

УДК 621.391

И.К. ВАСИЛЬЕВА, А.В. ПОПОВ

*Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", Украина***СИНТЕЗ АЛГОРИТМА РАСПОЗНАВАНИЯ ГИДРОМЕТЕОРОЛОГИЧЕСКИХ  
ОБЪЕКТОВ ПО РАДИОЛОКАЦИОННЫМ ДАННЫМ**

Синтезируется алгоритм параллельного принятия решений с позиций теории информации и статистической теории принятия решений на векторном пространстве поляризационных параметров. На первой ступени процедуры различения в результате применения алгоритма с накоплением информации и отбраковыванием классов по критерию достаточного количества информации формируется множество решений. На второй ступени алгоритма решение в пользу определенного класса принимается по критерию максимума апостериорной вероятности. Алгоритм апробирован на данных моделирования радиолокационных сигналов, отраженных специфическими классами гидрометеорологических объектов.

**классификация, мера информации, апостериорная вероятность, поляризационные параметры****Введение**

Одним из путей решения проблемы обеспечения безопасности полетов самолетов и вертолетов гражданской авиации в сложных метеоусловиях является разработка эффективных алгоритмов определения характеристик гидрометеорологических образований (ГМО) по данным дистанционного зондирования. При синтезе алгоритма распознавания ГМО встают задачи выбора информативных параметров сигналов, обеспечивающих классификацию заданного множества сигналов, и построения решающего правила в пространстве информативных параметров. Исходя из необходимости обеспечения оперативного обнаружения и распознавания опасных для полетов ГМО (мощных кучево-дождевых облаков, областей формирования града, грозоопасных зон), к алгоритму предъявляются требования минимизации времени на принятие решения и максимизации достоверности распознавания классов ГМО. В качестве объектов радиолокационного наблюдения рассматриваются следующие классы ГМО [1] (в скобках указан соответствующий номер класса,  $k$ ): облака капельные – не грозовые, CD (1), кучево-дождевые, Сб (2), грозовые, Сб-s (3), облака кристаллические – ледяные иглы, CN (4), и пластинки, CP (5), осадки – дождь (6), гроза (7), ледяная крупа (8), град (9) и снег (10).

Классификация, или распознавание (при неполной априорной информации о классах), представляет собой отнесение исследуемого объекта, заданного в виде совокупности наблюдений, к одному из взаимоисключающих классов. Существует однозначное отображение выборки наблюдений, являющейся конечным числовым множеством  $\{X\}$  на множество классов  $\{A\} = \{a_1, a_2, \dots, a_k\}$ ,  $\{A\} \leftarrow \{X\}$ . Если заменить классы их номерами, то классификацию можно представить как отображение наблюдений на конечное множество натуральных чисел  $\{1, 2, \dots, K\} \leftarrow \{X\}$ . Ввиду числовой природы множеств такое отображение отождествляется с функцией  $k = d(X)$ , принимающей целочисленные значения  $k = 1, 2, \dots, K$ .

**Целью работы** являлась оптимизация преобразования  $d(X)$  по критерию минимума суммарного количества наблюдений (определяемого объемом обучающих и контрольной выборок и размерностью признакового пространства), необходимого для обеспечения требуемого уровня достоверности распознавания грозо- и градоопасных ГМО.

При синтезе алгоритма распознавания применялась программа моделирования полей обратного однократного рассеяния в приближении Рэлея поли-

дисперсных, хаотически ориентированных сферонидальных частиц ГМО [2].

