

А.Ю. КАПЛИН, А.А. КОРОТИН, А.В. НАЗАРОВ, В.Л. ЯКИМОВ  
**АЛГОРИТМ КЛАССИФИКАЦИИ ГРУППОВЫХ ТОЧЕЧНЫХ  
ОБЪЕКТОВ С НЕУПОРЯДОЧЕННЫМИ ЭЛЕМЕНТАМИ НА  
ОСНОВЕ ВЕРОЯТНОСТНОЙ МЕРЫ БЛИЗОСТИ**

---

*Каплин А.Ю., Коротин А.А., Назаров А.В., Якимов В.Л.* Алгоритм классификации групповых точечных объектов с неупорядоченными элементами на основе вероятностной меры близости.

**Аннотация.** Представлен алгоритм классификации групповых точечных объектов (ГТО), основанный на сравнительном анализе фрагментов искаженных образов и шаблонов ГТО. В качестве фрагментов использованы последовательности элементов ГТО различной длины. В качестве признаков классификации выступают попарные и угловые межточечные расстояния. При решении задачи классификации используется вероятностная мера близости, задаваемая экспертом с помощью функции принадлежности и закона распределения вероятности дискретных значений признаков классифицируемых объектов. Алгоритм включает следующие этапы: поиск и сравнение состава фрагментов искаженных образов и шаблонов ГТО; формирование вероятностной оценки близости искаженного образа ГТО и каждого шаблона в пространстве рассматриваемых признаков по результатам анализа каждого фрагмента; накопление полученных вероятностей по результатам анализа всех фрагментов искаженного образа; ранжирование полученных вероятностей отнесения искаженного образа к шаблонам ГТО; определение наиболее вероятного шаблона. В алгоритме предусмотрена возможность уточнения класса искаженного образа ГТО за счет использования логических правил и аналитических выражений рассматриваемой предметной области. Приведены пример и результаты применения данного алгоритма для решения задачи классификации реальных ГТО на основе анализа их фрагментов в виде последовательностей из двух и трех элементов.

**Ключевые слова:** групповой точечный объект, классификация, вероятностная мера близости.

---

**1. Введение.** Широкий класс задач обработки информации в современных информационных системах может быть связан с извлечением информации из изображений, представленных в виде компактного множества изолированных друг от друга точечных отметок — групповых точечных объектов (ГТО), обладающих формой и внутренней структурой, причем точки ГТО могут быть объектами различного типа [1, 2]. Одной из таких задач является идентификация ГТО, для решения которой необходимо разработать алгоритм классификации, позволяющий принять решение о классе ГТО с требуемой достоверностью [3].

На рисунке 1 представлены два различных примера ГТО для которых решение задачи идентификации имеет схожий характер и элементами которых являются: а) набор точек одной из проекций фазового пространства динамической системы ориентации малого космического аппарата, характеризующий ее техническое состояние и полученный на основе значений временного ряда  $H_n$  телеметрируемого па-

раметра ориентации на Землю, где  $u$  — номер отсчета телеметрируемого параметра [4]; б) набор точек, образующий точечную модель одного из трех кораблей, полученный в результате обработки радиолокационного изображения [5, 6]. Для указанных примеров точки ГТО соответствовали: областям высокой и низкой плотности фазовых траекторий (рисунок 1а); областям радиолокационного изображения с различными уровнями яркости (рисунок 1б). В первом случае объект ГТО образован точками двух типов («▲», «▼»), а во втором — точками трех типов («■», «●», «○»).

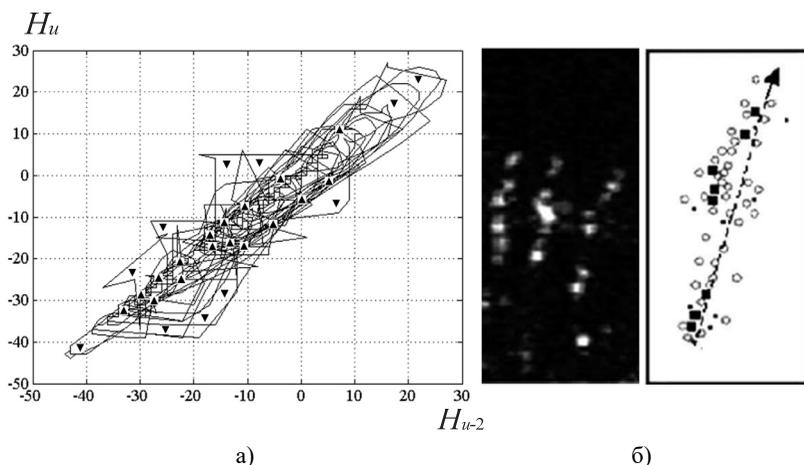


Рис. 1. Примеры ГТО из области: а) технической диагностики [4]; б) радиолокации [5, 6]

В общем случае решение задачи классификации ГТО может быть затруднено вследствие различного рода искажений, отсутствия ключевых точек (элементов ГТО), ложного определения типа элементов, неупорядоченности элементов и т. д.

Существуют различные подходы к определению ключевых элементов ГТО, формированию множества признаков классификации и решению задачи их распознавания. Наиболее обширной является категория алгоритмов обработки и распознавания объектов на изображении [3, 5, 7, 8]. При этом важнейшей задачей, предшествующей распознаванию ГТО на изображении, является задача достоверного обнаружения ключевых элементов ГТО. Для ее решения используются подходы на основе преобразований исходного изображения, позволяющих получить инвариантные признаки обнаруживаемых объектов; на основе алгоритмов детекции границ и различных фильтров и т. д. [9].

Большой класс методов обработки изображений и сигналов для классификации ГТО предлагает контурный анализ и его приложения [10]. Рассмотрение ГТО в виде квантернионных сигналов позволяет эффективно решать задачу их обнаружения на основе согласованной фильтрации [2, 10]. Так как ГТО представляют собой геометрические объекты, то в качестве признаков их классификации обычно используют различные межточечные и угловые расстояния [11, 12, 13]. После обнаружения элементов ГТО и определения множества признаков их классификации для решения задачи распознавания ГТО могут быть использованы различные методы, которые можно разделить на две большие группы, определяемые как геометрический (дискриминантный) и синтаксический (структурный) подходы [10].

