространственной сегментацией отдельных видеок кадров.....	48
2. Методы анализа изображений в условиях неполной информации.....	55
2.1 Метод нормализации следящего типа для перспективы и его применение при нормализации проективных преобразований.....	55
2.2 Нормализация проективно искаженных изображений с использованием инвариантных отображений.....	60
2.3 Нормализация на базе отношения четырех точек.....	64
2.4 Нормализация и классификация путем установления соответствий систем фрагментов на основе голосования как основы компонентного описания мер сопоставления изображений.....	70
2.5 Построение и анализ свойств функциональных признаков.....	79
3. Исследование и разработка методов распознавания и анализа графических сцен.....	85
3.1 Анализ существующих методов и подходов анализа сцен.....	85
3.1.1 Семантический анализ изображений.....	90
3.1.1.1 Общее описание семантического анализа изображений.....	90
3.1.1.2 Грамматика на основе простых фрагментов изображения.....	92

3.1.1.3 Грамматика на основе простых элементов.....	92
3.2 Общая проблематика задач обработки изображений.....	94
3.2.1 Семантическая нейронная сеть.....	95
3.2.2 Описание принципов разработанного подхода.....	102
3.3 Описание архитектуры и API программного комплекса.....	104
3.3.1 Общее описание архитектуры приложения.....	104
3.3.2 Используемые платформы и программные средства.....	105
3.3.2.1 Языковое решение.....	105
3.3.2.2 Графоподобная база данных.....	105
3.3.2.3 Иерархия основных интерфейсов механизма распознавания.....	106
3.3.3 Графический интерфейс и руководство пользователя.....	112
Выводы.....	118
Список использованных источников.....	121

РЕФЕРАТ

Розроблено методи та моделі контекстної кластеризації відеоданих; методи та моделі мікроситуаційного аналізу даних; методи нормалізації зображень та методи побудови інваріантних до геометричних перетворень ознак із використанням проєкцій зображення. Проведено дослідження та оцінка ефективності різних підходів і методів розпізнавання зображень у задачі комплексного аналізу графічних сцен. Обрана методологічна база, а саме семантичний підхід до аналізу зображень. Семантичний підхід до розпізнавання зображень являється значно більше гнучким, ніж традиційний статистичний, та припускає побудову інваріантних систем розпізнавання зображень, що не залежать від предметної області. Дані переваги мають велике значення для задачі аналізу графічних сцен, оскільки предметна область вихідних зображень не можуть бути обмежена. Узагальнений підхід до використання мір схожості при розв'язанні задач аналізу мультимедійних даних. Також в результаті були розроблені гібридні методи статистично-семантичного аналізу та розпізнавання графічних сцен, що не залежать від предметної області, та дозволяють виконувати обробку графічної інформації у режимі реального часу. розроблено алгоритм розпізнавання та аналізу сцен на базі семантичної нейронної мережі.

ABSTRACT

Context clusterization videodata methods and models; micro-situation data analysis methods and models; images normalization methods and construction invariant to geometrical transformations signs methods with the use image projections are worked out. A study and estimation efficiency of different approaches and artificial perception methods are undertaken in the complex analysis task of the graphic stages. A methodological base is chosen, namely semantic going near the images analysis. The semantic going near artificial perception is considerably anymore flexible, than traditional statistical, and assumes the construction of invariant IUS, that does not depend on a subject domain. These advantages matter very much for the graphic analysis stages, as a subject domain of initial images can not be limit. Generalized going near using of likeness measures for thr tasks of multimedia data analysis solved. Also the hybrid methods of statistically-semantic analysis and recognition of the graphic stages, that does not depend on a subject domain, and allow to execute treatment of graphic information real-time, were worked out as a result. The algorithm of recognition and analysis of the stages is worked out based on semantic neuron network.

ВВЕДЕНИЕ

Неявные (внутренние) связи, заложенные в исходные данные, и заданные (внешние) соотношения, определяемые решаемой задачей, представляют собой основную причину изучения моделей и методов, фактически управляющих обобщением (огрублением) или, напротив, конкретизацией (детализацией) информации. Решение подобных задач лежит, как правило, в рамках направления исследований, называемого грануляционным исчислением (Granular computing) [1, 2, 3]. Под гранулами понимается совокупность элементов, объединенных неразличимостью, сходством, близостью или функциональностью) [4, 5]. Одной из наиболее удобных моделей грануляции для получения, прежде всего, строгих теоретических результатов является представление гранул классами эквивалентностей или толерантностей, что нашло развитие в виде изучения мультиалгебраических систем, оперирующих именно фактор-множествами. Принимая во внимание, что грануляционное исчисление находится в стадии становления [3] проанализируем тенденции его развития. рассмотрим основные подходы, ориентированные на собственно грануляцию, изучение и учет свойств элементов гранул, а также на учет взаимодействия, в том числе и стратифицированного, гранул и их семейств, т.е. на учет внутренних, внешних и контекстных свойств группировок данных. Особое внимание уделим способам сравнения фактор-множеств, что является необходимым и, собственно, для получения и оценки процедур грануляции, и главное – для стратифицированного анализа данных с учетом степени детализации их представления.

Недостаточность, неопределенность, неоднозначность, неполнота, противоречивость информации, а зачастую и избыточность – основные факторы, определяющие пути моделирования объектов, процессов и явлений [6, 7]. Природа этих факторов разнообразна, это (наряду с собственно сложностью объекта исследования обычно) – погрешности регистрации данных, в частности, при использовании различных шкал, при косвенных измерениях, при существенных изменениях условий регистрации, при влиянии внешних возмущающих воздей-

ствий [7, 8]. Нередко устранение, по крайней степени, некоторое нивелирование влияния этих факторов возможно за счет обобщения (огрубления) данных при обработке с последующей конкретизацией (детализацией) при анализе и интерпретации.

Неразличимость и сходство, математической экспликацией которых являются, вообще говоря, отношения эквивалентности и толерантности, выступают в качестве основных инструментов при разработке моделей и методов грануляции [2, 9, 10]. Вместе с тем различие предметных областей и целей моделирования привели к возникновению целого ряда подходов [1, 3, 11, 12], удовлетворяющих, однако, общим принципам. На основе ретроспективного анализа можно резюмировать: в настоящее время методы теории грануляции – суть междисциплинарных исследований, активно развивавшиеся прежде всего в рамках интервального анализа [13], «грубых» (аппроксимационных) множеств (rough sets) [12, 14, 15], вычислительного интеллекта, математического аппарата баз данных с контекстным поиском [16, 17], универсального шкалирования, кластерных методов и т.п. [18, 19].

В основу интервального анализа [20] положено представление любого числа оценкой снизу и оценкой сверху, образующими интервальное число: $[x] \times [x_1, x_2]$. Это дает возможность учитывать погрешности любой природы при задании исходных данных и погрешности, вносимые обработкой этих данных. В итоге создаются предпосылки для достаточно корректного учета ряда моделей неоднозначности и неточности, индуцируемых используемыми методами преобразований информации. Базовый инструмент исследований – интервальные расширения, т.е. представление функций $f: \mathbb{R}^n \rightarrow \mathbb{R}$ в виде $[f]: [\mathbb{R}^n] \rightarrow [\mathbb{R}]$, где $[f]([x]) \supseteq \{f(y) : y \in [x]\}$ [21]. Значительные успехи в данном направлении получены в основном для учета распространяющихся погрешностей вычислений: изучены структура множества решений системы линейных алгебраических уравнений с интервальными коэффициентами, итерационные методы их решения, а также систем нелинейных алгебраических уравнений [20]

и т.д. Большой интерес представляют результаты, полученные при синтезе моделей и решении задач оптимального размещения [22]. Однако при недостаточно формализованных источниках неопределенности (трактуя ее в широком смысле) информации вопросы конструктивного применения интервального анализа в задачах грануляции остаются достаточно проблематичными.

Можно считать, что аппарат «грубых» множеств [23, 24] является естественным обобщением интервального анализа. Это объясняется следующим образом: произвольное множество некоторого универсума представляется двумя оценками – аппроксимацией снизу и аппроксимацией сверху, т.е. совокупностью элементов разбиения или покрытия, которые полностью принадлежат ему или имеют непустое пересечение.

Формально это означает следующее. Пусть $A = \langle U, R \rangle$ – аппроксимационное пространство (информационная система по З. Павлаку [15]), где U – непустое множество объектов, описываемых конечными множествами атрибутов, R – некоторое отношение (обычно эквивалентности или толерантности), индуцированное значениями этих атрибутов. Тогда, если фактор-множество имеет вид $U/R = \{[x]_1, [x]_2, \dots, [x]_n\}$, то произвольное множество $Z \subseteq U$ может характеризоваться аппроксимацией снизу

$$A_*(Z) = \{x \in U : [x]_\alpha \subseteq Z\} = \cup_{[x]_\alpha \subseteq Z} [x]_\alpha$$

и сверху

$$A^*(Z) = \{x \in U : [x]_\alpha \cap Z \neq \emptyset\} = \cup_{[x]_\alpha \cap Z \neq \emptyset} [x]_\alpha.$$

Границей аппроксимации (фактически областью неопределенности) является разность множеств $\partial Z = A^*(Z) / A_*(Z)$. Если $\partial Z = \emptyset$ т.е. $A_*(Z) = A^*(Z)$, то множество Z является полностью определяемым в аппроксимационном пространстве A классами толерантностей или эквивалентностей, в противном случае – нет. В этом случае возможны четыре ситуации для множества Z , оно:

– грубо определяемое в A , т.е. $A_*(Z) \neq \emptyset$ и $A^*(Z) \neq U$;

- внешне неопределяемое в A , т.е. $A_*(Z) \neq \emptyset$ и $A^*(Z) = U$;;
- внутренне неопределяемое, т.е. $A_*(Z) = \emptyset$ и $A^*(Z) \neq U$;;
- неопределяемое в A , т.е. $A_*(Z) = \emptyset$ и $A^*(Z) = U$.

Точность факторизованного представления множества Z грубыми множествами дает оценка $\gamma_R(Z) = \text{card } A_*(Z) / \text{card } A^*(Z) \in [0, 1]$, которая по сути – отношение количества классов фактор-множества U / R явно определяющих Z к количеству классов, возможно определяющих множество $Z \subseteq U$. Важными для практики характеристиками грубых множеств являются редукты, т.е. сокращенные наборы отношений $R' \subseteq R$, продуцирующие то же фактор-множество U / R . Они определяют достаточный набор признаков аппроксимационного пространства A . Если этот набор одновременно является и необходимым, то он задает ядро аппроксимационного пространства. Однако, несмотря на достаточно универсальный характер аппарата грубых множеств, они могут использоваться лишь для конечномерных представлений данных, и остаются открытыми вопросы, связанные, в частности, с трансформациями фактор-множеств, их изоморфизмом. Отсюда следует, что грубые множества пока не могут использоваться в качестве универсального математического аппарата грануляционного исчисления.

В качестве еще одного обобщения грануляции, можно указать общую теорию нечетких множеств [4, 5, 25]. Как правило, грануляция представляет собой итеративную алгоритмическую процедуру последовательного выделения частей различного уровня общности и согласования уровней абстракции и конкретизации при анализе неочевидно структурированных динамических систем. Отношения между гранулами описываются с помощью нечетких графов и нечетких логических правил «если..., то ...». При этом нередко используются методы многокритериальной оптимизации, а в качестве критериев – различные формальные и содержательные критерии, например, критерии информационной важности, точности прогнозирования и т.д. [4, 5]. Грануляционное исчисление в подобных формах дает множественное многоуровневое нечеткое опи-

сание объектов, процессов и явлений, но его унифицированное использование затруднено вследствие необходимости учета различных функций принадлежности.

Перспективным направлением развития грануляционного исчисления является исследование мультиалгебраических систем [26-29], которые возникли на базе теории компараторной идентификации [30] путем распространения положений традиционных алгебраических систем на фактор-множества.

Теория компараторной идентификации заключается в восстановлении различных операторов по их предикатным представлениям. Общая схема выглядит следующим образом: определенному классу операторов ставится в соответствие некоторый класс предикатов. Связь между оператором и предикатом выражается в том, что оператор является неотъемлемой частью структуры предиката. Необходимо найти экспериментально проверяемые условия или свойства предиката, которые определяли бы его структуру, т.е. являлись бы характеристическими. Тогда проверка их в эксперименте позволяет выяснить не только структуру предиката, но и структуру, и параметры неизвестного оператора, т.е. решить задачу идентификации. Формально это означает следующее. Допустим, компараторная модель реализована в виде бинарного отношения или предиката

$$E(x_1, x_2) = D(F[x_1], F[x_2]), \quad (1.1)$$

где x_1, x_2 – элементы множества входных сигналов,

$y_1 = F[x_1], y_2 = F[x_2]$ – элементы множества выходных сигналов B ,

D – стандартный предикат равенства, заданный на декартовом квадрате множества B .

Из выражения (1.1) видно, каким образом в структуру предиката включается неизвестный оператор F . В рамках компараторной идентификации необходимо по экспериментальным свойствам предиката E восстанавливать структуру F .

При изучении мультиалгебраических систем в общем виде изучены мультиотношения и мультиоперации, найдены условия их существования мультимоделей и мультиалгебраических систем, их изоморфизма ; проведено всесторонне изучение свойств [26-29]. Однако исследования проводились в общем виде, и поэтому их практическое использование нередко затруднено. Иначе говоря, необходимо исследовать конкретные математические структуры, распространяя их свойства на фактор-множества. Одной из таких структур и являются мультигруппы.

Итак, грануляция информации позволяет представлять нечто «целое» семействами классов, каждый из которых, в свою очередь, является множеством элементов, неразличимых с точки зрения внутренних (сущностных), внешних (структурных) или контекстных (предметно-ориентированных) свойств. В основу математического инструментария могут быть положены n -арные отношения, обобщение свойств которых обеспечит методологию и методику информационного анализа неочевидно структурированных динамических систем, например, в виде мультигрупп, носителями в которых являются фактор-множества.

При анализе средств сравнения фактор-множеств отметим, что во-первых, известные соотношения могут использоваться лишь для конечных множеств с мерой – мощностью подмножеств, т.е. требуется получение новых метрик для произвольных измеримых множеств. Во-вторых, поведение известных метрик позволяет с достаточной степенью адекватности сопоставлять лишь близкие разбиения. Это объясняется тем, что в их основу, в той или иной форме, положено сходство классов эквивалентности (как правило, в виде их пересечений). Таким образом, новые метрики в некоторой кумулятивной форме должны учитывать и меру сходства классов эквивалентности, и меру их различия. В-третьих, вычислительные модели известных метрик имеют достаточно высокую сложность, что при обработке фактор-множеств с большим числом элементов может существенно снижать их практическую ценность. Наконец, особо следует подчеркнуть острую потребность в исследованиях метрических

свойств стратифицированной грануляции, что позволит определять разумные пределы огрубления или детализации информации, а в конечном итоге – создать некоторые предпосылки для адекватного представления ее семантики.

Природа возникновения мультигрупп на фактор-множествах, как правило, не известна. Однако существует ряд задач, когда имеется их ясный физический смысл. В этом плане наиболее характерными и часто встречающимися на практике являются индуцированные перемещением в пространстве объекта и/или датчика геометрические деформации плоскости. Традиционно подобные преобразования моделируются группами преобразований плоскости [31, 32]. В любом случае возникает вопрос об инструментах, позволяющих обрабатывать фактор-множества. Оказывается что методы математической морфологии весьма удобны для этих целей, поскольку каждый класс эквивалентности можно трактовать, как бинарное множество.

На основе базовых операций и специфического выбора прототипа вводятся новые операции. Среди производных операций бинарной морфологии принято выделять операцию «взятия внутреннейности» $\gamma_B(A) = \delta_B(\varepsilon_B(A))$ и операцию замыкания $\varphi_B(A) = \varepsilon_B(\delta_B(A))$.

Применительно к визуальной информации, начиная с алгоритмизации преобразований средних осей, активно развиваются методы и алгоритмы построения скелетонов изображений, что направлено на автоматизацию анализа формы объектов. Существенный интерес представляют и вопросы сжатия информации, поскольку одним из основных критериев построения скелетонов изображений является требование обратимости, откуда следует возможность восстановления исходного бинарного изображения. Известен ряд и достаточно нетрадиционных применений скелетонов, в частности, при синтезе визуальной информации. В качестве одного из развиваемых аспектов скелетонизации выступают алгоритмы утолщения и утончения линий, которые используются и при построении контурных препаратов. Вместе с тем, несмотря на весьма значительные результаты в данной области, вопросам использования скелетонов в

задачах поиска параметров мультигрупп уделено недостаточное внимание

Современное состояние и тенденции развития информационных технологий подчеркивают непрерывный рост числа задач, в которых требуется поиск разумного компромисса между степенью детализации обрабатываемых данных и адекватностью отражения в моделях реальных объектов, процессов и явлений. При этом существенно расширяется и перечень проблемно-ориентированных областей, в которых грануляция информации играет доминирующую роль практически на всех этапах регистрации, приема-передачи, обработки, кодирования, архивирования и интерпретации данных. В этом плане уже достаточно традиционной задачей, решение которой, к сожалению, еще далеко от завершения, является синтез моделей, основанных на факторизации информации.

Проведенный анализ основных тенденций развития грануляционных подходов позволяет заключить, что перспективным направлением исследований в области теории и практики математического моделирования, является развитие моделей и методов в рамках мультиалгебраических систем. Основное внимание на данном этапе должно уделяться исследованию конкретных математических структур, в том числе и мультигруппам, изучению которых посвящена диссертационная работа.

Одним из важнейших механизмов возникновения фактор-множеств являются n -арные отношения, предопределенные природой исходных данных или привнесенные требованиями прикладной задачи. В связи с этим необходимо не только уметь анализировать индуцированные мультиалгебраические системы, но и иметь возможности управления процессами факторизации. Особую роль при этом играют метрические свойства фактор-множеств, изучению которых следует уделить самое пристальное внимание.

Из вышеизложенного можно заключить, что для синтеза эффективных моделей факторизации на основе мультигрупп необходимо:

– провести синтез, обоснование и исследование тернарных моделей, индуцированных произвольными n -арными отношениями;

– найти необходимые и достаточные условия существования мультигрупп на множествах произвольной природы;

– исследовать метрические свойства фактор-множеств как базы компораторных методов анализа и интерпретации данных, представленных в виде различных разбиений;

– исследовать специфику реализации и использования мультиалгебраических систем в виде мультигрупп, создать и внедрить исследовательские и специализированные программные средства.

РАЗДЕЛ 1

ГРУППОВЫЕ СТРУКТУРЫ НА ФАКТОР-МНОЖЕСТВАХ В ЗАДАЧАХ
КЛАССИФИКАЦИИ

Пусть на декартовом произведении $A \times B$ произвольных множеств A, B задано отношение $T(x, y)$. Если оно дифункционально, т.е. $T = TT^{-1}T$, то оно индуцирует и на области определения, и на области значений отношения эквивалентности, фактор-множества по которым равномощны. В случае $\text{card } \{A\} < \infty, \text{card } \{B\} < \infty$, из дифункциональности следует

$$\left. \begin{array}{l} T(a_i, b_j) = 1 \\ T(a_k, b_j) = 1 \\ T(a_k, b_l) = 1 \end{array} \right\} \Rightarrow T(a_i, b_l) = 1.$$

Иначе говоря, при задании отношения $T(x, y)$ матрицей A_T , если на местах $(i, j), (k, j), (k, l)$ стоят 1, то и на месте (i, l) тоже будет 1. Ясно, что эти «четырехугольники» перенумерацией элементов множеств A и B могут быть трансформированы в единичные 2×2 «квадраты» и помещены в любое место матрицы A_T . Очевидно и то, что подобные матрицы могут быть приведены к блочному виду

$$\left(\begin{array}{cccccccccccc} \overbrace{I \dots I}^{m_1} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & & & & & & & & \vdots \\ I \dots I & 0 & \dots & 0 \\ 0 & 0 & I \dots I & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & & & & & & \vdots \\ \vdots & \vdots & 1 \dots I & 0 & \dots & 0 \\ \dots & \vdots \\ \vdots & \vdots & & & & & 1 \dots 1 & 0 \dots 0 & & & & \vdots \\ \vdots & \vdots & & & & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & & & & & 1 \dots 1 & 0 \dots 0 & & & & \vdots \\ \vdots & \vdots & & & & & 0 \dots 0 & 0 \dots 0 & & & & \vdots \\ \vdots & \vdots & & & & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 \dots 0 & \dots & \dots & \dots & \dots & \dots & 0 \dots 0 & 0 \dots 0 & \dots & \dots & 0 \dots 0 & \vdots \end{array} \right),$$

где блоки состоят только из 1, имеют размеры (n_i, m_i) , $i = \overline{1, r}$ и $\sum_{i=1}^r n_i = n$, $\sum_{i=1}^r m_i = m$, $n = \text{card}\{A\}$, $m = \text{card}\{B\}$.

Поскольку любая перенумерация – это действие некоторой перестановки S_n , где n – число элементов, любую $n \times m$ матрицу можно трансформировать применением двух перестановок S_n и S_m . Результат этой трансформации произвольной матрицы A будем обозначать $S_{n,m}(A)$. В результате справедливо следующее

Утверждение 1. Для любого конечного дифункционального отношения $T(x, y)$ всегда найдутся две перестановки, для которых $S_{n,m}(A_T)$ имеет блочный вид.

Отметим, что отношение дифункциональности представляет собой модель $\langle \{A \times B\}, T(x, y) \rangle$ (будем называть ее дифункциональной). Но из вышесказанного следует, что на классах эквивалентности порождается мультимодель, в которой мультиотношение единственно и представляет собой простейшую эквивалентность – эквивалентность равенства

$$I(x, y) = \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases}$$

Таким образом, справедливо

Следствие 1. Любая дифункциональная модель всегда индуцирует мультимодель в виде эквивалентности равенства.

При этом, если у двух дифункциональных моделей число блоков в матрицах $S_{n_1, m_1}(A_{T_1})$ и $S_{n_2, m_2}(A_{T_2})$ совпадает (а природа множеств, образующих их носитель, несущественна), то индуцируется одна и та же мультимодель в виде эквивалентности равенства, что фактически означает их алгебраическую неразличимость.

Введем определения.

Определение 1. Порядком отношения дифункциональности T и соответствующей ей модели дифункциональности назовем число блоков в матрице $S_{n,m}(A_T)$ и будем его обозначать $\pi(T)$.

Определение 2. Два отношения дифункциональности T_1 и T_2 и соответствующие им модели будем называть изоморфными, если они порождают одну и ту же мультимодель в виде эквивалентности равенства.

Таким образом, можно сформулировать

Следствие 2. Две дифункциональности T_1 и T_2 (дифункциональные модели) изоморфны тогда и только тогда, когда $\pi(T_1) = \pi(T_2)$, т.е. $T_1 \square T_2 \Leftrightarrow \pi(T_1) = \pi(T_2)$.

Рассмотрим, при каких условиях произвольное тернарное отношение $F(x, y, z)$ на $A \times B \times C$ индуцирует мультисистему с единым носителем. Поскольку произвольное тернарное отношение представляет собой семейства бинарных, среди которых на уровне мультисистем только дифункциональные отношения формируют единый носитель с точностью до природы элементов, то естественно предполагать, что именно это свойство должно быть определяющим.

Определение 3. Тернарное отношение $F(x, y, z)$ на $A \times B \times C$ является усеченно-дифункциональным по аргументу x , если соответствующее сечение (при любом фиксированном $x \in A$ отношение $F(x, y, z)$ трактуется как бинарное отношение, заданное на $B \times C$) является отношением дифункциональности.

Определение 4. Тернарное отношение на $A \times B \times C$ называется вполне-дифункциональным, если оно является усеченно-дифункциональным по каждому из своих аргументов.

Лемма 1. Пусть $T_1(x, y)$ и $T_2(x, y)$ – дифункциональные отношения на $A \times B$. Тогда если $\forall x_1, x_2 \in A, \forall y \in B: T_1(x_1, y) = T_1(x_2, y) \Leftrightarrow T_2(x_1, y) = T_2(x_2, y)$, то $T_1 \sim T_2$.

Доказательство. Рассмотрим необходимость. Если $\forall y \in B$ для каких-то

$x_1, x_2 \in A$ имеет место равенство $T_1(x_1, y) = T_1(x_2, y)$, то элементы x_1 и x_2 принадлежат одному классу эквивалентности, которые индуцируются отношением T_1 на множестве A , т.е. $x_1 \overset{T_1}{\sim} x_2$ ($\overset{T_1}{\sim}$ — эквивалентность относительно T_1), а из этого следует: $T_2(x_1, y) = T_2(x_2, y)$, т.е. $x_1 \overset{T_2}{\sim} x_2$. Значит, класс эквивалентности A_1 / T_1 содержащий элементы x_1 и x_2 , принадлежит классу A_1 / T_2 , содержащему эти же элементы. Но условие леммы справедливо и в обратную сторону, следовательно, $A_1 / T_1 \subset A_1 / T_2$, $A_1 / T_2 \subset A_1 / T_1$ или $A_1 / T_1 = A_1 / T_2$. Таким образом, разбиения, индуцируемые на множестве A дифункциональностями T_1 и T_2 , совпадают, откуда следует совпадение мощностей разбиений $\text{card}\{A_1 / T_1\} = \text{card}\{A_1 / T_2\}$. Следовательно $T_1 \sim T_2$.

Достаточность очевидна: условие леммы 1 означает совпадение разбиений $\{A / T_1\}$ и $\{A / T_2\}$, но дифункциональности T_1 и T_2 , в случае их изоморфизма, имеют одинаковые разбиения, что и требовалось.

Замечание 1. В случае конечности множеств A и B условие леммы 1 означает равенство количества единичных блоков в матрицах A_{T_1} и A_{T_2} , т.е. $\pi(T_1) = \pi(T_2)$. Эта ситуация рассмотрена в силу практической важности. Однако, если множества A и B произвольны, то в условиях одной области определения двух отношений дифункциональности конечностью можно пренебречь, и лемма 1 справедлива в общем случае. Таким образом, для двух отношений дифункциональности, имеющих одну область определения, критерием изоморфизма является условие леммы 1.

Лемма 1 позволяет получить условия, при которых произвольное тернарное отношение, заданное на декартовом произведении произвольных множеств A, B и C индуцирует мультиалгебраическую систему с единым носителем, т.е. мультиотношение, определенное на декартовом кубе некоторого множества, что является необходимым для формирования мультигрупп.

Утверждение 2. Произвольное тернарное отношение $F(x, y, z)$ на $A \times B \times C$ имеет все изоморфные друг другу сечения тогда и только тогда, когда

оно вполне-дифункционально, и для любых $x_1, x_2 \in A$, $y_1, y_2 \in B$ и $z \in C$ имеют место условия:

$$\text{i) } F(x_1, y_1, z) = F(x_1, y_2, z) \Leftrightarrow F(x_2, y_1, z) = F(x_2, y_2, z);$$

$$\text{ii) } F(x_1, y_1, z) = F(x_2, y_1, z) \Leftrightarrow F(x_1, y_2, z) = F(x_2, y_2, z).$$

Доказательство. Рассмотрим достаточность. Нетрудно заметить, что если тернарное отношение $F(x, y, z)$ является вполне-дифункциональным, то все сечения являются отношениями дифункциональности, заданными на одной области определения.

Из условия i) следует, что в нем присутствуют два сечения $F(x_1, y, z)$ и $F(x_2, y, z)$, полностью удовлетворяющие условиям леммы 1, т.к. условие i) – это фактически условие леммы 1 для этих сечений. Но поскольку это критерий изоморфизма, то для любых $x_1, x_2 \in A$ имеем изоморфизм $F(x_1, y, z) \sim F(x_2, y, z)$. Рассуждая совершенно аналогично, из условия ii) имеем: $F(x, y_1, z) \sim F(x, y_2, z)$.

Если обозначить разбиения, индуцируемые некоторым отношением F на множествах A, B и C соответственно через $\{A/F\}, \{B/F\}, \{C/F\}$, то из $F(x_1, y, z) \sim F(x_2, y, z)$ следует, что мощности разбиений $\{B/F\}$ и $\{C/F\}$ совпадают и не зависят от выбора элемента x , т.е. $\text{card}\{B/F\} = \text{card}\{C/F\}$. Совершенно аналогично из условия $F(x, y_1, z) \sim F(x, y_2, z)$ имеем $\text{card}\{A/F\} = \text{card}\{C/F\}$, но тогда $\text{card}\{A/F\} = \text{card}\{B/F\} = \text{card}\{C/F\}$, что и означает вместе с вполне-дифункциональностью отношения $F(x, y, z)$ изоморфизм всех его сечений. Достаточность доказана.

Перейдем к доказательству необходимости. Если все сечения изоморфны, то вполне-дифункциональность выполняется, поскольку $\text{card}\{A/F\} = \text{card}\{B/F\} = \text{card}\{C/F\}$, откуда следует существование взаимнооднозначного соответствия между множествами классов эквивалентностей, что порождает дифункциональности в каждом из сечений. Условия i) и ii) будут выполняться как критерии изоморфизма. Доказательство утверждения завер-

шено.

1.1 Мультиалгебраические системы с единым носителем в виде декартового куба

Пусть дано n -арное отношение S , заданное на декартовом произведении произвольных множеств $A_1 \times \dots \times A_n$. Разобьем набор множеств $\{A_i\}_{i=1}^n$ на k непустых частей следующим образом: в первую часть входит A_1 и какие-то элементы из оставшихся, во вторую часть входит минимальный из номеров A_i , который не вошел в первую часть и какие-то элементы из остатка, не вошедшие в первую часть, в третью часть – минимальный из номеров A_i , не вошедший в первые две части и какие-то элементы остатка и т.д. Тогда нетрудно понять, что на декартовом произведении множеств $B_1 \times \dots \times B_k$, где B_j представляет собой декартово произведение множеств A_i , входящих в j -ую часть и расположенных в порядке возрастания номеров, заданным n -арным отношением S индуцируется некоторое k -арное отношение, зависящее от способа разбиения на k частей. Все такие разбиения образуют некоторое множество, элемент которого будем обозначать p_k .