### Обучение

Обучение является составной частью процесса классификации и имеет своей целью формирование эталонных описаний классов, форма которых определяется способом их использования в решающих правилах. В теории статистических решений все виды решающих правил основаны на формировании отношения правдоподобия  $L$  и его сравнения с порогом  $c$ , значение которого определяется выбранным критерием качества:

$$L = \frac{f_n(x_1, \dots, x_n | a_2)}{f_n(x_1, \dots, x_n | a_1)} \geq c, \quad (1)$$

где  $f_n(x_1, \dots, x_n | a_j)$  – условная совместная  $n$ -мерная плотность вероятности выборочных значений  $x_1, \dots, x_n$  при условии их принадлежности к классу  $a_j$ . Поскольку в статистическом распознавании эти плотности, в принципе, не известны, то в решающее правило (1) подставляются оценки плотности вероятности  $\hat{f}_n(x_1, \dots, x_n | a_j)$ , получаемые в процессе обучения. Таким образом, в решающем правиле с порогом  $c$  сравнивается оценка отношения правдоподобия  $\hat{L}$ .

Решить проблемы, связанные с недостаточностью статистической информации и отсутствием в полном объеме экспериментальных данных ДЗ для специфических типов ГМО позволяет математическое моделирование радиолокационных сигналов, отраженных от ГМО. С помощью программы моделирования [2] для каждого класса ГМО определялась полная поляризационная матрица рассеяния  $\hat{S}$  [4], на основании которой рассчитывались элементы ковариационной матрицы рассеяния  $\hat{R}$  [5] и другие поляриметрические признаки (ПП) [2, 5, 6]. Для непараметрического оценивания законов распределения ПП гистограммным методом объем выборки

составлял 20000 отсчетов. В качестве статистической модели ПП было выбрано Sb-распределение Джонсона [7] как наиболее универсальная аппроксимация оценок плотностей вероятностей ПП. Оценки параметров законов распределения Джонсона  $\varepsilon$ ,  $\lambda$ ,  $\eta$ ,  $\gamma$ , найденные по методике [8], являлись эталонными описаниями классов. Статистическая модель  $N$ -мерного вектора ПП  $\vec{x}$  для каждого класса объектов  $a_k$  имела вид [9]:

$$f(\vec{x}(t_j) | a) = (2\pi)^{-MN/2} \|R\|^{-1/2} \times \prod_{j=1}^M \prod_{i=1}^N \left( \frac{\eta_i \lambda_i}{(x_i(t_j) - \varepsilon_i)(\varepsilon_i + \lambda_i - x_i(t_j))} \right) \times \exp \left[ -\frac{1}{2} \sum_{j,k=1}^M \sum_{i,l=1}^N R_{ijkl}^{-1} \left( \gamma_i + \eta_i \ln \frac{x_i(t_j) - \varepsilon_i}{\varepsilon_i + \lambda_i - x_i(t_j)} \right) \times \left( \gamma_l + \eta_l \ln \frac{x_l(t_k) - \varepsilon_l}{\varepsilon_l + \lambda_l - x_l(t_k)} \right) \right], \quad (2)$$

где  $M$  – количество отсчетов;  $R$  – корреляционная матрица, элементы которой определяются как

$$\hat{R}_{il} = \frac{1}{M} \sum_{j=1}^M z_i(t_j) z_l(t_j),$$

где  $z = \gamma + \eta \ln[(x - \varepsilon)/(\varepsilon + \lambda - x)]$  – нелинейное преобразование случайной величины  $x$  с Sb-распределением Джонсона в случайную величину  $z$ , распределенную по нормальному закону с нулевым средним и единичной дисперсией.

### Выбор признакового пространства

В общем виде задача формирования признакового пространства ставится, исходя из требований к распознающей системе в целом. При этом от размерности  $p$  признакового пространства в значительной степени зависит вычислительная сложность процедур обучения и принятия решения, достоверность классификации, затраты на измерение необходимых характеристик объектов. Первоначальный набор признаков формируется из числа доступных измерению характеристик объекта  $Y_1, \dots, Y_g$ , отражающих его наиболее существенные для классифи-