В разработанном алгоритме для решения задачи классификации используется информация о сформированных моделях (шаблонах) ГТО в пространстве рассматриваемых признаков (межточечных и угловых расстояний), а основными критериями при его разработке были следующие: простота реализации и возможность использования современных технологий поиска в базах данных большой размерности, возможность получения именно вероятностной оценки отнесения текущего искаженного образа к каждому из множества шаблонов ГТО, удобная возможность использования при классификации дополнительной априорной информации в виде уточняющих экспертных правил.

**2. Постановка задачи классификации ГТО.** Решение задачи классификации ГТО можно разбить на три этапа:

- выделение признаков классификации ГТО, формирование базы шаблонов с массивами значений рассмотренных признаков;
- поиск в базе и соотнесение текущего искаженного образа ГТО со множеством шаблонов в пространстве рассматриваемых признаков;
- формирование вероятностной оценки для каждого претендента из базы шаблонов, ранжирование претендентов в соответствии с полученной вероятностной оценкой, принятие решения о наиболее вероятном претенденте.

В ходе решения задачи классификации необходимо синтезировать алгоритм  $f$ :

$$Y = f(\hat{S}, S_1, S_2, \dots, S_N), \quad (1)$$

где  $\hat{S} = (\hat{z}_1, \hat{z}_2, \dots, \hat{z}_L)$  — входной искаженный образ ГТО;  $\hat{z}_l$  — признаки искаженного образа ГТО;  $l=1 \dots L$  — номер признака искаженного ГТО;  $L$  — количество признаков искаженного образа ГТО;

$\mathbf{S}_j = (z_{j,1}, z_{j,2}, \dots, z_{j,N_j})$  —  $j$ -й шаблон ГТО;  $z_{j,m}$  — признак  $j$ -го шаблона ГТО;  $m = 1 \dots N_j$  — номер признака  $j$ -го шаблона ГТО;  $N_j$  — количество признаков  $j$ -го шаблона ГТО;  $N$  — количество шаблонов ГТО;  $\mathbf{Y}$  — код шаблона ГТО (номер класса), поставленный в соответствие искаженному образу  $\hat{\mathbf{S}}$ . При этом алгоритм классификации  $f$  реализуется на основе вычисления меры близости между искаженным образом и шаблонами ГТО в пространстве указанных признаков [3, 10].

В общем случае признаки ГТО могут быть не равноценны и вносить различный вклад в значение вероятности отнесения искаженного образа  $\hat{\mathbf{S}}$  к шаблону  $\mathbf{S}_j$  [14]. Так как образы ГТО искажены, а имеющиеся в базе шаблоны ГТО пересекаются в пространстве рассматриваемых признаков и зачастую имеют, одинаковые фрагменты, используем вероятностную меру для оценки близости искаженного образа и множества шаблонов [15, 16]. В отсутствии статистической информации о распределении значений признаков, зависимости этих распределений от внешних факторов (например, для наземных ГТО размещение элементов может сильно зависеть от рельефа и характера местности) и их статистической независимости, для выполнения такой процедуры можно использовать функции принадлежности, задаваемые экспертом [17]. Вероятностный метод построения функций принадлежности основан на сходстве понятий нечеткости и вероятности [17, 18]. При формировании функции принадлежности необходимо учесть, что маловероятное и наиболее вероятное событие должно иметь соответственно малую и максимальную степень принадлежности [17, 18, 19]. Будем считать признаки ГТО  $z_{j,m}$  и  $\hat{z}_i$  дискретными случайными величинами.

На рисунке 2 представлены функция принадлежности и закон распределения дискретной случайной величины  $z_{j,m}$  — одного из  $m$  признаков  $j$ -го шаблона ГТО. Вероятность  $P_{j,m}$  является условной вероятностью отнесения искаженного образа ГТО к  $j$ -му шаблону при условии, что значение его признака  $\hat{z}_i$  равно некоторому дискретному значению на шкале признака  $z_{j,m}$ . Так как случайная величина  $z_{j,m}$  дискретная и принимает строго фиксированные значения в пределах от  $z_{j,m,\min}$  до  $z_{j,m,\max}$ , то ее закон распределения и максимальное значение вероятности  $P_{j,m,\max}$  должны быть одинаковы для всех признаков шаблонов ГТО, а сумма вероятностей  $P_{j,m}$  по всем возможным дискретным

значениям величины  $z_{j,m}$  должна быть равна 1, что позволит обеспечить равнозначность признаков классификации (рисунок 2).

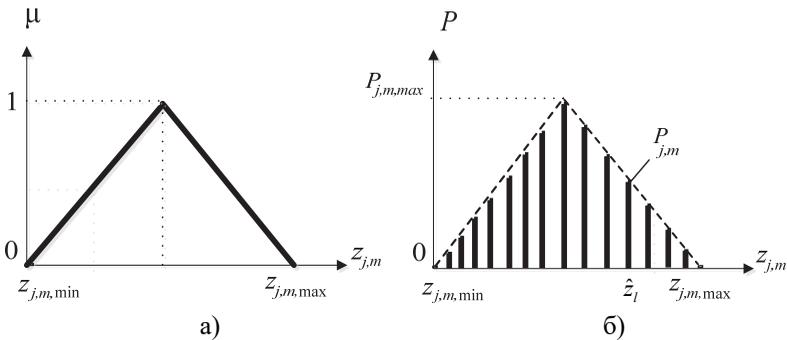


Рис. 2. Функция принадлежности (а) и закон распределения дискретной случайной величины  $z_{j,m}$  (б)

