

А.Ю. КАПЛИН, А.А. КОРОТИН, А.В. НАЗАРОВ, В.Л. ЯКИМОВ
**АЛГОРИТМ КЛАССИФИКАЦИИ И ВОССТАНОВЛЕНИЯ
ИСКАЖЕННЫХ N -МЕРНЫХ ГРУППОВЫХ ТОЧЕЧНЫХ
ОБЪЕКТОВ НА ОСНОВЕ КОМБИНАТОРНОГО ПОИСКА
ФРАГМЕНТОВ**

Каплин А.Ю., Коротин А.А., Назаров А.В., Якимов В.Л. Алгоритм классификации и восстановления искаженных N -мерных групповых точечных объектов на основе комбинаторного поиска фрагментов.

Аннотация. Представлен алгоритм классификации образцов многомерных групповых точечных объектов. Поиск осуществляется на основе комбинаторного поиска соразмерных фрагментов матриц попарных отношений на множестве шаблонов. Решение об отнесении образца к тому или иному шаблону принимается по критерию минимума евклидова расстояния. Представленный подход к распознаванию позволяет синтезировать инвариантные (относительно вращения, масштабирования или смещения системы координат) описания вторичных признаков, а также использовать достаточно мощный инструментарий теории многомерного и метрического шкалирования в компенсации искажений распознанных образов групповых точечных объектов. В алгоритме реализована процедура статистических испытаний Монте-Карло, в рамках которого каждая точка случайным образом размещенная в предполагаемой окрестности искомого координат проверяется по условию минимума квадратического показателя сходства. Приведены пример и результаты использования алгоритма для идентификации и восстановления искаженных и подвергнутых воздействию координатных шумов радиоизображений, представленных выборкой шаблонов «блестящих» точек.

Ключевые слова: групповой точечный объект, классификация, многомерное шкалирование.

1. Введение. Интенсивное развитие радио и оптических локационных систем позволили получать изображения объектов локации с высоким разрешением. Так в широкополосных радиолокационных системах с синтезированной апертурой антенны появилась возможность получать радиоизображения с разрешением $0,25, \dots, 1$ м. [1-3]. Увеличение степени детализации изображений и экспоненциальный рост информативности контекстных признаков затрудняет реализацию методов вторичной обработки в масштабе времени, близком к реальному.

При локационном наблюдении удаленных объектов отраженный сигнал изменяется в большом динамическом диапазоне. На объекте существуют отдельные точки и области, создающие сильное отраженное излучение, намного превышающее сигнал от соседних областей — «блестящие» точки. Их совокупность и взаимное расположение позволяют эффективно распознавать объекты и определять параметры их движения относительно центра масс.

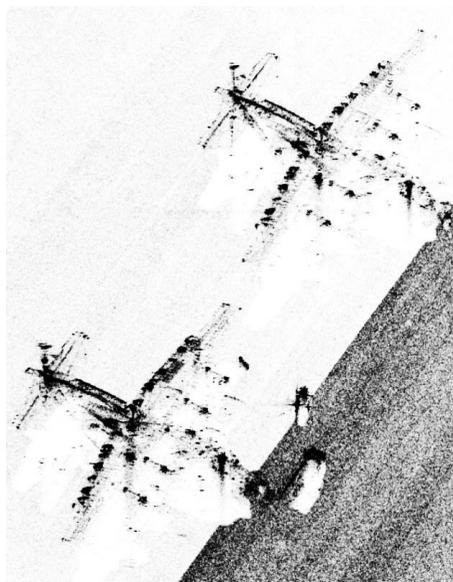
Особенностью функционирования средств формирования радиоизображений является обязательное использование не всегда до-

ступной априорной информации для пространственного сглаживания яркостных шумов. Ситуация характерна для радиолокационных изображений со спекл-шумами, где дисперсия шума достаточно высока, что приводит к сильному маскированию «полезных» точечных отражателей на радиолокационном изображении. Это усложняет обнаружение точечных и рассредоточенных объектов, а задачу их классификации, например, по одиночному изображению без учета контекстных признаков делает практически невозможной [4]. В условиях воздействия помех это в конечном итоге приводит к снижению разрешающей способности, точности определения координат и снижению оперативности и достоверности функционирования каналов получения некоординатной информации. Одним из подходов к решению этой проблемы является разработка алгоритмов параллельной автоматической идентификации групп сосредоточенных целей и восстановления радиолокационных изображений объектов, представленных групповыми точечными объектами (ГТО).

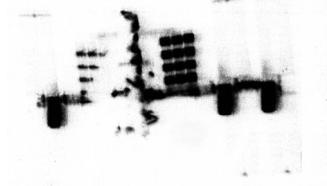
В силу ограниченности разрешающей способности датчиков в зависимости от условий наблюдения ГТО представляют часто встречающийся на практике класс изображений. Хотя отдельную точку нельзя охарактеризовать формой, закономерности расположения точек в составе ГТО дают возможность рассматривать его как некоторый образ и говорить о его форме как о форме некоторого сплошного изображения, ассоциированного с ГТО.

Для последнего времени характерен постоянно возрастающий интерес к многомерным ГТО [5, 6]. Например, учет яркостной составляющей при идентификации звезды по яркостному портрету окружающих ее звезд в пределах машинного кадра астродатчика или при обеспечении радиолокационного наблюдения порядка групповых точечных целей позволяет значительно повысить достоверность распознавания [7]. Введение дополнительных размерностей усиливает степень ортогональности зашумленных ГТО разных классов [8, 9]. В конечном счете, это приводит к росту достоверности результатов классификации в координатных или спекл-шумах, но связано со значительным увеличением объема требуемых вычислений.

Помимо естественного представления групп радиолокационных отметок радиолокационных систем или «ярких» точек квантово-оптических систем (рисунок 1, а-б), в виде ГТО можно параметризовать широкий класс объектов различной физической природы.



T=2008-04-03 16:37:18,066
 L(km) =413,483
 A(grad)=265,8931
 H(grad)=55,9728



T=2008-04-03 16:37:20,963
 L(km) =425,164
 A(grad)=265,4007
 H(grad)=53,5314

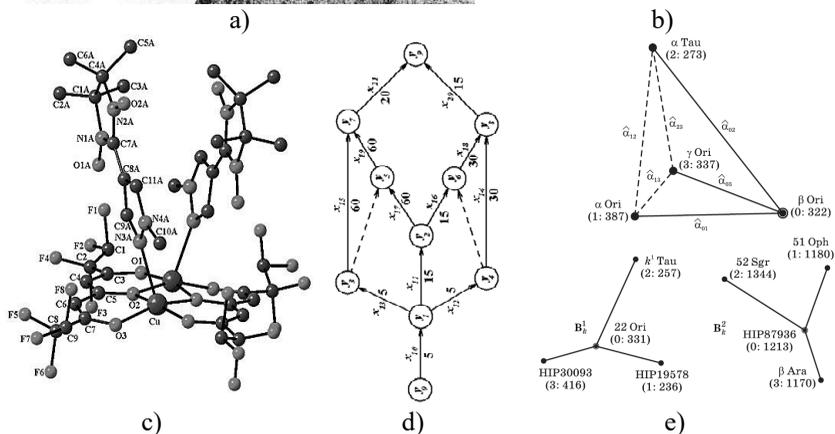
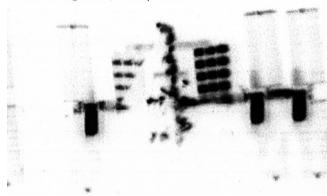


Рис.1. Примеры классифицируемых объектов, представимых в классе ГТО: *a)* радиоизображение [10], *b)* изображение МКС в ИК-диапазоне, *c)* сложное соединение, *d)* логико-вероятностная модель устойчивости, *e)* схема конфигурации звезд

Для этого выбирается m измеряемых характеристик явления или объекта, подлежащего классификации. Образец представляется как группа k точек в m -мерном пространстве классификационных призна-

ков. Классифицируемый процесс может быть представлен графом переходов состояний с уникальным количеством вершин и множеством длин переходов, схемой сложного соединения или схемой точечных особенностей сцены (рисунок 1, с-е).