Определение 5. Фактор-отношением k -го порядка данного произвольного n -арного отношения S будем называть k -арное отношение, индуцируемое исходным при помощи некоторого разбиения p_k области определения $A_1 \times \dots \times A_n$ на k частей и будем его обозначать $\Phi(p_k) / S$.

Теперь может быть сформулировано и доказано следующее

Утверждение 3. Пусть дано произвольное n -арное отношение S , заданное на декартовом произведении произвольных множеств $A_1 \times \dots \times A_n$. Тогда им индуцируется тернарное мультиотношение с единым носителем (на декартовом кубе некоторого множества) в том и только в том случае, когда существует раз-

биение p_3 набора множеств $\{A_i\}_{i=1}^n$ такое, что соответствующие ему фактор-отношение третьего порядка $\Phi(p_3)/S$ удовлетворяет как тернарное отношение условиям утверждения 2.

Доказательство. Сначала рассмотрим достаточность. Ясно, что если существует некоторое разбиение p_3 набора A_1, A_2, \dots, A_n , то соответствующее ему фактор-отношение третьего порядка, $\Phi(p_3)/S$, индуцируемое отношением S , задано на декартовом произведении множеств $B_1 \times B_2 \times B_3$, где $B_i = A_{i_1} \times \dots \times A_{i_{r_i}}$, где $i \in \{1, 2, 3\}$, а набор $A_{i_1}, A_{i_2}, \dots, A_{i_{r_i}}$ входит в набор A_1, A_2, \dots, A_n в порядке возрастания номеров, т.е. $i_1, i_2 < \dots < i_{r_i}$, r_i – мощность i -ой части разбиения p_3 , при этом $r_1, r_2, r_3 > 0$, $r_1 + r_2 + r_3 = n$.

Поскольку тернарное отношение $\Phi(p_3)/S$ удовлетворяет условиям утверждения 2, оно имеет все изоморфные друг другу сечения. При этом тернарное отношение $\Phi(p_3)/S$ индуцирует на каждом из множеств B_1, B_2, B_3 равномощные разбиения или фактически одно разбиение, с точностью до природы и состава элементов. Тогда будут иметь место равенства

$$\text{card}\{B_1 / [\Phi(p_3) / S]\} = \text{card}\{B_2 / [\Phi(p_3) / S]\} = \text{card}\{B_3 / [\Phi(p_3) / S]\}, \quad (1.2)$$

из которых следует, что множества $\{B_i / [\Phi(p_3) / S]\}$ можно взаимно-однозначно отобразить на некоторое множество C , на декартовом кубе C^3 которого задано тернарное мультиотношение, индуцируемое $\Phi(p_3)/S$ и фактически порожденное произвольным отношением S . Достаточность доказана.

Перейдем к доказательству необходимости. Если S порождает некоторое тернарное мультиотношение на декартовом кубе какого-то множества классов эквивалентности C , то существует некоторое разбиение p_3 набора элементов A_1, A_2, \dots, A_n . Это означает выполнение равенств (1.2), для соответствующих разбиению p_3 множеств B_1, B_2, B_3 . Но из утверждения 2 следует, что эти равен-

ства влекут изоморфизм всех сечений тернарного мультиотношения, индуцируемого S , т.е. выполнения условий: вполне-дифункциональности и условий i), ii) утверждения 2, что завершает доказательство утверждения.

Таким образом, тернарное мультиотношение с единым носителем может порождаться не только исходным тернарным отношением, но и произвольным n -арным. При этом важно выделить следующее

Следствие 3. Любое n -арное отношение индуцирует на классах эквивалентности все отношения по арности от 1 до n .

Формирование любой мультиалгебраической системы требует экспериментальной проверки условий ее существования, т.е. необходимо уметь оперировать данными непосредственно на уровне исходного n -арного отношения. В связи с этим полученные выше результаты представим в явном (будем говорить, координатном) виде.

Введем обозначения. Рассмотрим кортеж элементов x_1, \dots, x_n . Его можно различными способами разбить на m частей с сохранением номера элемента x_i . На рис. 1.1 схематично показана такая выборка.

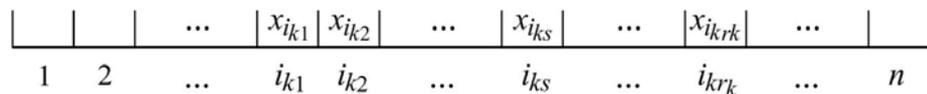


Рис. 1.1 – Элемент разбиения кортежа

На рис. 1.1 выделены позиции, которые занимают элементы $x_{i_{k1}}, x_{i_{k2}}, \dots, x_{i_{krk}}$, причем такая индексация подчеркивает, что важны не сами элементы кортежа, а номера аргументов, входящих в k -ю часть: это номера – $i_{k1}, i_{k2}, \dots, i_{krk}$. Введем преобразование $f_k(\bar{x}) = \{x_{i_{k1}}, x_{i_{k2}}, \dots, x_{i_{krk}}\}$, где \bar{x} есть кортеж x_1, \dots, x_n . Это преобразование $f_k(\bar{x})$ осуществляет выборку элементов с фиксацией их номеров. Например, рассмотрим разбиение кортежа $\bar{x} = \{x_1, x_2, x_3, x_4, x_5\}$ на две части $\{x_1, x_3, x_5\}$ и $\{x_2, x_4\}$, т.е. $f_1(\bar{x}) = \{x_1, x_3, x_5\}$, $f_2(\bar{x}) = \{x_2, x_4\}$, $\bar{x} = \{f_1(\bar{x}), f_2(\bar{x})\}$. Кроме того, если $\bar{x}, \bar{y} \in A_1 \times A_2 \times A_3 \times A_4 \times A_5$, то

$\{f_1(\bar{x}), f_2(\bar{y})\} = \{x_1, y_2, x_3, y_4, x_5\}$. В общем случае в соответствии с проведенными рассуждениями имеем

$$\forall \bar{x}_1, \bar{x}_2, \dots, \bar{x}_m \in A_1 \times \dots \times A_n \Rightarrow \{f_1(\bar{x}_1), f_2(\bar{x}_2), \dots, f_m(\bar{x}_m)\} \in A_1 \times \dots \times A_n.$$

Теперь можно вернуться к координатной формулировке искомым условий.

Лемма 2. Если для n -арного отношения S , заданного на декартовом произведении произвольных множеств $A_1 \times A_2 \times \dots \times A_n$ существует разбиение p_3 такое, что

$$B_1 = A_{i_{11}} \times A_{i_{12}} \times \dots \times A_{i_{1r_1}},$$

$$B_2 = A_{i_{21}} \times A_{i_{22}} \times \dots \times A_{i_{2r_2}},$$

$$B_3 = A_{i_{13}} \times A_{i_{13}} \times \dots \times A_{i_{3r_3}},$$

где $\bar{n} = \{f_1(\bar{n}), f_2(\bar{n}), f_3(\bar{n})\}$, $\bar{n} = \{1, 2, \dots, n\}$, $r_1 + r_2 + r_3 = n$, и для любых $\bar{x}, \bar{y}_1, \bar{y}_2, \bar{z}_1, \bar{z}_2 \in A_1 \times A_2 \times \dots \times A_n$ выполняется импликация

$$\left. \begin{array}{l} S(f_1(\bar{x}), f_2(\bar{y}_1), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}), f_2(\bar{y}_2), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}), f_2(\bar{y}_2), f_3(\bar{z}_2)) = 1 \end{array} \right\} \Rightarrow S(f_1(\bar{x}), f_2(\bar{y}_1), f_3(\bar{z}_2)) = 1,$$

тогда индуцированное тернарное отношение $\Phi(p_3)/S$ на $B_1 \times B_2 \times B_3$ является усеченно-дифункциональным по первому аргументу $\bar{x} \in B_1$.

Справедливость леммы очевидна. Из определения преобразований f_1, f_2, f_3 с учетом импликации в условии леммы имеем: $\Phi(p_3)/S$ – дифункциональность.

Нетрудно заметить, что если будут справедливы импликации по другим аргументам, то из леммы 2 будет следовать усеченная дифункциональность $\Phi(p_3)/S$ по остальным аргументам, т.е. индуцированное тернарное отношение

$\Phi(p_3)/S$ является вполне-дифункциональным. В итоге можно привести координатную формулировку утверждения 3.

Утверждение 4. Произвольное n -арное отношение S индуцирует тернарную мультисистему с единым носителем, если существует разбиение p_3 , удовлетворяющее условиям

i) для любых $\bar{x}, \bar{y}_1, \bar{y}_2, \bar{z}_1, \bar{z}_2 \in A_1 \times A_2 \times \dots \times A_n$ справедлива импликация

$$\left. \begin{array}{l} S(f_1(\bar{x}), f_2(\bar{y}_1), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}), f_2(\bar{y}_2), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}), f_2(\bar{y}_2), f_3(\bar{z}_2)) = 1 \end{array} \right\} \Rightarrow S(f_1(\bar{x}), f_2(\bar{y}_1), f_3(\bar{z}_2)) = 1;$$

ii) для любых $\bar{x}_1, \bar{x}_2, \bar{y}, \bar{z}_1, \bar{z}_2 \in A_1 \times A_2 \times \dots \times A_n$ справедлива импликация

$$\left. \begin{array}{l} S(f_1(\bar{x}_1), f_2(\bar{y}), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}_2), f_2(\bar{y}), f_3(\bar{z}_1)) = 1; \\ S(f_1(\bar{x}_2), f_2(\bar{y}), f_3(\bar{z}_2)) = 1 \end{array} \right\} \Rightarrow S(f_1(\bar{x}_1), f_2(\bar{y}), f_3(\bar{z}_2)) = 1;$$

iii) для любых $\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2, \bar{z} \in A_1 \times A_2 \times \dots \times A_n$ справедлива импликация

$$\left. \begin{array}{l} S(f_1(\bar{x}_1), f_2(\bar{y}_1), f_3(\bar{z})) = 1; \\ S(f_1(\bar{x}_2), f_2(\bar{y}_1), f_3(\bar{z})) = 1; \\ S(f_1(\bar{x}_2), f_2(\bar{y}_2), f_3(\bar{z})) = 1 \end{array} \right\} \Rightarrow S(f_1(\bar{x}_1), f_2(\bar{y}_2), f_3(\bar{z})) = 1;$$

iv) для любых $\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2, \bar{z} \in A_1 \times A_2 \times \dots \times A_n$ имеют место

$$\begin{aligned} S(f_1(\bar{x}_1), f_2(\bar{y}_1), f_3(\bar{z})) = S(f_1(\bar{x}_1), f_2(\bar{y}_2), f_3(\bar{z})) &\Leftrightarrow \\ S(f_1(\bar{x}_2), f_2(\bar{y}_1), f_3(\bar{z})) = S(f_1(\bar{x}_2), f_2(\bar{y}_2), f_3(\bar{z})), & \\ S(f_1(\bar{x}_1), f_2(\bar{y}_1), f_3(\bar{z})) = S(f_1(\bar{x}_2), f_2(\bar{y}_1), f_3(\bar{z})) &\Leftrightarrow \\ S(f_1(\bar{x}_1), f_2(\bar{y}_2), f_3(\bar{z})) = S(f_1(\bar{x}_2), f_2(\bar{y}_2), f_3(\bar{z})). & \end{aligned}$$

Справедливость утверждения достаточно очевидна. Импликации i) – iii) рассмотрены выше, а условия iv) есть ни что иное, как условия утверждения 2 об изоморфизме сечений тернарного отношения.

Подчеркнем практическую важность утверждения 4, поскольку оно в явном виде указывает путь проведения эксперимента при построении и обоснова-

торое не допускает дальнейшей процедуры факторизации, т.е. на каждом месте отношения $M(P_k)/F$ классы эквивалентностей состоят из одного элемента. Сформулируем и докажем этот факт более точно.

Обозначим через $\{B_m/F\}_{m=1}^k$ классы эквивалентностей, которые индуцируются на множествах $\{B_m\}_{m=1}^k$ за счет отношения F , т.е. мультиотношение $M(P_k)/F$ задано на декартовом произведении $B_1/F \times B_2/F \times \dots \times B_k/F$. Справедливо следующее

Утверждение 5. Если для какого-либо фиксированного $m \in \{1, \dots, k\}$ найдутся два элемента $b_m(F), b'_m(F) \in B_m/F$, для которых имеет место равенство

$$M(P_k)/F(b_1(F), \dots, b_m(F), \dots, b_k(F)) = M(P_k)/F(b_1(F), \dots, b'_m(F), \dots, b_k(F)), \quad (1.4)$$

где $\{b_1(F), \dots, b_{m-1}(F), b_{m+1}(F), \dots, b_k(F)\}$ – произвольный набор, принадлежащий декартовому произведению $B_1/F \times \dots \times B_{m-1}/F \times B_{m+1}/F \times \dots \times B_k/F$, то $b_m(F) = b'_m(F)$.

Доказательство. Допустим, выполняется равенство (1.4). При этом $b_m(F)$ и $b'_m(F)$ – это классы эквивалентностей. Возьмем по представителю этих классов $b_{m_{i_m}} \in b_m(F)$ и $b'_{m_{i_m}} \in b'_m(F)$. Аналогично выберем произвольные представители из остальных классов, т.е. $b_{1_{i_1}} \in b_1(F), \dots, b_{m-1_{i_{m-1}}} \in b_{m-1}(F) \in b_{m+1_{i_{m+1}}} \in b_{m+1}(F), \dots, b_{k_{i_k}} \in b_k(F)$. Тогда справедливость (1.4) означает, что имеет место равенство (1.3) для отношения F , т.е.

$$F(b_{1_{i_1}}, \dots, b_{m_{i_m}}, \dots, b_{k_{i_k}}) = F(b_{1_{i_1}}, \dots, b'_{m_{i_m}}, \dots, b_{k_{i_k}})$$

или $b_{m_{i_m}} \mathcal{L} b'_{m_{i_m}}$, т.е. элементы $b_{m_{i_m}}, b'_{m_{i_m}}$ принадлежат обоим классам эквивалентностей $b_m(F)$ и $b'_m(F)$. Таким образом, они пересекаются:

$b_m(F) \cap b'_m(F) \neq \emptyset$, но выполнение последнего равенства для классов эквивалентностей означает их совпадение. Теорема доказана.

Из утверждения 5 следует важное обстоятельство, состоящее в том, что для произвольного n -арного отношения S классы эквивалентностей, формирующиеся на декартовых произведениях множеств из набора A_1, \dots, A_n , являются максимально полными. Мультисистемы, индуцирующиеся отношением S и в общем случае тоже являющиеся отношениями, «различают» на каждом месте своих аргументов все элементы и в этом смысле процедура факторизации является окончательной.

1.2 Аксиоматика мультигрупп

Одна из реализаций мультисистем – распространение свойств традиционных математических структур на фактор-множества. При изучении различных информационных сред важное значение имеет обратимость тех или иных искусственных или заложенных природой свойств, определяющих их функциональность, адекватность и т.п. Иначе говоря, мультигруппы как инструмент изучения разбиений множеств произвольной природы представляют собой объект, имеющий и теоретическое, и практическое значение. Поскольку любая бинарная операция равносильна тернарному отношению, начнем с анализа свойств мультигрупп.

Лемма 3. Произвольное тернарное отношение S , заданное на декартовом кубе непустого множества G элементов произвольной природы индуцирует групповую операцию, относительно которой становится группой, если оно удовлетворяет следующим условиям

i) $\forall a, b \in G \exists! c \in G : S(a, b, c) = 1$;

ii)

$$\forall a, b, c, c_1, c_2, c_3, c_4 \in G : S(a, b, c_1) = S(c_1, c, c_2) = S(b, c, c_3) = S(a, c_3, c_4) = 1 \Rightarrow S(c_1, c, c_4) = 1$$

;

$$\text{iii) } \forall a \in G \exists e : S(e, a, a) = 1;$$

$$\text{iv) } \forall a \in G \exists a^{-1} \in G : S(a^{-1}, a, e) = 1.$$

Индукцируемая отношением S группа будет абелевой, если

$$\text{v) } \forall a, b \in G : S(a, b, c) = 1 \Rightarrow S(b, a, c) = 1.$$

Доказательство. Нетрудно заметить, что свойство i) – это описание закона композиции на языке тернарного отношения S . Действительно, если для произвольных a и b из G найдется единственный элемент c из G , удовлетворяющий равенству $S(a, b, c) = 1$, то его можно сопоставить элементам a и b из G , считая это равенство эквивалентным групповой операции $ab = c$.

Рассмотрим теперь ассоциативность. Если имеет место импликация ii), то рассмотренная эквивалентность тернарного отношения и групповой операции влечет эквивалентность пар равенств $ab = c_1$ и

$$S(a, b, c_1) = 1, \quad (1.5)$$

$c_1 c = c_2$ и

$$S(c_1, c, c_2) = 1, \quad (1.6)$$

что сразу означает:

$$(ab)c = c_2. \quad (1.7)$$

Аналогично, равенство $bc = c_3$ эквивалентно отношению

$$S(b, c, c_3) = 1, \quad (1.8)$$

а равенство $ac_3 = c_4$ эквивалентно равенству

$$S(a, c_3, c_4) = 1, \quad (1.9)$$

т.е. получаем, что

$$a(bc) = c_4. \quad (1.10)$$

Но по свойству ii) из (1.5), (1.6), (1.8), (1.9) вытекает, что $S(c_1, c, c_4) = 1$. Сравним это равенство с (1.6): первые два аргумента совпадают, т.е. из свойства i) следует единственность элемента из G , стоящего на третьем месте, т.е.

$c_2 = c_4$, но тогда с учетом (1.7) и (1.10) получим $(ab)c = a(bc)$, что и требовалось.

Теперь нетрудно понять, что выполнение свойства iii) означает эквивалентность равенств $ea = a$ и $S(e, a, a) = 1$, что означает существование левой единицы.

Аналогично, свойство iv) – это эквивалентность равенств $a^{-1}a = e$ и $S(a^{-1}, a, e) = 1$, что говорит о существовании левого обратного элемента.

Наконец, свойство v) устанавливает коммутативность индуцируемой операцией S операции, т.к. равенство $S(a, b, c) = 1$ эквивалентно равенству $ab = c$, а равенство $S(b, a, c) = 1$ эквивалентно равенству $ba = c$, откуда следует $ab = ba$, т.е. группа абелева. Лемма доказана.

Полученные выше результаты позволяют найти те свойства произвольного n -арного отношения, которые обеспечивают индуцирование им мультигруппы. В утверждении 4 найдены условия, при которых n -арное отношение S индуцирует тернарную мультисистему с единым носителем. Из утверждения 5. следует, что именно на этом этапе окончательно формируются классы эквивалентностей, на которых может формироваться групповая операция. Остается распространить результаты леммы 3 на n -арные отношения.

Используя нотацию леммы 2, тернарная мультисистема $\Phi(p_3)/S$ (далее будем ее обозначать ее через T) задана на декартовом произведении классов $B_1/T \times B_2/T \times B_3/T$. Если выполняются условия утверждения 4, то между классами эквивалентностей взаимно-однозначные соответствия

$$\begin{array}{ccc} B_1/T & \xleftrightarrow{\alpha} & B_2/T \\ \gamma \searrow & & \nearrow \beta \\ & B_3/T & \end{array}$$

$\alpha: B_1/T \leftrightarrow B_2/T$, $\beta: B_2/T \leftrightarrow B_3/T$, $\gamma: B_1/T \leftrightarrow B_3/T$, т.е. справедлива схема

Абстрагируясь (что часто имеет прямой практический смысл) от природы элементов множеств $B_i/T, i = \overline{1,3}$, можно анализировать одно множество B , на

декартовом кубе которого индуцируется тернарное мультиотношение. Если выбрать произвольный элемент $\bar{x} = \{x_1, \dots, x_n\}$ из множества $A_1 \times \dots \times A_n$, на котором задано n -арное отношение S , то отображение $f_1(\bar{x})$ переводит его в множество B_1 , т.е. $f_1(\bar{x}) \in B_1$. При этом элемент $f_1(\bar{x})$ принадлежит определенному классу эквивалентностей B_1 / T , который в свою очередь соответствует определенному элементу множества B . Аналогичная ситуация имеет место для множеств B_2 и B_3 , что можно схематично представить в виде рис. 1.2.

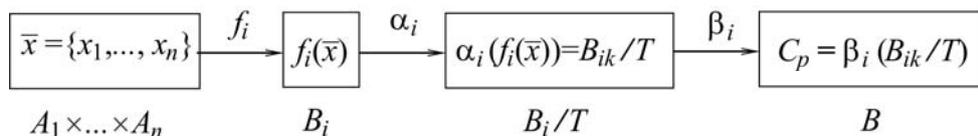


Рис. 1.2 – Индуцирование тернарной мультисистемы ($i = \overline{1,3}$)

Здесь полагается, что отображения α_i ставят в соответствие элементу множества B_i класс, которому он принадлежит, и каждое множество $B_i / T = \{B_{i_1} / T, \dots, B_{i_k} / T, \dots, B_{i_s} / T\}$ состоит из s элементов, а каждый класс за счет отображения β_i переходит в элемент множества B , состоящего из элементов $C_i, \dots, C_p, \dots, C_s$.

Ясно, что можно проследить и обратную цепочку. По произвольному элементу $C_p \in B$ для любого $i \in \{1, 2, 3\}$ можно установить ему соответствующий класс B_{ik} / T ; в множестве B_i / T выбрать любой представитель этого класса $f_i(\bar{x})$ и установить произвольный представитель \bar{x} . Очевидно, что обратные отображения не все однозначны, а именно β_i^{-1} – однозначное отображение в отличие от многозначных отображений α_i^{-1} и f_i^{-1} , но можно выбирать любой элемент образа α_i^{-1} и f_i^{-1} . В результате реализуема схема, представленная на рис. 3 для любых пар номеров i и j ($i, j \in \{1, 2, 3\}$, $i \neq j$, т.к. при $i = j$ возникающие отображения не представляют интереса для продуцирования тернарной мультигруппы).

Как видно из схемы рис. 1.3, в виде суперпозиции отображений индуцируются многозначные отображения ψ_{ij} и φ_{ij} , имеющие вид

$$\psi_{ij} : A_1 \times \dots \times A_n \rightarrow A_1 \times \dots \times A_n, \quad (1.11)$$

$$\psi_{ij} = f_i \circ \alpha_i \circ \beta_j \circ \beta_j^{-1} \circ \alpha_j^{-1} \circ f_j^{-1}, \quad (1.12)$$

$$\varphi_{ij} : B_i \rightarrow B_j, \quad (1.13)$$

$$\varphi_{ij} = \alpha_i \circ \beta_j \circ \beta_j^{-1} \circ \alpha_j^{-1}, \quad (1.14)$$

где « \circ » – символ суперпозиции отображений.

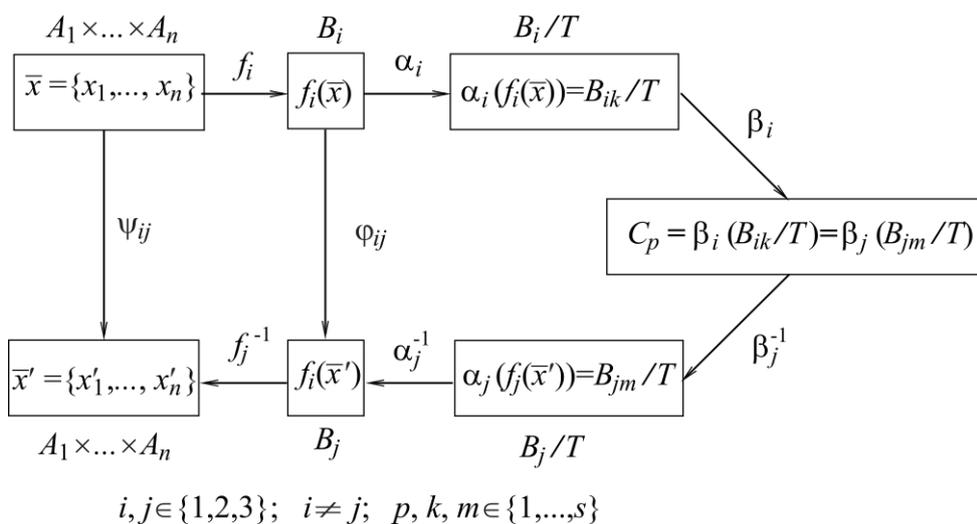


Рис. 1.3 – Индуцирование тернарной мультисистемы

Можно указать ряд свойств, которыми эти отображения обладают.

$$\text{Свойство 1. } \forall \bar{x} \in A_1 \times \dots \times A_n \Rightarrow \psi_{ij}(\bar{x}) = f_j^{-1}(\varphi_{ij}(f_i(\bar{x}))).$$

Это свойство указывает на связь между отображениями ψ_{ij} и φ_{ij} и которое очевидно вытекает из равенств (1.12) и (1.14), указывающих на то, что $\psi_{ij} = f_i \circ \varphi_{ij} \circ f_j^{-1}$, но эта суперпозиция эквивалентна импликации в свойстве 1.

Свойство 2. Отображения φ_{ij} позволяют корректным образом осуществ-

лять любые перестановки аргументов в тернарной мультисистеме, индуцируемой произвольным n -арным отношением S .

Действительно, поскольку имеет место (1.11) и (1.13), то, например, при $i=1, j=2$ для любых $\bar{x}_1, \bar{x}_2, \bar{x}_3 \in A_1 \times \dots \times A_n$ корректной является запись: $S(f_i(\bar{x}_1), \varphi_{12}(f_1(\bar{x}_2)), f_3(\bar{x}_3))$, т.к. $\varphi_{12}(f_1(\bar{x}_2))$ принадлежит B_2 . Важно подчеркнуть следующее: для любого сечения по аргументам x_1, \dots, x_n , осуществляемого отображениями f_i , при помощи отображений φ_{ij} , в тернарной мультисистеме $\Phi(p_3)/S$ его можно поместить на любые два оставшихся места. Подобные операции важны при формулировке условий существования мультигруппы на языке произвольного n -арного отношения S , которое ее индуцирует.

Свойство 3. Для любого $\bar{x} \in A_1 \times \dots \times A_n$ элементы $f_i(\bar{x})$ и $\varphi_{ij}(f_i(\bar{x}))$ соответствуют одному элементу множества B или для них выполняется равенство

$$\beta_j(\alpha_j(f_i(\bar{x}))) = \beta_j(\alpha_j(\varphi_{ij}(f_i(\bar{x}))))$$

Это свойство очевидным образом следует из рис. 2 и определений используемых отображений. На основе рассмотренных свойств можно сформулировать

Лемма 4. Отображения $\varphi_{ij} : B_i \rightarrow B_j$ позволяют любое «усечение» аргументов x_1, \dots, x_n переставлять с i -го места на j -е в тернарной мультисистеме $\Phi(p_3)/S$, и при этом элементы $f_i(\bar{x}) \in B_i$ и $\varphi_{ij}(f_i(\bar{x})) \in B$ принадлежат классам эквивалентностей B_{i_k}/T и B_{j_m}/T соответственно множеств B_i, B_j , которые переходят при отображениях β_i и β_j в один и тот элемент множества B , т.е. выполняется равенство $C_p = \beta_i(B_{i_k}/T) = \beta_j(B_{j_m}/T)$, где $f_i(\bar{x}) \in B_{i_k}/T$, $\varphi_{ij}(f_i(\bar{x})) \in B_{j_m}/T$, $C_p \in B$, $p \in \{1, 2, \dots, s\}$, $i, j \in \{1, 2, 3\}$ $i \neq j$.

1.3 Условия существования мультигрупп

В настоящем пункте сформулируем и докажем итоговый результат отвечающий на основной вопрос: при каких условиях произвольное n -арное отношение S индуцирует мультигруппу.

Теорема (о существовании мультигрупп). Пусть произвольное n -арное отношение S , заданное на декартовом произведении $A_1 \times A_2 \times \dots \times A_n$ множеств произвольной природы, удовлетворяет условиям утверждения 4 и набору свойств:

i) $\forall \bar{x}, \bar{y} \in A_1 \times \dots \times A_n \exists \bar{z} : S(f_1(\bar{x}), f_2(\bar{y}), f_3(\bar{z})) = 1$, если
 $\exists \bar{z}' : S(f_1(\bar{x}), f_2(\bar{y}), f_3(\bar{z}')) = 1$, то

$$\beta_3(\alpha_3(f_3(\bar{z}))) = \beta_3(\alpha_3(f_3(\bar{z}')));$$

ii) $\forall \bar{x}, \bar{y}, \bar{z}, \bar{z}_1, \bar{z}_2, \bar{z}_3, \bar{z}_4 \in A_1 \times \dots \times A_n$ таких, что

$$\begin{cases} S(f_1(\bar{x}), f_2(\bar{y}), f_3(\bar{z}_1)) = 1; \\ S(\varphi_{31}(f_3(\bar{z}_1), f_2(\bar{z}), f_3(\bar{z}_2))) = 1; \\ S(\varphi_{21}(f_2(\bar{y}), f_2(\bar{z}), f_3(\bar{z}_3))) = 1; \\ S(f_2(\bar{y}), \varphi_{32}(f_3(\bar{z}_3), f_3(\bar{z}_4))) = 1, \end{cases}$$

следует $S(\varphi_{31}(f_3(\bar{z}_1), f_2(\bar{z}), f_3(\bar{z}_4))) = 1$;

iii) $\forall \bar{x} \in A_1 \times \dots \times A_n \exists \bar{e} \in A_1 \times \dots \times A_n : S(f_1(\bar{e}), f_2(\bar{x}), \varphi_{23}(f_2(\bar{x}))) = 1$;

iv) $\forall \bar{x} \in A_1 \times \dots \times A_n \exists \bar{x}^{-1} \in A_1 \times \dots \times A_n : S(f_1(\bar{x}^{-1}), f_2(\bar{x}), \varphi_{13}(f_2(\bar{e}))) = 1$.