Элементы ГТО, в общем случае, не упорядочены, а номера признаков ГТО  $l$  и  $m$  — могут не совпадать [1]. Это требует процедуры сравнения искаженного образа ГТО по каждому признаку  $\hat{z}_l$  с каждым  $j$ -м шаблоном по каждому признаку  $z_{j,m}$  и получения вероятностей  $P_{j,l,m}$  отнесения искаженного образа ГТО по признакам  $\hat{z}_l$  к  $j$ -м шаблонам по признакам  $z_{j,m}$ . Анализируя вероятности  $P_{j,l,m}$ , найдем максимальное значение вероятности  $P_{j,l}$  отнесения искаженного образа ГТО по каждому признаку  $\hat{z}_l$  к каждому  $j$ -му шаблону, используя следующее правило:

$$P_{j,l} = \max_m \{P_{j,l,m}\}. \quad (2)$$

Допуская статистическую независимость признаков  $\hat{z}_l$ , оценим вероятности  $P_j$  отнесения искаженного образа ГТО к  $j$ -м шаблонам по совокупности  $L$  признаков:

$$P_j = \frac{1}{L} \sum_{l=1}^L P_{j,l}. \quad (3)$$

Решением задачи классификации (1) будет номер шаблона  $\xi$ , соответствующий значению максимальной вероятности  $P_j$ :

$$Y = \xi \Big| P_\xi = \max_j \{P_j\} \quad (4)$$

Для оценки качества классификации используем показатель достоверности  $G$ :

$$G = \frac{\sum_{g=1}^{N_p} \delta(\mathbf{Y}_g^*, \mathbf{Y}_g)}{N_p}, G \geq G_{\text{зад.}}; \quad (5)$$

$$\delta(\mathbf{Y}_g^*, \mathbf{Y}_g) = \begin{cases} 1, \mathbf{Y}_g = \mathbf{Y}_g^*; \\ 0, \mathbf{Y}_g \neq \mathbf{Y}_g^*, \end{cases}$$

где  $\mathbf{Y}_g^*$  — код (номер) шаблона ГТО на выходе классификатора при подаче на его вход искаженного образа ГТО  $\hat{\mathbf{S}}_g$ ;  $\mathbf{Y}_g$  — истинное значение кода шаблона ГТО при подаче на его вход искаженного образа  $\hat{\mathbf{S}}_g$ ;  $g$  — номер тестового примера ГТО,  $g=1 \dots N_p$ ;  $N_p$  — количество примеров ГТО, используемых для тестирования;  $G_{\text{зад.}}$  — заданное значение достоверности классификации;  $\delta(\mathbf{Y}_g^*, \mathbf{Y}_g)$  — символ Кронекера.

Рассмотрим в качестве признаков ГТО попарные расстояния между элементами (евклидовы расстояния), а также углы, вычисляемые на основе теоремы косинусов с использованием полученных расстояний [12, 13].

**3. Формирование базы шаблонов ГТО.** В качестве шаблонов ГТО будем рассматривать наборы элементов различного типа: 1, 2, 3... (рисунок 3). Положение элементов ГТО задано координатами  $x_i, y_i$ ,  $i=1 \dots n$ ,  $i$  — номер элемента,  $n$  — количество элементов. Величина  $n$  для различных шаблонов ГТО может различаться. Количество шаблонов  $N$  определяется требованиями к задаче классификации. Каждый шаблон можно представить в виде совокупности признаков, размещенных, к примеру, в виде таблицы.

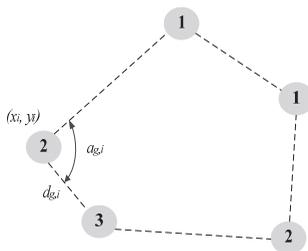


Рис. 3. Шаблон ГТО

Для формирования базы шаблонов выполним следующую последовательность действий:

*Шаг 1.* Сформируем шаблоны ГТО  $S_j$  в виде набора элементов различного типа: 1, 2, 3, ... (рисунок 3).

*Шаг 2.* Выбираем шаблон ГТО  $S_j$  и формируем последовательности элементов различной длины: «11232», «1122», «123»,... Для каждого  $i$ -го элемента вычисляем расстояния  $d_{j,i}$  и углы  $\alpha_{j,i}$ , а также ставим в соответствие минимальную и максимальную границу признаков  $d_{j,i,\min}$ ,  $\alpha_{j,i,\min}$ ,  $d_{j,i,\max}$ ,  $\alpha_{j,i,\max}$ , используя дополнительные экспертные знания о предметной области [13]. Например:  $d_{j,i,\min} = 0$ ,  $\alpha_{j,i,\min} = \alpha_{j,i} - 45^\circ$ ,  $d_{j,i,\max} = 2d_{j,i}$ ,  $\alpha_{j,i,\max} = \alpha_{j,i} + 45^\circ$ . Углы  $\alpha_{j,i}$  определяем на основе сторон образованных элементами ГТО треугольников по теореме косинусов. Таким образом, описываем каждый шаблон  $S_j$  последовательностями элементов и соответствующих им строк в массиве значений признаков:  $(d_{j,i}, \alpha_{j,i}, d_{j,i,\min}, \alpha_{j,i,\min}, d_{j,i,\max}, \alpha_{j,i,\max})$ .

*Шаг 3.* Выполним шаг 2 для каждого  $j$ -го шаблона и пополним данными таблицу 1.

Таблица 1. Пример базы шаблонов ГТО

Шаблоны ГТО $S_j$	Последовательности $i$ -х элементов ГТО		Признаки $d_{j,i}$	Признаки $\alpha_{j,i}$	Граничные значения признаков				Доп. эксперт. правила $\Lambda_{j,q}$
	Номер послед. $q$	Послед.			$d_{j,i,\min}$	$\alpha_{j,i,\min}$	$d_{j,i,\max}$	$\alpha_{j,i,\max}$	
$S_1$	1	123123	5,5,5...	30,40,23, ...	0,0,0, ...	$\alpha-45$	10,10, 10,...	$\alpha+45$	$\Lambda_{1,1}$
	2	1231	10,3,5...	30,50,23, ...	0,0,0, ...	$\alpha-45$	20,6, 10,...	$\alpha+45$	$\Lambda_{1,2}$
	...	...	...	...	...	...	...	...	...
$S_2$	1	122122	10,5,5...	10,40,25, ...	0,0,0, ...	$\alpha-45$	20,10, 10,...	$\alpha+45$	$\Lambda_{2,1}$
	2	2212	10,4,5...	20,50,2, ...	0,0,0, ...	$\alpha-45$	20,8, 10,...	$\alpha+45$	–
	3	1221	10,2,5...	30,60,8, ...	0,0,0, ...	$\alpha-45$	20,4, 10,...	$\alpha+45$	$\Lambda_{2,3}$
	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
$S_N$	...	...	...	...	...	...	...	...	...