Во всех случаях информация, позволяющая достоверно сличить входной образец (процесс) с одним из представителей классов (шаблонов) содержится в уникальном взаимном расположении k точек (вершин графов) и характеристиках самих точек ГТО.

Далее осуществляется преобразование в ту или иную систему признаков, обеспечивающую надежное распознавание, например, за счет использования сходства или различия вычисленных в заданной метрике. Заключительной операцией, которую должна выполнять любая система машинного видения, от одномерной системы технического зрения до систем анализа процессов, представимых графами локации в радио- или оптическом диапазоне, является распознавание класса объекта (процесса) и восстановление искаженных элементов распознанного образца.

2. Обоснование и особенности алгоритма. Алгоритм относится к классу метрических методов распознавания с эталонами. Критерием принятия классификационного решения является минимум расстояния между треугольными полуматрицами смежности (межэлементных расстояний) классифицируемого образца и фрагмента шаблона в заданной метрике.

Отличительными особенностями представленного алгоритма от подходов, описанных в источниках [4-18] являются:

1. Переход от пространства истинных координат точек ГТО в пространство его межэлементных (межточечных) расстояний с последующей нормировкой расстояний в диапазон $[0, 1]$. Иными словами

m -мерный ГТО представляется n -мерным объектом, где $n = \frac{N!}{2(n-2)!}$,

N – исходное количество точек ГТО, что позволяет синтезировать инвариантные (например, относительно вращения, масштабирования или смещения системы координат) описания признаков.

2. Отсутствие ограничений на мерность m представления ГТО. Дополнительные свойства точек ГТО представляются в метрических шкалах дополнительных ортогональных проекций. Мерность n пространства, в котором вычисляются расстояния между ГТО зависит только от количества N точек в образце ГТО.

3. Метрика при вычислении расстояний между точками ГТО — есть в общем случае метрика Минковского со степенью q :

$$d_{ij} = \left(\sum_{r=1}^n |y_{ir} - y_{jr}|^p \right)^{\frac{q}{p}}, \quad (1)$$

где n — размерность нормированного пространства межэлементных расстояний ГТО, i — индекс образца, представленного к распознаванию, j — индекс шаблона, r — индекс размерности, на практике применяется метрика «*block city*» при $p = 1$: $d_{ij} = \sum_r |y_{ir} - y_{jr}|$, или евкли-

дова метрика при $p = 2$: $d_{ij} = \left(\sum_r (y_{ir} - y_{jr})^2 \right)^{\frac{1}{2}}$, в которой расстояния

не зависят от направления осей, в результате чего возможны любые вращения осей координат. Так как пространство признаков ГТО может быть анизотропно, то можно уточнить:

$$d_{ij} = \left(\sum_{r=1}^n \lambda_{rs}^2 (y_{ir} - y_{jr})^2 \right)^{\frac{1}{2}}, \forall s: \sum_{r=1}^n \lambda_{rs}^2 = 1, s = \overline{1, L}, \text{ где } L \text{ — количество шаб-$$

лонов (мощность алфавита классов), при этом $\lambda_{rs} = \frac{1}{\sigma_{rs}}$, где σ_{rs} —

среднеквадратическое отклонение признака (межэлементного расстояния y_r) для образцов, отнесенных к шаблону s . Величина λ_{rs} при этом имеет ясный физический смысл: чем меньше рассеяние признака по шаблонам, тем больше его классифицирующее значение (ценность). Каждый образец полуматрицы расстояний путем последовательной конкатенацией строк преобразовывается в вектор. При выборе показателя степени в равенстве (1) для вычисления расстояний между точками ГТО следует учитывать, что увеличение p приводит к возрастанию роли наибольших расстояний. Выбор меры расстояния с большой степенью q дает резкое выделение шаблонов с наименьшими расстояниями. Обоснованный выбор указанных величин можно сделать с помощью практического опыта в определенных классах задач. Сказанное относится и к выбору весовых коэффициентов λ_{rs} .

4. Требования инвариантности к изменениям масштаба сравниваемых при классификации фрагментов ГТО обуславливают необходимость нормировки компонент полуматриц межэлементных рас-

стояний образцов X' и соразмерных фрагментов шаблонов X . Приведение данных к единичному масштабу $[0, 1]$ обеспечивается нормировкой каждого значения переменных на диапазон разброса их значений. В простейшем варианте — это линейное преобразование

$$x_i^{norm} = \frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}}(b - a) + a, \text{ в единичный отрезок: } x_i^{norm} \in [0, 1], \text{ где}$$

a и b — границы нормированного диапазона ($a=0, b=1$). Обратная линейная денормировка осуществляется в соответствии с преобразованием:

$$x_i = \frac{x_i^{norm} - a}{b - a}(x_{i,\max} - x_{i,\min}) + x_{i,\min}.$$

Линейная нормировка оптимальна, когда значения переменной x_i плотно заполняют определенный интервал. Если в данных имеются относительно редкие выбросы, намного превышающие типичный разброс, именно эти выбросы определяют согласно предыдущей формуле масштаб нормировки. Это приведет к тому, что основная масса значений нормированной переменной x_i^{norm} сосредоточится вблизи нуля (рисунок 2, а) [12].

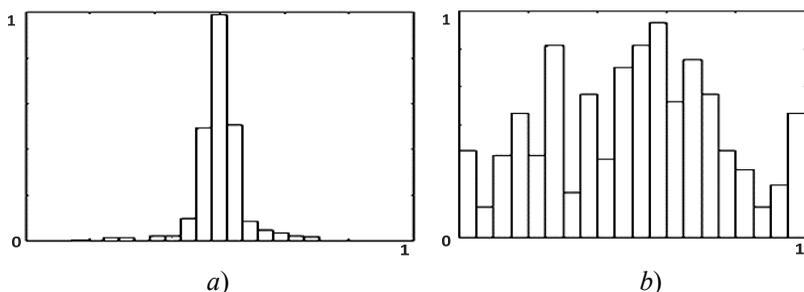


Рис. 2. Гистограмма: а) значений переменной при наличии редких, но больших по амплитуде отклонений от среднего, б) после нелинейной нормировки, использующей логистическую функцию

При нормировке следует ориентироваться не на экстремальные значения, а на типичные, т.е. статистические характеристики данных, такие как среднее и дисперсия. В этом случае основная масса данных будет иметь единичный масштаб, т.е. типичные значения всех переменных будут сравнимы (рисунок 2, б) [19]. Как видно из рисунка, распределение значений после такого нелинейного преобразования гораздо ближе к равномерному. Преобразование $x_i \rightarrow x_i^{norm}, i = \overline{1, n}$ в

диапазон $[0, 1]$ осуществляется в соответствии с выражениями, содержащими функцию нелинейного преобразования:

$$x_i^{norm} = \frac{1}{1 + e^{-\left(\frac{x_i - x_i^*}{\sigma_i}\right)}}, x_i^* = \frac{1}{n} \sum_{i=1}^n x_i, \sigma_i^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - x_i^*)^2.$$

В ряде модификаций классификационных моделей ГТО нелинейная нормировка данных $x_i \rightarrow x_i^{norm}, i = \overline{1, n}$ может быть осуществлена в диапазон $[-1, 1]$ в соответствии с выражениями:

$$x_i^{norm} = \frac{2}{1 + e^{-\left(\frac{x_i - x_i^*}{\sigma_i}\right)}} - 1, x_i^* = \frac{1}{n} \sum_{i=1}^n x_i, \sigma_i^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - x_i^*)^2,$$

где n — число компонент вектора, составленного из элементов матрицы межэлементных расстояний; x_i^* — среднее значение i -го расстояния; σ_i — стандартное отклонение значения i -го расстояния.