кации свойства. В качестве такого набора рекомендуется использовать множество ПП, т.к. они тесно связаны со структурными характеристиками ГМО (формой, размерами, фазовым состоянием частиц, их преимущественной ориентацией). На следующем этапе формируется новый набор  $X_1, \dots, X_p$ ;  $p < g$ . Традиционные способы формирования новых признаков в условиях полного априорного знания основаны на максимизации некоторой функции  $J(Y_1, \dots, Y_g)$ , называемой критерием и обычно понимаемой как некоторое расстояние между классами в признаковом пространстве с координатами  $Y_1, \dots, Y_g$ . В других случаях критерий  $J(Y_1, \dots, Y_g)$  выражает диаметр или объем области, занимаемый классом в признаковом пространстве, и новые признаки формируются путем минимизации критерия. В дискриминантном анализе принят критерий, сущность которого состоит в совместной минимизации внутриклассового разброса наблюдений и максимизации межклассового расстояния [3, 10].

В теории информации в качестве критерия используют дивергенцию [11], представляющую собой меру различимости множества классов по вектору ПП. Локальная дивергенция [12] определяется как мера различимости класса  $a_k$  на фоне остальных классов из множества  $A = \{a_k\}$ :

$$J(k) = \int_R (f(\bar{x}|a_k) - f(\bar{x}|\tilde{a}_k)) \ln \frac{f(\bar{x}|a_k)}{f(\bar{x}|\tilde{a}_k)} d\bar{x}, \quad (3)$$

где  $R$  – выборочное пространство;  $f(\bar{x}|a_k)$  – плотность вероятности реализации случайного вектора ПП  $\bar{x}$  для класса  $a_k$ ,  $f(\bar{x}|\tilde{a}_k)$  – смесь плотностей вероятностей для всех классов, за исключением  $a_k$ :

$$f(\bar{x}|\tilde{a}_k) = \frac{1}{1 - P(a_k)} \sum_{i \neq k=1}^N f(a_i, \bar{x}),$$

где  $P(a_k)$  – априорная вероятность класса  $a_k$ .

Информативность ПП для распознавания класса  $a_k$  определяется сравнением  $J(k)$  с величиной минимально достаточной дивергенции  $\hat{\Theta}_k$  (МДД) [12]:

$$\hat{\Theta}_k = \left[ \frac{[1 - P(a_k)]^2}{P(a_k)^2} - 1 \right] \ln \frac{[1 - P(a_k)]^2}{P(a_k)^2}. \quad (4)$$

При  $J(k) \geq \hat{\Theta}_k$  возможно распознать класс  $a_k$  на фоне других классов по данному ПП.

Среднее значение  $J(k)$  представляет собой меру различимости множества классов  $\{A\}$  на совокупности ПП  $\{\xi\}$  (обобщенную дивергенцию [12]):

$$J(\{\xi\}) = \sum_{k=1}^N P(a_k) J(k) \quad (5)$$

и используется для сравнения информативности различных подмножеств совокупности ПП сигнала.

В работе [1] были проанализированы информационные возможности ПП ГМО по критерию МДД. При этом для ПП, сильно зависящих от структурных факторов ГМО, использовались статистические модели, полученные посредством интегрирования по диапазону возможных значений влияющего фактора. По результатам анализа выбрано 13 ПП [2], имеющих наибольшие показатели дивергенции:  $Z_h$  – радиолокационная отражаемость;  $Y_2, |Y_3|$  – структурные параметры, характеризующие геометрические и диэлектрические свойства частиц;  $H_{dr}$  – дифференциальная отражаемость градового сигнала;  $Z_{dr}$  – дифференциальная отражаемость;  $TAN$  – коэффициент анизотропии;  $HGM$  – характеристический угол объекта;  $AN$  – параметр асимметрии частиц;  $I_1, |I_2|$  – параметры, учитывающие особенности распределения частиц по размерам, формам и по углу наклона;  $ELV, ORH$  – углы эллиптичности и ориентации для  $H$  и  $V$ - поляризованных сигналов, соответственно,  $DPS$  – дифференциальный фазовый сдвиг.