тогда оно индуцирует мультисистему, образующую группу на классах эквивалентностей.

Доказательство. Из утверждения 4 следует, что n -арное отношение S индуцирует тернарную мультисистему $\Phi(p_3)/S$, заданную на декартовом кубе множества B , представляющего собой множество классов эквивалентностей. Если оно как тернарное отношение удовлетворяет условиям леммы 3, то на множестве B будет индуцироваться групповая операция, а само множество B относительно этой операции будет образовывать группу на классах эквива-

лентностей. Покажем, что тернарное отношение $\Phi(p_3)/S$ удовлетворяет условиям леммы 3, и теорема будет доказана.

Выше показано, что суперпозиция отображений $f_i, \alpha_i, \beta_i, i = \overline{1,3}$ вида $F_i(\bar{x}) = \beta_i(\alpha_i(f_i(\bar{x})))$ ставит в соответствие произвольному элементу $\bar{x} \in A_1 \times \dots \times A_n$ класс эквивалентностей, которому он принадлежит, т.е. некоторый элемент множества B . В свою очередь, введенные ранее отображения φ_{ij} ($i, j \in \{1,2,3\}, i \neq j$) позволяют переставлять $f_i(\bar{x})$ как аргументы отношения S с сохранением классов эквивалентностей, что дает возможность переставлять аргументы тернарного отношения $\Phi(p_3)/S$. Именно это позволяет решить поставленную задачу.

Действительно, рассмотрим свойство i). Обозначим $F_1(\bar{x}) = \beta_1(\alpha_1(f_1(\bar{x}))) = a$, $F_2(\bar{x}) = \beta_2(\alpha_2(f_2(\bar{x}))) = b$, $F_3(\bar{x}) = \beta_3(\alpha_3(f_3(\bar{x}))) = c$. Тогда произвольность элементов \bar{x}, \bar{y} будет означать произвольность элементов a и b как элементов множества B . А для мультитернарного отношения $\Phi(p_3)/S$, как нетрудно видеть, из первого свойства следует, что для любых $a, b \in B$ найдётся $c \in B$, для которого $[\Phi(p_3)/S](a, b, c) = 1$. При этом единственность элемента $c \in B$ обеспечивается второй частью свойства i), поскольку из равенств $F_3(\bar{z}) = c$, $F_3(\bar{z}') = c'$ следует, что для $c' \in B$ выполняется $[\Phi(p_3)/S](a, b, c') = 1$, то $c = c'$. Тем самым групповая операция на множестве B задана.

Ассоциативность этой операции определяется свойством ii). Введем по аналогии обозначения $F_1(\bar{x}) = a$, $F_2(\bar{y}) = b$, $F_3(\bar{z}_1) = c_1$, $F_2(\bar{z}) = c$, $F_2(\bar{z}_2) = c_2$, $F_3(\bar{z}_3) = c_3$, $F_3(\bar{z}_4) = c_4$. Далее отмечаем, что из указанных выше свойств отображений φ_{ij} вытекает, что например, $\phi_{31}(f_3(\bar{z}_1))$ соответствует таким элементам множества $A_1 \times \dots \times A_n$ при многозначном отображении f_1^{-1} , которые лежат в классе эквивалентностей, обозначенном в множестве B через C_1 . В этом случае, например, из первого равенства свойства ii) для тернарного отношения

$\Phi(p_3)/S$ следует $[\Phi(p_3)/S](c_1, c, c_2) = 1$ (здесь использовано понятное свойство: если для отношения S имеет место равенство $S(f_1(\bar{x}), f_2(\bar{y}), f_3(\bar{z})) = 1$, то при введенных обозначениях $[\Phi(p_3)/S](a, b, c) = 1$). Теперь нетрудно видеть, что равенства свойства ii) для отношения $\Phi(p_3)/S$ означают

$$[\Phi(p_3)/S](a, b, c_1) = [\Phi(p_3)/S](c_1, c, c_2) = [\Phi(p_3)/S](b, c, c_3) = [\Phi(p_3)/S](a, c_3, c_4) = 1$$

Но в этом случае справедлива импликация свойства ii), а для $\Phi(p_3)/S$ выполняется равенство $[\Phi(p_3)/S](c_1, c, c_4) = 1$. Окончательно можно сделать вывод о выполнении для $\Phi(p_3)/S$ свойства ii) леммы 3, что означает ассоциативность групповой операции на B – множестве классов эквивалентностей.

Аналогично, при обозначениях $F_2(\bar{e}) = e$, $F_2(\bar{x}) = a$ будут иметь место свойства ii) и iv) леммы 3 в применении к отношению $\Phi(p_3)/S$, т.е. в итоге доказано, что мультитернарное отношение $\Phi(p_3)/S$, индуцируемое произвольным n -арным отношением S удовлетворяет условиям леммы 3, что завершает доказательство теоремы.

Замечание 2. Если необходимо обеспечить коммутативность групповой операции, то n -арное отношение S должно обладать свойством: $\forall \bar{x}, \bar{y} \in A_1 \times \dots \times A_n$ из равенства $S(f_1(\bar{x}), f_2(\bar{y}), f_3(\bar{z})) = 1$ следует $S(\varphi_{21}(f_1(\bar{y})), \varphi_{12}(f_1(\bar{x})), f_3(\bar{z})) = 1$, что легко показать, используя схему доказательства теоремы о существовании мультигрупп.

1.4 Примеры продуцирования мультигрупп

Рассмотрим множество целых чисел \mathbb{R}^+ и некоторое отношение $S \subseteq (\mathbb{R}^+)^4$. Зафиксируем $k \in \mathbb{R}$ и определим это 4-арное отношение в виде:

$$S(n_1, n_2, n_3, n_4) = \begin{cases} 1, & (n_1 + n_2) \bmod k = n_3 \bmod k; \\ 0, & (n_1 + n_2) \bmod k \neq n_3 \bmod k, \end{cases} \quad (1.15)$$

где $a \bmod b \equiv a - \lfloor a/b \rfloor \times b$, $\lfloor \cdot \rfloor$ – ближайшее целое, не превосходящее заданное. Отметим, что заданное отношение, во-первых, не зависит от аргумента n_4 , а, во-вторых, если кортеж $\langle n_1, n_2, n_3, n_4 \rangle$ разбить, например, в виде $p_3 = \{(n_1, n_4), (n_2), (n_3)\}$, и рассмотреть тернарное отношение $\Phi(p_3)/S$, индуцируемое S и заданное на декартовом произведении множеств $B_1 \times B_2 \times B_3$, где $B_1 = (\mathbb{R}^+)^2$, $B_2 = \mathbb{R}^+$, $B_3 = \mathbb{R}^+$, а также ввести функции $f_1(\bar{n}) = \{n_1, n_2\}$, $f_2(\bar{n}) = \{n_2\}$, $f_3(\bar{n}) = \{n_3\}$, то будет иметь место равенство

$$[\Phi(p_3)/S]\{(n_1, n_4), n_2, n_3\} = S(f_1(\bar{n}), f_2(\bar{n}), f_3(\bar{n})). \quad (1.16)$$

При этом тернарное отношение $\Phi(p_3)/S$, задающееся равенством (1.16) с учетом (1.15), будет формировать на множествах B_1, B_2 и B_3 классы эквивалентностей следующего вида.

Во-первых, на множестве $B_1 = (\mathbb{R}^+)^2$ – это целочисленные пары (n_1, n_4) , $(n'_1, n'_4) \in B_1$, для которых

$$n_1 \bmod k = n'_1 \bmod k. \quad (1.17)$$

Действительно, если имеет место (1.17), то для каких-то $q_1, q_2 \in \mathbb{R}^+$ выполняются равенства $n_1 = q_1 k + r_1$, $n'_1 = q'_1 k + r_1$ ($r_1 \in \mathbb{R}_0^{k-1} = \{0, 1, \dots, k-1\}$). Но тогда для любых $n_2, n_3 \in \mathbb{R}^+$ справедлива цепочка равенств:

$$\begin{aligned} (n_1 + n_2) \bmod k &= (q_1 k + r_1 + q_2 k + r_2) \bmod k = \bmod k (r_1 + r_2; k) = \\ &= (q'_1 k + r_1 + q'_2 k + r_2) \bmod k = (n'_1 + n_2) \bmod k, \end{aligned} \quad (1.18)$$

где $n_2 = q_2k + r_2$, $r_2 \in \mathbb{R}_0^{k-1}$.

Теперь с учетом определения 4-арного отношения S (равенство (1.15)) и связи между $\Phi(p_3)/S$ и S (равенство (1.16)) легко видеть: если имеет место (1.17), т.е. остатки у чисел n_1 и n'_1 при делении на k совпадают, то и для отношения S выполняется

$$S(n_1, n_2, n_3, n_4) \equiv S(n'_1, n_2, n_3, n_4) \quad (1.19)$$

при любых $n_2, n_3, n_4 \in \mathbb{Z}^+$, а для тернарного отношения $\Phi(p_3)/S$ справедливо

$$[\Phi(p_3)/S]\{(n_1, n_4), n_2, n_3\} \equiv [\Phi(p_3)/S]\{(n'_1, n'_4), n_2, n_3\} \quad (1.20)$$

при любых элементах $n_2, n_3, n'_4, n_4 \in \mathbb{R}^+$.

Обратное утверждение также справедливо, если выполняются тождества (1.19) или (1.20), которые равносильны, что видно из (1.15) и (1.16), то остатки при делении на k для n_1 и n'_1 должны быть равны. Показывается это от противного: если $n_1 = q_1k + r_1$, $n'_1 = q'_1k + r'_1$ и $r_1 \neq r'_1$, где q_1, q'_1 , и $r_1, r'_1 \in \mathbb{R}_0^{k-1}$, то при $n_2 = 0$, а $n_3 = r_1$ будем иметь цепочку равенств

$$(n_1 + n_2) \bmod k = (q_1k + r_1) \bmod k = r_1 \bmod k = n_3 \bmod k.$$

С учетом равенства (1.15) получим

$$S(n_1, 0, r_1, n_4) = 1 \quad (1.21)$$

при любых $n_4 \in \mathbb{R}^+$, т.к. отношение S не зависит от n_4 . Аналогично будет выполняться

$$[\Phi(p_3)/S]\{(n_1, n_4), 0, r_1\} = 1. \quad (1.22)$$

С другой стороны, имеем $(n'_1 + n_2) \bmod k = (q'_1 k + r'_1) \bmod k = r' \bmod k \neq n_3 \bmod k$ т.к. $n_3 = r_1 \neq r'_1$. Отсюда следует

$$S(n'_1, 0, r_1, n_4) = 0$$

и

$$[\Phi(p_3) / S] \{(n'_1, n'_4), 0, r\} = 0, \quad (1)$$

что объясняется совершенно аналогичными рассуждениями. Однако, равенства (1.21) и (1.23), а также равенства (1.18) и (1.22) противоречат исходным тождествам (1.19) и (1.24) соответственно. Таким образом, показано, что тернарное отношение $\Phi(p_3) / S$ индуцирует на B_1 классы эквивалентности среди пар $(n_1, n_4), (n'_1, n'_4) \in B_1, n_1, n'_1, n'_4, n_4 \in \mathbb{R}^+$, для которых имеет место равенство (1.17).

Во-вторых, аналогично показывается, что на множествах $B_2 = B_3 = \mathbb{R}^+$ отношение $\Phi(p_3) / S$ индуцирует классы эквивалентности среди натуральных чисел, у которых остатки при делении на k равны, т.е. опять имеет место равенство (1.17).

Таким образом, тернарное отношение $\Phi(p_3) / S$ индуцирует тернарное мультиотношение $M(p_3) / S$, заданное на декартовом кубе множестве классов эквивалентностей, представляющее собой набор остатков при целочисленном делении на k , т.е. конечное множество \mathbb{R}^+ . Мультиотношение $M(p_3) / S$ как тернарное отношение на $(\mathbb{R}_0^{k-1})^3$ задает на \mathbb{R}_0^{k-1} бинарную операцию по следующим правилам

$$r_1 \oplus r_2 = r_3 \Leftrightarrow [M(p_3) / S](r_1, r_2, r_3) = 1 \Leftrightarrow S(n_1, n_2, n_3, n_4) = 1$$

где \oplus – операция сложения по модулю k , $n_1 = q_1 k + r_1$, $n_2 = q_2 k + r_2$, $n_3 = q_3 k + r_3$, $n_4 = q_4 k + r_4$, $q_1, q_2, q_3, q_4 \in \mathbb{R}^+$, $r_1, r_2, r_3, r_4 \in \mathbb{R}_0^{k-1}$.

Если следовать терминологии алгебраических систем [26], тройка

$\langle A, F, \Phi \rangle$, где A – произвольное множество (носитель системы), F – набор отношений и Φ – набор операций, заданных на данном носителе (в зависимости от арности на нем или на его декартовых степенях) является алгебраической системой. При этом: если $\Phi = \emptyset$, то имеем модель, если $F = \emptyset$ – алгебру. В рассмотренном случае, изначально имея модель $\langle A, F, \Phi \rangle$, в силу естественной эквивалентности, заложенной «природой» в отношении S , а не привнесенной «извне» (как это, как правило, происходит при рассмотрении процессов факторизации в традиционных алгебраических структурах) получаем алгебру $\langle \mathbf{R}_0^{k-1}, \emptyset, \oplus \rangle$, носителем которой стали классы эквивалентностей, образующие известную алгебраическую структуру – циклическую абелеву группу порядка k . В нашей терминологии это мультигруппа, индуцируемая 4-арным отношением S .

Для полноты рассмотрения необходимо сделать несколько замечаний. С одной стороны проанализированное разбиение p_3 кортежа $\{n_1, n_2, n_3, n_4\}$, очевидно, не единственное, при котором отношением S индуцируется мультигруппа. Например, таким же свойством обладают и разбиения $p_3^1 = \{(n_1), (n_2, n_4), (n_3)\}$, $p_3^2 = \{(n_1), (n_2), (n_3, n_4)\}$, что легко доказывается проведением выше изложенных рассуждений. С другой стороны, не любое разбиение на три части приводит к индуцированию мультигруппы. Если рассмотреть разбиение $p_3^3 = \{(n_1, n_2), (n_3), (n_4)\}$, то становится ясным, что тернарное отношение $\Phi(p_3^3) / S$ будет задано на том же носителе, что и отношение $\Phi(p_3) / S$, т.е. на $B_1 \times B_2 \times B_3$. Однако классы эквивалентностей уже изменятся. Используя (14) и связь между отношением S и $\Phi(p_3) / S$ получаем $[\Phi(p_3^3) / S] \{(n_1, n_2), n_3, n_4\} = S(f_1(\bar{n}), f_2(\bar{n}), f_3(\bar{n}))$, где $f_1(\bar{n}) = \{n_1, n_2\}$, $f_2(\bar{n}) = \{n_3\}$, $f_3(\bar{n}) = \{n_4\}$.

Нетрудно заметить, что классы эквивалентностей на B_1 представляют собой точки в $\mathbf{R}^+ \times \mathbf{R}^+$, удовлетворяющие условию: $(n_1, n_2), (n'_1, n'_2)$ принадлежат

одному классу, если

$$(n_1 + n_1) \bmod k = (n'_1 + n'_2) \bmod k .$$

На B_2 – это множество \mathbb{R}_0^{k-1} , а на B_3 – это один класс: все множество \mathbb{R}^+ . В результате у тернарного мультиотношения $M(p_3)/S$ нет единого носителя и групповая операция, а значит, и мультигруппа индуцироваться не могут.

Таким образом, разбиение, индуцирующие мультигруппу, не единственно, но не все разбиения приводят к требуемому результату. Более того, следует отметить, что не все 4-арные отношения индуцируют мультигруппу, т.е. существуют такие, которые при любом разбиении кортежа из четырёх аргументов на три части ее не индуцируют. Действительно, рассмотрим 4-арное отношение, задающееся равенством

$$S(n_1, n_2, n_3, n_4) = \begin{cases} 1, & (n_1 + n_2) \bmod k_1 = n_3 \bmod k_2; \\ 0, & (n_1 + n_2) \bmod k_1 \neq n_3 \bmod k_2, \end{cases}$$

где $n_1, n_2, n_3, n_4 \in \mathbb{R}^+$, $k_1, k_2 \in \mathbb{R}$ и $k_1 \neq k_2$.

Используя предыдущие рассуждения, нетрудно убедиться, что при рассмотренных разбиениях p_3 (p_3^1, p_3^2) продуцируются классы эквивалентностей, характеризующиеся равенством остатка при целочисленном делении на k_1 и k_2 , т.е. для мультиотношений не формируется единый носитель, а следовательно, и мультигруппа. Для разбиений, которые явно зависят от n_4 , наблюдается еще более существенный эффект: по n_4 образуется один класс эквивалентностей – \mathbb{R}^+ , что в итоге препятствует формированию единого носителя для мультиотношения. Наконец, заметим, что примеры 4-арных отношений могут быть легко трансформированы в n -арные путем продолжения аргументов (аналогично элементу n_4), доводя отношение S до любой арности.

Таким образом, приведенные примеры, с одной стороны, имеют теорети-

ческий смысл, поскольку явно указывают на существование функциональных изучаемых объектов. С другой стороны, они имеют и практический смысл, т.к. свидетельствуют о необходимости экспериментальной проверки сформулированных условий существования мультигрупп.

1.5 Анализ условий возникновения групповых преобразований на фактормножествах при обработке видео

В информационно-аналитических системах на произвольных фактормножествах могут возникать групповые преобразования, генезис которых далеко неочевиден. Однако существует целый ряд случаев, когда эти преобразования естественным образом заложены в процесс построения классов эквивалентностей. Типичным примером являются часто встречающиеся на практике индуцированные перемещением в пространстве объекта и/или датчика геометрические деформации плоскости, которые связаны с принятием решения о сходстве изображений. Традиционно – в силу природы возникновения – подобные преобразования моделируются группами преобразований плоскости. В поле зрения $D \subset R^2$ видеодатчика математические модели преобразований, связывающих изображения $V(x, y)$ и $V_0(x, y)$, в самом общем виде задаются соотношением [31]

$$V(x, y) = k(x, y)V_0[T(x, y)] + r(x, y), \quad (1.25)$$

где

$k(x, y)$ – нелинейные мультипликативные искажения яркости;

$r(x, y)$ – аддитивные изменения фона (включая случайные помехи);

$T(x, y)$ – оператор геометрических преобразований фактически носителя изображения объекта.

На операторы $T(x, y)$ обычно накладываются следующие ограничения:

- геометрические преобразования полностью не выводят изображение объекта за пределы поля зрения;
- геометрические преобразования невырождены, т.е. если

$$T(x, y) = (T^1(x, y), T^2(x, y)),$$

то якобиан J преобразования $T(x, y)$ отличен от нуля

$$J = \det \begin{pmatrix} \frac{\partial T^1(x, y)}{\partial x} & \frac{\partial T^1(x, y)}{\partial y} \\ \frac{\partial T^2(x, y)}{\partial x} & \frac{\partial T^2(x, y)}{\partial y} \end{pmatrix} \neq 0.$$

На практике в качестве основной модели многопараметрических преобразований традиционно используется группа аффинных преобразований плоскости. Это объясняется разумным компромиссом между достаточной простотой обработки (преобразования линейны) и адекватностью реальным условиям наблюдения (при расстоянии до объекта, существенно превосходящем фокусное расстояние видеодатчика, перспективными искажениями можно пренебречь).

На практике любые визуальные данные содержат не только информацию об исследуемом объекте, но и информацию о фоновых составляющих, которые должны устраняться при анализе. Итак, в поле зрения наблюдается изображение $V(x, y)$

$$V(x, y) = \mu(x, y)[\lambda(x, y)V_0(x, y) + [1 - \lambda(x, y)]\Phi(x, y)] + \eta(x, y), \quad (1.26)$$

где

$\mu(x, y)$ – мультипликативные помехи;

$\lambda(x, y)$ – характеристическая функция носителя изображения объекта;

$V_0(x, y)$ – изображение обрабатываемого объекта;

$\Phi(x, y)$ – фоновая составляющая;

$\eta(x, y)$ – аддитивные случайные помехи.

Отметим, что характеристическая функция

$$\lambda(x, y) = \begin{cases} 1, & V_0(x, y) \neq 0, \\ 0, & V_0(x, y) = 0 \end{cases}$$

фактически определяется результатом сегментации видеокадра, и точность обработки и интерпретации в основном определяется именно точностью выделения объектов из фона.

Пусть в некоторые моменты времени t', t'' в поле зрения $D \subset \mathbb{R}^2$ зафиксированы два изображения одного и того же объекта (рис.1.4). Будем полагать, что они связаны аффинным преобразованием плоскости $V'(z) = g \circ V''(z)$, где $z \in \mathbb{R}^2$, $g \in G = \text{Aff}(2, \mathbb{R})$. Тогда действие группы G имеет вид

$$V'(z) = V''(Az) + b, \quad (1.27)$$

где

$A \in Gl(2, \mathbb{R})$ – невырожденная матрица второго порядка;

$b \in \mathbb{R}^2$ – вектор параллельного переноса.

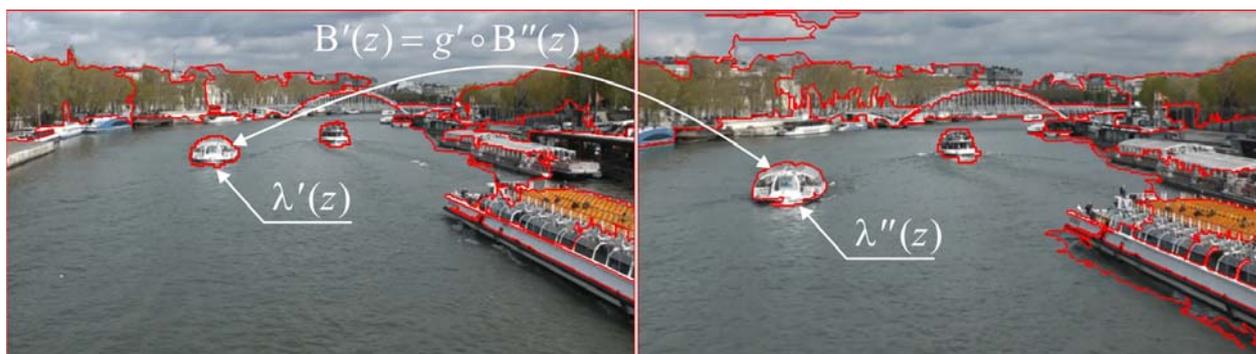


Рис. 1.4 – Пример геометрических преобразований изображений

На рис. 1.5 приведен пример видеопоследовательности длительностью в секунду (медиаконтейнер m2ts, формат AVCHD, разрешение 1920×1080 , частота 29,97 кадр/сек), на которой присутствуют геометрические преобразования объектов (например, движение автобуса в левом нижнем углу). Принимая во внимание (1.25) – (1.27), нетрудно понять, устойчивое автоматическое выделение аффинно-эквивалентных точек обеспечивает простую алгоритмизацию поиска параметров геометрических преобразований, т.к. они могут непосредственно находиться из системы линейных уравнений

$$\begin{cases} x'_k = a_{11}x''_k + a_{12}y''_k + b_1, \\ y'_k = a_{21}x''_k + a_{22}y''_k + b_2, \end{cases} \quad k = \overline{1, K}, \quad K \geq 3. \quad (1.28)$$



Рис. 1.5 – Пример видеоряда с групповыми преобразованиями

С другой стороны, следует подчеркнуть, что аффинные преобразования фактически действуют на носителе, т.е. задача выделения объектов сопряжена с сегментацией каждого (или с пропусками) изображения из видеоряда. Но, во-первых, сегментация – суть построения фактор-множеств, а во-вторых, различия этих фактор-множеств – необходимое условие, сопутствующее изменению сюжета. Таким образом, можно констатировать, что при анализе видео для уче-

та групповых свойств фактор-множества, индуцируемые традиционной пространственной сегментацией представляют собой один из основных объектов изучения, по крайней мере, с точки зрения исследования «пространственного содержания» визуальной информации. Рассмотрим пути получения разбиений поля зрения.

Допустим, поле зрения D разбито на K областей s_β , $\beta = \overline{1, K}$ с количеством элементов в каждой n_β . Поскольку рассматриваются только разбиения, то $D = \bigcup_\beta s_\beta$, и $\forall \beta', \beta'' \in \{1, 2, \dots, K\} : \beta' \neq \beta'' \Rightarrow s_{\beta'} \cap s_{\beta''} = \emptyset$. Пользуясь положениями двухфакторного дисперсионного анализа, полагая $z \in R^2$, рассмотрим средние в каждой области

$$m_\beta = \frac{1}{n_\beta} \sum_{z \in s_\beta} z$$

и общее по полю зрения среднее

$$m = \frac{1}{n} \sum_{z \in D} z,$$

где $n = \text{card } D = \sum_\beta n_\beta$.

Тогда, определяя «внутриобластную» и «межобластную» дисперсии соответственно

$$d' = \sum_{\beta=1}^K \sum_{z \in s_\beta} \|z - m_\beta\|^2,$$

$$d'' = \sum_{z \in D} \|z - m\|^2,$$

где $\|\circ\|^2$ – евклидова норма,

получаем критерий

$$C = (d'' - d') / d'.$$

Очевидно, что чем больше значение C , тем лучше области разделены, а внутренние элементы более «концентрированы» и наоборот. Таким образом, пересчитывая C по каждой сегментированной области и вводя обозначения

$$C^* = \frac{1}{n} \sum_{\beta=1}^K C_{\beta} \text{card } s_{\beta}, \quad (1.29)$$

получаем критерий: необходимо минимизировать C^* по всем возможным разбиениям для фиксированного числа областей. Действительно, если сегментация удачна (области однородны), значения C для каждой области малы, а следовательно, и C^* стремится к нулю. Для сегментации каждого видеокadra, фактически перехода к его «пространственному содержанию» в работе использованы процедуры минимизации (1.29) в виде процедур, изложенных в [33, 34]. Следует отметить, что результаты сегментации, а именно – установление соответствия между ними в два различных момента времени, и позволяют формировать переопределенные системы (1.28) для поиска параметров групповых преобразований.

1.6 Сегментация видеопотоков во времени на основе сравнения факторно-множеств, индуцированных пространственной сегментацией отдельных видеокadров

Прежде, чем перейти к экспериментальному анализу метрической сегментации видеопотоков во времени, т.е. их разбиении (в идеале) на семантически эквивалентные последовательности видеокadров, рассмотрим этот процесс во общем виде.

Рассмотрим видео(поток) $\Phi = \{B_1(x, y), B_2(x, y), \dots, B_K(x, y)\}$. Далее пространственные координаты $x, y \in D$, если это не мешает пониманию ситуации,

будем опускать, акцентируя внимание на дискретном времени $k = \overline{1, K}$. Если $i < j$ и $B_i, B_j \in \Phi$, то видеосегмент (интервал) $S[B_i, B_j]$ на видеопотоке Φ – это множество упорядоченных по времени изображений: $\{B_k \in \Phi : k = \overline{i, j}\}$. Тем самым, начало произвольного сегмента – B_i , его окончание – B_j , а множество всех интервалов – $S(\Phi)$.

Пусть $P_i \subset \mathbb{R}^p$ – вектор признаков, характеризующий изображение B_i , тогда каждый интервал $S[B_i, B_j]$ описывается набором $\Pi_{ij} = \{(B_i, P_i), (B_{i+1}, P_{i+1}), \dots, (B_j, P_j)\}$. Используя «мгновенное» (между двумя последовательными кадрами) изменение изображений, т.е. $\pi_k = \rho(P_{k-1}, P_k)$, ($k = \overline{2, K}$), ρ – некоторая мера сходства или метрика на множестве признаков, изменение признаков на интервале $S[B_i, B_j]$ можно представить в виде

$$\Delta(\Pi_{ij}) = f(\rho(P_i, P_k), \pi_k), k = \overline{i+1, j},$$

где функция $f(\circ)$ учитывает сходство (отличие) начала сегмента с текущим кадром и сходство (отличие) двух последовательных во времени изображений.

Если значение $\Delta(\Pi_{ij})$ меньше наперед заданного или найденного из каких либо априорных соображений порога ε , то $S[B_i, B_j]$ является простейшим сегментом (или просто сегментом) видеопотока Φ . На базе сегментов складываются последовательности видеок кадров, соответствующие семантическим понятиям «события», «сюжета» и т.п. Число уровней стратификации и структура каждого из них существенно зависят от решаемых задач анализа видеопотоков. Выделим лишь понятие «видеогруппы» $F_m \neq \emptyset$, индуцирующей разбиение видеопотока, т.е.

$$F_m = \cup_l S[B_{il}, B_{jl}],$$

$$\Phi = \bigcup_m F_m,$$

$$F'_m \cap F''_m = \emptyset.$$

Вообще говоря, с семантической точки зрения сегменты могут быть не-последовательными во времени (например, при чередовании сюжетов или наличии рекламы). Именно поэтому видеогруппа является основным объектом анализа видеопотока, а в качестве основной единицы обработки выступает сегмент. Не снижая общности рассмотрения, будем полагать, что сегмент и видеогруппа совпадают, объединяя несвязные во времени, но близкие в заданной мере сходства сегменты в один. Нередко основным описанием сегмента $S[B_i, B_j]$ является некоторый ключевой кадр $B_k \in S[B_i, B_j]$, который либо выбирается как самый характерный, либо синтезируется на основе некоторого усреднения в признаковом или сигнальном пространстве.