Среднее количество информации, приносимой каждой компонентой x_i , равно энтропии распределения значений этой компоненты $H(x_i)$. Если эти значения сосредоточены в относительно небольшой области единичного интервала, информационное содержание такой компоненты мало. Общий принцип предобработки данных, как известно, состоит в максимизации энтропии обучающих выборок. Использование более «вытянутой» сжимающей функции (в отличие от гиперболического тангенса: $f(x) = \frac{e^{ax} - e^{-ax}}{e^{ax} + e^{-ax}}$) позволяет получить более равномерное распределение нормированных значений в соответствующем диапазоне. Из диапазона $[0, 1]$, денормировка осуществляется по формуле $x_i = x_i^* - \frac{\sigma_i}{2} \ln\left(\frac{2}{x_i^{norm}} - 1\right)$, из диапазона $[-1, 1]$

по формуле $x_i = x_i^* - \frac{\sigma_i}{2} \ln\left(\frac{2}{x_i^{norm}} - 1\right)$, из диапазона $[-1, 1]$

$$x_i = x_i^* - \frac{\sigma_i}{2} \ln\left(\frac{2}{x_i^{norm} + 1} - 1\right).$$

5. Метрика при вычислении расстояний между полуматрицами межэлементных расстояний образцов X' и соразмерных фрагментов

шаблонов X — всегда квадратичная: $p = 2, q = 1, d_{ij} = \left(\sum_r (x_{ir} - x_{jr})^2\right)^{\frac{1}{2}}$,

где x — компоненты полуматрицы межэлементных расстояний. При ненормализованных векторах X и X' в качестве меры сходства может служить скалярное произведение: $d(x, x') = 1 - x \cdot x' = 1 - \|x\| \|x'\| \cos(x, x')$,

при нормализованных сравниваемых векторах, показатель аналогичен евклидовому расстоянию:

$$d(x, x') = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \dots + (x_n - x'_n)^2} = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2}.$$

При выполнении операции нормализации: $\overline{x_i} = \frac{x_i}{\|x\|}, i = \overline{1, n}$, где n — количество точек в сравниваемых попарных расстояний фрагмента и образца, векторы X и X' преобразуются в единичные с тем же направлением:

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

6. Введение в алгоритм заключительного этапа восстановления искаженных (отсутствующих) элементов ГТО позволяет на основе развитого инструментария теории многомерного метрического шкалирования повысить информативность вторичной обработки изображений.

Рассмотрим синтез классификационной модели, реализующей решающее правило отнесения каждого входного вектора $X_i \in D_j$ из конечного множества входных векторов к одному из элементов множества шаблонов на основе теории статистических решений как наиболее теоретически обоснованной теорией распознавания образов используемой в ходе вторичной обработки радио- и оптических изображений.

При отнесении i -го образца X_i полуматрицы межэлементных расстояний, представленного к распознаванию к j -му шаблону D_j вычисляются расстояния d_{ij} , и решения принимаются в форме: $X_i \in D_j$, при:

$$|d_{is} - d_{ij}| > \varepsilon, s = 1, \dots, L; s \neq i, \quad (2)$$

где ε — порог распознавания ($\varepsilon > 0$).

В методах статистических решений для распознавания различения с помощью логарифма отношения правдоподобия используется следующее правило [20]:

$$\begin{aligned} \text{при } \ln \frac{f(X / D_j)}{f(X / D_s)} > \ln \delta, X \in D_j, \\ \text{при } \ln \frac{f(X / D_j)}{f(X / D_s)} < \ln \delta, X \in D_s, \end{aligned} \quad (3)$$

где $\ln \delta$ — логарифмический порог классификации. При условии, что все компоненты конкатенированного вектора межэлементных расстояний

ний x_1, x_2, \dots, x_n , где n — размерность нормированного пространства межэлементных расстояний ГТО, независимы и распределены по нормальному закону:

$$f(X / D_j) = \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \frac{(x_i - m_{ji})^2}{\sigma_{ij}^2} \right\}.$$

В соответствии с условием (3) при

$$\ln \frac{f(X / D_j)}{f(X / D_s)} = \frac{1}{2} \left\{ \sum_{i=1}^n \frac{(x_i - m_{si})^2}{\sigma_{si}^2} - \sum_{i=1}^n \frac{(x_i - m_{ji})^2}{\sigma_{ji}^2} \right\} > \ln \delta, X \in D_j. \quad \text{Послед-}$$

нее соотношение можно записать так: $|d_{is} - d_{ij}| > 2 \ln \delta, X \in D_j$, где:

$$d_{is} = \sum_{i=1}^n \frac{(x_i - m_{si})^2}{\sigma_{si}^2}; d_{ij} = \sum_{i=1}^n \frac{(x_i - m_{ij})^2}{\sigma_{ij}^2}.$$

Таким образом, алгоритм классификации искаженных n -мерных групповых точечных объектов на основе комбинаторного поиска фрагментов полуматриц межэлементных расстояний совпадает с алгоритмами по методу максимального правдоподобия, если используется квадратичная мера расстояния в пространстве компонент конкатенированного вектора межэлементных расстояний и весовые коэффициенты $\lambda_{ij} = \frac{1}{\sigma_{ij}}$.

В соответствии с условиями (2), (3) порог распознавания $\varepsilon = 2 \ln \delta$.

3. Описание алгоритма. Исходными данными алгоритма являются:

1. Предположение о размерности R метрического пространства ГТО. Помимо трех пространственных координат, дополнительными ортогональными проекциями могут быть параметризованы различные характеристики точек ГТО из выборки исходного описания данных.

2. Исходная выборка шаблонов ГТО — как алфавит L классов классифицируемых образцов ГТО.

3. Параметры p и q метрики расстояний между элементами ГТО:

$$\hat{d}_{ij} = \left(\sum_{r=1}^R |y_{ir} - y_{jr}|^p \right)^{\frac{q}{p}}, \quad \text{где } R \text{ — размерность пространства признаков}$$

ГТО, r — индекс размерности,

4. Образец ГТО (индекс образца ГТО в исходной выборке), подлежащий классификации. Для простоты верификации, в качестве образца использован один из шаблонов, подвергнутый искажениям.

5. Параметры искажения образца ГТО, подлежащего распознаванию: *a*) среднеквадратическое отклонение шумовой аддитивной составляющей к координатам точек ГТО (в % от линейных усредненных по выборке размеров ГТО); *b*) количество удаляемых элементов в распознаваемом образце, координаты которых необходимо восстановить (в % от мощности — числа элементов ГТО); *c*) угол поворота заданных осей координат пространства вложения распознаваемого образца ГТО; *d*) коэффициента изменения масштаба.

6. Количество точек — статистических испытаний по восстановлению координат пропущенных элементов ГТО методом Монте-Карло.

7. Технологический параметр — начальное значение, присваиваемое переменным, участвующим в процедурах поиска минимального расстояния в заданной метрике.

Алгоритм классификации включает 14 шагов:

1. Задание исходных данных алгоритма в соответствии с описаниями пп.1-3, 6, 7 перечня исходных данных.