### Принятие решений

**Построение решающего правила.** Согласно теории статистических решений все решающие правила для  $K \geq 2$  классов основаны на сравнении отношений правдоподобия  $L_{kn}$  между собой или с

определенными порогами. При использовании байесовского, минимаксного критериев, либо критерия максимума апостериорной вероятности [3, 10] в решающее правило вводятся априорные данные, но знание априорных вероятностей классов и функции потерь являются в статистическом распознавании скорее исключением, чем правилом. Кроме того, средний риск мало пригоден для оценки качества классификации, когда важно знать ее достоверность. При отсутствии априорной информации о вероятностях состояний и потерях используются критерии Неймана-Пирсона, максимального правдоподобия и последовательного анализа Вальда [3, 10]. С позиции получаемого количества информации процедура принятия решения заключается в сравнении количества информации в пользу гипотезы  $H_k$  (принятие класса  $a_k$ ) с порогами  $h_k^{(K)}$  и  $s_k^{(K)}$ . Если  $I_k(\bar{x}) \geq h_k^{(K)}$ , то гипотеза  $H_k$  принимается, если  $I_k(\bar{x}) \leq s_k^{(K)}$ , то гипотеза  $H_k$  отвергается, если  $s_k^{(K)} < I_k(\bar{x}) < h_k^{(K)}$ , то решение о классе  $a_k$  не выносится. В зависимости от того, какая из мер количества информации используется в процедуре принятия решения, различают критерии Котельникова, Шеннона, Кульбака, Байеса и Фишера [13].

Количество информации – неопределенность, которую удастся снять при трансформации априорного распределения  $P_k$  в апостериорное  $Q_k$ :

$$Q_k = \frac{f(\bar{x}|a_k)P_k}{\sum_{i=1}^K f(\bar{x}|a_i)P_i}.$$

Мера количества информации Кульбака [11, 13]:

$$I_k = -\ln \frac{1-P_k}{P_k} + \ln \frac{1-Q_k}{Q_k}. \quad (6)$$

Минимально достаточное количество информации (МДКИ) – такое количество информации, которое необходимо, чтобы скомпенсировать неопределенность, содержащуюся в заданном классе сигналов. Уровень МДКИ (по Кульбаку) определяется как:

$$\psi_k = 2 \ln[(1-P_k)/P_k]. \quad (7)$$

Для случая равновероятного появления любого из заданных классов сигналов ( $P_k = 1/K$ ) выражение (7) принимает вид

$$\psi_k = 2 \ln(K-1). \quad (8)$$

Статистическая теория принятия решения, основанная на критериях стоимости принятия той или иной гипотезы, и теория информации в задаче распознавания по критерию получения максимума количества информации равнозначны, однако информационный способ анализа систем распознавания обеспечивает наиболее общую постановку задачи с четко определенными ограничениями, налагаемыми на решение, определяет структуру оптимальной системы и позволяет оценить качество ее работы.

**Методы принятия решения.** Если имеется фиксированное число  $\nu_0$  независимых наблюдений над совокупностью параметров  $\bar{\xi}(t)$  и при всех  $j \neq k = 1, \dots, N$  выполняется условие  $I_j(\bar{x}) < \psi_j$ , где  $\psi_j$  – МДКИ для класса  $a_j$ , тогда под непоследовательным анализом понимают такой анализ, который приводит хотя бы на одном из  $\nu_0$  этапов наблюдения к такому количеству информации о классе  $a_k$ , которое не меньше МДКИ  $I_k(\bar{x}) \geq \psi_k$  [13]. Под последовательным методом анализа понимают такой анализ, который приводит на  $\nu$ -м шаге наблюдения к количеству информации о  $k$ -м классе не меньше минимально достаточного  $I_k^{(\nu)}(\bar{x}) \geq \psi_k$  при условии, что на всех  $\nu-1$  этапах наблюдения получаемое количество информации было меньше минимально достаточного [13]. Таким образом, метод осуществляет принятие решения последовательно с учетом информации, накопленной на предыдущих интервалах наблюдения.