На рис. 1.6 приведена иллюстрация видеопотока Φ , сегментов $S[B_i, B_j]$, ключевых кадров B_k и «пространственных» фактор-множеств R_k им соответствующих на основе критерия (1.26).

Таким образом, подводя некоторый промежуточный итог, можно конкретизировать схему выделения фактор-множеств с мультигрупповой структурой.

На первом этапе необходимо разбить видеопоток $\Phi = \{B_1, B_2, \dots, B_K\}$ на сегменты $S_l[B_i, B_j]$, $l = 1, 2, \dots, L$ такие, что $\Phi = \bigcup_m S_m$, $S_{m'} \cap S_{m''} = \emptyset$. В качестве критерия будем использовать различие между фактор-множествами R_k , $m = \overline{1, K}$, индуцированных сегментацией изображений. Полагая в качестве меры множества его мощность, имеем

$$\rho(R', R'') = \sum_{i=1}^n \sum_{j=1}^m \text{card}(r'_i \Delta r''_j) \text{card}(r'_i \cap r''_j), \quad (1.30)$$

где $R' = \{r'_i\}_{i=1}^m$, $R'' = \{r''_j\}_{j=1}^n$ – разбиения изображений $B', B'' \in \Phi$. Принятие решения о том, что изображение принадлежит текущему

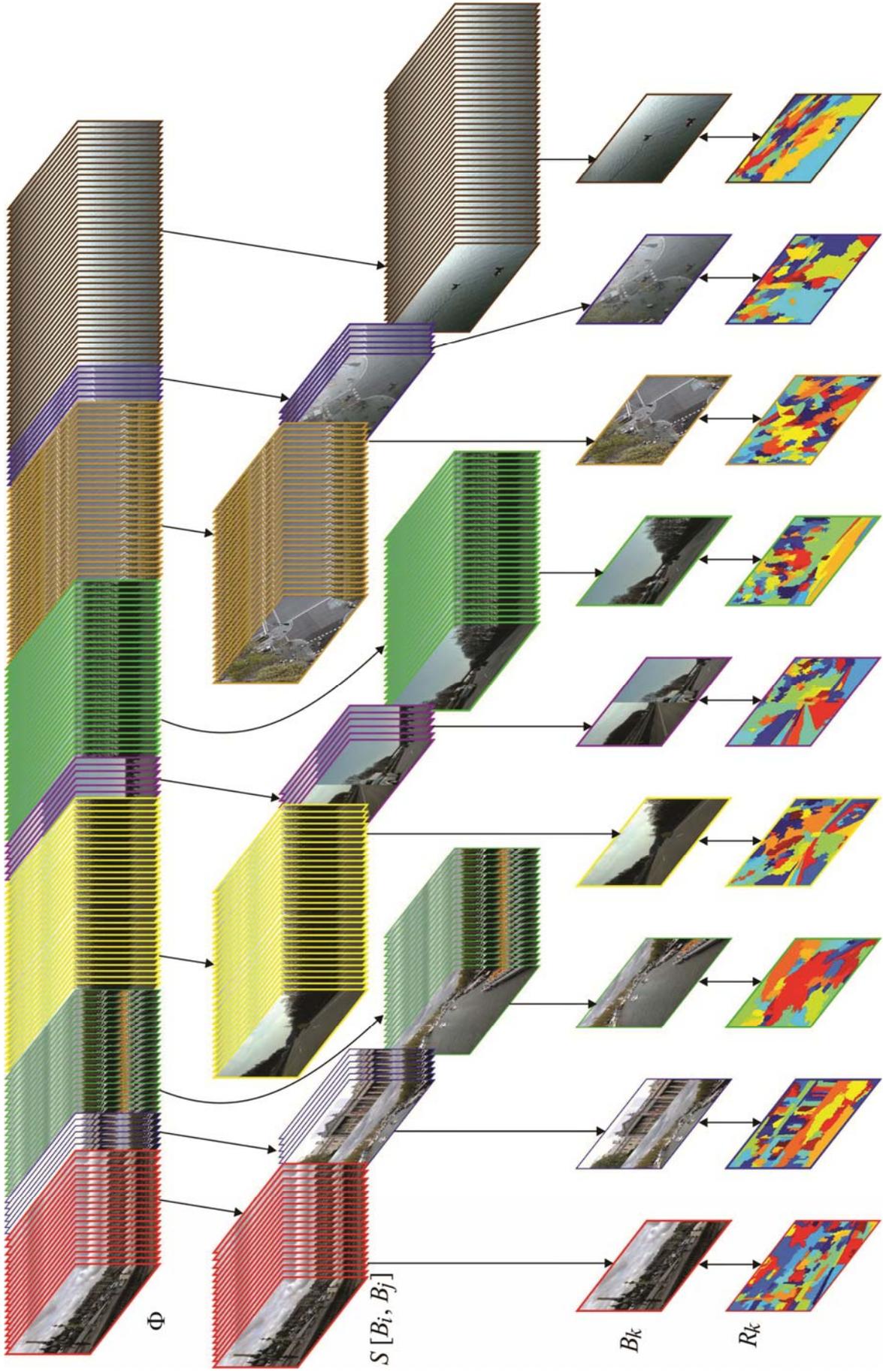


Рис. 1.6 – Структура видеопотока

сегменту осуществляется на основе (1.30) фактически путем анализа одномерного временного ряда. Таким образом, имеем $\Delta(\Pi_{ij}) = f(\rho(P_i, P_k), \pi_k), k = \overline{i+1, j}$. При этом, как правило, используются методы, использующие контрольный следящий сигнал (трекинг-сигнал), являющийся индикатором изменений контролируемых сигналов [35]. Существует ряд форм следящего сигнала, при этом его выход за определенные границы свидетельствует о возникших изменениях.

На практике наиболее широкое распространение получили эвристические процедуры такие, как методы Чоу, Робертса-Рида, Брауна, Тригга-Лича, Шоуна и т.п. [36, 37]. В условиях дефицита априорной и текущей информации о характеристиках и свойствах контролируемого сигнала, что является типичным для видеопоследовательностей, предпочтение следует, естественно, отдавать подобным методам [35, 36]. В [36] подчеркивается, что при идентификации процессов с достаточно плавным дрейфом, преимущества имеет метод Брауна. Резкие скачки лучше определяются с помощью следящего сигнала Тригга-Лича.

Цель экспериментальных исследований данного подраздела – изучение специфики сегментации видео для определения видеосегментов, на которых существует групповая связь между объектами и, соответственно, между фактор-множествами. Все эксперименты проводились с видеорядами в 1500 кадров, что соответствует длительности в одну минуту каждый при исходном разрешении 1920×1080 пикселей (медиаконтейнер m2ts, формат AVC-HD, частота 29,97 кадр/сек). Типичный пример видеоряда приведен на рис. 1.6: он содержит 9 сегментов, 3 из которых фактически являются плавными переходами от одного к другому.

Очевидно, что качество исходных изображений оказывают существенное влияние на индуцируемые фактор-множества. На рис. 1.7 приведены примеры наиболее часто используемых разрешений (приближенных к стандарту 16:9), которые получены из HDTV путем передискретизации: слева-

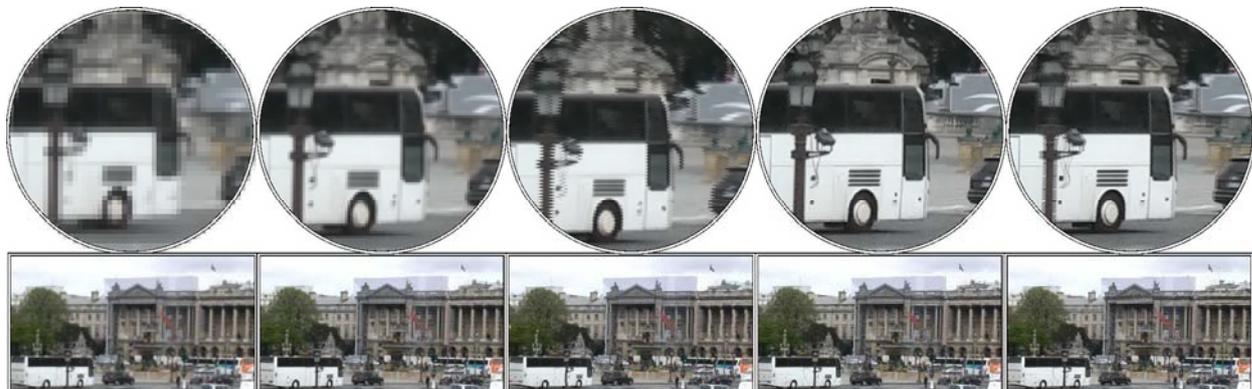


Рис. 1.7 – Пример изображений с различным разрешением

направо показаны изображения 246×136 , 480×272 , 720×412 , 1280×720 , 1920×1080 пикселей. Рис. 1.8 иллюстрирует изменение фактор-множеств (парно, последовательно во времени), полученных одним и тем же алгоритмом сегментации изображений. Видно, что можно найти разумный компромисс между вычислительной сложностью и адекватностью результатов, в частности, заложенные в видеоряд сегменты наиболее четко выделяются при разрешении 480×272 и 720×412 пикселей. Сегменты практически не выделяются при разрешении 246×136 пикселей и выделение ряда из них затруднительно (в случаях плавного изменения) при разрешении 1280×720 и 1920×1080 элементов, что одновременно связано и с существенным повышением объемов обрабатываемых данных.

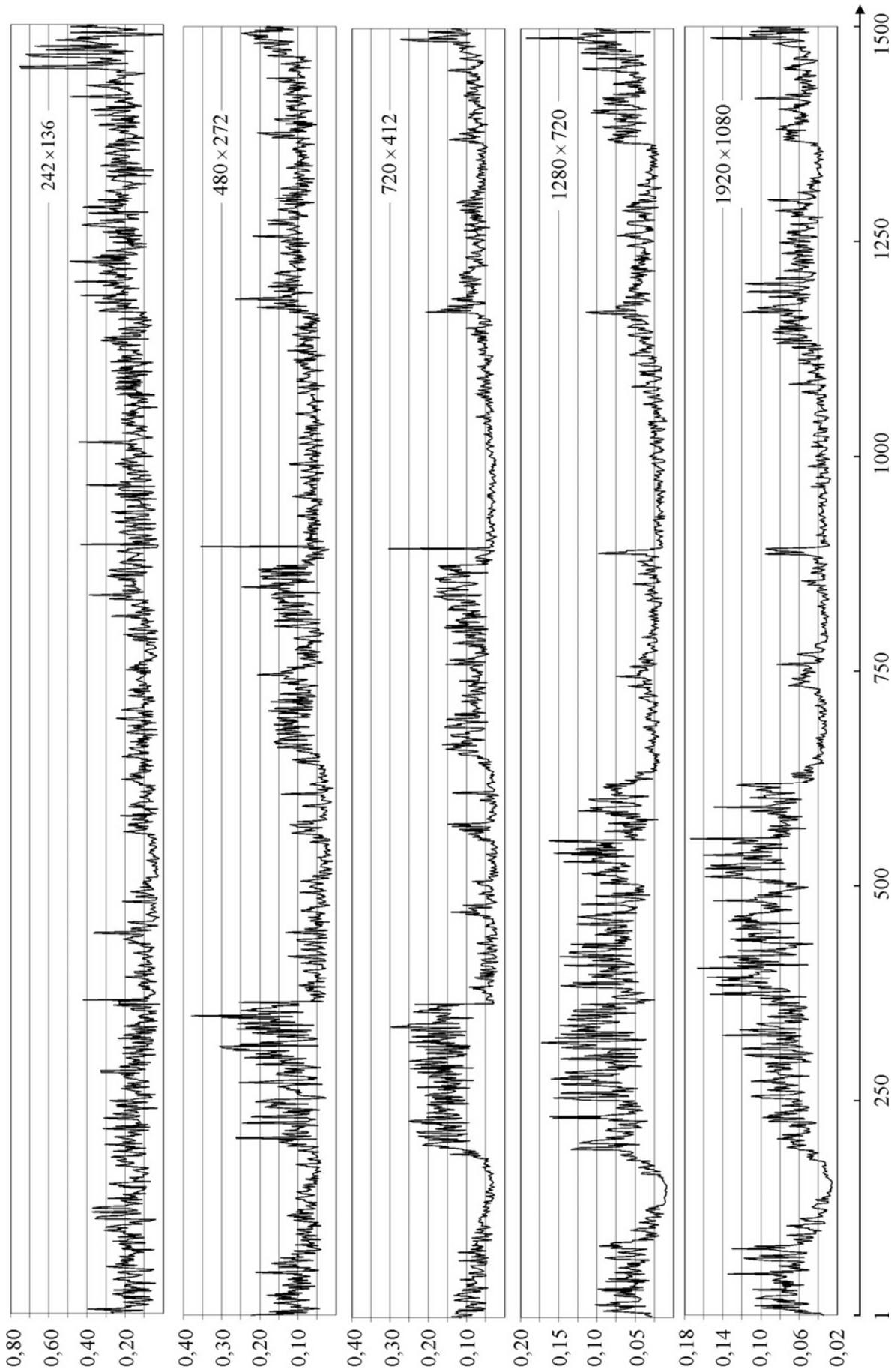


Рис. 1.8 – Изменение фактор-множеств при различном разрешении видеоряда, приведенного на рис. 4.3

РАЗДЕЛ 2
МЕТОДЫ АНАЛИЗА ИЗОБРАЖЕНИЙ
В УСЛОВИЯХ НЕПОЛНОЙ ИНФОРМАЦИИ

2.1. Метод нормализации следящего типа для перспективы и его применение при нормализации проективных преобразований

Прежде всего отметим, что сложность нормализации проективных преобразований в большей степени вызвана присутствием нелинейных искажений перспективы, которая и обуславливает условия возникновения неполной информации. При этом каждая точка изображения при преобразовании перспективы вида отображается как

$$B_0(x, y) = B\left(\frac{x}{\alpha x + \beta y + 1}, \frac{y}{\alpha x + \beta y + 1}\right), \quad (2.1)$$

где B и B_0 – исследуемые изображения,

x, y – координаты исходного изображения,

α, β – параметры преобразования.

Пусть $dP = \begin{vmatrix} \frac{\partial X}{\partial x} & \frac{\partial X}{\partial y} \\ \frac{\partial Y}{\partial x} & \frac{\partial Y}{\partial y} \end{vmatrix}$, где $X = \frac{x}{\alpha x + \beta y + 1}$, $Y = \frac{y}{\alpha x + \beta y + 1}$. Известно, что опре-

делитель $\det(dP)$ дает информацию о преобразовании точки, а $|\det(dP)|$ отображает коэффициент сжатия этой точки. Поскольку

$$\begin{aligned} \frac{\partial X}{\partial x} &= \frac{1 + \beta y}{(\alpha x + \beta y + 1)^2}; & \frac{\partial X}{\partial y} &= \frac{-\beta x}{(\alpha x + \beta y + 1)^2}; \\ \frac{\partial Y}{\partial x} &= \frac{-\alpha y}{(\alpha x + \beta y + 1)^2}; & \frac{\partial Y}{\partial y} &= \frac{1 + \alpha x}{(\alpha x + \beta y + 1)^2}, \end{aligned}$$

$$\det(dP) = \frac{1}{(\alpha x + \beta y + 1)^3}. \quad (2.2)$$

Тогда из выражения (2.2) следует, что если входное B и эталонное B_0 изображения отличаются только преобразованием перспективы вида (2.1), которое не выводит изображение за пределы поля зрения D , то интегрирование по полю зрения для $B(x, y)$ дает

$$\iint_D B(x, y) dx dy = \iint_D \frac{B_0(x, y) dx dy}{|(\alpha x + \beta y + 1)^3|}. \quad (2.3)$$

На основе разложения проективной группы

$$|\det \Pi| = |\det A| |\det P_\phi| \quad (2.4)$$

и с учетом того, что $|\det A|$ не зависит от x и y , а $|\det P_\phi|$ – зависит, можно определить критерий присутствия проективных преобразований.

Если $|\det P_\phi| = 1$, то преобразования перспективы отсутствуют.

Рассмотрим для входного и искаженного изображений отношение $\frac{E_B}{S}$, где функционал $E_B = \iint_D B(x, y) dx dy$ отражает энергию изображения $B(x, y)$,

а $S = \iint_D \text{sign}(B(x, y)) dx dy$ – площадь изображения $B(x, y)$, где $\text{sign}(B(x, y)) = 1$,

если точка $B(x, y)$ не относится к фону и $\text{sign}(B(x, y)) = 0$ в противном случае.

Аналогично определим энергию и площадь эталонного изображения $B_0(x, y)$

как $E_{B_0} = \iint_D B_0(x, y) dx dy$ и $S_0 = \iint_D \text{sign}(B_0(x, y)) dx dy$.

Пусть $(x, y)_{B_0}$ – координаты точки эталонного изображения, $(x, y)_B$ – координаты образа точки $(x, y)_{B_0}$. Тогда координаты точки $(x, y)_{B_0}$ вместе со своей бесконечно малой окрестностью связаны с координатами своего образа соотношением в рамках аффинной модели:

$$(x, y)_B = |\det A|(x, y)_{B_0}.$$

Проинтегрируем по всей области, в которой находится рассматриваемое изображение. В результате получим функциональную зависимость энергий входного и эталонного изображений:

$$\iint_D B(x, y) dx dy = \iint_D |\det A| B_0(x, y) dx dy.$$

Т.к. $|\det A|$ не зависит от x и y , то его можно вынести за интеграл, в результате получим

$$\iint_D B(x, y) dx dy = |\det A| \iint_D B_0(x, y) dx dy.$$

Подставив полученные результаты в отношение $\frac{E_B}{S}$, получим

$$\frac{E_B}{S} = \frac{|\det A| \iint_D B_0(x, y) dx dy}{|\det A| \iint_D \text{sign}(B_0(x, y)) dx dy} = \frac{\iint_D B_0(x, y) dx dy}{\iint_D \text{sign}(B_0(x, y)) dx dy}.$$

Видно, что для аффинных преобразований выполняется условие:

$$\frac{E_B}{S} = \frac{E_{B_0}}{S_0}. \quad (2.5)$$

Действуя по аналогии, исходя из (2.3), получим соответствующие выражения для преобразования перспективы:

$$\frac{E_B}{S} = \frac{\iint |\det P_\phi| B_0(x, y) dx dy}{\iint |\det P_\phi| \text{sign}(B_0(x, y)) dx dy},$$

$$\frac{E_{B_0}}{S_0} = \frac{\iint B_0(x, y) dx dy}{\iint \text{sign}(B_0(x, y)) dx dy},$$

откуда можно сделать вывод, что для преобразования перспективы:

$$\frac{E_B}{S} \neq \frac{E_{B_0}}{S_0}. \quad (2.6)$$

Тогда можно предложить критерий идентификации (присутствия) проективных преобразований. На основании того, что перспектива, в отличие от аффинных преобразований, вносит изменение в перераспределение энергии на изображении, построим функционал вида:

$$\Phi_1 = SE_{B_0} - S_0E_B, \quad (2.7)$$

где S и S_0 – площади изображений $B(x, y)$ и $B_0(x, y)$ соответственно.

Исходя из (2.3), для нормализации преобразований целесообразно будет ввести также функционал Φ_2 вида:

$$\Phi_2 = E_B - \iint_D \frac{B_0(x, y) dx dy}{|(\alpha x + \beta y + 1)^3|}. \quad (2.8)$$

Исходя из результатов исследования свойств аффинной и проективной группы, сформулируем критерий присутствия на изображении проективных искажений на основе значений функционалов Φ_1, Φ_2 .

Таким образом, выполнение условия

$$\Phi_1 \Phi_2 < \varepsilon, \quad (2.9)$$

где $\varepsilon > 0$ – некоторый порог.

Равенство нулю хотя бы одного из функционалов (2.7), (2.8) позволяет констатировать принадлежность искажений на изображении к аффинной группе, в противном случае – к проективной. В целях упрощения также допустимо использование одного из функционалов Φ_1 или Φ_2 .

На основании предложенного критерия (2.9) можно осуществить нормализацию проективных преобразований, которая заключается в многократном выполнении преобразования перспективы с некоторым шагом с вычислением функционала (2.9). Преобразования продолжаются до тех пор, пока значение функционала не станет равным нулю. Этот принцип построения соответствует нормализатору следящего типа [31, 38, 44].

Проективная группа представляется в виде композиции аффинного и перспективного преобразований, не влияющего на параметры аффинного. В свою очередь, перспективное преобразование представляется в виде композиции коммутативных преобразований однопараметрической перспективы вдоль осей OX и OY . Тогда можно записать

$$\Pi = AP_\phi(h) = AP_x(\alpha)P_y(\beta) = AP_y(\beta)P_x(\alpha).$$

Таким образом, можно найти параметры перспективы корреляционным методом по двум параметрам, после чего перейти к новому изображению, которое отличается от эталонного только аффинным преобразованием. Как резуль-

тат разработки приведем метод нормализации перспективных преобразований изображения с использованием критерия (2.9). Этапы метода:

1. Задаем первоначальные значения для α, β .
2. Вычисляем функционал вида (2.9).
3. Если критерий (2.9) не выполнен, увеличиваем один из параметров α, β на шаг $\Delta\alpha, \Delta\beta$ и переходим к п.2, иначе α, β являются искомыми значениями.
4. Нормализуем изображение.

Этот метод реализует следящий принцип нормализации [31, 44]. Построение параллельного нормализатора, где параметры определяются в соответствии с математическим выражением в явном виде, в данной ситуации затруднено вследствие отсутствия надежных и точных методов определения перспективы и нелинейности самих преобразований.

На практике условие равенства нулю вычисляемых функционалов из-за погрешности вычислений заменяют на условие вида $|\Phi_1 \Phi_2| < \varepsilon$, где $\varepsilon > 0$ – некоторый порог.

2.2. Нормализация проективно искаженных изображений с использованием инвариантных отображений

Развивая результаты предыдущего подраздела, можно предложить другие способы построения нормализаторов проективной группы на основе разложения в аффинную и перспективу.

Представим эталон $B_0(x, y)$ в виде набора (вектора) инвариантов $\chi = [\chi_1, \dots, \chi_v]$, вычисленных по всему полю зрения, где v – число вычисленных инвариантов, $\chi_k \in \mathbb{R}^1$, $k = \overline{1, v}$ [31, 39]. Несомненным достоинством инвариантов интегрального типа есть устойчивость к действию флуктуационного шума. Таким путем достигаем независимости от аффинных преобразований, и сводим задачу исключительно к нормализации перспективы.

Метод нормализации на основе инвариантов χ можно описать следующим образом:

1. Вычисляем инвариантные признаки χ^{et} для эталона и χ – для входного изображения.

2. Проверяем совпадение (эквивалентность) χ и χ^{et} , например, путем вычисления их отклонения в некоторой метрике ρ для числовых векторов. Сохранение аффинной инвариантности, которое проверяется в виде выполнения условия $\rho(\chi, \chi^{et}) \leq \varepsilon$, подтверждает, что изображение искажено только аффинной группой. Противная ситуация, т.е. при невыполнении условия $\rho(\chi, \chi^{et}) \leq \varepsilon$, является сигналом присутствия перспективных искажений, и их необходимо нормализовать.

3. Нормализуем перспективные искажения путем применения разработанного корреляционного метода нормализации следящего типа для перспективы с параметрами α, β и шагами дискретизации $\Delta\alpha, \Delta\beta$, добиваясь достижения близости к нулю соответствующего функционала. Строим инвариантные к аффинным преобразованиям признаки для измененного изображения.

Пример работы предложенного метода приведен на рис. 2.1. Исходное изображение искажалось масштабированием $\lambda = 1.5$, $\mu = 1$ и поворотом $\theta = 34^\circ$. Далее были применены перспективные искажения вдоль горизонтальной оси с параметрами $\alpha = 0.03$, $\beta = 0$. Допустимая погрешность отклонения признаков изображения от эталона составила 20%.

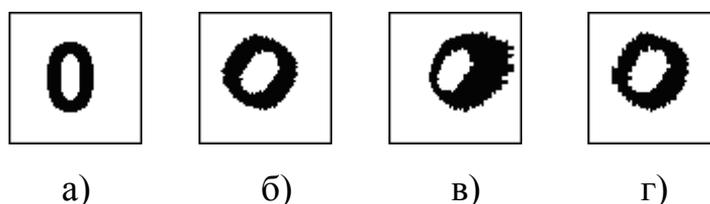


Рис. 2.1 – Нормализация перспективы: а – эталонное изображение; б – изображение, искаженное аффинными преобразованиями; в – изображение, искаженное проективными преобразованиями; г – изображение после нормализации искажений перспективы

На рис. 2.1а показано эталонное изображение, рис. 2.1б, в представляет результаты аффинных и перспективных искажений. На рис. 2.1г показан результат сведения проективно-искаженного изображения к аффинно искаженному.

Заметим, что точность определения инвариантных признаков зависит от сложности изображения (количества информации). На рис. 2.2 приведено полутоновое изображение более сложного плана, его искаженный вид (параметры $\lambda = 1.5$, $\mu = 1.5$, $m = 3$, $n = -4$). Как видим, значения признаков достаточно близки, различия же объясняются искажениями дискретизации, особенно по краям.

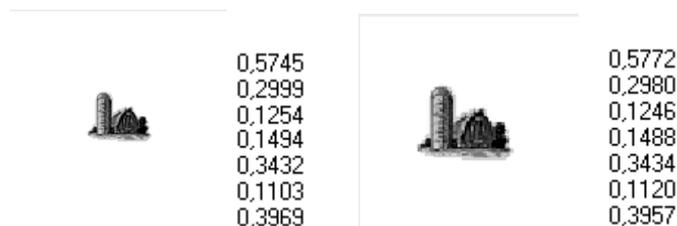


Рис. 2.2 – Пример значений инвариантных признаков

Другим направлением достижения инвариантности к аффинным преобразованиям есть использование геометрических (пространственных) инвариантов, построенных на вычислении инвариантных соотношений для совокупностей некоторых окрестностей точек изображения [40]. Для получения геометрических инвариантов необходимо применение детекторов характерных точек изображения.

Рассмотрим инварианты аффинной группы в виде коэффициентов ξ, η как отношение четырех точек в аффинном базисе, т.е. [41, 42].

$$\begin{cases} x_4 = \xi(x_3 - x_1) + \eta(x_2 - x_1) + x_1, \\ y_4 = \xi(y_3 - y_1) + \eta(y_2 - y_1) + y_1, \end{cases} \quad (2.10)$$

где $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ – произвольные четыре точки изображения (точки базиса $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ должны быть неколлинеарными).

После применения аффинного преобразования T имеем

$$\begin{cases} Tx_4 = \xi(Tx_3 - Tx_1) + \eta(Tx_2 - Tx_1) + Tx_1, \\ Ty_4 = \xi(Ty_3 - Ty_1) + \eta(Ty_2 - Ty_1) + Ty_1, \end{cases} \quad (2.11)$$

причем значения ξ, η остаются неизменными.

Решение системы (2.11) в явном виде:

$$\begin{cases} \xi = \frac{(x_2y_4 - y_1x_2 - y_4x_1 - x_4y_2 + y_1x_4 + y_2x_1)}{(y_2x_1 - y_2x_3 + y_1x_3 - y_1x_2 + y_3x_2 - y_3x_1)}, \\ \eta = \frac{(x_1y_4 - y_4x_3 - y_1x_4 + y_3x_4 + y_1x_3 - y_3x_1)}{(y_2x_1 - y_2x_3 + y_1x_3 - y_1x_2 + y_3x_2 - y_3x_1)}. \end{cases}$$

Недостатком построения и применения инвариантных соотношений (2.10, 2.11) является необходимость поиска групповых соответствий точек исходного и эталонного изображения, что является трудоемкой процедурой с точки зрения количества вычислительных операций [40, 43].

Используя инвариантные признаки характерных точек изображения в виде множества Ω векторов $\omega = (\xi, \eta), \omega \in \Omega$ на основе вычисления (2.10, 2.11) можно построить наборы инвариантов путем анализа подмножеств точек структурного описания. Примером могут быть функционалы, отражающие интегральные свойства описания в виде вектора $\omega_\Sigma = (\sum \xi, \sum \eta)$, компонентами которого есть суммы инвариантов. Другими способами построения инвариантных функционалов есть вычисление, например, максимума или минимума от некоторой функции, заданной на множестве значений $\{\xi, \eta\}$.

Метод нормализации на основе структурных инвариантных описаний $\{\xi, \eta\}$ опишем так:

1. Вычисляем инвариантные описания эталона $\{\omega^{et}\}$ и входного изображения $\{\omega\}$.

2. Проверяем совпадение (эквивалентность) множеств $\{\omega\}$ и $\{\omega^{et}\}$, например, путем поиска соответствующих значений и подсчета их количества. Сохранение аффинной инвариантности, которое проверяется в виде выполнения условия $\rho(\{\omega\}, \{\omega^{et}\}) \leq \varepsilon$ для некоторой метрики или меры подобия на множествах, подтверждает, что изображение искажено только аффинной группой. Обратная ситуация дает возможность предположить присутствие перспективных искажений, и их необходимо нормализовать.