2. Выборка и искажение образца ГТО в соответствии с описаниями пп.4, 5 перечня исходных данных алгоритма.

3. Вычислить матрицу Δ' попарных расстояний точек y_1, \dots, y_N образца ГТО в изначально определенной метрике сходства/различия:

$$\begin{bmatrix} 0 & D'(y'_2/y'_1) & D'(y'_j/y'_1) & D'(y'_N/y'_1) \\ D'(y'_1/y'_2) & 0 & D'(y'_j/y'_2) & D'(y'_N/y'_2) \\ D'(y'_1/y'_j) & D'(y'_2/y'_j) & 0 & D'(y'_N/y'_j) \\ D'(y'_1/y'_N) & D'(y'_2/y'_N) & D'(y'_j/y'_N) & 0 \end{bmatrix}, \text{ где в общем случае}$$

$$D'(y'_j/y'_i) = \left(\sum_{r=1}^R |y'_{ir} - y'_{jr}|^p \right)^{\frac{q}{p}}, R — \text{размерность пространства вложения}$$

образца ГТО. В экспериментах метрика пространства ГТО и пространства компонент вектора межэлементных расстояний — квадратичные.

4. На основе Δ' путем последовательного составления строк вдоль ненулевой размерности сформировать вектор элементов треугольной полуматрицы попарных расстояний образца:

$$X' = [D'(y'_2/y'_1), \dots, D'(y'_j/y'_1), \dots, D'(y'_N/y'_1), D'(y'_j/y'_2), \dots,$$

$$D'(y'_N/y'_2), \dots, D'(y'_N/y'_j)] = [x'_1, \dots, x'_n], \quad n = \frac{N!}{2(N-2)!},$$

где N — количество точек образца.

5. Нормировка компонент вектора элементов треугольной полуматрицы попарных расстояний образца X' путем приведения данных к единичному масштабу $[0, 1]$:

$$x_i^{norm} = \frac{1}{1 + e^{-\left(\frac{x_i - x_i^*}{\sigma_i}\right)}}, \quad x_i^* \equiv \frac{1}{n} \sum_{i=1}^n x_i, \quad \sigma_i^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (x_i - x_i^*)^2,$$

где n — число компонент вектора, составленного из элементов матрицы межэлементных расстояний; x_i^* — среднее значение i -го расстояния; σ_i — стандартное отклонение значения i -го расстояния.

Далее для каждого l -го шаблона ГТО, $l = 1, \dots, L$ исходной выборки данных выполнить шаги 6-13:

6. Проверка условия «Количество точек шаблона ГТО больше или равно количеству точек образца ГТО, представленного к распознаванию», если нет, то переход к следующему шаблону выборки.

7. При выполнении условия (шаг 6) — создание упорядоченного массива индексов для формирования множества фрагментов данного шаблона, содержащих количество точек — равное количеству точек образца ГТО.

8. Формирование множества K сочетаний точек шаблона по количеству N точек образца ГТО $K = \frac{M!}{N!(M-N)!}$, где M — количество точек шаблона ГТО.

Для каждого k -го сочетания, $k = 1, \dots, K$, содержащего N точек шаблона ГТО выполняются шаги 9-12:

9. Вычислить матрицу Δ попарных расстояний точек y_1, \dots, y_N фрагмента шаблона в изначально заданной метрике сходства/различия:

$$\begin{bmatrix} 0 & D(y_2/y_1) & D(y_j/y_1) & D(y_N/y_1) \\ D(y_1/y_2) & 0 & D(y_j/y_2) & D(y_N/y_2) \\ \dots & \dots & \dots & \dots \\ D(y_1/y_i) & D(y_2/y_i) & 0 & D(y_N/y_i) \\ \dots & \dots & \dots & \dots \\ D(y_1/y_N) & D(y_2/y_N) & D(y_j/y_N) & 0 \end{bmatrix}, \quad \text{где } D(y_j/y_i) = \left(\sum_{r=1}^R |y_{ir} - y_{jr}|^p \right)^{\frac{q}{p}},$$

R — размерность пространства вложения фрагментируемого ГТО.

10. На основе Δ сформировать вектор элементов треугольной полуматрицы попарных расстояний фрагмента шаблона:

$$X = [D(y_2/y_1), \dots, D(y_j/y_1), \dots, D(y_N/y_1), D(y_j/y_2), \dots, D(y_N/y_2), \dots, D(y_N/y_i)] = [x_1, \dots, x_n], \quad n = \frac{N!}{2(N-2)!},$$

где N — количество фрагмента шаблона.

11. Нормировка компонент вектора элементов треугольной полуматрицы попарных расстояний фрагмента шаблона X путем приведения данных к единичному масштабу $[0, 1]$:

$$x_i^{norm} = \frac{1}{1 + e^{-\left(\frac{x_i - x_i^*}{\sigma_i}\right)}}, \quad x_i^* \equiv \frac{1}{n} \sum_{i=1}^n x_i, \quad \sigma_i^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (x_i - x_i^*)^2,$$

где n — число компонент вектора, составленного из элементов матрицы межэлементных расстояний; x_i^* — среднее значение i -го расстояния; σ_i — стандартное отклонение значения i -го расстояния.

12. Вычислить квадратичный показатель сходства E_k^l между вектором элементов треугольной полуматрицы попарных расстояний образца X' и соразмерным вектором элементов треугольной полуматрицы попарных расстояний фрагмента X шаблона, где $k = 1, \dots, K$ — количество сочетаний из M по N точек l -го шаблона ГТО, $l = 1, \dots, L$, L — общее количество шаблонов. Показатель вычисляется как:

$$E_k^l = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \dots + (x_n - x'_n)^2} = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2},$$

где n — число компонент вектора, составленного из треугольной полуматрицы межэлементных расстояний между точками классифицируемого образца и соразмерного k -го фрагмента l -го шаблона ГТО.

13. Вычислить $S = K * L$ квадратичных показателей сходства по всем K комбинациям всех L шаблонов исходной выборки.

14. Принятие решения C об индексе искомого шаблона, соответствующего искаженному образцу по критерию минимума показателя:

$$E_k^l = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \dots + (x_n - x'_n)^2} = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2}.$$

$C = \arg \max_l (\arg \max_k (E_k^l))$, $k = \overline{1, K}$, $l = \overline{1, L}$, где K — количество ком-

бинаций точек в k -м фрагменте l -го шаблона ГТО соразмерного распознаваемому образцу, L — общее число шаблонов выборки.

Блок-схема алгоритма представлена на рисунке 3.