Количество последовательностей  $N_{\text{посл.}}$  по каждому шаблону связано с количеством элементов ГТО  $n$  следующим образом:

$$N_{\text{посл.}} = \sum_{i=2}^n \frac{n!}{i!(n-i)!} = 2^n - n - 1. \quad (6)$$

Классификация ГТО на основе представленной в таблице информации сводится к решению задачи оптимального поиска в базе данных большой размерности. С учетом возможностей современных ПЭВМ осуществить поиск оптимального решения в представленной таблице за приемлемое время при использовании небольших ГТО ( $n \leq 20$ ) не представляет большого труда. Учитывая, что при описании сложных ГТО ( $n > 100$ ) можно ограничиться рассмотрением небольшого количества ключевых объектов, данный подход может быть применим для решения широкого класса задач. Тем не менее при больших  $n$ , значимой проблемой может стать низкая оперативность решения задачи классификации.

Оперативность можно повысить, ограничившись рассмотрением последовательностей из двух и трех элементов ГТО. Причем, если в первом случае в качестве признаков классификации ГТО могут быть использованы лишь попарные расстояния между элементами (таблица 2), то во втором случае их количество можно расширить за счет рассмотрения углов между образованными сторонами треугольников (таблица 3), что должно благоприятно отразиться на достоверности классификации.

Количество последовательностей  $N_{\text{посл.}}$  по каждому шаблону в таблице 2 связано с количеством элементов ГТО  $n$  следующим образом:

$$N_{\text{посл.}} = \frac{n!}{2!(n-2)!}, \quad (7)$$

а в таблице 3 соответственно:

$$N_{\text{посл.}} = \frac{n!}{3!(n-3)!}. \quad (8)$$

Таблица 2. Пример базы шаблонов ГТО

Шаблоны ГТО $S_j$	Последовательности $i$ -х элементов ГТО		Признаки $d_{j,i}$	Граничные значения признаков		Доп. эксперт. правила $\Lambda_{j,q}$
	Номер послед. $q$	Послед.		$d_{j,i,min}$	$d_{j,i,max}$	
$S_1$	1	12	5	0	10	$\Lambda_{1,1}$
	2	13	10	0	20	$\Lambda_{1,2}$
	3	23	10	0	20	–
	4	14	5	0	10	–
	...		...	...	...	....
$S_2$	1	12	6	0	12	$\Lambda_{2,1}$
	2	13	16	0	32	–
	...		...	...	...	...
...	...		...	...	...	...
$S_N$	...		...	...	...	...

Таблица 3. Пример базы шаблонов ГТО

Шаблоны ГТО $S_j$	Последовательности элементов ГТО		Признаки $d_{j,i}$	Признаки $\alpha_{j,i}$	Граничные значения признаков				Доп. эксперт. правила $\Lambda_{j,q}$
	Номер послед. $q$	Послед.			$d_{j,i,min}$	$\alpha_{j,i,min}$	$d_{j,i,max}$	$\alpha_{j,i,max}$	
$S_1$	1	123	5,5,5	30	0,0,0	$\alpha-45$	10,10,10	$\alpha+45$	$\Lambda_{1,1}$
	2	124	10,3,5	50	0,0,0	$\alpha-45$	10,6,10	$\alpha+45$	$\Lambda_{1,2}$
	...		...	...	...	...	...	...	....
$S_2$	1	123	10,5,5	40	0,0,0	$\alpha-45$	20,10,10	$\alpha+45$	$\Lambda_{2,1}$
	2	134	10,4,5	25	0,0,0	$\alpha-45$	20,8,10	$\alpha+45$	0
	...		...	...	...	...	...	...	...
...	...		...	...	...	...	...	...	...
$S_N$	...		...	...	...	...	...	...	...

Учитывая значительный размер представленных таблиц, в базе шаблонов ГТО могут храниться лишь координаты их элементов и экспертные правила, а все представленные в таблицах 1-3 последовательности элементов различной длины, попарные расстояния, углы, а так-

же граничные значения признаков могут быть вычислены в процессе реализации алгоритма классификации.

В качестве экспертных правил могут выступать различные логические условия, позволяющие улучшить решение задачи классификации. К примеру, если элементы ГТО представляют собой разнотипные объекты, размещенные на местности, то в качестве таких правил могут быть использованы логические условия размещения данных объектов в зависимости от характера местности, что позволит уточнить координаты элементов ГТО и улучшить решение задачи классификации.

**4. Последовательный алгоритм распознавания искаженного образа и формирования вероятностной оценки отнесения его к множеству шаблонов ГТО.** Рассмотрим одну из реализаций алгоритма классификации ГТО в пространстве перечисленных выше признаков с учетом (1-5).

*Шаг 1.* Формируем набор искаженных неполных образов ГТО с ложными элементами из имеющейся базы шаблонов при различных среднеквадратичных отклонениях координат их элементов  $\sigma$  и углах поворота  $\varphi$ , задаваемых в пределах, указанных экспертом.

*Шаг 2.* Вводим переменную  $g$  – номер искаженного образа ГТО,  $g = 1$ . Выбираем  $\hat{S}_g$  (рисунок 4).

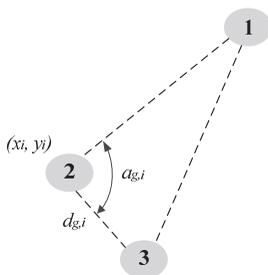


Рис. 4. Пример искаженного образа ГТО

*Шаг 3.* Вводим переменную  $j$  — номер анализируемого шаблона,  $j = 1$ . Выбираем шаблон  $S_j$  из базы шаблонов.

*Шаг 4.* Вводим переменные:  $r$  — номер анализируемой последовательности искаженного образа ГТО,  $r=1 \dots N_r$ , где  $N_r$  — количество последовательностей искаженного образа, присваиваем  $r=0$ ;  $q$  — номер анализируемой последовательности шаблона  $S_j$ ,  $q=1 \dots N_{j,q}$ , где  $N_{j,q}$  — количество последовательностей шаблона  $S_j$ , присваиваем  $q=1$ ;  $P_{j,r}$  — максимальную вероятность отнесения  $g$ -го искаженного образа  $\hat{S}_g$  к шаблону  $S_j$  по результатам анализа последовательности  $A$  и

всех последовательностей шаблона  $S_j$ ;  $P_{\max}$  — промежуточное значение вероятности  $P_{j,r}$ , присваиваем  $P_{\max}=0$ . Выбираем  $r$ -ю последовательность элементов искаженного образа ГТО, обозначаем ее  $A$ , определяем признаки  $(d_{g,i}, \alpha_{g,i})$  последовательности  $A$ .