3. Нормализуем перспективные искажения путем применения разработанного корреляционного метода нормализации следящего типа для перспективы с параметрами α, β и шагами дискретизации $\Delta\alpha, \Delta\beta$, добиваясь максимизации количества соответствующих описаний. Строим инвариантные к аффинным преобразованиям описания для измененного изображения.

2.3. Нормализация на базе отношения четырех точек

Одним из существующих методов построения функционалов для нормализации в условиях действия проективной группы является определение соответствующих образов четырех точек на эталонном и входном изображениях. При этом три из этих точек не должны лежать на одной прямой (условие невырожденности), поскольку сложное отношение четырех точек является главным инвариантом в проективной геометрии.

Основным недостатком такого подхода является точность, поскольку наличие координатной погрешности при сопоставлении точек может оказаться решающим фактором для оценивания параметров нормализации. В практических ситуациях часто возникает необходимость определения соответствующих точек человеком-оператором в интерактивном режиме. Чаще всего этот метод применяется для анализа прямоугольных изображений, в которых в качестве

четырёх точек выступают соответствующие углы изображения, нахождение которых можно автоматизировать.

Поиск параметров нормализации по четырем точкам формально можно осуществить, опираясь на связь между точками в эталонном и преобразованном изображениях (рис. 2.3).

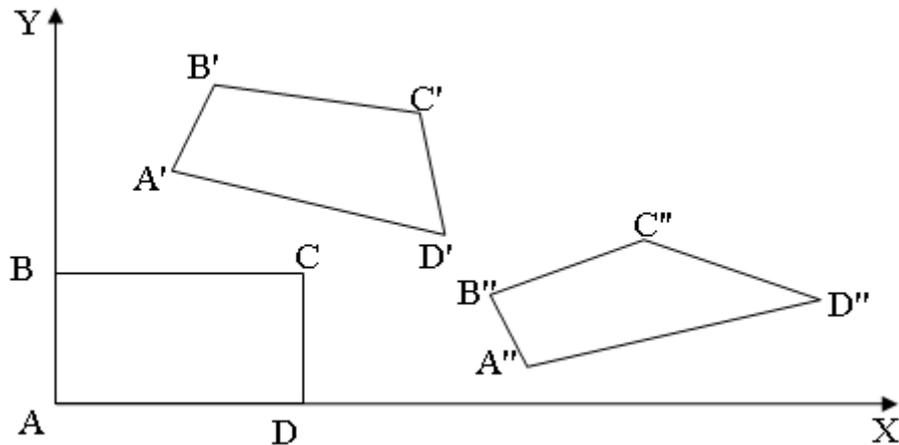


Рис. 2.3 – Проективные искажения: ABCD – исходная фигура, A'B'C'D' и A''B''C''D'' – преобразованные фигуры

Параметры проективного преобразования в этом случае для точек фигур ABCD и A'B'C'D' определяются по формулам:

$$\left\{ \begin{array}{l} p_{13} = x_{A'}; \\ p_{31} = \frac{d_1(y_{C'} - y_{B'}) - d_2(x_{C'} - x_{B'})}{AD(d_3 - d_4)}; \\ p_{11} = \frac{x_{D'}(p_{31}AB + 1) - p_{13}}{AB}; \\ p_{21} = \frac{y_{D'}(p_{31}AB + 1) - p_{23}}{AB}; \end{array} \right. \quad \left\{ \begin{array}{l} p_{23} = y_{A'}; \\ p_{32} = \frac{d_1(y_{C'} - y_{D'}) - d_2(x_{C'} - x_{D'})}{AB(d_4 - d_3)}; \\ p_{12} = \frac{x_{B'}(p_{32}AD + 1) - p_{13}}{AD}; \\ p_{22} = \frac{y_{D'}(p_{23}AD + 1) - p_{23}}{AD}, \end{array} \right. \quad (2.12)$$

$$\text{где } d_1 = x_{D'} - x_{C'} + x_{B'} - p_{13};$$

$$d_2 = y_{D'} - y_{C'} + y_{B'} - p_{23};$$

$$d_3 = (x_{C'} - x_{D'})(y_{C'} - y_{B'});$$

$$d_4 = (y_{C'} - y_{D'})(x_{C'} - x_{B'});$$

AB, AD – длины отрезков.

Тогда исходное изображение преобразуется к виду множества функций

$$L(x, y, \sigma) = B(x, y) \otimes G(x, y, \sigma), \quad (2.13)$$

где $B(x, y)$ – функция изображения;

$G(x, y, \sigma)$ – гауссиан, с параметром σ ;

\otimes – символ свертки.

Величина σ изменяется по правилам геометрической прогрессии в пределах $[\sigma_0, \sigma_0 \times 2^{\log_2(\min(m,n))^{-3}}]$, где m, n – высота и ширина исходного изображения $B(x, y)$. Знаменатель прогрессии определяется выражением $2^{\frac{1}{S}}$, где S – произвольная константа, подбираемая экспериментальным путем.

Далее на базе системы функций L формируется множество вида

$$D(x, y, \sigma_i) = L(x, y, \sigma_{i+1}) - L(x, y, \sigma_i). \quad (2.14)$$

Выражение (2.14) представляет собой аппроксимацию оператора Лапласа Δ^2 . Локальные экстремумы (2.14) формируют множество характерных точек.

Для определения более точного положения каждой точки в пространстве изображения осуществляют разложение функции $D(x, y, \sigma)$ в ряд Тейлора в характерных точках:

$$D(c) = D + \frac{\partial D}{\partial c} c + \frac{1}{2} c^T \frac{\partial^2 D}{\partial c^2} c, \quad (2.15)$$

где $c = (x, y)$ – смещение от выбранной характерной точки z^* в пределах фиксированного масштаба σ . Местоположение \hat{c} экстремума функции $D(c)$ по

смещению определяется путем приравнивания нулю производной от $D(c)$ относительно c .

Поскольку специфика данной задачи подразумевает сопоставление точек в целях нормализации (например, для двух четырехугольных изображений на однородном фоне), краевые точки могут нести особый интерес и этот этап может быть модернизирован в обратную сторону для поиска именно краевых точек.

Формирование характерных точек осуществляется путем исследования соотношения между главными значениями кривизны в направлениях x, y . Определение параметров кривизны сводится к нахождению собственных значений

матрицы гессиана $H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$ второго порядка, где D_{uv} – вторая производная функции D по переменным u, v .

Собственные значения α_1, α_2 матрицы H , вычисленной для координат характерных точек, пропорциональны величинам основных искривлений (пространственной кривизне) функции D при фиксированном σ . Точки, для которых одно из направлений кривизны значительно превышает другое, отбрасываются. Практическая реализация этого этапа может быть выполнена с помощью решения системы уравнений: $\alpha_1 \alpha_2 = \text{Det}(H)$, $\alpha_1 + \alpha_2 = \text{Tr}(H)$, где Tr – след, а Det – определитель матрицы. Оценочная функция имеет вид $R = \text{det}(H) - k \text{Tr}(H)^2$, где k – эмпирическая константа, $k = 0.04 - 0.06$. Краевыми точками считаем те из них, для которых $R < 0$.

Каждой характерной точке назначается одна или несколько ориентаций на основе локальных значений градиента изображения. Для достижения инвариантности к вращениям относительно этой ориентации в дальнейшем строятся дескрипторы точек. Для изображения $L(x, y, \sigma)$ при фиксированном значении σ вычисляются амплитуда $m(x, y)$ и ориентация $\theta(x, y)$ градиента для характерной точки с координатами (x, y)

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2},$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}.$$

Признаковая информация содержится в массиве направлений $T = \{t_i\}$, $t_i = (x, y, m, \theta, \sigma)$, вычисленном для множества характерных точек. В окрестности 16×16 каждой точки при фиксированном масштабе анализируется множество из 16 непересекающихся фрагментов 4×4 , образующих разбиение. Для каждого из фрагментов, в свою очередь, строится 8-мистолбцовая гистограмма направлений по переменной θ , охватывающая 360 градусов. Одно направление соответствует диапазону в 40 градусов. Значениями дескриптора являются суммы амплитуд $v_i = \sum_{q=i} m_q$. В результате для каждой характерной точки имеем вектор v (дескриптор), содержащий $16 \times 8 = 128$ значений направлений.

Для обеспечения инвариантности к поворотам направления нормализуются путем сдвига относительно направления характерной точки, а для обеспечения устойчивости к изменению освещения гистограммы нормализуются путем делению компонент на норму вектора.

Таким образом, на выходе SIFT образуется набор дескрипторов $v = \{v_i\}$ из 128 значений для каждой из значимых (прошедших отбор) точек. На рис. 2.4 приведен пример отфильтрованных характерных точек, отражающих местоположение углов на изображении.

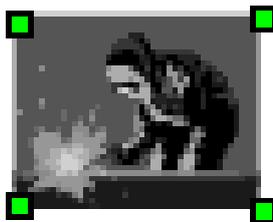


Рис. 2.4 – Характерные угловые точки изображения

Проблема выбора соответствующих точек может быть решена за счет по-

иска близких в некотором смысле дескрипторов точек.

На рис. 2.5 показана ситуация многообразия нескольких характерных точек изображения, большинство из которых локализованы во внутренней области объектов. В этой ситуации нормализация может быть затруднена вследствие нахождения ложных соответствий, которые могут возникнуть из-за нахождения лишних ХТ внутри однородных областей.



Рис. 2.5 – Характерные точки изображения как результат работы классического метода SIFT ($\delta = 0.03$)

Более эффективной выглядит поиск соответствий точек на границах изображения, которые могут быть выделены предложенной модификацией при анализе параметров кривизны ХТ. Рис. 2.6 демонстрирует обратный пример анализа кривизны, при котором остаются только краевые точки изображения.



Рис. 2.6 – Характерные угловые точки изображения как результат работы модифицированного метода SIFT ($\delta = 0.03$)

Необходимо отметить, что выражения (2.12) являются весьма идеализированным случаем, поскольку цена погрешности при нормализации проективных искажений чрезвычайно высока. Например, на рис. 2.7 показана ситуация,

при которой соответствующие точки указаны пользователем вручную. Найденные согласно (2.12) значения искажений составили $\alpha = -0.0255$, $\beta = 0.01107$, нормализованное изображение также приведено на рис. 2.7. Заданные параметры искажений при этом составили $\alpha = -0.02$, $\beta = 0.01$.

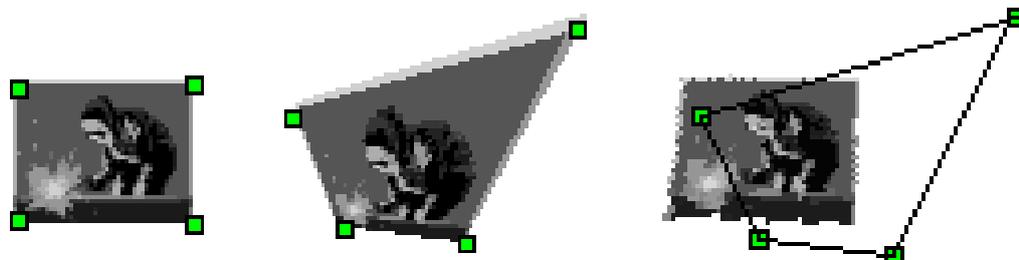


Рис. 2.7 – Пример нормализации по четырем точкам

2.4. Нормализация и классификация путем установления соответствий систем фрагментов на основе голосования как основы компонентного описания мер сопоставления изображений

Рассмотрим пример применения корреляционного подхода к оценке одного из наиболее простых геометрических преобразований – смещения на основе сопоставления отдельных фрагментов изображений, представляющих структурное описание. Будем анализировать оценивание (нормализацию) смещения применительно к задаче классификации с заранее заданным числом эталонов. Корреляционные методы приобрели популярность в системах компьютерного зрения из-за высокой надежности, хорошей работоспособности в широком диапазоне внешних условий, а также вследствие высокой помехозащищенности относительно аддитивного шума. Корреляционными считают все подходы, так или иначе основанные на построении меры сходства $\psi(B, B_0)$ анализируемого изображения $B(x, y) \in W$, $(x, y) \in D$ и эталона $B_0(x, y) \in W_0$, $W_0 \subset W$ (W – множество изображений, W_0 – множество эталонов, и оптимизации ее значения на множестве W_0 и/или на множестве преобразований g из группы G).

Разработка модификаций корреляционного подхода привела к появлению методов, обладающих наряду с устойчивостью к аддитивным шумам также достаточно хорошей помехозащищенностью и к действию помех локального типа. Этот путь связан с анализом фрагментов D_i изображения, когда область определения D представляется в виде $D = \cup D_i$. В методе частных корреляций при определении координат объектов предполагается построение иерархической меры, значение которой вычисляется в два этапа: сначала для соответствующих фрагментов, а затем на основе этого – результирующее сходство.

Эффективным путем установления соответствия между множествами фрагментов, аппроксимирующими распознаваемый объект и эталоны, является голосование. Методы голосования получили распространение при распознавании визуальных объектов путем анализа множества локальных признаков [40]. Представляется целесообразным осуществить построение корреляционных методов на основе голосования фрагментов или их признаков, обеспечивая тем самым устойчивость и гибкость принятия решения относительно локальных помех и фона.

Фрагментация может быть выполнена несколькими способами, имеющими свои преимущества в определенных ситуациях. В методе частных корреляций в целях использования оптимальных соответствий фрагменты строятся путем включения всех точек объекта (рис. 2.8а), в то же время в других ситуациях в интерактивном режиме можно построить более простую и часто более эффективную для целей распознавания систему информативных фрагментов (рис. 2.8б).

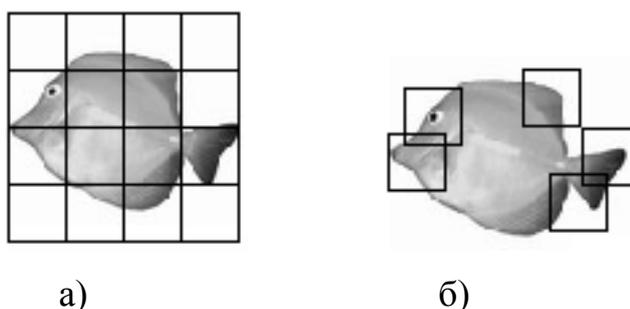


Рис. 2.8 – Примеры систем фрагментов: а – фиксированная; б – произвольная

Особенностью применения голосования в корреляционных методах по сравнению с признаковыми подходами есть то, что пространственное соответствие фрагментов анализируемого изображения и эталона строго фиксируется, что в целом позволяет существенно упростить обработку.

Пусть эталонные изображения представлены множествами информативных фрагментов $B_0^j = \{b_i^j\}, i = \overline{1, s_j}$, $W_0 = \{B_0^j\}, j \in J$, где J – множество классов (эталонов), s_j – количество фрагментов j -го эталона.

На пространственное размещение фрагментов эталонов в общем случае не будем накладывать особых ограничений. Фрагменты могут выбираться либо с помощью автоматической процедуры путем построения и оптимизации откликов локальных фильтров, либо случайным образом, либо в интерактивном режиме.

В сравнительном аспекте можно проанализировать несколько разновидностей систем фрагментов, каждая из которых имеет особенности при реализации (рис. 2.8). Поэтому формализацию классификации путем голосования фрагментов осуществим следующим образом.

Реализуем построение на анализируемом изображении B системы фрагментов, аналогичных эталону с номером j . Затем проведем оптимизацию на множестве эталонов значения некоторой функции $\psi(\cdot)$, определяющей количество или долю голосов, отданных j -й класс. Значение функции $\psi(\cdot)$ отображает степень соответствия двух изображений, представленных множествами фрагментов. Аргументами функции $\psi(\cdot)$ являются: тип системы фрагментов, параметр g геометрических преобразований, номер j эталона. Решение принимается в соответствии с максимальным значением $\psi(\cdot)$. При фиксированной системе фрагментов $\psi(\cdot)$ есть функцией двух переменных: j и g . Класс изображения можно определить как

$$j^* = \arg \max_{j \in J} \max_{g \in G} \psi(B(j), B_0^j g), \quad (2.16)$$

где $B(j)$ – система фрагментов, построенная на изображении B в соответствии с фрагментным представлением эталона из класса j ;

$B_0^j g$ – аналогичная система, полученная на преобразованном под действием геометрического преобразования эталоне B_0^j .

В результате решения (2.16) наряду с оценкой j^* параметра класса получаем также и значение оценки g^* для параметра преобразования g . Учитывая, что в корреляционных подходах решение задач оптимизации вида (2.16) осуществляется, как правило, путем полного перебора возможных значений параметров, порядок поиска максимума принципиального значения не имеет. Решение будет аналогичным, если в (2.16) вначале найти максимум по переменной j , а затем – по переменной g . Для конкретности ограничимся рассмотрением случая, когда G есть группа двумерных смещений с множеством значений в пределах поля зрения.

Введем понятие функции $\xi(i, j)$ соответствия для пары фрагментов с номерами i и j . В общем случае сравниваемые фрагменты могут иметь разный размер и типы, но в целях упрощения будем рассматривать фрагменты одинакового вида. В корреляционных методах функция $\xi(i, j)$ обычно вычисляется как сходство-различие фрагментов в соответствии с некоторой мерой. В качестве примера $\xi(i, j)$ можно использовать одну из метрик $\rho(i, j)$ для векторных пространств, например, простую в вычислительном плане метрику суммы модуля разностей

$$\xi(i, j) = -\rho(i, j) = -\sum_{k=1}^K |b_k(i) - b_k^j|, \quad (2.17)$$

где k – номер пикселя изображения внутри фрагмента;

K – количество точек фрагмента;
 $\rho(i, j)$ – расстояние для пары фрагментов (i, j) , а знак минус используется для того, чтобы величина $\xi(i, j)$ росла с увеличением сходства. Другой вариант достижения этого свойства – построение функции соответствия вида $\xi(i, j) = \exp[-q\rho(i, j)]$, где q – константа.

Диапазон изменения значений функции $\xi(i, j)$ в виде (2.17) можно считать известным в связи с фиксированным диапазоном яркости изображения. При необходимости этот диапазон можно привести к интервалу $[0, 1]$.

Рассмотрим теперь возможные подходы к построению меры сходства ψ путем голосования на основе функции ξ .

Совместное голосование. Значение функции $\psi(\cdot)$ определим в виде

$$\psi(B(j), B_0^j g) = \frac{1}{s_j} \sum_{i=1}^{s_j} \eta_g(i, j), \quad (2.18)$$

где предикат $\eta_g(i, j)$ соответствия i -го фрагмента изображения эталону j -го класса определяется в виде бинарной модели

$$\eta_g(i, j) = \begin{cases} 1, & \xi_g(i, j) \geq \varepsilon, \\ 0, & \xi_g(i, j) < \varepsilon, \end{cases} \quad (2.19)$$

где $\xi_g(i, j)$ – функция соответствия эталону j с учетом параметра преобразования g ;

ε – значение, которое задает порог значимости величины $\xi_g(\cdot)$.

Нормировка в (2.18) обеспечивает независимость значения голосующей функции от количества фрагментов эталона, если величины s_j различны. В соответствии с выражением (2.18) значения функции $\psi(\cdot) \in [0, 1]$.

Следует отметить, что значения, полученные в соответствии с соотношениями (2.18), (2.19), и далее используемые при определении максимума, из-за отсечения важной информации могут оказаться незначительными по величине и в результате привести к ложному определению класса. Дело в том, что вычислением в (2.18), (2.19) не ограничивается размер коалиции фрагментов, на основе которой принимается решение. В результате решение может быть принято даже по одному фрагменту, что недопустимо с практической точки зрения.

Чтобы избежать этих трудностей, на величину максимума в (2.16) должно быть наложено дополнительное ограничение вида

$$\max_{j \in J} \max_{g \in G} \psi(B(j), B_0^j g) \geq \varepsilon_m, \quad (2.20)$$

т.е. суммарное оптимальное сходство фрагментов приобретает значимость только тогда, когда оно больше некоторой величины доверия ε_m . Из неравенства (2.20) следует, что в процессе вычисления максимума необходимо дополнительно проверять условие $\psi(B(j), B_0^j g) \geq \varepsilon_m$.

Учитывая, что значения $\psi(\cdot)$ принадлежат отрезку $[0, 1]$, значение порога ε_m со статистической точки зрения при равноценных фрагментах должно быть выбрано в виде $\varepsilon_m = 0,5$, хотя возможны ситуации, когда степень доверия задается условием $\varepsilon_m < 0,5$. В таком случае решение может быть принято по коалиции наиболее важных фрагментов.

Независимое голосование. Рассмотрим теперь систему из s фрагментов, представляющую собой объединение множеств информативных фрагментов всех классов, $s = \sum_{j=1}^J s_j$ – количество фрагментов системы. Построим идентичную систему из s фрагментов для изображения и эталонов. Частным случаем есть система одинаковых в пространственном смысле фрагментов для всех эталонов,

т.е. когда $s_k = s_j$, $k \neq j$, а координаты и размеры фрагментов разных эталонов совпадают. В таком случае выполнено $\forall k s = s_k$.

Вначале для каждого из фрагментов полученной системы определим номер класса, для которого достигается максимальное значение величины $\xi_g(i, j)$, т.е. вычислим

$$\varphi_g [i] = \arg \max_{j \in J} \xi_g(i, j). \quad (2.21)$$

Значение $\varphi_g [i]$ зависит от параметра g . Затем на полученном множестве значений $\varphi_g [i]$ сформируем распределение (гистограмму) $p_g[j]$ голосов, отданных элементами множества $\{\varphi_g [i]\}$ за каждый из классов

$$p_g[j] = \#\{i, \varphi_g [i] = j\}, \quad (2.22)$$

где символ $\#$ обозначает мощность (количество элементов) множества.

Номер класса определим в соответствии с максимумом по параметру g на множестве гистограмм

$$j^* = \arg \max_{g \in G} \{p_g[j]\}. \quad (2.23)$$

Ограничение вида (5) для данного типа голосования имеет вид

$$\max_{g \in G} \{p_g[j]\} \geq \varepsilon_p, \quad (2.24)$$

где ε_p – порог для величины максимума, выражающийся в значимом количестве голосов.

Последовательность действий (2.21) – (2.23) фактически задает серию возможных построений конкретных алгоритмов.

Принципиальное отличие независимого и совместного видов голосования состоит в том, что при независимом голосовании каждый фрагмент самостоятельно формирует голос, а затем по множеству этих голосов принимается решение. При совместном голосовании голос за класс каждый раз формируется по результатам анализа всех фрагментов системы.

Одним из вариантов является множественное голосование одновременно за несколько классов, что в целом приводит к сглаживанию итоговой гистограммы и повышению устойчивости решения при действии помех. При этом появляется необходимость использования дополнительных параметров порогов для значений функции $\xi(i,j)$, связанных с ограничением количества одновременно голосующих элементов.

Заметим, что построение независимого голосования в случае различающихся для эталонов систем фрагментов основано на несколько большем общем количестве фрагментов, чем совместное голосование, и, соответственно, имеет большую информативность, что должно обеспечить и большую надежность решений.

Голосование с поддержкой. Общеизвестным фактом можно считать то, что проведение голосования с учетом голосов соседних фрагментов или ближайших по значению характерных признаков повышает надежность отдельного локального решения за счет увеличения объема используемой информации [40]. Здесь наряду с признаками близлежащей окрестности можно использовать структурные отношения. В результате гистограмма голосов приобретает ярко выраженный максимум, а процедура распознавания становится более устойчивой относительно действия фоновых помех.

При соблюдении корреляционных принципов обработки реализация поддержки сводится к анализу множества фрагментов, в пространственном плане соседних с анализируемыми. Схематично элементы окрестности при восьми-связности можно представить в виде рис. 2.9, где анализируемый фрагмент

обозначен ХП (характерный признак), а соседние с ним фрагменты окрестности для конкретности изложения пронумерованы числами от 1 до 8.

1	2	3
8	ХП	4
7	6	5

Рис. 2.9 – Расположение фрагментов поддержки

В большинстве применений размер фрагмента равен 5×5 . Таким образом, путем расширения области построения и анализа признаков общий размер поля изображения, участвующего в принятии локального решения, возрастает до 15×15 . Схематично представленное на рис. 2.9. расположение фрагментов-соседей лишь иллюстрирует схему анализа. В принципе, допускаются произвольные конфигурации построения и расположения фрагментов, например, в виде прямоугольников или квадратов, пересекающихся между собой или полученных путем сканирования с некоторым шагом.

Заметим, что введенная дополнительная обработка окрестности в любом случае увеличивает итоговое время анализа, поэтому используемые процедуры должны быть по возможности простыми.

Таким образом, для окрестности анализируемого фрагмента с номером i имеем n дополнительных фрагментов в виде совокупности $U^i = (b_1, \dots, b_n)^i$ (на рис. 2.9 $n = 8$). Множество фрагментов из окрестности фрагмента j -го эталона, относительно которого принимается локальное решение, обозначим $U^j = \{b_v^0\}^j, v = \overline{1, n}$. Эквивалентность окрестностей U^i, U^j может быть установлена на основе предиката $\eta_U[U^i, U^j]$, подтверждающего или опровергающего локальное решение в пользу класса с номером j ,

$$\eta_U[U^i, U^j] = \begin{cases} 1, & \xi_U(U^i, U^j) \geq \varepsilon_U, \\ 0, & \xi_U(U^i, U^j) < \varepsilon_U, \end{cases} \quad (2.25)$$

где функция $\xi_U(\cdot)$ характеризует величину сходства (соответствия) двух систем фрагментов U^i, U^j ;

ε_U – порог значимости сходства.

В общем случае функция $\xi_U[U^i, U^j]$ может быть построена на основе иной меры, чем функция соответствия фрагментов $\xi(i, j)$ из выражения (2.17). С другой стороны, с целью унификации можно использовать и одинаковые соотношения, например,

$$\xi_U(U^i, U^j) = - \sum_{l=1}^8 \sum_{k=1}^K |b_{kl}(i) - b_{kl}^j|, \quad (2.26)$$

где b_{kl} – пиксель фрагмента поддержки с номером l .

2.5. Построение и анализ свойств функциональных признаков

В частном случае соотношения для анализа искажений перспективы один из параметров α, β может принимать значение 0, т.е. искажение осуществляется вдоль одной из осей (абсцисс или ординат) и описывается как [41, 42]:

$$\begin{cases} x' = \frac{x}{\alpha x + 1}, \\ y' = \frac{y}{\alpha x + 1}; \end{cases} \quad \begin{cases} x' = \frac{x}{\beta y + 1}, \\ y' = \frac{y}{\beta y + 1}, \end{cases} \quad (2.27)$$

где x, y – координаты исходного изображения;

x', y' – координаты преобразованного изображения.

Определим область интереса $C = \{(x, y) | (x, y) \in I_B(x, y)\}$ как носитель объекта в виде множества точек, принадлежащих анализируемому объекту с функцией яркости $I_B(x, y)$. Для бинарных изображений, полученных в результате сегментации, область интереса содержит непосредственно точки объекта. Осуществим дискретное интегрирование по области интереса C и рассмотрим выражение, структурно соответствующее искажениям перспективы, и включающее функционалы, вычисленные путем интегрирования.

Формальный вид предлагаемых признаков o_1, o_2 запишем следующим образом:

$$\begin{cases} o_1 = \frac{S_x}{S_x + S_y}, \\ o_2 = \frac{S_y}{S_x + S_y}, \end{cases} \quad (2.28)$$

$$\text{где } S_x = \sum_{(x, y) \in I_B(x, y)} x;$$

$$S_y = \sum_{(x, y) \in I_B(x, y)} y, \quad (x, y) \in I_B(x, y).$$

Необходимо отметить, что (2.28) может быть реализовано как для системы координат с центром в физическом центре области $V(x, y)$, так и в техническом классическом представлении с центром в верхнем левом углу. В первом случае при полной симметрии объекта по обеим осям возможно получение нуля в знаменателе, однако, множество таких изображений (например, круг) достаточно ограничено. Механизм построения признаков (2.82) согласуется с построением координатных характеристик изображения объекта как полезного сигнала.

Проанализируем область возможных значений, которые могут принимать функционалы (2.28). Нетрудно заметить, что значения o_1, o_2 в любом случае

зависят от координат силуэта объекта, а также от того, в какой части поля зрения расположен объект.

Для примера рассмотрим бинарное изображение (рис. 2.10а) размера 64x64 и его смещенный образ (рис. 2.10б).

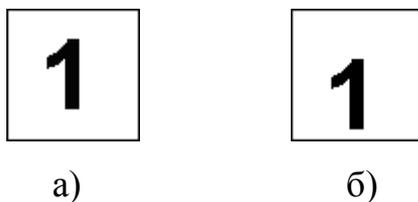


Рис. 2.10 – Пример изображения: а – эталонное; б – входное (смещенное)

Пусть $x, y \geq 0$ (центр координат изображения находится в верхнем левом углу поля зрения) для изображения рис. 2.10а получаем $o_1 = 0,4911$, $o_2 = 0,4921$. Смещенный образ рис. 2.10б характеризуется значениями $o_1 = 0,4210$, $o_2 = 0,5647$.

Далее с целью упрощения будем полагать: $x \geq 0$, $y \geq 0$, что всегда достижимо на этапе предварительной обработки и дает возможность оценить степень искажений рассматриваемого изображения с использованием функционалов o_1, o_2 . Эти условия также исключают ситуацию равенства знаменателя (2.28) нулю.