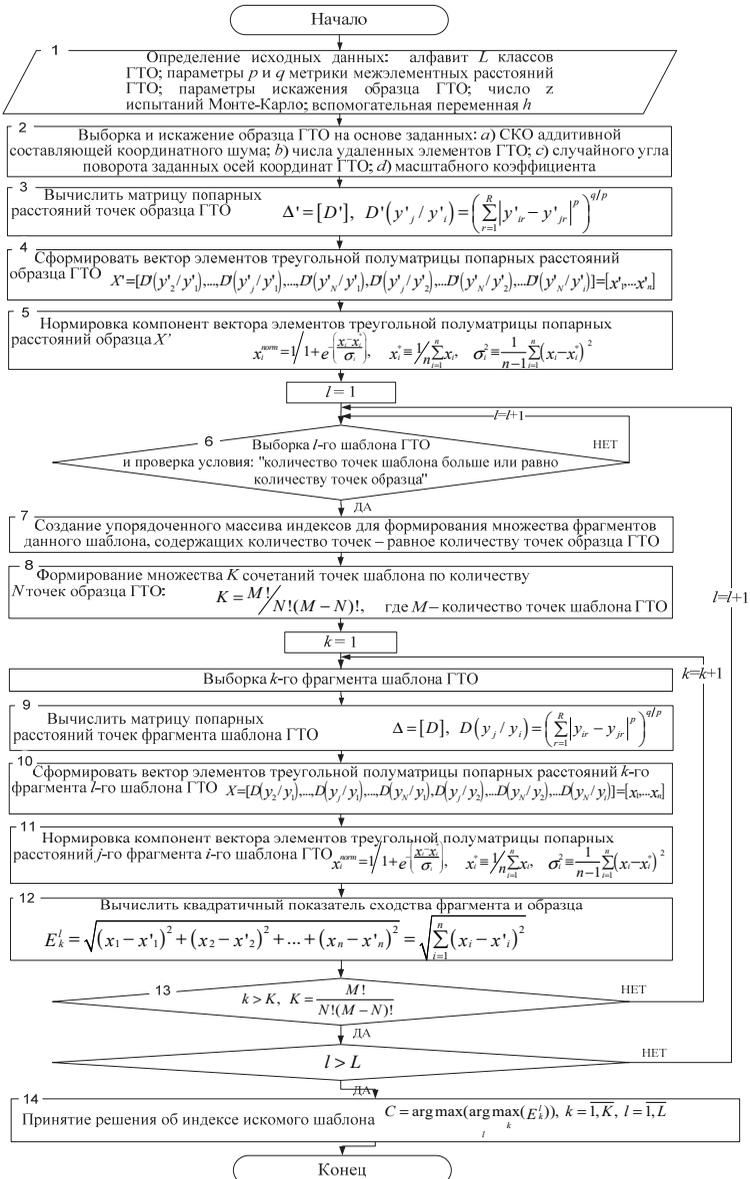


Рис. 3. Блок-схема алгоритма классификации и восстановления искаженных n -мерных групповых точечных объектов на основе комбинаторного поиска фрагментов

4. Пример реализации алгоритма. Появление радиолокационных станций с синтезированной апертурой, способных в высоком темпе осуществлять сбор данных об обстановке на больших участках подстилающей поверхности влечет необходимость получения в реальном масштабе времени достоверных результатов обработки такой информации. Особый интерес представляют две задачи: автоматическое обнаружение точечных целей и выделение групп среди обнаруженных объектов такого вида.

Достоверность алгоритма классификации образцов ГТО экспериментально исследована по базе 110 шаблонов, каждый из которых описывает контрастированное радиоизображение распределенных по площади (400×350 м) объектов. По оси Z (рисунок 4) в шкале целых чисел параметризован тип точки ГТО.

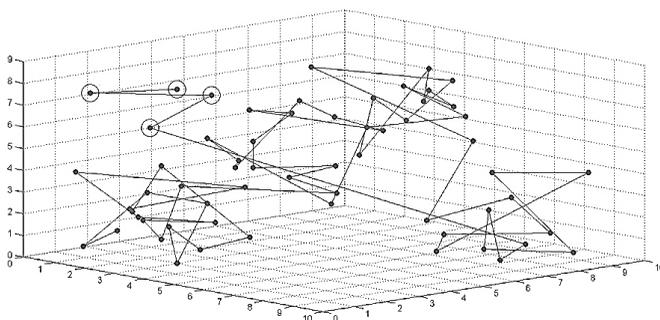


Рис. 4. Пример шаблона ГТО «блестящих» точек радиоизображения с параметризованной характеристикой вида точки и масштабированными осями в интервале [0,10]

На рисунках 5-7 представлены результаты численных экспериментов по классификации искаженных образцов ГТО из состава выборки.

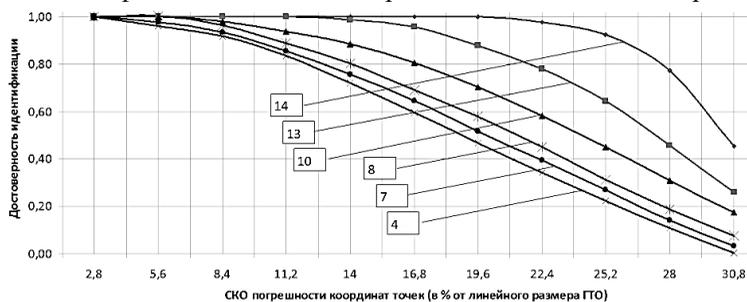


Рис. 5. Зависимость достоверности идентификации шаблона ГТО от среднеквадратической погрешности (в %) координат его точек (в одной плоскости) при различном количестве элементов ГТО (от 4 до 14)

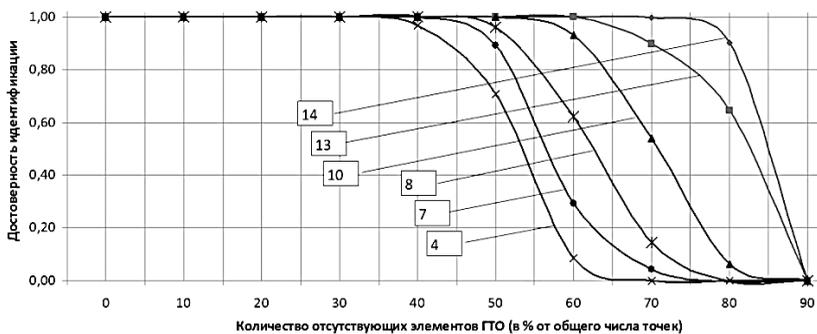


Рис. 6. Зависимость достоверности идентификации шаблона ГТО от количества отсутствующих элементов (в % от общего числа точек) при различном количестве элементов распознаваемого образца (от 4 до 14)

Достоверность идентификации P (вероятность правильной классификации) образцов определена частотным способом $P = N_1/N$, где N — общее число реализаций алгоритма классификации, N_1 — количество реализаций алгоритма, в которых образец классифицирован верно.

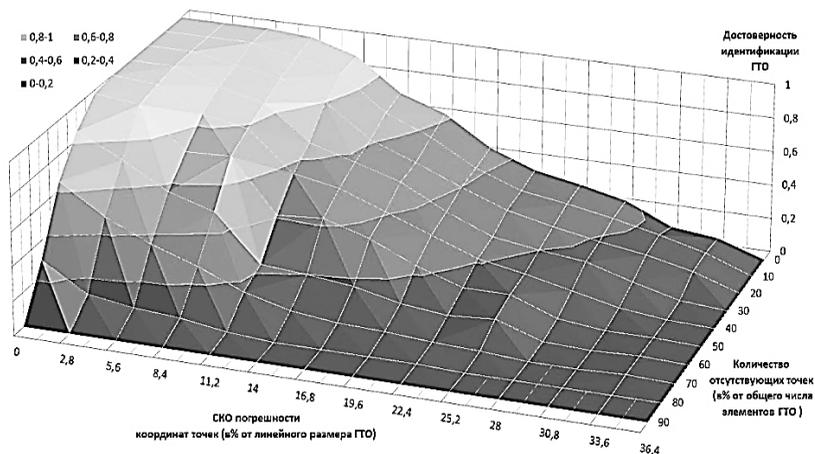


Рис. 7. Зависимость достоверности идентификации шаблона ГТО от среднеквадратической погрешности координат его точек при различном количестве отсутствующих элементов

5. Восстановление элементов искаженного образца. Восстановление элементов искаженного образца ГТО осуществляется на основе априорных данных относительно матрицы попарных расстояний между точками неискаженного шаблона. Индекс шаблона становится

известен по завершении этапа классификации на шаге 14 алгоритма (рисунок 3).