В алгоритме с накоплением информации решение выносится на основании критерия МДКИ [13]. Процедура принятия решения следующая: если  $I_k \geq \psi_k$ , то гипотеза  $H_k$  (о принятии класса  $a_k$ ) принимается, если  $I_k \leq -\psi_k$ , то гипотеза  $H_k$  от-

вергается, если  $-\psi_k < I_k < \psi_k$ , то решение о гипотезе  $H_k$  не выносится.

Для повышения гибкости процедуры разбиения выборочного пространства на множества информативности можно предположить, что априорные вероятности появления различных видов сигналов не остаются постоянными. Так, если на 1-м цикле наблюдений  $P_k^{(1)} = 1/K$  (полная апостериорная неопределенность), то на каждом следующем цикле в качестве априорной вероятности класса  $a_k$  можно принять значение апостериорной вероятности, вычисленное на предыдущем шаге:  $P_k^{(v)} = Q_k^{(v-1)}$ .

Видоизменение процедуры принятия решения заключается в перераспределении априорных вероятностей при исключении одного класса из дальнейшего рассмотрения. Предположим, что на 1-м цикле наблюдения было принято решение, исключающее из дальнейшего рассмотрения класс  $a_{r_1}$ . Тогда на 2-м цикле наблюдений априорная вероятность класса  $a_{r_1}$   $P_{r_1}^{(2)}$  равна 0. Оставшиеся  $K-1$  априорных вероятностей  $P_1^{(1)} \dots P_K^{(1)}$  необходимо пересчитать на новую группу вероятностей  $P_1^{(2)} \dots P_K^{(2)}$ , за исключением  $P_{r_1}^{(2)} = 0$ . Пусть на  $m-1$  цикле наблюдений исключается класс  $a_{r_{m-1}}$ . В этом случае образуется новая группа вероятностей  $P_k^{(m)}$ , которая должна удовлетворять условиям:

$$\sum_{k=1}^K P_k^{(m)} = 1, \quad P_k^{(m)} \leq 0,5 \quad (k = 1, \dots, K; \quad k \neq r_1, \dots, r_{m-1}).$$

В силу (8), обязательным является выполнение условия:  $K \geq 3$ , поэтому, когда число оставшихся классов равно 3, следует завершить процедуру распознавания с отбрасыванием классов и перейти к последовательному методу принятия решения.

Процесс принятия решения о наличии или исключении какого-либо класса рекомендуется принудительно затягивать на несколько циклов для уменьшения вероятности ошибочного решения,

обусловленной случайными выбросами ПП.

Поскольку при последовательном анализе размер выборки является случайной величиной, то даже при сравнительно малых средних значениях длительности процедуры возможны случаи недопустимо больших размеров выборки. Типичным примером компромиссного решения для распределения длительности процедуры является усеченный последовательный анализ, при котором устанавливается максимальное значение объема выборки  $n_{\max}$ , при достижении которого последовательная процедура заканчивается и соответствующее отношение правдоподобия сравнивается не с двумя порогами,  $c_1$  и  $c_2$ , а только с одним,  $c_{yc}$ , в результате чего обязательно принимается одно из решений.