К особенностям функциональных признаков (2.28) следует отнести их чувствительность к преобразованиям смещений объекта изображения в пределах поля зрения. По этой причине для получения представления эталонного изображения необходимо вначале компенсировать смещение, например, нормировкой по центрам тяжести. С учетом того, что моделируемые искажения выполняются именно относительно центра поля зрения (а, значит, и относительно его центра тяжести), все искажения, не изменяющие центр тяжести (поворот, масштабирование и косой сдвиг) не будут влиять на значения функционалов (2.28).

На рис. 2.11 приведены значения функционалов (2.28) для эталонного и для нормализованного относительно центра поля зрения изображений.

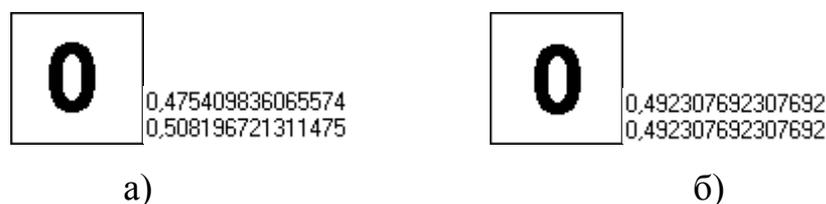


Рис. 2.11 – Значения функционалов o_1, o_2 : а – для входного изображения; б – для нормализованного по центрам тяжести изображения

Как видим, значения во втором случае (рис. 2.11б) более близки между собой, что указывает на равномерное распределение энергии изображения по полю зрения после нормализации смещений. Конкретно центры тяжести изображений рис. 2.11 имеют значения (29; 31) и (32; 32) соответственно.

По результатам проведенного анализа можно считать, что функциональные признаки (2.28) при оценке перспективных искажений могут быть применены как альтернатива функционалам центров тяжести.

Проанализируем влияние однопараметрических перспективных искажений (2.27) на предложенные признаки (2.28). Для этого исследуем зависимость значений o_1, o_2 при расположении центров отсчета x, y вдоль оси, соответствующей диагонали изображения. Сравнение поведения значений (2.28) эталона и изображения позволяет оценить возможность отслеживания параметров искажений перспективы с помощью (2.28).

На рис. 2.12а представлено эталонное изображение автомобильного номера размером 256x256 пикселей (с диагональю длиной 362 пикселя), а на рис. 2.12б представлен график зависимости значений функционалов o_1, o_2 от расположения центра x, y .

Точка разрыва на графике соответствует центру тяжести изображения, который в данном примере находится вблизи точки (120; 120). За исключением

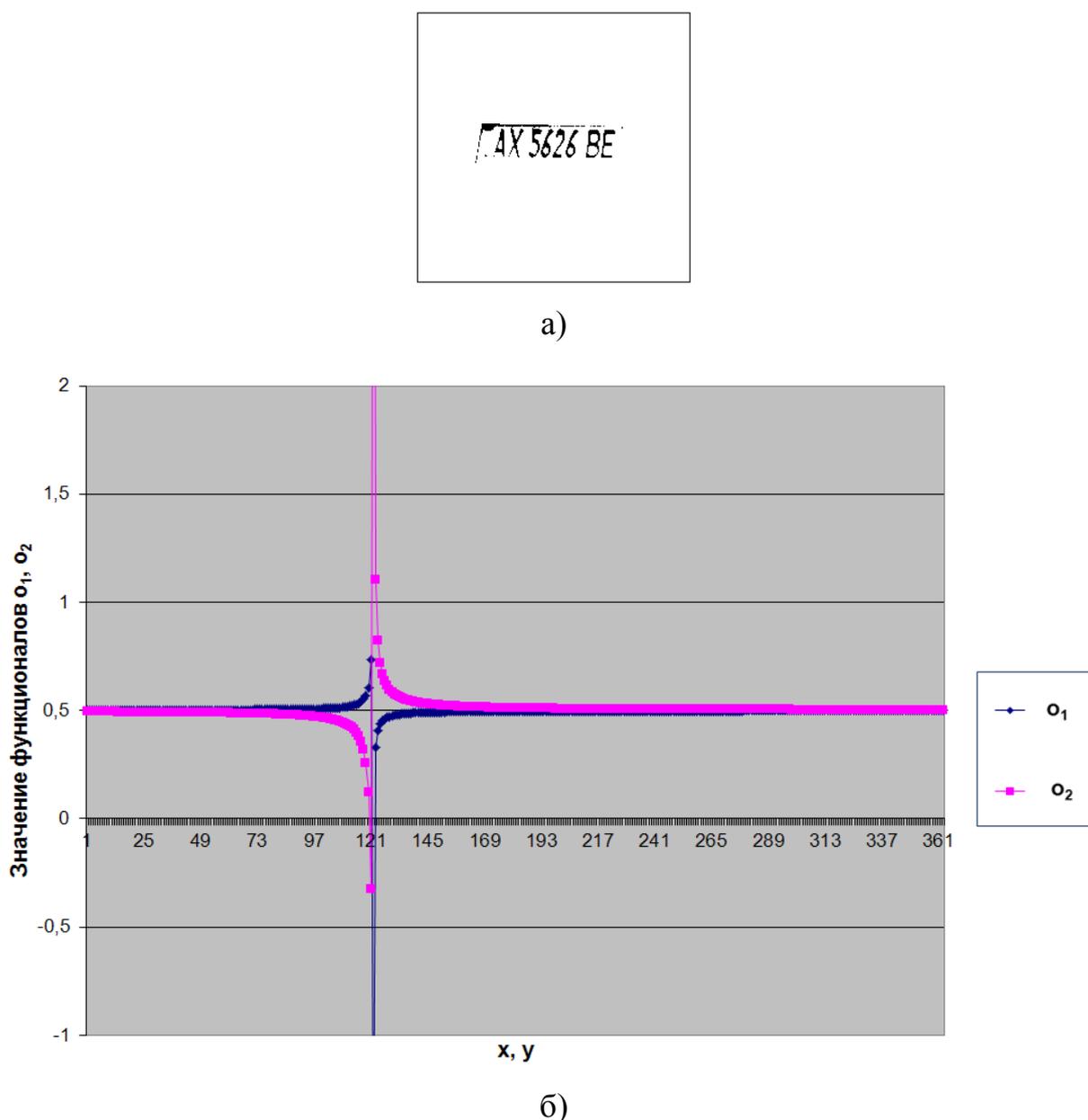
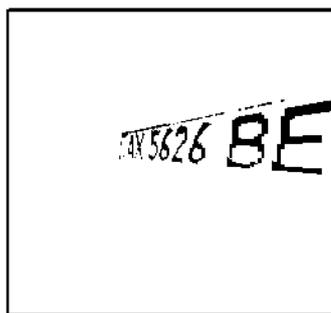


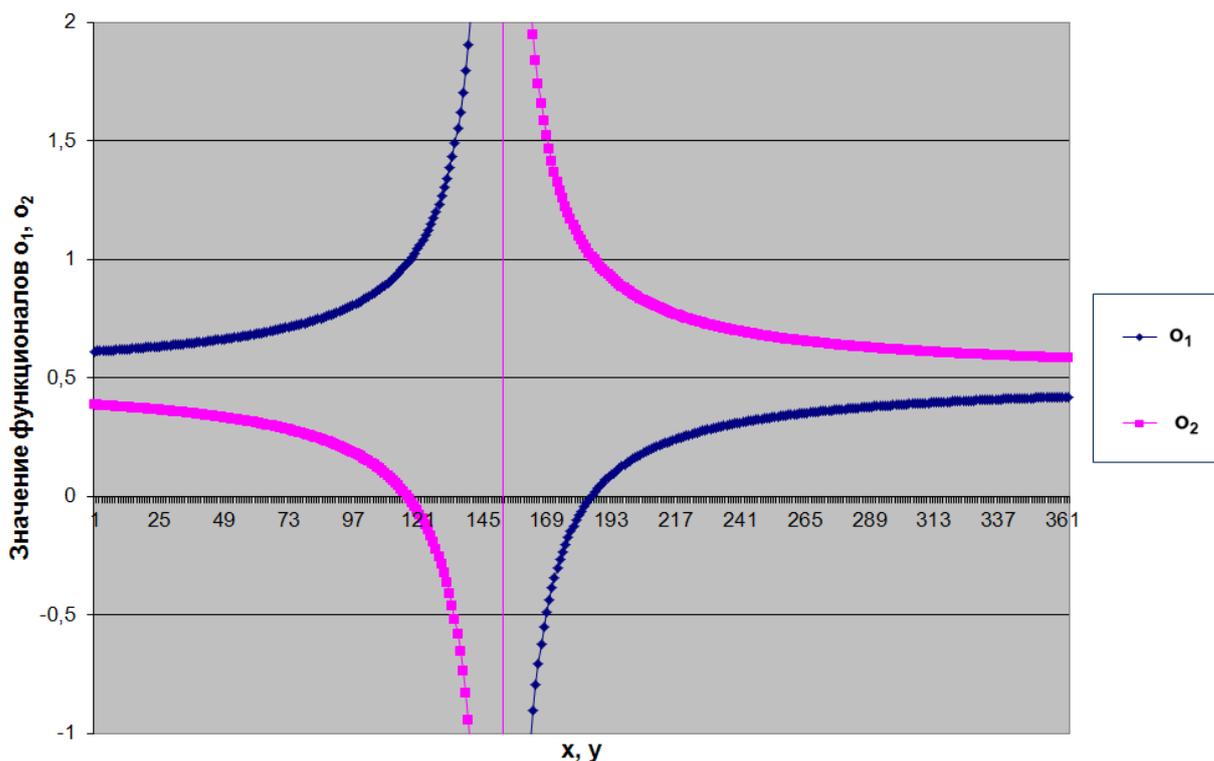
Рис. 2.12 – Эталонное изображение (а) и график зависимости функциональных признаков от центров отсчета (б)

интервала вблизи центра тяжести, значения o_1, o_2 близки к значению 0.5, что видно на рис. 2.12б.

Исследуем теперь значения o_1, o_2 при влиянии перспективных преобразований (рис. 2.13) вида P_x . При воздействии на эталон перспективных искажений тип графика не претерпевает существенных изменений (кривые гиперболического типа), меняются лишь параметры. А именно: точка центра тяжести смещается, разрыв приобретает больший размах, значения o_1, o_2 удаляются от эталонного значения, близкого к 0.5 (рис. 2.13б).



а)



б)

Рис. 2.13 – Входное изображение (а) и график зависимости O_1, O_2 от влияния искажений перспективы (б)

Анализируя значения функционалов (2.2) и графические представления рис. 2.12, 2.13, можно сделать вывод о том, что значения O_1, O_2 отражают распределение энергии функции яркости на изображении и размещение самого объекта в поле зрения по отношению к эталонным данным. Это дает возможность оценить наличие искажений перспективы ввиду их непосредственного влияния на поведение функционалов O_1, O_2 .

РАЗДЕЛ 3

ИССЛЕДОВАНИЕ И РАЗРАБОТКА МЕТОДОВ РАСПОЗНАВАНИЯ И АНАЛИЗА ГРАФИЧЕСКИХ СЦЕН

3.1 Анализ существующих методов и подходов анализа сцен

Задача анализа сцены и получение ее структурного описания является очень важной задачей, но порой оказывается очень трудной. Приемлемая для всех формализация этой проблемы, едва лишь выделяется из работ, выполненных к настоящему моменту. Были предприняты попытки заимствовать понятия из теории формальных языков и построить лингвистическую модель для анализа сцен. При этом сцена рассматривается, как выражение на языке, грамматика которого определяется допустимыми структурными отношениями. Исходя из этой формулировки, анализ сцен можно рассматривать как процесс грамматического разбора, в результате которого описание всей сцены получается в виде композиций взаимосвязных подсцен.

Процесс анализа сцены является по существу процессом упрощения – сложный объект (исходное изображение) превращается в более простой с помощью нескольких последовательных операций. Единственным этапом этой последовательности является превращение исходного изображения в контурный рисунок. Контурное изображение обычно содержит меньше информации, чем исходное, и нет гарантии, что потерянная информация является не существенной. Однако, примем, что преобразование изображения в контурный рисунок является при некоторых обстоятельствах полезным, и сосредоточим свое внимание на способах осуществления этой информации.

Контурное изображение может быть получено из исходного путем выделения областей, содержащих резкие переходы от темного к светлому, и подавления областей с примерно однородной интенсивностью. Другими словами, контуры есть края, а края — это по определению переходы между

двумя существенно различными интенсивностями. На языке функций интенсивности край — это область плоскости (X, Y) , где велик градиент функции $g(x, y)$. Для получения контурного рисунка, таким образом, требуется оценка величины модуля градиента функции. Эту величину можно вычислить, если известны производные этой функции по каким-либо двум ортогональным направлениям. Следовательно, нужно только выбрать два ортогональных направления и способ приближенного вычисления (одномерной) производной, чтобы иметь все необходимые составляющие алгоритма получения контурных рисунков.

Также при решении многих задач анализа сцен анализирующая система сталкивается с той или иной формой следующего простого вопроса: содержит ли сцена изображение заранее заданного объекта? Метод, традиционно используемый для ответа на вопросы этого типа, называется сравнением с эталоном. Рассмотрим сначала сравнение с эталоном на простых примерах с бинарными изображениями, а затем распространим наше обсуждение на общий случай. Далее мы будем вести рассмотрение, пользуясь примером. Предположим, что у нас есть градиентное изображение, на котором представлены простые геометрические тела в виде контуров, и мы хотели бы выяснить, имеется ли на этом изображении треугольник. Очень простой подход к решению этой задачи может заключаться в следующем: нужно построить эталон или трафарет, и просматривать через него последовательно все изображение. Если мы найдем такую позицию, при которой «отверстие» в эталоне заполнено белым, можно будет сделать вывод, что в этом месте обнаружен треугольник. Сразу же возникает возражение против такой процедуры: любая достаточно большая сплошная область белого цвета может быть ошибочно принята за треугольник. [45].

Таким образом, процедура анализа сцен сводится к распознаванию различных областей, образов, признаков на сцене.

Возможность распознавания опирается на схожесть однотипных объектов. Несмотря на то, что все предметы и ситуации уникальны в строгом

смысле, между некоторыми из них всегда можно найти сходства по тому или иному признаку. Отсюда возникает понятие классификации — разбиения всего множества объектов на непересекающиеся подмножества — классы, элементы которых имеют некоторые схожие свойства, отличающие их от элементов других классов. И, таким образом, задачей распознавания является отнесение рассматриваемых объектов или явлений по их описанию к нужным классам.

Однако до сих пор существует область, которая не поддается автоматизации с помощью современной компьютерной техники. Речь идет о распознавании графической информации, а именно изображений и видео.

Технологии распознавания и анализа графических сцен имеют широкое применение в современных компьютерных системах, и находят все большее применение в повседневной жизни. В решении таких задачи очень важным является нахождение всех возможных признаков, описывающие объекты распознавания (явления). На сегодняшний день не существует способов автоматической генерации признаков необходимых для распознавания объектов. Пока что, это под силу только человеку. Поэтому, выбор признаков можно осуществлять, имея представление об их общих свойствах. Признаки могут подразделяться на:

Детерминированные признаки - это такие характеристики объектов или явлений, которые имеют конкретные и постоянные числовые значения.

Вероятностные признаки - это характеристики объекта (явления), носящие случайный характер.

Логические признаки распознавания - это характеристики объекта или явления, представленные в виде элементарных высказываний об истинности (“да”, “нет” или “истина”, “ложь”).

Структурные признаки - производные (то есть, элементарные, не производимые из других элементарных признаков) элементы (символы), примитивы изображения объекта распознавания.

Важную роль в определении признаков играют методы анализа и выявления их. Одним из таких методов является семантический анализ.

Несмотря на свою востребованность практически во всех областях жизни человека, семантический анализ является одной из сложнейших математических задач. Вся сложность заключается в том, чтобы “научить” компьютер правильно трактовать необходимые образы.

Методика построения семантического описания заключается в семантической интерпретации выделенных на этапе геометрического особых точек, структурных элементов контура и их групп. Каждая предметная область, в которой решается задача накопления и поиска графической информации, имеет свои методики семантического описания изображений различной степени формализации.

Итак, семантическое описание изображения может быть получено путем семантической интерпретации геометрического описания, для семантической интерпретации в базе знаний должны храниться геометрические описания соответствующих понятий предметной области.

Основой метода семантического анализа является геометрическая модель объекта. Геометрическая модель объекта - есть геометрическое описание контурного изображения объекта, формализованное на семантической сети. На этапе семантического описания элементы геометрической модели объекта интерпретируются семантически значимыми понятиями, например, точка перегиба может стать внутренним углом глаза, а дуга - складкой нижнего века и т.п. Геометрическую модель объекта называется семантической интерпретацией, если ее элементы имеют семантически значимые названия.

Построение аннотации изображения на основе семантически интерпретированной геометрической модели объекта заключается в реализации простых вычислительных процедур, описывающих зависимости между атрибутами семантически значимых элементов модели и семантическими признаками.

База знаний, используемая в процессе семантического анализа, должна включать четыре группы знаний:

- знания геометрических примитивов, используются на этапе геометрического описания изображения объекта, обеспечивают возможность интерпретации частей контура как простейших геометрических фигур и линий;
- знания структуры изображаемых объектов и их классов, описывают множество моделей классов, используются на этапе семантической интерпретации геометрической модели объекта;
- знания лингвистики, описывают соответствие нечетких понятий и числовых величин, обеспечивают возможность построения геометрического и семантического описаний в терминах нечетких понятий естественного языка;
- знаний структуры базы данных, в которой хранятся аннотации изображений, обеспечивают возможность работы с базой изображений, в том числе дополнения его аннотациями новых изображений.

Необходимо разработать структуру базы знаний и представить ее на семантической сети.

Базовым методом вывода знаний на семантической сети является метод вывода путем сопоставления ее фрагментов с целью нахождения общих и отличительных элементов сети. Метод сопоставления фрагментов семантической сети состоит в переборе всех или некоторых возможных вариантов сопоставления их вершин и луг с целью выбора лучшего варианта.

Существует 4 процесса функционирования модели памяти, которые подлежат моделированию на нечеткой семантической сети:

1. Ввод в базу знаний информации о новом объекте - первичное обучение;
2. Переработка информации в базе знаний - упрощение семантической сети, вывод новых знаний;
3. Ввод в базу знаний дополнительной информации об известном объекте - вторичное обучение;
4. Извлечение информации из базы знаний - сопоставление описания неизвестного объекта с описаниями известных объектов из базы знаний.

Процесс переработки информации в базе знаний или упрощения семантической сети (2-ой процесс) предполагает замену ее фрагмента одной

вершиной, которой сопоставлено более сложное понятие, объект или явление. Замена происходит в результате целенаправленного поиска фрагмента семантической сети, описывающего понятие, в исходной семантической сети. Для поиска заданного фрагмента необходимо применять метод сопоставления фрагментов [46], [47].

Ввод в базу знаний дополнительной информации об известном объекте (3-ий процесс) с целью дообучения системы осуществляется в результате предъявления на вход системы нового изображения известного объекта. Метод сопоставления фрагментов семантической сети применяется для сравнения старого и нового описаний объекта, после чего старое описание модифицируется с учетом отличительных моментов нового описания.

Извлечение информации из базы знаний с целью распознавания изображения некоторого объекта (4-й процесс) предполагает предъявление на вход системы изображения неизвестного объекта. Изображение подвергается описанию на семантической сети, полученное описание поочередно сравнивается с описаниями всех известных системе объектов, определяется наиболее близкое описание, после чего делается вывод о типе изображенного объекта. Метод сопоставления фрагментов используется для сопоставления описаний известного и неизвестного объектов.

3.1.1 Семантический анализ изображений

3.1.1.1 Общее описание семантического анализа изображений

Семантический анализ изображений относительно новое и перспективное направление в распознавании изображений, которое позволяет создавать решения задачи распознавания изображений, независимые от предметной области. Особенностью данного подхода является способ представления изображения в виде словаря графических лексем, а основу семантического аппарата распознавания представляет базовая грамматика графических слов, за счет которой достигается гибкость и инвариантность подход, поскольку

словарь может быть изменен от задачи к задаче. Нарботки в области семантических нейронных сетей распознавания речи могут быть использованы в качестве аппарата распознавания изображений, однако для этого необходимы значительные доработки аппарата семантических нейронных сетей [48].

Для построения систем распознавания изображения, подобных человеческой, необходимо задание базы знаний, состоящей из изображений, и словарей фрагментов, которые в некоторой комбинации будут ассоциироваться с теми или иными изображениями. Важно не забывать, что сами ассоциированные изображения могут быть фрагментами в распознавании еще больших сцен. То есть, фактически, необходима разработка математического и логического аппарата многоуровневой рефлексивной ассоциативной памяти.

Семантическая сеть представляет собой ориентированный граф, вершины которого представляют собой понятия, а ребра (дуги) - определяют семантические отношения между этими понятиями. Примером такой сети можно считать гипертекст с его якорями и гиперсвязями (но это лишь частный простой случай). Граф может быть древо-подобным, но может содержать и циклические структуры.

Семантическая сеть дает возможность охарактеризовать отношения между понятиями. Понятия представляются графическими лексемами. Простым примером является иерархическая сеть, где понятия являются таксономическими терминами и единственным отношением является *type-of* (относится к типу) (гипонимные отношения - *hyponymous relationship*, когда одно понятие включает в себя другие понятия). Более сложные семантические сети включают в себя отношения различных типов, такие как яркость, текстура и цвет.

Понятие, которое семантически связано с другим понятием, отмечают RT и используются для семантических отношений. Связанные понятия могут, например, выражать взаимоисключающие отношения, усиливающие или подавляющие отношения, причинно-следственные связи, отношения положения или парадигматические отношения. Семантические отношения

отражают отношения в графическом языке, который зависит от предметной области [49].

3.1.1.2 Грамматика на основе простых фрагментов изображения

Наиболее гибкой и полной является грамматика, которая строится на основе множества простых фрагментов изображения. На рисунке 3.1 изображен пример подобной грамматики.



Рисунок 3.1 – Пример словаря базовых графических фрагментов

Подобный словарь не зависит от предметной области, поэтому приемлемо его применение в решении разнообразных задач распознавания. Однако, недостатком подобного словаря является его недетерминированность и, как следствие, бесконечная длина грамматики.

Альтернативным вариантом построения грамматики для системы семантического распознавания изображений является определение конечного набора элементарных геометрических фигур.

3.1.1.3 Грамматика на основе простых элементов

Точка. Под точкой понимается некоторая окружность, столь малого диаметра, что в системе счисления распознаваемого изображения размерами окружности можно пренебречь, важно лишь положение центра окружности.

Прямые.

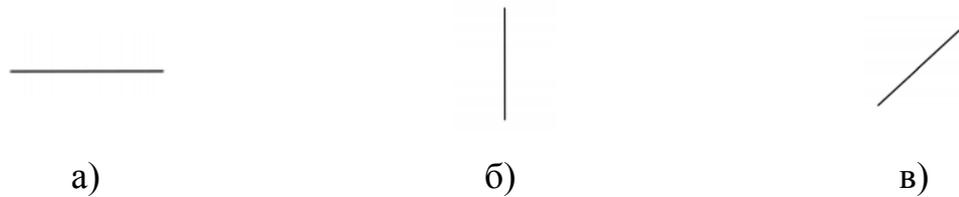


Рисунок 3.2 - Прямые линии с четко заданным направлением

Под прямыми линиями понимают фрагментарные приближения на исходном изображении, которые схожи с прямыми. Важно не забывать, что различные положения линии вызывают ассоциации с разными лексемами грамматики.

Ломаные.

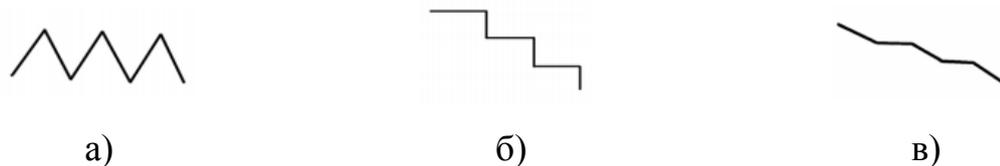


Рисунок 3.3 - Ломаные линии

На рисунке 3.3 изображены следующие элементы:

- а - простая ломаная линия под углом 45 градусов;
- б - простая ломаная линия под углом 90 градусов;
- в - многоугольная ломаная линия под углом 135 градусов.

Кривые из геометрических элементов окружности.

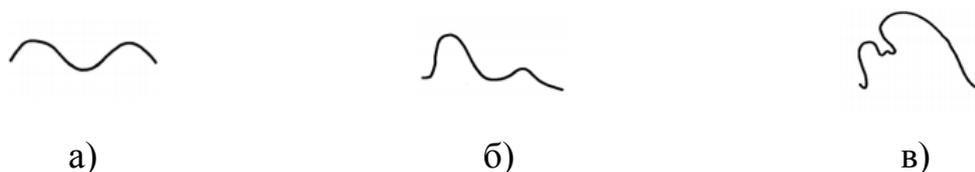


Рисунок 3.4 - Кривые из геометрических элементов окружности

На рисунке 3.4 изображены следующие элементы:

- а - простая кривая из геометрических элементов окружности;
- б - усложненная кривая из свободных элементов;
- в - усложненная кривая из различных комбинаций данных элементов.

Смешанная линия.

Для обозначения конкретных детерминированных сложных элементов могут применяться конкретные лексемы сложного вида, как на рисунке 3.5.



Рисунок 3.5 - Сложная смешанная линия

3.2 Общая проблематика задач обработки изображений

Решение многих проблем науки и практики приводит к необходимости извлечения полезной информации из различного рода многомерных данных, а именно оптических изображений. Такие задачи возникают в очень многих областях знаний: в медицине, радио-, тепло- и гидролокации, обеспечении наблюдения и безопасности, дорожных ситуациях, исследовании Космоса и Земли, телевидении и прочих. Характерно, что эти задачи приходится решать при наличии различного рода мешающих факторов – помех, мешающих изображениям, переменчивости условий наблюдения, динамики наблюдаемого объекта, взаимного перемещения приемника и объекта. Полезный сигнал может быть очень слаб по отношению к помехам и визуально неразличим на фоне мешающих изображений.

Нередко объем исходных данных очень велик (мониторинг дорожного движения, массовые медицинские обследования), они поступают с большой скоростью и требуют обработки в режиме реального времени. Поэтому методы, которые работают в теории оказываются абсолютно несостоятельными на практике, поскольку потребляют слишком большой объем вычислительно мощности компьютерных систем. Другим значительным недостатком алгоритмов, построенных на математических моделях, является невозможность

распараллелить их работу на нескольких вычислительных процессорах, вследствие монолитности структуры [50].

В теории распознавания существует два принципиально разных подхода к распознаванию изображений:

- статистический подход, использующий традиционные: корреляционные, признаковые и семантические методы;
- семантический подход, использует разнообразные методы построения семантических сетей.

Статистический подход, практически, изжил себя на сегодняшний день, хотя именно статистический подход дал начальный толчок развития науке распознавания изображений.

Семантический подход изначально применялся для анализа текстовой информации. Однако с развитием вычислительной техники становится понятно, что семантический подход – наиболее перспективный подход к распознаванию изображений на сегодняшний день.

3.2.1 Семантическая нейронная сеть

В сети фон Неймана существуют ограничения на топологию связей - постулируем отсутствие таких ограничений. Обработываются только логические значения – постулируем обработку нечетких значений. Все нейроны в сети фон Неймана синхронизируются тактами. В отличие от сети фон Неймана, в семантической нейронной сети отсутствуют ограничения на топологию нейронов. Это приводит к невозможности относительной адресации нейронов, как это делал фон Нейман. В этом случае, необходимо ввести абсолютную адресацию. Каждый нейрон должен иметь некоторый уникальный идентификатор, зная который можно получить к нему прямой доступ. Разумеется, нейроны, которые взаимодействуют друг с другом через аксоны-дендриты, должны иметь идентификаторы друг друга. Относительную

адресацию можно моделировать, введя специфичность нейронов, аналогично тому, как это реализовано в биологических нейронных сетях.

В первоначальном описании семантической нейронной сети отсутствует описание способностей к саморефлексии и самомодификации. Конечно, можно сказать, что семантическая нейронная сеть унаследовала эти способности от ее прототипа - сети фон Неймана. Да, это так, идея саморефлексии унаследована. Но от простой идеи далеко до практической реализации. Итак, требуется обеспечить анализ и модификацию структуры одного фрагмента нейронной сети другим ее фрагментом.

Допустим, что нечетких данных, передаваемых от нейрона к нейрону, будет для этого недостаточно. Постулируем наличие указателя на нейрон. Этот указатель будет просто уникальным номером - идентификатором нейрона в хранилище нейронов. Пусть нейроны в состоянии обрабатывать не только нечеткие данные, но и указателями друг на друга. Очевидно, что совершенно реально реализовать это технически.

Теперь необходимо осмыслить, что есть такой указатель. Указатель на нейрон представляет собой виртуальную связь, которая не реализована в виде дендрита или аксона. Пусть в сконструированной виртуальной реальности, нейроны будут взаимодействовать друг с другом не только посредством передачи сигналов через аксоны-дендриты, но и с помощью паранормальных эффектов. Итак, нейрон будет иметь сигнальные входы, сигнальные выходы и набор виртуальных связей с другими нейронами. Нейрон будет иметь возможность взаимодействовать с другими нейронами, владея их указателями, но, не имея с ними сигнальных связей. Так же очевидно отличие между указателями и сигнальными связями. Сигнальные связи являются двухсторонними структурными образованиями связанными как с источником, так и с приемником сигнала. Указатель же односторонен. Владелец идентификатора некоторого нейрона в состоянии инициировать взаимодействие с этим нейроном. Определить наличие указателя на некоторый нейрон технически невозможно без проведения поиска такого указателя. Связи,

проявленные в виде аксонов и дендритов можно рассматривать как долговременную память системы, инвариантную к контексту. Сигналы и указатели, обрабатываемые нейронами - как сверхоперативную информацию, зависящую от текущего контекста. Потеря сигналов или указателей (по аналогии с эпилептическим припадком) не должна повлиять на проявленную структуру и привести к изменениям в долговременной памяти или личности.