Классически восстановление каждой необнаруженной в образце точки ГТО возможно по уравнениям дальности на основе матрицы попарных расстояний неискаженного образца. Метод имеет недостаток, связанный с ухудшением точности восстановления при выборе опорных точек находящихся вблизи прямой в R -мерном пространстве, а также с некорректным требованием учета искажений опорных точек.

Восстановление элементов искаженного образца возможно с помощью методов многомерного шкалирования, где в качестве исходных данных использует симметричные матрицы близости с одним набором данных. В результаты анализа каждый объект представляется как точка в R мерном пространстве. Основателем современного направления методов многомерного шкалирования считается В. Торгерсон [21], который предложил алгоритм на основе работы Т. Юнга и А. Хаусхолдера [22]. Торгерсон предположил, что оценки различий равны расстояниям в евклидовом пространстве небольшой размерности R : $\delta_{ij} = \hat{d}_{ij}$, где $\hat{d}_{ij}^2 = \sum_{r=1}^R (x_{ir} - x_{jr})^2$. Далее на основе исход-

ной матрицы близостей Δ рассчитывается матрица с двойным центрированием Δ^* , каждый элемент которой равен

$$\delta_{ij}^* = -\frac{1}{2}(\delta_{ij}^2 - \delta_{i\cdot}^2 - \delta_{\cdot j}^2 + \delta_{\cdot\cdot}^2) \quad , \quad \text{где} \quad \delta_{i\cdot}^2 = \frac{1}{J} \sum_{j=1}^J \delta_{ij}^2 \quad , \quad \delta_{\cdot j}^2 = \frac{1}{I} \sum_{i=1}^I \delta_{ij}^2 \quad ,$$

$$\delta_{\cdot\cdot}^2 = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J \delta_{ij}^2 \quad . \quad \text{Доказано, что каждый элемент полученной матрицы}$$

Δ^* будет равен $\delta_{ij}^* = \sum_{r=1}^R x_{ir} x_{jr}$. Предположение о том, что оценки различия равны расстояниям является очень жестким. Менее жесткой является следующая модель, в которой различия соответствуют расстояниям с точности до некоторой аддитивной константы $\delta_{ij} = \hat{d}_{ij} + c$. Проблема первоначальной оценки аддитивной константы в литературе по многомерному шкалированию называется «проблемой аддитивной константы» [23].

Значения известных точек искаженного образца ГТО могут использоваться в качестве стартовых значений параметров модели многомерного шкалирования, основанных на итеративных процедурах по аналогии с использованием описанного выше метрическим алгоритмом Торгерсона, который также может быть использован для оценки стартовых значений параметров модели.

Минимизируемый при многомерном шкалировании функционал сложным образом зависит от nL переменных, имеет огромное количество локальных минимумов, и его вычисление довольно трудоемко. Поэтому многие алгоритмы многомерного шкалирования основаны на итерационном размещении точек по одному. В настоящем алгоритме использован метод статистических испытаний Монте-Карло, в рамках которого каждая точка случайным образом размещенная в предполагаемой площади искомым координат проверяется по условию минимума

$$\text{разностей показателей сходства вида } W(s'_j / s'_i) = \left(\sum_{r=1}^R |s'_{ir} - s'_{jr}|^2 \right)^{\frac{1}{2}},$$

$R=3$ — размерность пространства вложения образца. Данная процедура включается 15-м, заключительным этапом в алгоритм, представленный на рисунке 3.

На рисунке 8 представлен пример восстановления координат 10 элементов по 4 известным элементам распознанного ГТО на плоскости.

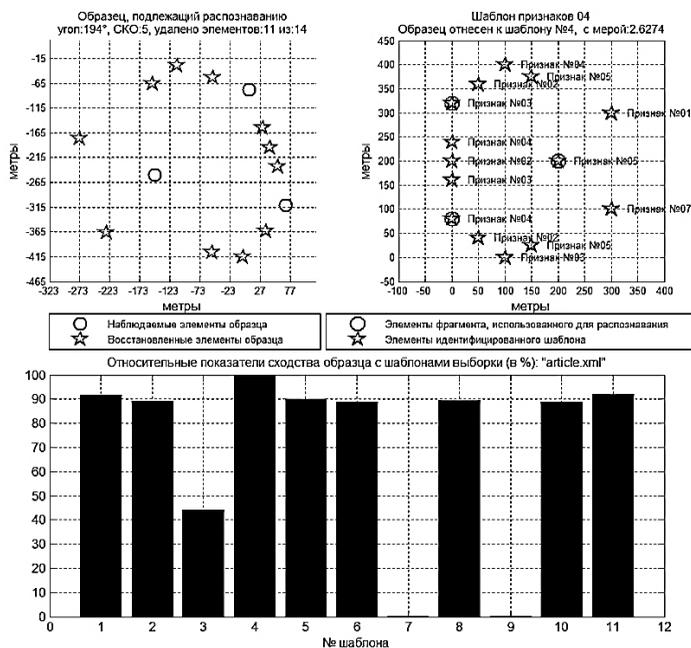


Рис. 8. Результат восстановления двумерных координат десяти точек искаженного образца шаблона по данным 4 наблюдаемых точек методом Монте-Карло (количество статистических испытаний 50000); относительные показатели сходства представлены для 11 наиболее вероятных шаблонов из 110

На рисунке 9 представлены зависимости евклидовых показателей погрешности восстановления W в зависимости от числа испытаний Монте-Карло.

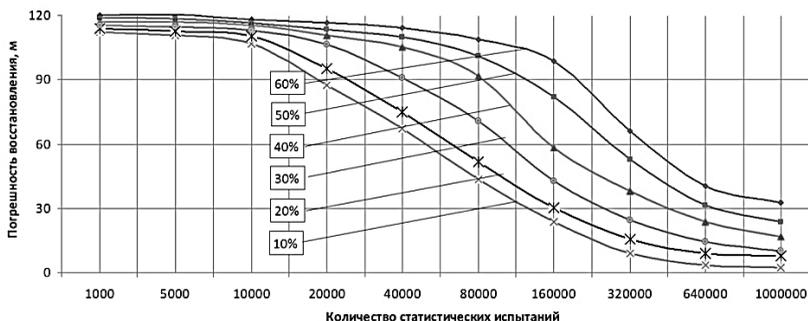


Рис. 9. Зависимости ненормированной среднеарифметической по выборке погрешности восстановления координат от числа статистических испытаний Монте-Карло при различном числе восстанавливаемых точек ГТО (погрешность образца 40 % от линейного размера ГТО)

При увеличении погрешности позиционирования точек образца и уменьшении числа известных элементов в искаженном образце погрешности восстановления координат монотонно растут.

6. Заключение. Одним из направлений повышения эффективности систем технического зрения радио- и оптического диапазона является разработка алгоритмов параллельного распознавания и восстановления радио, оптических локационных изображений целей, представленных многомерными групповыми точечными объектами, включающих инвариантные к неафинным искажениям процедуры идентификации и многомерного метрического шкалирования как способа структурной верификации конечного результата.

Достоинством представленного алгоритма является возможность реализации многомерных масштабируемых процедур классификации/восстановления образов, представимых групповыми точечными объектами, используя при этом пространство той минимальной мерности, которая позволяет достоверно классифицировать зашумленные и искаженные (поворот, смещение, прореживание) образцы ГТО.

Возможным направлением развития данного класса алгоритмов в части восстановления элементов искаженного образа является использование карт Кохонена или карт Сэммона, позволяющих аппроксимировать многомерные отображения распределения точек на N -мерную поверхность с регулярным или случайным соседством между узлами. При обучении на вход подаются строки фрагментов матрицы

попарных расстояний неискаженного шаблона, на выходе формируются представления истинных координат с учетом ограничений и начальных условий относительно изменения масштаба, смещения центра ГТО или поворота осей его координат.