### Анализ методов принятия решения

Анализ информативности ПП ГМО по критерию МДД показал неэффективность классического (непоследовательного) метода принятия решения. Для сравнительной оценки двух стратегий последовательной процедуры использовались массивы 13 указанных выше ПП; объем контрольной выборки для каждого класса составлял 19500 отсчетов. Среднее число отсчетов, необходимых для принятия решения  $\bar{n}$ , и результаты распознавания по  $Z_{dr}$  и  $I_1$  приведены в табл. 1, 2. При переходе от алгоритма с накоплением информации к алгоритму с отбрасыванием классов вероятность правильной классификации существенно не меняется, а  $\bar{n}$  заметно уменьшается; таким образом, использование алгоритма с отбрасыванием классов позволяет сократить время на принятие решения, что является одним из основных требований при проектировании бортовых метеонавигационных систем. В то же время недостаточная информативность одиночных ПП обуславливает необходимость увеличения размерности признакового пространства. Для уменьшения длительности последовательной процедуры налагается ограничение на количество циклов: если по достижению  $n_{\max}$  решение в пользу какого-либо из 3-х классов

Таблица 1

Оценки вероятностей  $\hat{P}(\gamma_j|a_k)$  принятия решения в пользу  $j$ -го класса при условии, что для распознавания предъявлен класс  $a_k$ , для алгоритма с накоплением информации по одиночному ПП

ПП	$a_k \backslash \gamma_j$	$\hat{P}(\gamma_j a_k)$										$\bar{n}$
		CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	
$Z_{dr}$	CD	0,811	4,26e-2	0	0	0	0	0	0	1,42e-2	0,132	31
	Cb	0,47	0,474	0	0	0	0	8,80e-4	8,80e-4	5,37e-2	0	17
	Cb-St	0	0	0,989	0	9,93e-3	0	4,52e-4	4,52e-4	0	0	9
	CN	0	0	1,35e-2	4,48e-3	1,79e-2	0,408	0,417	0,139	0	0	87
	CP	0	0	0,712	0	0,272	0	1,39e-2	1,73e-3	0	0	17
	дождь	0	0	1,85e-2	2,31e-3	1,85e-2	0,157	0,677	0,127	0	0	45
	гроза	0	0	0,26	0	0,41	0	0,329	1,05e-3	0	0	21
	крупа	0	0	5,13e-2	0	2,32e-2	1,22e-3	0,498	0,421	4,88e-3	0	24
	град	0	5,65e-2	1,62e-2	0	3,23e-3	0	0,17	0,378	0,376	0	32
	снег	0,917	6,21e-2	0	0	0	0	0	0	1,49e-2	6,29e-3	15
$I_1$	CD	0,948	2,61e-2	0	0	0	0	0	0	0	2,55e-2	6
	Cb	1,46e-2	0,924	0	0	0	0	0	0	3,65e-4	6,09e-2	7
	Cb-St	0	0	0,956	0	3,66e-2	0	7,32e-3	0	0	0	20
	CN	0	0	0	0,269	0	0,654	0	0	6,99e-2	6,99e-3	68
	CP	0	0	0,303	0	0,625	0	7,24e-2	0	0	0	33
	дождь	0	0	0	1,84e-2	0	0,632	0	9,20e-3	0,34	0	60
	гроза	0	0	1,03e-2	0	5,16e-3	0	0,873	0,11	1,72e-3	0	34
	крупа	0	0	0	0	0	0,35	0	0,155	0,495	0	69
	град	0	0	0	2,70e-3	0	1,22e-2	0	0	0,869	0,116	26
	снег	0,857	9,77e-2	0	0	0	0	0	0	3,85e-4	4,50e-2	7

Таблица 2

Оценки вероятностей  $\hat{P}(\gamma_j|a_k)$  принятия решения в пользу  $j$ -го класса при условии, что для распознавания предъявлен класс  $a_k$ , для последовательного алгоритма с отбрасыванием классов по одиночному ПП