Наличие таких указателей приводит к возможностям косвенного взаимодействия. Аналогом из традиционных языков программирования является двойная или тройная разадресация. В этом случае некоторый нейрон-1, имеющий виртуальную связь с нейроном-2 в состоянии взаимодействовать с нейроном-3 при условии, что нейрон-2 владеет указателем на нейрон-3. Это приводит к широким возможностям взаимодействия одного нейрона с другим нейроном без непосредственного контакта как через проявленные, так и через виртуальные связи.

Для обеспечения саморефлексии в сети можно ввести нейроны, выполняющие функции анализа и изменения структуры сети. Введем нейроны рецепторы, реагирующие на структурные элементы конструируемой нейронной сети. Такие нейроны рецепторы могут переходить в возбужденное состояние при выполнении некоторых определенных условий, описывающих наличие или отсутствие нейронов и связей между нейронами с определенными характеристиками. Так же введем нейроны эффекторы, которые, будучи переведены в возбужденное состояние, выполняют некоторую модификацию структуры нейронной сети.

Для обеспечения полноты системы необходимо обеспечить самоприменимость нейронов рецепторов и нейронов эффекторов. Нейрон рецептор должен иметь возможность анализировать другие нейроны рецепторы, в том числе и того же типа что он сам. Нейроны эффекторы должны уметь модифицировать другие нейроны эффекторы, а не только нейроны, выполняющие обработку сигналов. Это возможно благодаря наличию указателей на нейроны, постулированных ранее [51], [52].

В качестве нейронов рецепторов можно постулировать нейроны, определяющие наличие или отсутствие нейронов определенных типов связанных с некоторым нейроном; или нейроны, определяющие наличие связи между заданными нейронами или наличие или отсутствие связи определенного типа у некоторого нейрона. В случае удовлетворения условия, нейрон рецептор принимает возбужденное состояние, в обратном случае – пассивное. В качестве нейронов – эффекторов можно постулировать нейроны, связывающие два нейрона некоторой связью, нейроны создающие другие нейроны, эффекторы, уничтожающие нейроны или связи. Если уровень возбуждения превысит некоторый порог, то нейрон эффектор выполнит свою функцию.

В системе был проведен эксперимент по синтезу структуры нейронной сети на основе внешнего задания средствами самой нейронной сети. В виртуальной машине, выполняющей семантическую нейронную сеть, была реализована поддержка следующих типов нейронов связанных с самоприменимостью:

- нейрон линковщик - связывает два нейрона сигнальной (реальной) связью, на которые указывают виртуальные указатели;
- нейрон копировщик - копирует виртуальную связь из одного нейрона в другой нейрон;
- нейрон репликатор - заставляет делиться нейрон, на который указывает его первая виртуальная связь. сохраняет виртуальную связь с новым нейроном во второй виртуальной связи;
- нейрон поглотитель нейрона - уничтожает нейрон, на который указывает виртуальная связь;
- нейрон поглотитель связи - уничтожает сигнальную связь между двумя нейронами, на которые указывают виртуальные связи;
- нейрон хранитель указателя - хранит указатель на некоторый нейрон, аналогичен обычной ячейке памяти компьютера.

Рассмотрим возможность восстановления семантической нейронной сети после повреждений. Очевидно, что структура нейронной сети определяет

индивидуальность интеллекта. Смерть нейрона означает забывание соответствующей ему информации. Восстановить информацию о наличии некоторого нейрона про розового слона без знания того, что система имеет знание об этом слоне, невозможно. Если мы постулируем, что наличие нейрона означает наличие знания о слоне, то отсутствие нейрона означает отсутствие этого знания. Уничтожение нейрона уничтожило знание. Уничтожение нейрона означает необратимое забывание. Следовательно, регенерация нейронной ткани по заданному в ДНК образцу представляется невозможной. Чтобы восстановить поврежденный нейрон необходимо заново обучиться той же информации.

Семантика нейрона определяется не его внутренним состоянием или внутренней сложностью. Она определяется связями нейрона с его соседями. Только соседи определяют семантику отдельного нейрона. Поэтому не важно, в какой конкретной точке пространства будет находиться ансамбль нейронов, описывающий понятие розового слона. Главное, чтобы имелись связи, позволяющие определять "слоновость" и "розовость".

Регенерация поврежденных рецепторов, эффекторов или других регулярных структур вполне возможна. При отсутствии повреждений в семантических структурах, связанных с поврежденными регулярными структурами, сохраняется возможность заново восстановить связи регенерированных регулярных структур. Даже если не удастся полностью восстановить точную копию регулярной структуры бывшей до регенерации, изоморфизм связей позволит выполнять вновь регенерированным нейронам ту же функциональную нагрузку.

Повреждения центральной нервной системы могут быть устранены во время повторного обучения. Если нейрон-розовый-слон будет уничтожен, то система забудет о таком слоне. Когда же в поле рецепторов попадет некоторый розовый слон, то будет создан новый нейрон в любом другом месте нейронной сети, связанный с соседями так, чтобы отражать новое понятие.

Регенерация нейронной сети, будет идти не в области повреждения, а в области сохранившейся ткани. Поврежденную область, для исключения

«эффекта эпилепсии», желательна дегенерировать и рассосать. В дальнейшем, по мере обучения, здоровая ткань, с увеличением количества вновь образованных нейронов, займет место поврежденной. Очевидно, расширение будет происходить только по мере обучения новой информации [53], [54].

Разработанный алгоритм семантической классификации изображений основан на трех основных идеях:

- 1) для классификации изображений используется спектральное признаковое пространство;
- 2) семантика полного изображения является производной от семантики сегментов изображения, что позволяет реализовать иерархическую процедуру классификации;
- 3) исключаются какие-либо априорные предположения о конфигурации семантического класса в пространстве признаков, классификация выполняется по достоверным прецедентам базы данных.

В контексте семантической классификации представлены новые алгоритмы адаптивной сегментации изображений, алгоритмы информативной оценки системы первичных признаков и формирования сложных вторичных признаков, алгоритмы нечеткой метрической классификации сегментов изображения, алгоритмы нечеткой иерархической классификации изображений по результатам сегментной классификации [55].

Типичной задачей обработки мультимедиа изображений является их классификация. Среди различных видов классификаций по уровню значимости и уровню сложности выделяется задача семантической классификации зрительных образов.

Анализ семантики является важной частью иерархической процедуры обработки изображений. В основании пирамиды лежат методы формирования первичной системы информативных признаков. Основное требование на данном этапе – обеспечить максимально-возможную инвариантность признаков к топологическим преобразованиям и высокое быстродействие в получении первичной информации. Первичные признаки представлены в пространстве

высокой размерности и имеют значительные различия по уровню информативности, эти обстоятельства препятствуют их непосредственному использованию в задаче классификации. На следующем уровне иерархии формируется система вторичных признаков с примерно одинаковым уровнем значимости. Исходной информацией при этом служит анализ накопленной базы данных, который производится всякий раз при добавлении новых данных. Главное требование к процедуре анализа состоит в том, чтобы подобрать оптимальную систему признаков и сократить размерность признакового пространства. В режиме обучения классифицирующая система, взаимодействуя с оператором, накапливает банк семантических понятий. Временные затраты на анализ базы данных на этом этапе не имеют решающего значения, поскольку не влияют на быстродействие системы в рабочем режиме. На верхнем уровне пирамиды, параметрическими или непараметрическими методами решается задача семантической классификации образов. Ответ может быть многозначным, поэтому необходимо ранжировать полученные решения, используя ту или иную оценку уровня значимости.

Задача семантического распознавания пересекается с задачей создания искусственного разума. Сущность подхода состоит в трансформации количественной информации любого вида к «текстовой форме». В процессе обучения реализуется выделение повторяющихся текстовых цепочек, которые трактуются как концептуальные понятия. Главная проблема использования данного подхода связана с адекватной трансформацией изображения к «текстовому» образу.

Завершающим этапом данной задачи является семантическая классификация и распознавание изображений. Это обусловлено тем обстоятельством, что база знаний наполняется человеком, а методология в дальнейшем опирается на базу.

В качестве первичной системы признаков используются спектральные коэффициенты яркости и цветности пикселей элементарных составляющих изображения.

3.2.2 Описание принципов разработанного подхода

Основу разработанного программного комплекса составляет алгоритм так называемой сегментной свертки изображения. Данный алгоритм построен на принципе разбора изображения исходя из содержимого базы знаний. Это означает, что программная система должна пытаться на основании собственной базы знаний составить изображение или его фрагмент максимально сопоставимый с изображением находящимся на обработке системой распознавания. Детализированный пример работы программного комплекса построенного на основании данного алгоритма представлен на рисунке 3.6.

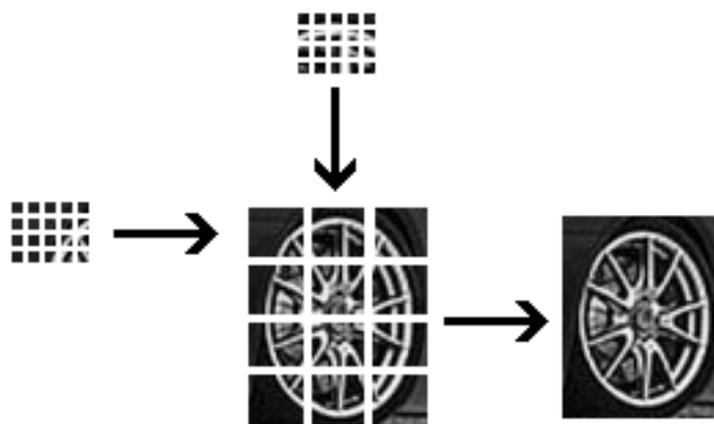


Рисунок 3.6 – Графическое представление алгоритма семантической свертки изображения

Таким образом, разработанный алгоритм позволяет «собрать» изображение из фрагментов графического словаря. При этом не имеет значения, с какой точки начинать сборку. Исходное изображение разбирается на фрагменты заданного размера и после этого в работу вступает обработчик семантической свертки изображения. Обработчик изображения должен загружать фрагменты изображения из базы данных и находить соответствия между фрагментами из базы знаний и фрагментами обрабатываемого изображения, учитывая порядок расположения каждого фрагмента и его соседей.

Основу обрабатывающего алгоритма составляет динамическая семантическая нейронная сеть. Суть данной структуры заключается в непостоянстве структуры сети. Основные сегменты сети хранятся в базе данных на жестком диске, а нужный сегмент сети, в зависимости от предметной области и связанных с нею областей, загружается в RAM вычислительной машины и участвует обработке данных. При этом ненужные и незначительные для определенных предметных областей сегменты сети не загружают CPU и RAM вычислительной станции.

Ясно, что при таком построении системы знаний необходима достаточно большая база знаний, однако такой подход обеспечивает гибкость программных решений, разработанных на его основе.

Подход на основании алгоритма графической свертки изображения обладает следующими преимуществами:

- гибкость программных решений, разработанных на его основе;
- независимость от предметной области;
- возможность описать любой объект максимально точно;
- данный подход основан на объектах естественного мира, а не на идеальных математических моделях, что позволяет повысить точность распознавания;
- возможность распараллелить обработку изображений.

В то же самое время данный алгоритм несет в себе ряд недостатков, которые обусловлены его преимуществами:

- необходимость наполнения большой базы знаний со сложными связями;
- необходимость разработки гибкой и конфигурируемой системы обучения;
- потребность в большом количестве вычислительных мощностей.

Однако, выгода, полученная от использования данного подхода при современном уровне вычислительной техники, несомненно, превышает затраты на нивелирование недостатков.

3.3 Описание архитектуры и API программного комплекса

3.3.1 Общее описание архитектуры приложения

Разработанное программное обеспечение представляет собой комплексный продукт, который включает в себя множество модульных подсистем, для решения задачи автоматизации распознавания графических сцен на всех этапах, от получения видеопотока до сохранения результатов в базу данных. В то же самое время модульная архитектура приложения обозначает, что каждый модуль имеет определенный программный интерфейс, который по желанию пользователя может иметь собственную реализацию для покрытия некоторых условий, специфичных для предметной области пользователя. Это обозначает, что различные модули системы могут менять свое поведение, в зависимости от условий задачи, а также конфигурации системы. Кроме того поведение некоторых модулей может быть дополнено или полностью изменено по желанию пользователя

Программный комплекс поставляется в виде готового настольного приложения, которое пользователь может использовать как готовое решение. Кроме того, за счет модульной архитектуры приложения существует возможность извлечь библиотеки, отвечающие за распознавание сцен, из поставляемого программного комплекса и использовать их в приложении пользователя, в качестве подключаемого модуля.

Общий алгоритм работы программного комплекса был разделен на 9 основных подсистем, которые изображены в порядке их задействованности в процессе семантического распознавания сцен на общей блок-схеме работы программного комплекса. Общая блок-схема работы программного комплекса изображена на рисунке 3.7

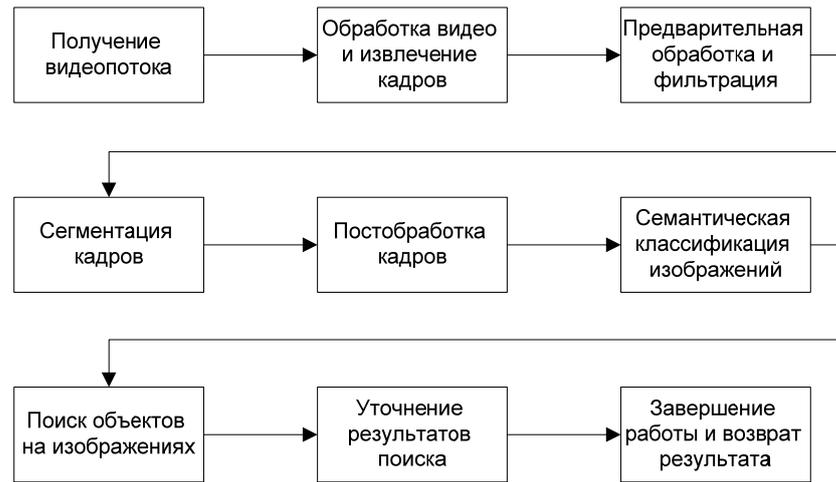


Рисунок 3.7 – Общая блок-схема работы программного комплекса

В целом, архитектура разработанного программного продукта представляет собой набор интерфейсов, которые образуют собою внутренние API и SPI приложения, каркас, объединяющий все блоки программного комплекса воедино и отвечающий за их последовательный вызов и передачу данных от блока к блоку, а также систему конфигурации приложения при помощи файлов формата XML.

Каждый блок, изображенный на схеме, представляет собой отдельный модуль, отвечающий за определенный аспект работы общего алгоритма программы.

3.3.2 Используемые платформы и программные средства

3.3.2.1 Языковое решение

В качестве основного языка для написания данного программного продукта был выбран язык программирования Java.

3.3.2.2 Графоподобная база данных

Поскольку база знаний приложения представляет собой набор сложных связей между относительно простыми компонентами, то наиболее естественной

структурой для подобной базы данных выступает граф, связывающий сегменты изображений в единое целое, то в качестве базы данных была выбрана база данных на графах. Наиболее эффективным решением для построения графоподобной базы данных на сегодняшний день является база данных Neo4J, которая представляет собой легковесную полноценную платформу для манипулирования данными, при этом база данных представляет собой направленный граф. Данная особенность Neo4J является очень эффективным и необходимым свойством для построения базы знаний, хранящей информацию семантической системы распознавания изображений.

Neo4j это высокоэффективной БД для хранения данных в виде графов. Вместо столбцов, строк или ключей и их значений, модель хранения данных представляет собой структуру, похожую на сеть - задаются узлы, их свойства и отношения между собой. БД обеспечивает прекрасную производительность и широкие возможности по масштабированию решения - один сервер может обслуживать несколько миллиардов объектов.

Neo4j поддерживает транзакции, индексацию данных и обладает хорошей масштабируемостью – база может обслуживаться кластером из нескольких серверов. Готовые биндинги распространяются для языков Clojure, Erlang, Gremlin, Groovy, Java, PHP, Python, Ruby и Scala. Код проекта написан на языке Java.

3.3.2.3 Иерархия основных интерфейсов механизма распознавания

Разработанный алгоритм и архитектура приложения требуют, чтобы иерархия классов представляла собой модульную структуру с возможностью подмены любого из элементов на альтернативную имплементацию. Данная особенность позволяет изменять поведение программы без переписывания исходного кода. Таким образом, основу приложения составляют API и

конвенции между интерфейсами приложения. Это позволяет достичь высокой гибкости и модульности приложения в целом.

Основу разработанного программного комплекса составляет подсистема распознавания образов, которая состоит из девяти модулей. Ее диаграмма классов изображена на рисунке 3.8.

Центральное управляющее место в иерархии классов системы распознавания изображений занимает контроллер процесса RecognitionContoroller. Данный интерфейс представляет обработчик, который играет управляющую роль в процессе выполнения алгоритма обработки и распознавания сцен. Первичной обработкой графических сцен занимается ImagePreparationHandler. Данный обработчик призван выполнить предварительную подготовку графической сцены для анализа, а именно: фильтрацию изображения от помех, удаление бликов, повышение контрастности изображения. Реализация ImagePreparationHandler, которая включена в поставку, называется CompsiteImagePreparationHandler. Данный класс позволяет выполнять настройку цепочки фильтров с помощью XML.

Для того чтобы оптимизировать процесс семантического распознавания изображений, разработанный программный комплекс позволяет выполнять предварительную сегментацию изображений. В дальнейшем распознавание может проводиться для каждого сегмента графической сцены в отдельности. Сегментацию графических сцен выполняет обработчик изображений ImageSegmntntationHandler.

После выполнения сегментации в работу вступает последний обработчик фазы подготовки изображения сцены ImagePreparationFinalizer. Данный обработчик выполняет дополнительную подстройку исходных изображений, чем завершает фазу подготовки изображения для работы модуля семантического распознавания изображений. Цель существования данного модуля – разделить подготовку изображения на две стадии.

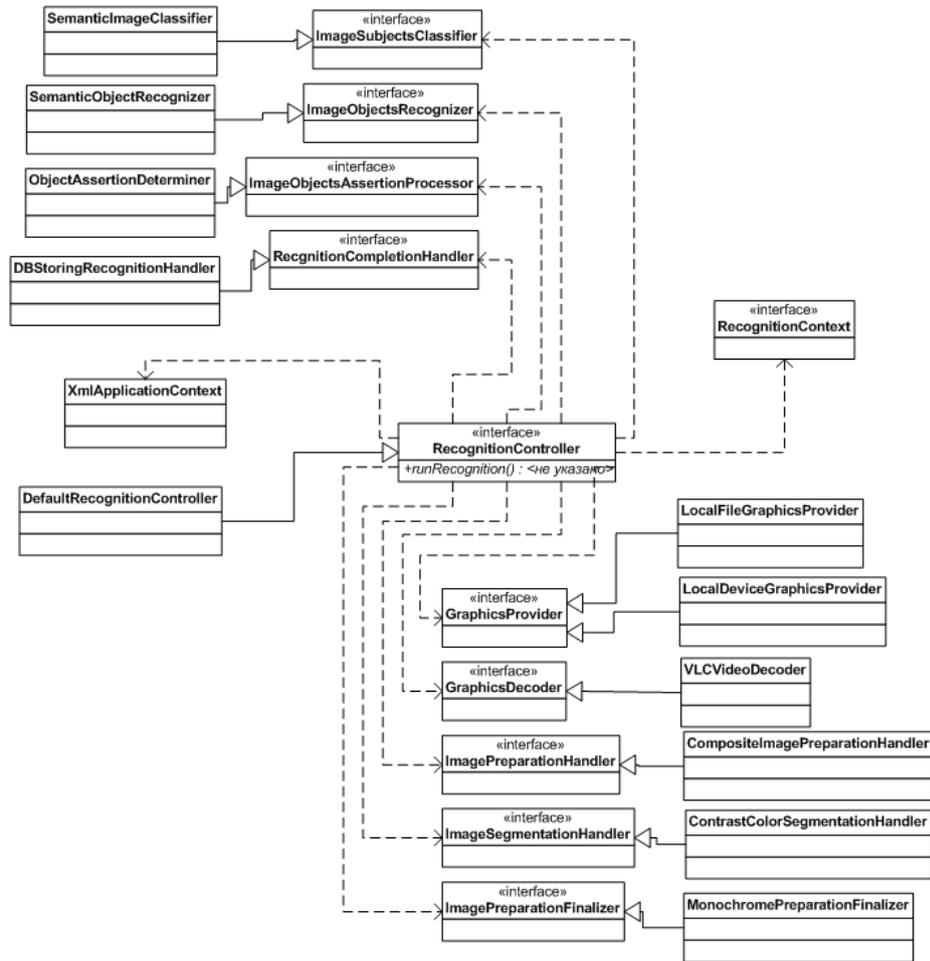


Рисунок 3.8 – Основной сегмент диаграммы классов – подсистема распознавания

Первый модуль в цепочке распознавания `ImageSubjectClassifier`, используя координаты сегментов, найденных ранее, выполняет разбивку изображения на сетку семантических фрагментов и выполняет их дальнейшую свертку и выбор предметной области для каждого сегмента. `SemanticImageClassifier` является базовой имплементацией интерфейса данного модуля. Он выполняет определение предметной области для объектов сегмента на основе базы знаний, путем произведения семантической свертки изображения. Сегмент изображения разрезается на фрагменты, далее для каждого фрагмента подбирается соответствующий фрагмент из базы данных и выбираются все предметные области, которые связаны с более чем 10% фрагментов сегмента. Информация о подходящих предметных областях будет использована в дальнейшем для распознавания объектов графической сцены.

Дальнейшую обработку изображения графической сцены выполняет модуль `ImageObjectRecognizer`. Данный модуль призван найти на изображении объекты, которые наиболее точно соответствуют предметным областям, определенным на предыдущем шаге. `SemanticObjectRecognizer` является базовой имплементацией `ImageObjectRecognizer`. Данный класс является программной реализацией принципов семантической свертки объектов графической сцены. `SemanticObjectRecognizer` выполняет разделение изображения на фрагменты, после этого, для сегментов, которые представляют наибольший интерес с точки зрения многообразия подходящих предметных областей запускается семантическая свертка изображения.

После того как этап поиска объектов завершен в работу включается обработчик последнего этапа фазы распознавания `ImageObjectAssertionProcessor`. Цель работы данного обработчика – уточнить результаты распознавания и выбрать один объект для каждого многообразия объектов, найденных на предыдущем этапе. Базовая имплементация данного процессора `ObjectAssertionDeterminer` выполняет уточнение результатов работы семантической системы распознавания на основании признаковых методов. Данный обработчик получает данные результатов распознавания с предыдущих этапов, а именно:

- исходное изображение;
- результаты сегментации изображения;
- предметные области объектов изображения, которые подходят под изображение сцены, исходя из данных базы знаний;
- информация о фрагментах объектов, которые были определены механизмом семантического распознавания объектов.

На основании собранных ранее данных, `ObjectAssertionDeterminer` выполняет выбор из базы знаний объектов, которые наиболее точно соответствуют фрагментам графической сцены.

Последний обработчик `RecognitionCompletionHandler` завершает этап распознавания. Пользователь может определить любые необходимые операции

для этого обработчика. Базовая имплементация данного обработчика `DBStoringRecognitionHandler` выполняет сохранения результатов распознавания в контексте распознавания, для вывода результатов пользователю. Кроме того, выполняется корректировка информации в базе знаний, на основании распознанных данных. Он выполняет обновление информации о предметных областях и фрагментах объектов, которые могут принадлежать тому или иному объекту. Таким образом, достигается изменчивость, приспособляемость и обучаемость разработанного программного комплекса.

Важную роль процесса распознавания составляет понятие контекста распознавания. Данный объект хранит всю информацию, получаемую в процессе распознавания, не только о текущей распознаваемой графической сцене, но и о всех графических сценах обработанных ранее, что позволяет строить контекстно-зависимые связи между объектами встречающимися на графических сценах разнесенных во времени. В разработанном программном комплексе контекст распознавания описывается интерфейсом `RecognitionContext`. Диаграмма классов контекста распознавания представлена на рисунке 3.9.

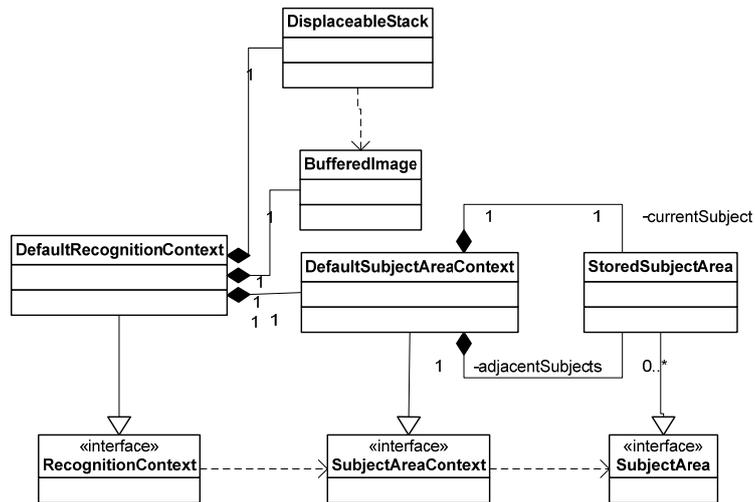


Рисунок 3.9 – Диаграмма классов контекста распознавания

Базовая реализация контекста распознавания хранит информацию о текущем изображении, а также об изображениях, распознанных за предыдущий отрезок времени. Данные изображения сохраняются в вытесняемом стеке `DisplaceableStack`, который разработан специально для построения данного

программного комплекса. Описанный стек хранит информацию об изображении до тех пор, пока информация об изображении представляет ценность для распознавания последующих кадров. Временной фрейм может быть настроен для каждой задачи распознавания в отдельности. Кроме того контекст распознавания хранит информацию о предметных областях, которые имеют отношение процессу распознавания. Информация о предметных областях сохраняется в контексте предметных областей `SubjectAreaContext`, точнее в его базовой имплементации `DefaultSubjectAreaContext`. Данный объект хранит информацию о предметной области, которая представляется наиболее релевантной объектам распознаваемой графической сцены, а также о всех смежных предметных областях. Информация о предметных областях загружается заранее из базы знаний и используется в процессе распознавания. Каждый узел данных в базе знаний имеет связь с одной или несколькими предметными областями, что позволяет представлять многообразие связей семантических данных.

В разработанном программном комплексе использует платформа Spring IOC, которая позволяет облегчить конфигурацию приложения. `XmlApplicationContext` является одним из вариантов имплементации контекста приложения, предоставляемых Spring IOC. Диаграмма классов модели и сервисов представлена на рисунке 3.10.

Основу модели данных приложения составляют `Node` и `SubjectArea`, которые являются гибкими структурами и могут представлять древообразные структуры с N-уровнем вложенности. `SubjectArea` и его реализация `StoredSubjectArea` хранят информацию о предметных областях. `Node` и его реализация `DataNode` хранят информацию о всех данных базы знаний необходимой для работы механизма семантического распознавания графических сцен. Такое положение вещей обеспечивается за счет того, что данных по большей части хранятся не в узлах графоподобной базы знаний, а в связях между узлами. Каждый узел хранит небольшой фрагмент изображения и

ссылки на узлы, которые входят в его состав, если такие узлы существуют, а также ссылки на узлы, в состав фрагментов изображения которых они входят.

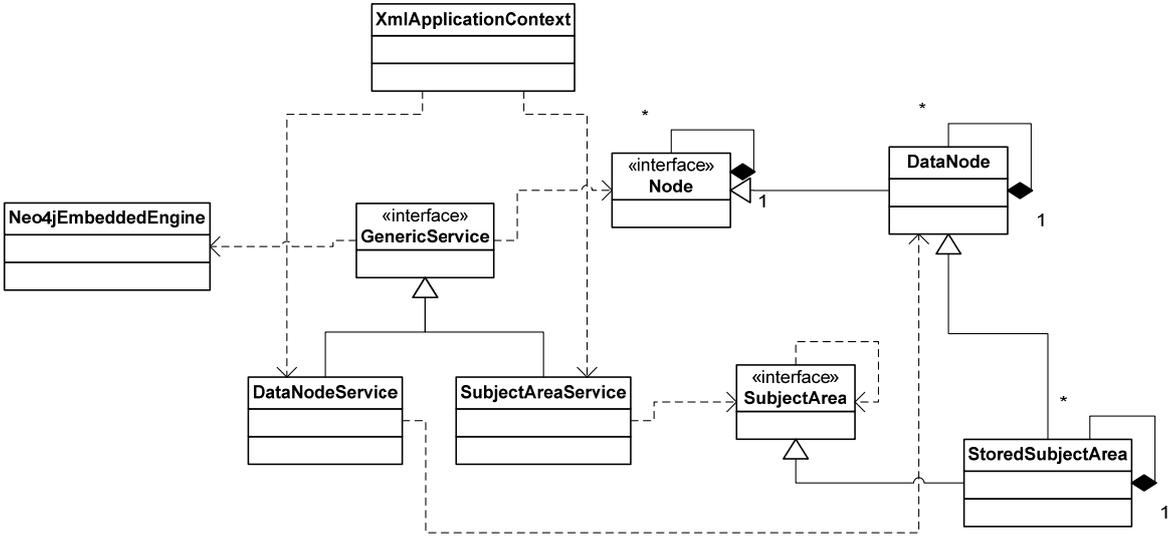


Рисунок 3.10 – Диаграмма классов модели и сервисов приложения

3.3.3 Графический интерфейс и руководство пользователя

Не смотря на то, что базовая поставляема конфигурация разработанного программного – подключаемая библиотека, для демонстрации возможностей разработанного продукта был разработан графический интерфейс приложения.