Это позволит повысить достоверность классификации зашумленных радиолокационных изображений при ограничениях на оперативность вторичной обработки локационной информации в реальном масштабе времени.

Литература

1. *Dostovalov M., Ermakov R., Moussiniants T.* Airborne Testing of Multi-channel Modes of the Perspective Spaceborne // Proceedings of 10th European Conference on Synthetic Aperture Radar (EUSAR 2014). 2014. pp. 1–4.
2. *Авдеев В.А., Бахолдин В.С., Гаерлиов Д.А. и др.* Комплекс экспериментов по приему отраженных земной поверхностью сигналов спутниковых радионавигационных систем ГЛОНАСС/GPS // Труды института прикладной астрономии РАН. 2012. № 23. С. 230–235.
3. *Romeiser R. et al.* First Analysis of TerraSAR-X Along-Track InSAR-Derived Current Fields // IEEE Trans. on Geoscience and Remote Sensing. 2010. vol. 48. no. 2. pp. 820–829.
4. *Верба В.С.* Авиационные комплексы радиолокационного дозора и наведения. Принципы построения, проблемы разработки и особенности функционирования // М.: Радиотехника. 2014. 528 с.
5. *Rozhentsov A.A., Bayev A.A.* Estimation of 3D images parameters set by disordered readouts // Proceedings of 9th Int. Conf. “Pattern Recognition and Image Analysis: New Information Technologies”. 2008. vol. 2. pp. 133–135.
6. *Фурман Я.А.* Точечные поля и групповые объекты // М.: Физматлит. 2015. 440 с.
7. *Хафизов Р.Г., Фурман Я.А.* Исследование влияния яркостной информации точек пространственного группового точечного объекта на устойчивость его проводочной модели // Вестник Марийского государственного технического университета. 2009. № 2(6). С. 35–43.
8. Комплекснозначные и гиперкомплексные системы в задачах обработки многомерных сигналов / Под ред. *Я.А. Фурмана* // М.: Физматлит. 2004. 298 с.
9. *Фурман Я.А., Хафизов Д.Г.* Распознавание групповых точечных объектов в трехмерном пространстве // Автотметрия. 2003. Том. 39. №1. С. 3–14.
10. *Doerry A.W., Dubbert D.F., Thompson M.E., Gutierrez V.D.* A portfolio of fine resolution Ka-band SAR images: Part I // SPIE Defense and Security Symposium. 2005.
11. *Фурман Я.А., Егوشина И.Л., Ерусланов Р.В.* Согласованная фильтрация зашумленных дискретных кватернионных сигналов // Журнал радиоэлектроники. 2012. №3. С. 1–35.
12. *Ипатов Ю.А., Кривецкий А.В.* Методы обнаружения и пространственной локализации групп точечных объектов // Кибернетика и программирование. 2014. № 6. С.17–25.
13. *Szeliski R.* Computer Vision: Algorithms and Applications // Springer. 2011. 812 p.
14. *Furman Y.A., Eruslanov R.V., Egoshina I.L.* Iterative Algorithm for angular matching of group point objects with a priori uncertainty of parameters // Pattern recognition and image analysis. 2013. vol. 23. no. 3. pp. 381–388.

15. *Furman Ya.A., Egoshina I.L., Eruslanov R.V.* Matching angular and vector descriptions of three-dimensional group point objects // Optoelectronics, Instrumentation and Data Processing. 2012. vol. 48. № 6. pp. 537–549.
16. *Novikov A.I., Sablina V.A., Nikiforov M.B., Loginov A.A.* Contour analysis and image superimposition task in computer vision systems // The 11th International Conference "Pattern Recognition and Image Analysis: New Information Technologies» (PRIA-11-2013). 2013. pp. 282–285.
17. *Furman Ya.A., Eruslanov R.V., Egoshina I.L.* Reconstruction of Images and Recognition of Polyhedral Objects // Pattern Recognition and Image Analysis. 2012. vol. 22. no. 1. pp. 196–209.
18. *Furman Ya.A., Eruslanov R.V., Lazarev A.O.* Calculation of the Vertex Coordinates of a Polyhedral Body by the Results of Coordinated Filtration of the Contours of Its Perspective Projections // Pattern Recognition and Image Analysis. 2012. vol. 22. no. 1. pp. 188–195.
19. *Ежов А.А., Шумский С.А.* Нейрокомпьютинг и его применения в экономике и бизнесе // М.: МИФИ. 1998. 224 с.
20. *Биргер И. А.* Техническая диагностика // М.: Машиностроение. 1978. 240 с.
21. *Torgerson W.S.* Theory and methods of scaling // New York: Wiley. 1958. 460 p.
22. *Young F.W., Householder A.S.* A note on multidimensional psychophysics // Psychometrika. 1941. vol. 6. pp. 331–333.
23. *Дэйвисон М.* Многомерное шкалирование. Методы наглядного представления данных // М.: Финансы и статистика. 1982. 254 с.

Каплин Александр Юрьевич — к-т техн. наук, заместитель генерального директора-генеральный конструктор, ОАО «Радиоавионика». Область научных интересов: системы управления и связи специального назначения, человеко-машинные системы, бортовая радиолокация и радионавигация. Число научных публикаций — 50. a.kaplin@list.ru; Троицкий пр., 4 лит. Б, Санкт-Петербург, 190103; р.т.: +7(812) 251-3875, Факс: +7(812) 251-2743.

Коротин Андрей Анатольевич — к-т техн. наук, директор научно-исследовательского центра, ОАО «Радиоавионика». Область научных интересов: системы управления и связи специального назначения, человеко-машинные системы, аппаратно-программные комплексы, защита космических аппаратов от радиации. Число научных публикаций — 30. kaa2805@mail.ru; Троицкий пр., 4 лит. Б, Санкт-Петербург, 190103; р.т.: +79119107595, Факс: +7(812) 251-2743.

Назаров Андрей Вячеславович — д-р техн. наук, доцент, начальник кафедры космической радиолокации и радионавигации, Военно-космическая академия имени А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: распознавание образов, нейросетевые технологии, моделирование распределенных систем, обработка сигналов в оптико-электронных информационных системах. Число научных публикаций — 100. naz-av@mail.ru; ул. Ждановская, 13, Санкт-Петербург, 197198; р.т.: +7(812) 347-95-33.

Якимов Виктор Леонидович — к-т техн. наук, доцент, заместитель начальника кафедры приемных устройств и радиоавтоматики, Военно-космическая академия имени А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: моделирование сложных систем, нейросетевые технологии, техническая диагностика. Число научных публикаций — 40. yakim78@yandex.ru; ул. Ждановская, 13, Санкт-Петербург, 197198; р.т.: +7(812)347-95-36.

A.Y. KAPLIN, A.A. KOROTIN, A.V. NAZAROV, V.L. YAKIMOV
**ALGORITHM OF CLASSIFICATION AND RECOVERY OF THE
 DISTORTED N -DIMENSIONAL GROUP POINTWISE OBJECTS
 ON THE BASIS OF COMBINATORIAL SEARCH OF FRAGMENTS**

Kaplin A.Y., Korotin A.A., Nazarov A.V., Yakimov V.L. **Algorithm of classification and recovery of the distorted N -dimensional group pointwise objects on the basis of combinatorial search of fragments.**

Abstract. The algorithm of classification of multidimensional group pointwise objects samples is presented. Search is carried out on the basis of combinatorial search of proportionate fragments of matrixes of pairwise relations on a set of templates. The decision on assignment of the sample to this or that template is made according to criterion of the minimum Euclidean distance. The presented approach to recognition allows one to synthesize invariant (concerning rotation, scaling or offset of system of co-ordinates) descriptions of secondary signs and to use quite a powerful toolkit of the theory of multidimensional and metric scaling in compensating distortions of the recognized group pointwise objects images. The algorithm implements a procedure of statistical tests of Monte-Carlo, within the frames of which each point, allocated in a random way in a prospective neighborhood of required coordinates, is checked by condition of the minimum of the quadratic similarity measure. The paper gives an example and the results of using the algorithm for identification and recovery of the distorted radio images exposed to coordinate noises and presented by sampling of templates of "brilliant" points.