*Шаг 5.* Выбираем  $q$ -ю последовательность из множества последовательностей шаблона  $S_j$ , обозначаем ее  $B$ .

*Шаг 6.* Сравниваем состав последовательностей  $B$  и  $A$ . Если они различаются, то увеличиваем  $q$  и переходим на шаг 5, иначе — переходим на шаг 7.

*Шаг 7.* Определяем для каждого перехода последовательности  $A$  вероятности  $(P_{d_{j,r,q,i}}, P_{\alpha_{j,r,q,i}})$  — отнесения его к шаблону  $S_j$  по каждому из признаков  $(d_{g,i}, \alpha_{g,i})$  на основе треугольных функций распределения вероятностей (рис. 2). Вероятности  $(P_{d_{j,r,q,i}}, P_{\alpha_{j,r,q,i}})$  могут быть уточнены путем использования корректирующих экспертных правил  $L_{j,q}$ .

*Шаг 8.* Определяем вероятности  $P_{j,r,q}$  отнесения образа  $\hat{S}_g$  к шаблону  $S_j$  по результатам анализа последовательностей  $A$  и  $B$ :

$$\begin{aligned} P_{j,r,q} &= \min_i \min_{d,\alpha} \{P_{d_{j,r,q,i}}, P_{\alpha_{j,r,q,i}}\} \\ P_{j,r,q,\max} &= \max\{P_{j,r,q}, P_{\max}\} \\ P_{\max} &= P_{j,r,q,\max} \end{aligned} \quad (9)$$

где  $P_{j,r,q,\max}$  — промежуточное значение вероятности  $P_{\max}$ . Если рассмотрены не все  $N_{j,q}$  последовательностей шаблона  $S_j$ , то увеличиваем  $q$  и переходим на шаг 5, иначе — определяем вероятность отнесения образа  $\hat{S}_g$  к шаблону  $S_j$  по результатам анализа последовательности  $A$  и всех  $N_{j,q}$  последовательностей шаблона:

$$P_{j,r} = P_{\max} \cdot \quad (10)$$

*Шаг 9.* Если рассмотрены не все  $N_r$  последовательностей искаженного образа  $\hat{S}_g$ , то увеличиваем  $r$  и переходим на шаг 4, иначе — определяем вероятность отнесения искаженного образа  $\hat{S}_g$  к шаблону  $S_j$  по результатам анализа всех  $N_r$  последовательностей  $\hat{S}_g$  и  $N_{j,q}$  последовательностей шаблона  $S_j$ :

$$P_j = \frac{1}{N_r} \sum_{r=1}^{N_r} P_{j,r}. \quad (11)$$

*Шаг 10.* Если рассмотрены не все  $N_j$  шаблонов, то увеличиваем  $j$  и выполняем шаги 3–9, иначе — нормируем, при необходимости, полученные вероятности  $\{P_j\}$ :

$$P_j = \frac{P_j}{\sum_{j=1}^{N_j} P_j}, \forall j, \quad (12)$$

ранжируем множество вероятностей  $\{P_j\}$ , принимаем решение о наиболее вероятном шаблоне с номером  $\xi$ :

$$Y_g = \xi \left| P_\xi = \max_j \{P_j\} \right. \quad (13)$$

*Шаг 11.* Выполняем шаги 2-10 для всех искаженных образов  $\hat{S}_g$  из сформированной выборки примеров. Оцениваем достоверность классификации  $G$  в соответствии с (5).

*Шаг 12.* Если  $G \geq G_{\text{зад}}$ , то заканчиваем выполнение алгоритма. Сформированная база шаблонов таблицы 2-3 с признаками и представленный алгоритм могут быть использованы для автоматической классификации новых искаженных образов. Иначе, если  $G < G_{\text{зад}}$ , осуществляем коррекцию таблиц 2-3.

Функцию принадлежности, а следовательно, и закон распределения вероятностей можно аппроксимировать полиномиальной зависимостью и набором коэффициентов. Используя статистические данные, можно получить коэффициенты этих полиномов с помощью метода наименьших квадратов таким образом, чтобы результат классификации был наиболее достоверным [20, 21].

При накоплении вероятностей (11) необходимо учесть различие шаблонов по количеству элементов: если имеется два шаблона ГТО с разным количеством элементов и одинаковой вероятностью отнесения  $P_j$  к искаженному образу ГТО, то нужно отдать предпочтение шаблону с меньшим количеством элементов. Реализовать данное логическое условие можно путем умножения значения вероятности отнесения  $P_j$  на множитель  $(n_j/n_m)$ , где  $n_j$  — количество элементов в искаженном образе ГТО,  $n_m$  — количество элементов в анализируемом шаблоне. С другой стороны, данный множитель можно рассматривать как априорную вероятность того, что обнаруженные элементы искаженного ГТО

являются элементами именно  $j$ -го шаблона. Данное условие является одним из примеров экспертных правил  $\Lambda_{j,q}$ , используемых в алгоритме.

**5. Результаты моделирования.** Рассмотренный алгоритм использован для решения задачи классификации искаженных образов наземных ГТО в виде набора элементов различного типа с прямоугольными координатами, полученных по результатам наблюдения в оптическом диапазоне. Элементы ГТО представляли собой объекты 12 типов, причем каждый тип кодировался числом от 1 до 12. Всего имелось 11 шаблонов ГТО, образующих базу шаблонов. Количество элементов в шаблонах ГТО составляло от 2 до 14. Особенностью некоторых шаблонов являлось наличие одинаковых фрагментов из нескольких элементов. Предъявлялись следующие требования к достоверности классификации ГТО:  $G \geq G_{\text{зад.}}$ ,  $G_{\text{зад.}} = 0.75$ .