Главное окно программы содержит информацию о найденных объектах, а также об открытых источниках видео. Основную часть окна занимает панель для отображения потока графических сцен. Главное окно программы отображено на рисунке 3.11.

Главное окно программы содержит основное управляющее меню программы, управляющие панели для: управления источниками графических сцен, отображения текущего активного потока графической информации, отображения лога найденных объектов.

Главное меню программы содержит четыре пункта, которые разделяют основные функции программы.

Меню «Файл» содержит подпункты:

- открыть:

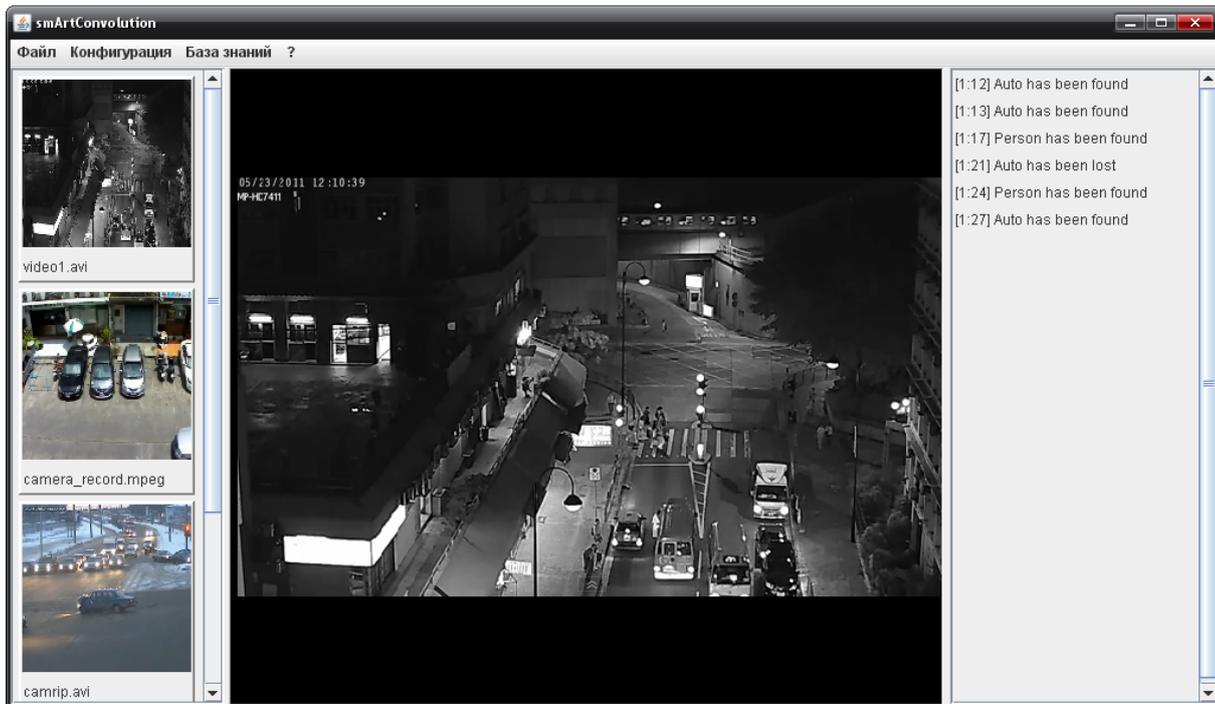


Рисунок 3.11 – Главное окно программы

- локальный файл – позволяет открыть источник графических данных из файла графической информации, размещенного на локальном жестком диске или переносном постоянном запоминающем устройстве;

- локальное устройство – локальное или удаленное устройство, которое смонтировано в системе, как локальное устройство для получения графической информации;

- приостановить поток данных – останавливает чтение графической информации из текущего активного потока данных;

- перезагрузить поток данных – перезапускает чтение графической информации из текущего активного потока данных; для локального файла происходит перезапуск чтения файла, для локального устройства выполняется перепоключение к устройству считывания графической информации;

- закрыть поток данных – закрывает активный поток данных и освобождает все связанные с ним ресурсы;
- закрыть все – закрывает все открытые потоки данных и освобождает все ресурсы, связанные с графическими данными;
- выход – освобождает ресурсы и закрывает окно программы.

Вид меню «Файл» отображен на рисунке 3.12.

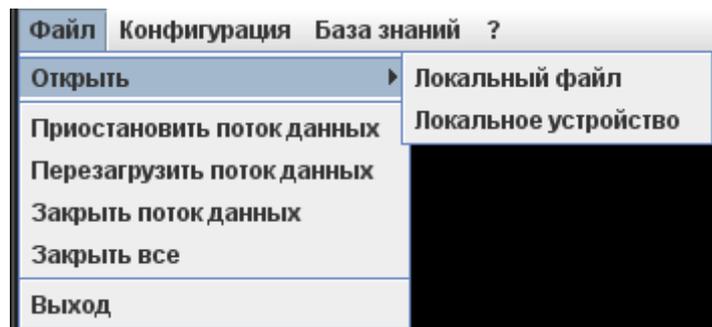


Рисунок 3.12 – Меню «Файл»

Меню «Конфигурация» содержит подпункты позволяющие произвести настройку переменных для алгоритмов распознавания:

- фильтры:
 - предварительной подготовки – открывает диалоговое окно для настройки типов фильтров, которые должны быть применены в цепочке фильтрации, а также для настройки констант фильтрации, как то: интенсивность фильтрации, размеры масок, коэффициенты гауссиана, и т.п;
 - заключительной подготовки – открывает диалоговое окно для настройки коэффициентов и констант цепочки процессоров окончательной обработки графической сцены. В базовой сборке программного комплекса данный пункт меню открывает диалоговое окно для установки пороговых значений, а также коэффициентов контрастности и яркости для модуля обесцвечивания и настройки контрастности изображения;
 - коэффициент модуля сегментации – данный пункт меню открывает диалоговое окно для настройки переменных модуля сегментации. В базовой сборке для сегментации изображений графических сцен используется модуль сегментации на основании алгоритма диаграмм яркости, который

разделяет изображение на сегменты, исходя из показателей соседства. Данное конфигурационное модальное окно позволяет настроить: минимальные и максимальные размеры кластеров, а также допустимые отклонения по цветовым и яркостным показателям сегментов;

о коэффициент семантического процессора – данный пункт меню позволяет открыть диалоговое окно для конфигурации переменных семантического процессора. Открываемое диалоговое окно позволяет настраивать составляющие модули семантического процессора, а также их параметры. Для постпроцессора уточнения распознанных объектов допускается настраивать точность совпадения с эталонными образцами объектов по цвето-яркостным и размерным характеристикам.

Вид меню «Конфигурация» отображен на рисунке 3.13.

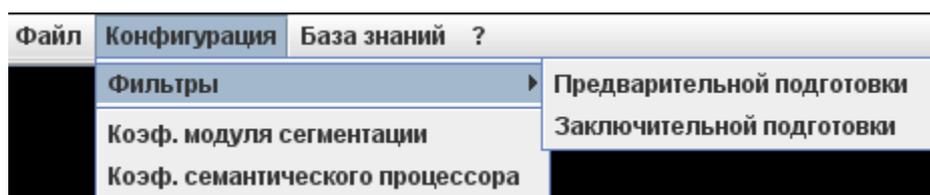


Рисунок 3.13 – Меню «Конфигурация»

Меню «База знаний» обеспечивает доступ к управлению одной из наиболее важных подсистем, разработанного программного комплекса, - подсистемы управления базой знаний. Данный пункт меню содержит следующие подпункты:

- открыть граф базы – данный подпункт открывает окно, позволяющее получить прямой доступ к чтению данных из базы знания Neo4j. В открываемом модальном окне может быть просмотрено содержимое базы данных в виде графа. Загрузка данных из базы производится постепенно, по мере необходимости отображения графических данных. Таким образом, изначально отображается коревой (нулевой) узел и его соседи, активным узлом также является корневой узел. Далее при активации любого другого узла, загружается детальная информация о нем и о связанных с ним узлах;

- редактор запросов открывает простое диалоговое окно, которое позволяет делать запросы к базе данных Neo4j на SQL-подобном языке, который позволяет описывать правила обхода графа и выбора данных из базы. Фактически открываемое диалоговое окно является браузером для базы данных приложения;

- предметные области – данный пункт меню открывает диалоговое окно для просмотра и ввода в базу знаний системы информации о предметных областях и связях между ними;

- объекты – этот пункт меню позволяет открывать диалоговое окно для просмотра и ввода в базу знаний системы информации об объектах, связях между ними, а также их отношениях к предметным областям. Необходимо отметить, что графическая составляющая привязывается к объектам путем обучения программного комплекса, а не в ручном режиме через открываемое данным пунктом меню диалоговое окно;

- обучение:

- коррекция – данный пункт меню открывает диалоговое окно для чтения предметных областей, объектов и связанных с ними фрагментов изображений, которые на данный момент хранятся в базе знаний. Открываемое диалоговое окно позволяет просматривать содержимое базы знаний и выполнять коррекцию базы знаний системы, путем изменения и удаления ошибочных фрагментов изображений в базе знаний, а также путем корректировки связи между фрагментами, относительно предметных областей и объектов;

- полуавтоматическое обучение – данный пункт меню открывает специализированное диалоговое окно для запуска процесса анализа и заполнения базы знаний в полуавтоматическом режиме. Заполнение базы знаний данными и будет представлять собой обучение программного комплекса. Для выполнения заполнения базы знаний этого пользователь будет обязан задать источник графических сцен;

о пункт меню «Конфигурация обучения» позволяет открыть диалоговое окно, в котором можно настроить параметры и коэффициенты обучения семантических сетей, включить или выключить некоторые этапы обучения, настроить коэффициенты для механизмов сегментации и процессора семантического разбора. Кроме того, данное диалоговое окно позволяет отфильтровать, какие предметные области и объекты будут доступны для работы в процессе обучения.

Вид меню «База знаний» отображен на рисунке 3.14.

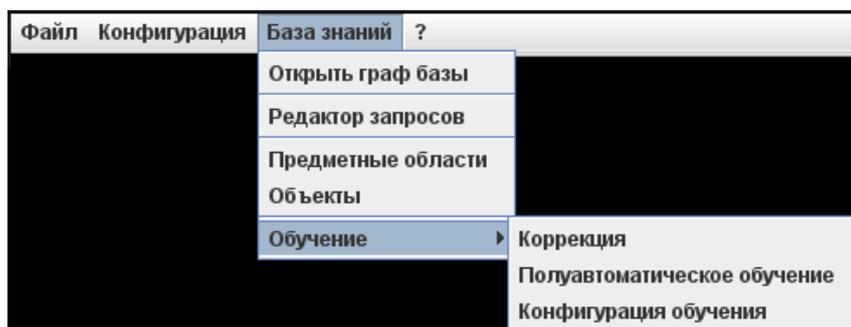


Рисунок 3.14 – Меню «Конфигурация»

Последний пункт меню «Помощь» содержит подпункты для доступа к помощи и общей информации о системе:

- помощь – данный пункт меню открывает диалоговое окно, которое содержит HTML документ с информацией о программе, возможностями работы, как выполнить то, или иное действие, как выполнить определенные настройки, как запустить обучение базы знаний и тому подобное;
- рекомендации по настройке – данный пункт меню диалоговое открывает окно, которое содержит HTML документ с информацией о том, как более эффективно настроить данный программный комплекс для работы в той или иной предметной области с различными типами видео и прочее;
- лицензия - данный пункт меню открывает диалоговое окно, которое содержит документ с информацией о лицензировании программы;
- о программе – открывает простое диалоговое окно с информацией о программе и ее авторах.

ВЫВОДЫ

Факторизация информации (например, в задачах распознавания образов) в том или ином признаковом пространстве концептуально является одним из основных подходов, лежащих в основе и обработки, и интерпретации данных. Применение классов эквивалентностей, выступающих в качестве объектов обобщения–детализации, приводит к необходимости решения целого класса задач. В первую очередь следует указать поиск и анализ внутренних эквивалентностей, установление правомочности (согласованности) использования совокупности эквивалентностей в том числе внешних, введение которых диктуется предметной областью, наконец, разумное определение контекстно-зависимых операций на классах эквивалентностей. В этом плане предложенные результаты являются основой для изучения групповых свойств, возникающих на разбиениях множеств произвольной природы.

В результате экспериментальных исследований показана эффективность предложенных моделей мультигрупповой факторизации данных. Рассмотрена зависимость метрического различия фактор-множеств при различных вариантах их анализа.

Двойная – во времени и пространстве – сегментация видеорядов создает предпосылки для выделения множеств изображений, на которых заданы мультигруппы, что дает возможность определять их параметры.

Получил дальнейшее развитие корреляционный метод нормализации следящего типа для перспективы на основе функционалов (2.7), (2.8), учитывающих перераспределение яркости изображения вследствие действия проективных преобразований. Преимуществом метода является возможность сведения проективно искаженного изображения к искажениям аффинной группы, для которой известны методы нормализации.

Получила дальнейшее развитие модель последовательной нормализации для проективной группы преобразований, которая основана на применении инвариантного к аффинной группе отображения структурного или

интегрального типа и нормализаторе для перспективы. В качестве примера сопоставляемых признаков использованы множества аффинных инвариантов и интегральные функционалы на базе моментов.

Применение инвариантов интегрального типа позволяет предположить наличие искажений перспективы и может быть использовано в качестве критерия их присутствия. Инвариантные описания структурного типа позволяет получить набор соответствующих точек на эталонном и входном изображениях, что дает возможность использовать их для нормализации проективной группы.

Усовершенствована модель нормализации, основанная на нахождении параметров проективного преобразования по четырем точкам. Предложена модификация метода нахождения структурных инвариантов с использованием SIFT, суть которой заключается в поиске точек на границе объекта, подлежащего нормализации, а не внутри его.

Предложены функционалы (2.28), отражающие распределение энергии функции яркости. Проведено исследование влияния геометрических искажений перспективы на изменение значений синтезированных признаков.

Разработанный программный комплекс прошел предварительные испытания на реальных данных и доказал, что распознавание, как способность обнаруживать в потоке информации, поступающей от органов чувств, признаки определённые объекты, закономерности и явления, доступна не только живым организмам.

Так как возможность распознавания опирается на схожесть однотипных объектов, то наиболее удачной является стратегия распознавания объектов на основании определенной базы знаний, а не на основании идеальных математических моделей, которые, тем не менее, не имеют связи с реальностью. Однако такой подход требует представления и хранения огромного количества данных, в подходящей форме, с чем с успехом справляется графоподобная база данных Neo4j. Применение легковесной платформы для манипулирования данными, хранящимися в виде

направленного графа, позволило создать динамическую семантическую нейронную сеть, что позволило сделать важный шаг в построении систем распознавания, конфигурируемых и независящих от предметных областей.

Несмотря на то, что выбранный подход требует значительной мощности вычислительных станций, современное развитие процессоров позволяет выполнять количество вычислительных операций в секунду, достаточное для построения эффективных систем распознавания. Кроме того, одним из важнейших свойств разработанной динамической семантической нейронной сети является возможность деления сети на множество сегментов и распараллеливания вычислительных процессов на множество математических процессов. Поэтому одной из важнейших задач, которые ставятся в качестве возможных вариантов развития данного программного комплекса, – разработка подключаемого модуля, который бы позволил использовать вычислительные процессоры GPU, которые на данный момент включаются сотнями в видеокарты.

Благодаря гибкости и независимости от предметной области, разработанный программный комплекс потенциально может быть использован для автоматизации процессов в разнообразнейших областях: медицина, обеспечении безопасности, контроле технологических процессов, автоматическом управлении. Однако, несмотря на значительные перспективы, применение комплекса потребует значительных затрат на первоначальном этапе внедрения, поскольку будет необходимо провести обучение системы и наполнить базу знаний программно-аппаратного комплекса данными, актуальными для предметной области и необходимыми для работы комплекса. Это потенциально позволит автоматизировать труд во множестве областей и снизить затраты в различных видах человеческой деятельности.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Bargiela A. Toward a theory of granular computing for human-centered information processing / A. Bargiela W. Pedrycz // IEEE Transactions on Fuzzy Systems. – 2008. – Vol. 16, No. 2. – P. 320-330.
2. Lin T.Y. Granular computing I: the concept of granulation and its formal model / T.Y. Lin // International Journal of Granular Computing, Rough Sets and Intelligent Systems. – 2009. – Vol. 1, No. 1. – P. 21-42.
3. Yao Y.Y. Granular computing: past, present, and future / Y.Y. Yao // Rough Sets and Knowledge Technology / G. Wang, T. Li, Jerzy, W. Grzymala-Busse, D. Miao, A. Skowron, Y.Yao (eds.). – Lecture Notes in Artificial Intelligence. – Berlin, Heidelberg: Springer-Verlag. – 2008. – Vol. 5009. – P. 27-28.
4. Zadeh L.A. Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems / L.A. Zadeh // Soft Computing. – 1998. – Vol. 2, No.1. – P. 23-25.
5. Zadeh L.A. Toward a logic of perceptions based on fuzzy logic / L.A. Zadeh // Discovering the World With Fuzzy Logic / V. Novák and I. Perfilieva (eds.) // Series: Studies in Fuzziness and Soft Computing. – Heidelberg: Physica-Verlag. – 2000. – P. 4-28.
6. Фор А. Восприятие и распознавание образов / А. Фор. // М.: Машиностроение, – 1989. – 272 с.
7. Demri S.P. Incomplete information: structure, inference, complexity / S.P. Demri, E.S. Orłowska // Series: Monographs in Theoretical Computer Science. – Berlin, Heidelberg: Springer-Verlag. – 2002. – 408 p.
8. Шлезингер М.И. Десять лекций по статистическому и структурному распознаванию / М.И. Шлезингер, В. Главач. – К.: Наук. думка, 2004. – 546 с.
9. Луганский А.М. Связь отношений эквивалентности и толерантности в метатекстовых информационно–аналитических системах / А.М.

- Луганский, В.П. Машталир, В.В. Шляхов // Радиоэлектроника и информатика. – 2004. – №1. – С. 100–103.
10. Zhang L. Fuzzy reasoning model under quotient space structure / L. Zhang, B. Zhang // Information Sciences – 2005. – Vol. 173, No. 3. – P. 353-364.
 11. Bargiela A. Granular computing: an introduction / A. Bargiela, W. Pedrycz. The Kluwer International Series in Engineering and Computer Science. – Boston, Dordrecht, London: Kluwer Academic Publishers. – 2002. – Vol. 717. – 478 p.
 12. Skowron A. Stepaniuk J. Information granules: towards foundations of granular computing / A. Skowron, J. Stepaniuk // International Journal of Intelligent Systems. – 2001. – Vol. 16, No. 1. – P. 57-85.
 13. Kreinovich V. Interval computation as an important part of granular computing: an introduction / V. Kreinovich // Handbook of granular computing / W. Pedrycz, A. Skowron, V. Kreinovich (eds.) Chichester: John Wiley and Sons Ltd. – 2008. – P. 3-32.
 14. Düntsch I. Rough set data analysis: a road to non-invasive knowledge discovery / I. Düntsch, G. Gediga // Primers Series Vol. 2. – Bangor: Methodos Publishers. – 2000. – 107 p.
 15. Pawlak Z. Rough sets: theoretical aspects of reasoning about data / Z. Pawlak. – Dordrecht: Kluwer Academic Publishers. – 1991. – 250 p.
 16. Каграманян А.Г. Метрические свойства грануляции информации / А.Г. Каграманян, В.П. Машталир, В.В. Шляхов // Бионика интеллекта. – 2007. – №1 (66). – С. 17-24.
 17. Yao Y.Y. Granular computing for data mining / Y.Y. Yao // Proceedings of SPIE Conference on Data Mining, Intrusion Detection, Information Assurance and Data Networks Security / B.V. Dasarathy (ed.). – Kissimmee. – 2006. – Vol. 6241 (624105). – 12 p.
 18. Ивахненко А.Г. Применение самоорганизации для разбиения множества данных на заранее не заданное число кластеров / А.Г. Ивахненко, Ю.В.

- Коппа, С.А. Петухова, М.А. Ивахненко // Автоматика. – 1985. – №5. – С. 93–123.
19. Jain A.K. Algorithms for clustering data / Jain A.K., Dubes R.C. – NJ: Prentice-Hall, Inc. – 1988. – 320 p.
20. Hansen E. Global optimization using interval analysis / E. Hansen, G.W. Walster. – NY: Marcel Dekker, Inc. – 1992. – 728 p.
21. Moore R.E. Introduction to interval analysis / R.E. Moore, R.B. Kearfott, M.J. Cloud. – Philadelphia: SIAM Press. – 2009. – 235 p.
22. Стоян Ю.Г. Выпуклые интервальные многоугольники / Ю.Г. Стоян // Доповіді НАН України. – 2000. – № 5. – С. 33-39.
23. Peters J.F, Skowron A, Synak P, Ramanna S, Rough sets and information granulation // Fuzzy Sets and Systems / T. Bilgiç, B. De Baets, O. Kaynak (eds.). – Lecture Notes in Artificial Intelligence. – Berlin, Heidelberg: Springer-Verlag. – 2003. – Vol. 2715. – P. 370-377.
24. Pedrycz W. Granular clustering: a granular signature of data / W. Pedrycz, A. Bargiela // IEEE Transactions on Systems, Man and Cybernetics. – Part B-Cybernetics. – 2002. – Vol. 32, No. 2. – P. 212-224.
25. Заде Л.А. Размытые множества и их применение в распознавании образов и кластер-анализе / Заде Л.А. // Классификация и кластер / Д.В. Райзин (ред.). – М: Мир.– 1980. – С. 208-247.
26. Машталир В.П. Обратимость логических матриц как свойство многозначности обработки информации / В.П. Машталир, В.В. Шляхов // Вестник ХГПУ (Новые решения в современных технологиях). – Харьков: ХГПУ. – 2000. – Вып. 119. – С. 12-15.
27. Машталир В.П. Свойства мультиалгебраических систем в задачах компаративного распознавания / В.П. Машталир, В.В. Шляхов // Кибернетика и системный анализ. – Киев – 2003. – №6. – С. 11-21.
28. Машталир В.П. Индуцированная согласованность отношений в задачах грануляции информации / В.П. Машталир, В.В. Шляхов // Бионика интеллекта. – 2006. – №1 (64). – С. 19-26.

29. Шляхов В.В. Характеристические свойства мультиалгебраических систем / В.В. Шляхов, С.В. Яковлев // Доповіді НАН України. – 2001. – №10. – С. 72-76.
30. Бондаренко М.Ф. Практические приложения компараторной идентификации линейных конечномерных объектов / М.Ф. Бондаренко, С.Ю. Шабанов-Кушнарченко, Ю.П. Шабанов-Кушнарченко // Бионика интеллекта. – 2009. – № 2(71). – С. 3-12.
31. Путятин Е.П. Обработка изображений в робототехнике / Е.П. Путятин, С.И. Аверин. – М.: Машиностроение. – 990. – 320 с.
32. Хадвигер Г. Лекции об объеме, площади поверхности и изопериметрии. / Г. Хадвигер. – М.: Наука, 1966. – 416 с.
33. Distance measures for image segmentation evaluation / X. Jiang, C. Marti, C. Irniger, H. Bunke // EURASIP Journal on Applied Signal Processing. – 2006. – Article ID 35909. – 10 p.
34. Image retrieval with segmentation-based query / A. Chupikov, D. Kinoshenko, V. Mashtalir, K. Shcherbinin // Adaptive Multimedia Retrieval: User, Context and Feedback / S. Marchand-Maillet et al. (eds.). – Lecture Notes in Computer Science. – Berlin-Heidelberg: Springer-Verlag. – 2007. – Vol. 4398. – P. 208-222.
35. Montgomery D.C. Forecasting and time series analysis / D.C. Montgomery, I.A. Johnson, J.S. Yardiner. – N.Y.: McGraw-Hill, 1990. – 394 p.
36. Льюис К.Д. Методы прогнозирования экономических показателей / Льюис К.Д.. – М.: Финансы и статистика, 1986. – 133 с.
37. Чуев Ю.В. Прогнозирование количественных характеристик процессов / Ю.В. Чуев, Ю.Б. Михайлов, В.И. Кузьмин. – М.: Сов. радио, 1975. – 400 с.
38. Любченко В.А. Нормализация перспективных преобразований проективно искаженных изображений / В.А. Любченко, Е.В. Яковлева, Е.О. Передрий // Вестник национального технического университета

- «ХПИ». Системный анализ, управление и информационные технологии. – 2008. – № 26. – С. 174–184.
39. A study of Zernike invariants for content-based image retrieval / P. Toharia, O.D. Robles, A. Rodriguez, L. Pastor // *Advances in Image and Video Technology. Lecture Notes in Computer Science*. – 2007. – P. 944–957.
40. Шапиро Л. Компьютерное зрение: пер. с англ. / Л. Шапиро, Дж. Стокман. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
41. Гороховатский В.А. Корреляционные методы распознавания изображений путем голосования систем фрагментов / В.А. Гороховатский, Е.О. Передрий // 6. *Радіоелектроніка. Інформатика. Управління*. – 2009. – № 1 (20). – С. 74–81.
42. Любченко В.А. Идентификация группы преобразований для систем компьютерного зрения / В.А. Любченко, Е.О. Передрий // *Сучасні інформаційні системи. Проблеми та тенденції розвитку: II міжнар. наук. конф., 2–5 окт. 2007 г.: зб. матер.* – Харків, 2007. – С. 343–344.
43. Путятин Е.П. Исследование инвариантных особенностей перспективных преобразований для задач компьютерного зрения / Е.П. Путятин, В.А. Любченко, Е.О. Передрий // *Информационные компьютерные технологии и системы: междунар. конф., 22–23 окт. 2008 г.: сб. науч. тр.* – Харьков, 2008. – Т. 5. – С. 169–171.
44. Projective methods of image recognition / Y. Putyatin, V. Gorohovatsky, A. Gorohovatsky, E. Peredriy // *Information Science and Computing*. – Sofia, 2008. – No. 5. – P. 37–43.
45. Дуда Р., Харт П., Распознавание образов и анализ сцен [Текст] пер. с англ. – Г.Г. Вайештейнв, А.М. Васьковский, под редакцией В.П. Стефанюка. – 509 с.
46. Вершинина В.В., О нечетком структурном описании визуального образа // *Сборник трудов молодых ученых.* – Рыбинск:РГАТА, 2000. С. 172-174.
47. Белоус Н.В., Борисенко В.П., Пономарев Ю.В. Технология интеграции и интеллектуального анализа данных в корпоративной автоматизированной

- системе для обеспечения надежности и эффективности функционирования объектов газотранспортной системы // Международный научный семинар им. Ю.Н.Руденко МЕТОДИЧЕСКИЕ ВОПРОСЫ ИССЛЕДОВАНИЯ НАДЕЖНОСТИ БОЛЬШИХ СИСТЕМ ЭНЕРГЕТИКИ
48. Антошук С.Г., Нутович А.А., Белоус Н.В. Семантическая обработка видеоинформации при управлении дорожным движением // Материалы международной научно-технической конференции «информационные системы и технологии в энергетике и жилищно-коммунальной сфере».
49. Semantic Networks, John F.Sowa [Электронный ресурс] - Режим доступа: www. URL: <http://www.jfsowa.com/pubs/semnet.htm> - Загл. с экрана.
50. Семантические сети в современном программировании [Текст]: тез. докл. науч.-практ. конф./ Слободяник М. Р., 15 Международный молодежный форум «Радиоэлектроника и молодежь в 21 веке». - Харьков, 2011, с.264-265
51. Исследование и анализ алгоритмов и разработка подхода классификации сложных объектов в реальном времени с использованием нейронных сетей [Текст]: тез. докл. науч.-практ. конф./ Бездельный Н.Ю., 16-й Международный молодежный форум «Радиоэлектроника и молодежь в XXI веке» сб. материалов форума. Т. 6. – Харьков: ХНУРЭ. 2012. –с.191-192
52. Bondarenko M.F., Bilous N.V., Kobzar G.A. itc. Intelligent Data Processing in Global monitoring for Environment and Security//ITHEA, Kiev-Sofia, 2010, First Edition. - ISBN: 978-954-16-0045-0 .
53. Семантическая нейронная сеть – [Электронный ресурс] - Режим доступа: www URL: http://www.itfru.ru/index.php/logika/730-robototehnika_semanticheskie_neironnye_seti_57669 - 14.11.2010г. - Загл. с экрана.
54. Белоус Н.В., Борисенко В.П., Пономарев Ю.В., Борисенко И.В. Методы и алгоритмы обнаружения объектов для системы управления промышленной безопасностью в рамках концепции создания

комплексной АСУ газотранспортной компанией // *Applicable Information Models*, ITHEA, Sofia, 2011

55. Борисенко В.П., Белоус Н.В. Технология комплексного интеллектуального анализа данных в корпоративной системе управления газотранспортной компании // *Актуальные проблемы механики, математики, информатики: сб. тез. науч.-практ. конф. (Пермь, 12–15 октября 2010 г.)* / гл. ред. В.И. Яковлев; Перм. гос. ун-т. – Пермь, 2010. – с. 51.