Keywords: group pointwise object, classification, multidimensional scaling.

Kaplin Alexander Yurievich — Ph.D., deputy director-general, general designer, Joint Stock Venture «Radioavionika». Research interests: control and communication systems of a special purpose, human - machine systems, on-board radar and navigation. The number of publications — 50. a.kaplin@list.ru; P.O.B. 111, St. Petersburg, 190103; office phone: +7(812) 2513875, Fax: +7(812) 251-2743.

Korotin Andrey Anatolievich — Ph.D., director of research center, Joint Stock Venture «Radioavionika». Research interests: control and communication systems of a special purpose, human-machine systems, hardware-software complexes, protection of spacecraft from the radiation. The number of publications — 30. kaa2805@mail.ru; P.O.B. 111, St. Petersburg, 190103; office phone: +79119107595, Fax: +7(812) 251-2743.

Nazarov Andrey Vyacheslavovich — Ph.D., Dr. Sci., associate professor, head of space radiolocation and a radio navigation department, Mozhaisky Military Space Academy. Research interests: pattern recognition, neural nets, modeling of distributed systems, signal processing in optical-electronic information systems. The number of publications — 100. nazav@mail.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 347-95-33.

Yakimov Victor Leonidovich — Ph.D., associate professor, deputy head of receiving devices and radio automatic equipment department, Mozhaisky Military Space Academy. Research interests: simulation of difficult systems, neural network technologies, technical diagnostics. The number of publications — 40. yakim78@yandex.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812)347-95-36.

References

1. Dostovalov M., Ermakov R., Moussinants T. Airborne Testing of Multi-channel Modes of the Perspective Spaceborne SAR proc. Proceedings of 10th European Conference on Synthetic Aperture Radar (EUSAR 2014). 2014. pp. 1–4.
2. Avdeev V.A., Bakholdin V.S., Gavrilov D.A. [Complex of experiments on reception of the signals of satellite radio navigational systems reflected by an earth surface]. *Trudy instituta prikladnoi astronomii RAN – Proceedings of applied astronomy Institute of RAS*. 2012. vol. 23. pp. 230-235. (In Russ.).
3. Romeiser R. et al. First Analysis of TerraSAR-X Along-Track InSAR-Derived Current Fields. *IEEE Trans. on Geoscience and Remote Sensing*. 2010. vol. 48. no. 2. pp. 820–829.
4. Verba V.S. *Aviatsionnye komplekсы radiolokatsionnogo dozora i navedeniia. Printsipy postroeniia, problemy razrabotki i osobennosti funktsionirovaniia* [Aviation complexes of radar-tracking patrol and induction. Principles of creation, a problem of development and a functioning singularity]. Moscow: Radiotekhnika Publ. 2014. 528 p. (In Russ.).
5. Rozhentsov A.A., Bayev A.A. Estimation of 3D images parameters set by disordered readouts. Proceedings of 9th Int. Conf. “Pattern Recognition and Image Analysis: New Information Technologies”. 2008. vol. 2. pp. 133–135.
6. Furman Ya.A. *Tochechnye polia i gruppovye ob’ekty* [Point field and group objects]. Moscow: Fizmatlit Publ. 2015. 440 p. (In Russ.).
7. Khafizov R.G., Furman Ya.A. [Research of influence of the brightness information of points of space group pointwise object on stability of its wire model]. *Vestnik Mariiskogo gosudarstvennogo tekhnicheskogo universiteta – Bulletin of Mari state technical university*. 2009. vol. 2(6). pp. 35–43. (In Russ.).
8. Furman Ya. A. *Kompleksnoznachnye i giperkompleksnye sistemy v zadachakh obrabotki mnogomernykh signalov* [Complex-valued and Hypercomplex systems in processing of multidimensional signals]. Moscow: Fizmatlit Publ. 2004. 298 p. (In Russ.).
9. Furman Ya. L., Khafizov D. G. [Recognition of group pointwise objects in three-dimensional space]. *Avtometriia – Autometry*. 2003. vol. 39. no. 1. pp. 3–14. (In Russ.).
10. Doerry A.W., Dubbert D.F., Thompson M.E., Gutierrez V.D. A portfolio of fine resolution Ka-band SAR images: Part I. SPIE Defense and Security Symposium. 2005.
11. Furman Ya.A., Egoshina I.L., Eruslanov R.V. [Matched filtering of noisy discrete quaternion signals]. *Zhurnal radioelektroniki – Magazine of radio electronics*. 2012. no. 3. pp. 1–35. (In Russ.).
12. Ipatov Iu.A., Krevetskii A.V. [Methods of detection and spatial localization of groups of point objects]. *Kibernetika i programmirovaniie – Cybernetics and programming*. 2014. vol. 6. pp. 17–25. (In Russ.).
13. Szeliski R. *Computer Vision: Algorithms and Applications*. Springer. 2011. 812p.
14. Furman Ya. A., Eruslanov R.V., Egoshina I.L. Iterative Algorithm for angular matching of group point objects with priory uncertainty of parameters. *Pattern recognition and image analysis*. 2013. vol. 23. no. 3. pp. 381–388.

15. Furman Ya.A., Egoshina I.L., Eruslanov R.V. Matching angular and vector descriptions of three-dimensional group point objects. *Optoelectronics, Instrumentation and Data Processing*. 2012. vol. 48. no. 6. P. 537–549.
16. Novikov A.I., Sablina V.A., Nikiforov M.B., Loginov A.A. Contour analysis and image superimposition task in computer vision systems. The 11th International Conference "Pattern Recognition and Image Analysis: New Information Technologies» (PRIA-11-2013). 2013. pp. 282–285.
17. Furman Ya.A., Eruslanov R.V., Egoshina I.L. Reconstruction of Images and Recognition of Polyhedral Objects. *Pattern Recognition and Image Analysis*. vol. 22. no. 1. 2012. pp. 196–209.
18. Furman Ya.A., Eruslanov R.V., Lazarev A.O. Calculation of the Vertex Coordinates of a Polyhedral Body by the Results of Coordinated Filtration of the Contours of Its Perspective Projections. *Pattern Recognition and Image Analysis*. 2012. vol. 22. no. 1. pp. 188–195.
19. Ezhov A.A., Shumskii S.A. Nejrokomputing i ego primenenija v jekonomike i biznese [Neurocomputing and its applications in economy and business]. Moscow: MIFI Publ. 1998. 224 p. (In Russ.).
20. Birger I. A. Tehnicheskaja diagnostika [Technical diagnostics]. Moscow: Mashinostroenie Publ. 1978. 240 p. (In Russ.).
21. Torgerson W.S. Theory and methods of scaling. New York: Wiley. 1958. 460 p.
22. Young F.W., Householder A.S. A note on multidimensional psychophysics. *Psychometrika*. 1941. vol. 6. pp. 331–333.
23. Deivison M. *Mnogomernoje shkalirovanie. Metody nagljadnogo predstavlenija dannyh* [Multidimensional scaling. Methods of evident data representation]. Moscow: Finansy i statistika Publ. 1982. 254 p. (In Russ.).