На рисунке 5 представлены изображения одного из шаблонов ГТО под номером «б» и соответствующего ему искаженного образа при наличии 40% элементов ГТО, из которых 20 % — ложных, среднеквадратическом отклонении координат элементов ГТО  $\sigma = 30$  м и угле поворота  $\varphi = 140$  град.

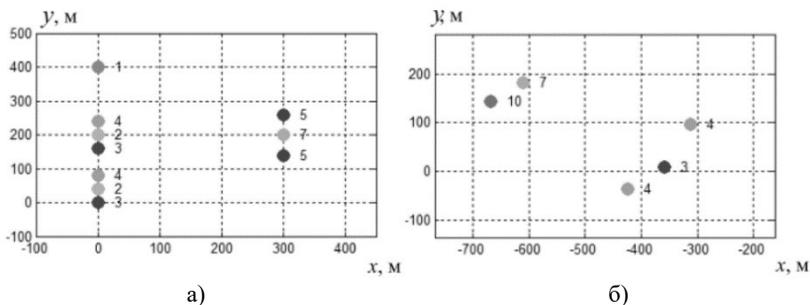


Рис. 5. Изображение ГТО: а) шаблона; б) искаженного образа

Анализ рисунка 6 показывает, что увеличение количества признаков ГТО часто позволяет получить более выраженный пик зависимости вероятности  $P_j$  от номера шаблона  $j$ , что, в конечном итоге, положительным образом отражается на повышении достоверности классификации  $G$  в широком диапазоне варьируемых параметров наблюдения.

В качестве признаков классификации были использованы: а) попарные расстояния между каждыми тремя элементами ГТО и один из углов образованного треугольника; б) попарные расстояния между элементами ГТО. В результате работы рассмотренного алгоритма получен набор вероятностей  $P_j$  — отнесения искаженного образа под номером «б» к каждому из 11 шаблонов (рисунок 6).

Данный вывод подтверждают зависимости, представленные на рисунках 7 и 8, полученные по результатам статистических экспериментов на всей базе шаблонов при изменении относительного количества пропущенных  $v$  и ложных элементов  $w$  в искаженных образах ГТО из базы, а также среднеквадратического отклонения координат элементов  $\sigma$ .

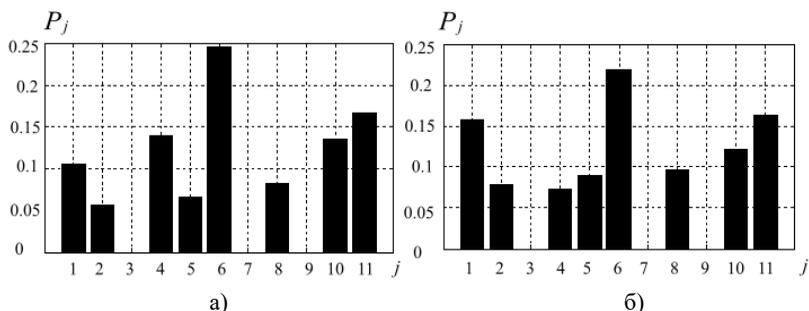


Рис. 6. Зависимость вероятности отнесения искаженного образа ГТО под номером «б» к каждому из 11 базовых шаблонов с использованием в качестве признаков: а) попарных расстояний между каждым тремя элементами ГТО и одного из углов образованного треугольника (таблица 2); б) попарных расстояний между элементами ГТО (таблица 3)

Параметры  $v$  и  $w$  определяются следующим образом:

$$v = \frac{n_{\text{ш}} - n_j}{n_{\text{ш}}} \times 100\%,$$

$$w = \frac{n_{\text{л}}}{n_j} \times 100\%,$$

где  $n_j$  — количество элементов в искаженном образе ГТО,  $n_{\text{л}}$  — количество ложных элементов в искаженном образе ГТО,  $n_{\text{ш}}$  — количество элементов в шаблоне ГТО, из которого синтезирован искаженный образ. Для указанных зависимостей определен доверительный интервал на значение достоверности  $G$  с доверительной вероятностью 0.99 по результатам 1000 экспериментов (отмечен на рисунках пунктиром).

Как показывают результаты, достоверность классификации значимо ухудшается, если в искаженном образе ГТО одновременно имеются пропущенные и ложные элементы, а также существует значимый разброс их координат. Тем не менее заданное значение достоверности

классификации  $G_{зад}=0.75$  было достигнуто в «плохих» ситуациях и достаточно широком диапазоне варьируемых параметров  $\nu$  и  $w$ .

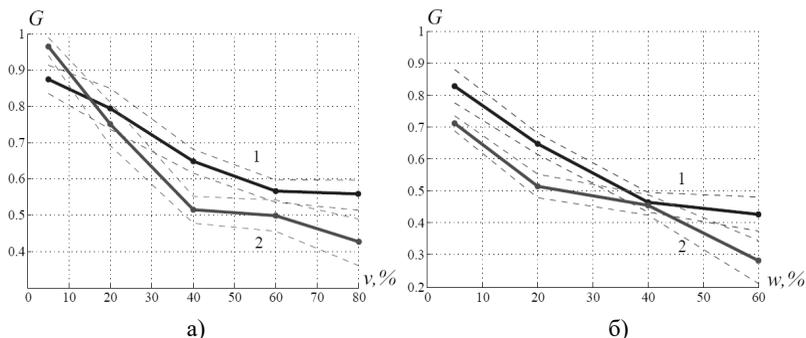


Рис. 7. Зависимости достоверности классификации ГТО на всей базе шаблонов  $G$ : а) от относительного количества пропущенных элементов в искаженных образах ГТО  $\nu$  при заданном среднеквадратическом отклонении  $\sigma = 30$ м и относительном количестве ложных элементов  $w = 20\%$ ; б) от относительного количества ложных элементов  $w$  в искаженных образах ГТО при заданном  $\sigma = 30$ м и  $\nu = 40\%$  и использовании в качестве признаков попарных расстояний между каждыми тремя элементами и одного из углов образованных треугольников (кривая 1) и попарных расстояний между элементами ГТО (кривая 2)

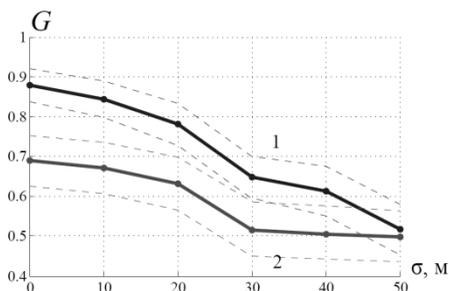


Рис. 8. Зависимость достоверности классификации ГТО на всей базе шаблонов  $G$  от среднеквадратического отклонения координат элементов  $\sigma$  при заданном относительном количестве пропущенных  $\nu = 40\%$  и ложных  $w = 20\%$  элементов с доверительными интервалами при использовании в качестве признаков попарных расстояний между каждыми тремя элементами и одного из углов образованных треугольников (кривая 1) и попарных расстояний между элементами ГТО (кривая 2)

Использование в качестве признаков попарных расстояний между каждыми тремя элементами и одного из углов образованных треугольников позволяет в целом получить лучший результат, чем при использовании только попарных расстояний. Второй подход может

ний. Второй подход может быть использован в тех случаях, когда в искаженном образе ГТО имеется в наличии лишь два элемента, а также в «хороших» ситуациях, характеризующихся малым количеством пропущенных и ложных элементов ГТО, а также малым значением среднеквадратического отклонения координат его элементов. Представленные результаты получены при значениях  $\sigma$ ,  $\nu$  и  $w$ , которые характерны для «плохих» ситуаций, что подтверждает возможность использования разработанного алгоритма и выбранных признаков классификации сложных ГТО в различных реальных применениях.

**6. Заключение.** Достоинством представленного алгоритма является возможность удобного включения в процедуру поиска не только дополнительных признаков, позволяющих уточнить класс ГТО, но и различных логических или эвристических правил, а также аналитических выражений, описывающих данную предметную область и позволяющих скорректировать нужным образом оценку значений признаков ГТО и повысить достоверность решения задачи классификации (1). Кроме того, имеется возможность встраивания в алгоритм адаптивных процедур, позволяющих уточнить результат классификации на множестве получаемых статистических данных. Основными направлениями совершенствования данного алгоритма являются: поиск эффективных признаков классификации ГТО; оценка влияния длины анализируемых последовательностей ГТО на достоверность классификации; оптимизация параметров используемых функций принадлежности; исследование различных способов получения оценок  $P_j$  и принятия решения о наиболее вероятном шаблоне.

### Литература

1. *Фурман Я.А., Роженцов А.А., Евдокимов А.О.* Распознавание групповых точечных объектов с неупорядоченными отметками // Автоматика. 2005. Т. 41. №1. С. 19–28.
2. *Фурман Я.А., Егояшина И.Л., Ерусланов Р.В.* Согласованная фильтрация зашумленных дискретных кватернионных сигналов // Журнал радиоэлектроники. 2012. № 3. С. 1–35.
3. *Роженцов А.А., Евдокимов А.О., Григорьев А.В.* Распознавание плоских изображений групповых точечных объектов при наличии ошибок обнаружения // Изв. высш. учебн. заведений: Приборостроение. 2006. Т. 49. № 4. С. 59–64.
4. *Мальцев Г.Н., Назаров А.В., Якимов В.Л.* Алгоритм реконструкции фазового пространства и его применение для создания прогнозных моделей // Информационно-управляющие системы. 2014. № 2. С. 33–39.
5. *Неронский Л.Б. и др.* Формирование точечных моделей объектов по комплексным РСА - изображениям // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 4. С. 158–164.
6. *Sharp R.* Jane's Fighting Ships, 1999-2000 // Jane's Information Group. 1990. 800 p.
7. *Дзенчарский Н.Л., Медведев М.В., Шлеймович М.П.* Поиск изображений с выделением особых точек на основе вейвлет-преобразования // Вестник Казанского государственного технологического университета. 2011. № 1. С. 131–135.
8. *Ипатов Ю.А., Кривецкий А.В.* Методы обнаружения и пространственной локализации групп точечных объектов // Кибернетика и программирование. 2014. № 6. С.17–25.
9. *Szeliski R.* Computer Vision: Algorithms and Applications // Springer. 2011. 812 p.

10. *Фурман Я.А.* Точечные поля и групповые объекты // М.: Физматлит. 2015. 440 с.
11. *Furman Y.A., Eruslanov R.V., Egoshina I.L.* Iterative Algorithm for angular matching of group point objects with apriori uncertainty of parameters // *Pattern recognition and image analysis*. 2013. vol. 23. no. 3. pp. 381–388.
12. *Воробьев С.Н., Лазарев И.В.* Алгоритм распознавания конфигураций звезд // *Информационно-управляющие системы*. 2008. №2. С. 2–8.
13. *Дубровкина М.В.* Векторно-нормализованный метод распознавания групповых точечных объектов произвольной формы // *Вестник Сумского государственного университета*. 2009. № 4. С. 32–38.
14. *Варшавский П.Р., Еремеев А.П.* Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений // *Искусственный интеллект и принятие решений*. 2009. №1. С. 45–57.
15. *Уздин Д.З.* О новом подходе в теории распознавания образов (состояний). Новые методы математической диагностики // М.: МАКС Пресс. 2012. 232 с.
16. *Осипов Г.С.* Методы искусственного интеллекта // М.: Физматлит. 2011. 296 с.
17. *Kandel A., Wyatt W.* Fuzzy sets, fuzzy algebra, and fuzzy statistics // *Proceedings of the IEEE*. 1978. vol. 66. no. 12. pp. 1619–1639.
18. *Зак Ю.А.* Принятие решений в условиях нечетких и размытых данных: Fuzzy-технологии // М.: Книжный дом «Либроком». 2013. 352 с.
19. *Смагин В.А., Парамонов И.Ю.* Вероятностный критерий оценивания нечеткой энтропии // *Информация и космос*. 2015. №2. С. 42–46.
20. *Бураков М.В., Брунов М.С.* Структурная идентификация нечеткой модели // *Труды СПИИРАН*. 2014. Вып. 3. С. 232–246.
21. *Ходашинский И. А.* Построение компактных и точных нечетких моделей на основе статистических информационных критериев // *Информатика и системы управления*. 2014. № 1(39). С. 99–107.

**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., д. 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., д. 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  
**CLASSIFICATION ALGORITHM OF GROUP POINT OBJECTS  
WITH UNORDERED ELEMENTS BASED ON CLOSENESS  
PROBABILITY MEASURE**

---

*Kaplin A.Y., Korotin A.A., Nazarov A.V., Yakimov V.L. Classification Algorithm of Group Point Objects with Unordered Elements based on Closeness Probability Measure.*

**Abstract.** The paper presents a classification algorithm of group point objects (GPO) based on the comparative analysis of fragments of distorted images and the GPO templates. The sequences of the GPO elements of different lengths are used as fragments. The paired and angular interdot distances are used as classification signs. The probability measure of closeness, set by the expert by means of the membership function and the distribution law of probability of discrete values of classified objects signs, is used in solving a classification task.

The algorithm includes the following stages: search and comparison of fragments composition of distorted images and the GPO templates; formation of a probable assessment of closeness of GPO distorted image and each template in space of the considered signs according to the analysis of each fragment; accumulation of the received probabilities on the basis of analysis results of all distorted image fragments; ranging of the received probabilities of classifying the distorted image as the GPO template; determination of the most probable template. The algorithm provides the possibility of specifying a GPO distorted image class using logical rules and analytical expressions of the considered data domain. The example and results of the algorithm application for solving a classification task of real GPO on the basis of the analysis of their fragments in the form of sequences from two and three elements are given.

**Keywords:** group point object, classification, probability measure of closeness.

---

**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) 251-3875, 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. Furman Ja.A., Rozhencov A.A., Evdokimov A.O. [Recognition of group point objects with ordered marks]. *Avtometriya – Avtometriya*. 2005. vol. 41. no. 1. pp. 19–28. (In Russ.).

2. Furman Ja.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.).
3. Rozhencov A.A., Evdokimov A.O., Grigor'ev A.V. [Recognition of flat images of group point objects in the presence of error detection]. *Izv. vyssh. uchebn. zavedenij: Priborostroenie – Proceedings of the higher educational institutions: Instrumentation*. 2006. vol. 49. no. 4. pp. 59–64. (In Russ.).
4. Maltsev G.N., Nazarov A.V., Yakimov V.L. [A reconstruction algorithm for a dynamic system phase space and its application for development of predictive models]. *Informacionno-upravljajushhie sistemy – Information and control systems*. 2014. vol. 2(69). pp. 33–39 (In Russ.).
5. Neronskiy L.B. et al. [Generation of object point models by SAR complex images]. *Sovremennye problemy distantsionnogo zondirovaniia Zemli iz kosmosa – Actual problems of remote sensing of the Earth from space*. 2010. vol. 7. no. 4. pp. 158–164. (In Russ.).
6. Sharp R. *Jane's Fighting Ships, 1999-2000*. Jane's Information Group. 1990. 800 p.
7. Dzencharkii N.L., Medvedev M.V., Shleimovich M.P. [Image search with the release of specific points on the basis of wavelet transform]. *Vestnik Kazanskogo gosudarstvennogo tekhnologicheskogo universiteta – Vestnik of the Kazan state technological university*. 2011. vol. 1. pp. 131–135. (In Russ.).
8. Ipatov Iu.A., Krevetskii A.V. [Methods of detection and spatial localization of groups of point objects]. *Kibernetika i programirovanie – Cybernetics and programming*. 2014. vol. 6. pp. 17–25. (In Russ.).
9. Szeliski R. *Computer Vision: Algorithms and Applications*. Springer. 2011. 812 p.
10. Furman Ja.A. *Tochechnye polia i gruppovye ob"ekty* [Point field and group objects]. Moscow: Fizmatlit Publ. 2015. 440 p. (In Russ.).
11. Furman Y.A., Eruslanov R.V., Egoshina I.L. [Iterative algorithm for angular matching of group point objects with apriori uncertainty of parameters]. *Pattern recognition and image analysis*. 2013. vol. 23. no. 3. pp. 381–388.
12. Vorob'ev S.N., Lazarev I.V. [Configurations recognition algorithm stars]. *Informacionno-upravljajushhie sistemy – Information and control systems*. 2008. vol. 2. pp. 2–8. (In Russ.).
13. Dubrovkina M.V. [Vector-normalized method for detection of group point objects of arbitrary shape]. *Vestnik Sumskogo gosudarstvennogo universiteta – Bulletin of the Sумы state university*. 2009. vol. 4. pp. 32–38. (In Russ.).
14. Varshavskij P.R., Eremeev A.P. [Modelling of reasoning based on precedents in intelligent decision support systems]. *Iskusstvennyj intellekt i prinjatие reshenij – Artificial intelligence and decision-making*. 2009. no. 1. pp. 45–57 (In Russ.).
15. Uzdin D.Z. *O novom podkhode v teorii raspoznavaniia obrazov (sostoianii). Nove metody matematicheskoi diagnostiki* [A New Approach to the Theory of Pattern Recognition (States). New Methods of Mathematical Diagnostics]. Moscow: MAKSPress Publ. 2012. 232 p. (In Russ.).
16. Osipov G.S. *Metody iskusstvennogo intellekta* [Methods of artificial intelligence]. Moscow: Fizmatlit Publ. 2011. 296 p. (In Russ.).
17. Kandel A., Byatt W. Fuzzy sets, fuzzy algebra, and fuzzy statistics. *Proceedings of the IEEE*. 1978. vol. 66. no. 12. pp. 1619–1639.
18. Zak Ju.A. *Prinjatие reshenij v usloviyah nechetkih i razmytyh dannyh: Fuzzy-tehnologii* [Decision-making in a fuzzy and blurry data: Fuzzy-technology]. Moscow: Knizhnyj dom «Librokom» Publ. 2013. 352 p. (In Russ.).
19. Smagin V.A., Paramonov I.Ju. [Probabilistic estimation of fuzzy entropy criterion] // *Informacija i kosmos – Information and Space*. 2015. no. 2. p. 42–46. (In Russ.).
20. Burakov M.V., Brunov M.S. [Structure identification of fuzzy model]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2014. vol. 3. pp. 232–246. (In Russ.).
21. Khodashinskii I. A. [Building compact and powerful fuzzy models based on statistical information criteria]. *Informatika i sistemy upravleniia – Information and control systems*. 2014. vol. 1(39). pp. 99–107. (In Russ.).