ПП	$a_k \backslash \gamma_j$	$\hat{P}(\gamma_j a_k)$											$\bar{n}$
		CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	смесь	
$Z_{dr}$	CD	0,679	9,7e-2	0	0	0	0	0	0	5,4e-3	0,198	2,1e-2	18
	Cb	0,345	0,572	6,2e-4	0	0	0	6,2e-4	1,2e-3	5,6e-2	2,4e-2	1,2e-3	12
	Cb-St	0	0	0,97	0	2,3e-2	0	3,7e-3	3,0e-3	0	0	0	7
	CN	0	0	1,8e-3	1,8e-3	1,6e-2	7,8e-2	0,152	0,474	1,8e-3	0	0,275	34
	CP	0	0	0,727	0	0,229	0	4,2e-2	2,0e-3	0	0	9,8e-4	10
	дождь	0	3,9e-3	1,2e-2	3,9e-3	2,6e-2	8,3e-2	0,273	0,513	2,6e-3	0	8,2e-2	25
	гроза	0	0	0,319	5,8e-4	0,314	1,2e-3	0,354	1,1e-2	0	0	0	11
	крупа	0	5,9e-3	3,4e-2	0	2,9e-2	4,2e-3	0,321	0,592	1,0e-2	0	4,2e-3	16
	град	1,1e-3	5,9e-2	4,5e-3	5,7e-3	7,9e-3	1,5e-2	6,1e-2	0,515	0,33	1,1e-3	0	22
	снег	0,809	0,113	0	0	0	0	0	0	8,8e-3	7,0e-2	0	11
$I_1$	CD	0,85	0,108	0	0	0	0	0	0	4,2e-2	0	5	
	Cb	7,4e-2	0,804	0	0	0	0	0	0	8,4e-4	0,121	0	5
	Cb-St	0	0	0,882	0	3,8e-2	0	5,1e-2	3,6e-3	0	0	2,6e-2	14
	CN	0	0	0	0,224	0	0,327	0	1,5e-2	0,308	1,3e-2	0,113	28
	CP	0	0	0,373	0	0,359	0	0,212	1,6e-2	0	0	4,1e-2	18
	дождь	0	0	1,4e-3	2,8e-2	0	0,27	1,4e-3	3,4e-2	0,548	1,8e-2	9,8e-2	28
	гроза	0	0	2,2e-2	0	5,9e-3	9,8e-4	0,656	0,26	1,6e-2	9,8e-4	3,8e-2	19
	крупа	0	0	0	4,3e-3	0	0,163	0	0,229	0,494	1,0e-2	9,9e-2	28
	град	0	2,5e-3	0	2,5e-3	0	1,0e-3	0	4,2e-3	0,684	0,285	1,2e-2	16
	снег	0,74	0,194	0	0	0	0	0	0	2,5e-4	6,6e-2	0	5

не принято, то для случая ГМО можно предположить наличие смеси из 2-х компонентов. При этом отбрасывается класс с наименьшим количеством информации, а количества информации о двух других классах,  $I_{a_1}$  и  $I_{a_2}$ , возможно интерпретировать как предварительную оценку их содержания в двухкомпонентной смеси:

$$V_1 \approx 100\% \cdot I_{a_1} / (I_{a_1} + I_{a_2}); V_2 = 100\% - V_1.$$

Пример такой интерпретации объектов распознавания по  $I_1$  при усечении алгоритма с отбрасыванием классов до  $n_{\max} = 30$  дан в табл. 3, где  $M$  – общее число принятых решений;  $m$  – число решений о наличии 2-х классов ГМО.

Путем различных сочетаний из 13 ПП по  $N$  ( $N = 2 \dots 10$ ) были сформированы  $N$ -мерные ПП (эталонные описания объектов). При этом в один вектор не объединялись сильно коррелированные ПП (коэффициент корреляции  $|r_{ij}| > 0,8$ ).

Лучшие результаты различения получены при использовании следующих векторных ПП:

$$P1 = (H_{dr}, I_1, AN); P2 = (Z_{dr}, AN, I_1, H_{dr});$$

$$P3 = (AN, I_1, |I_2|, Z_h); P4 = (Z_h, I_1).$$

Однако, ряд классов ГМО, в частности, сильная гроза (интенсивностью  $R = 55$  мм/ч) и сухая крупа по данным признакам классифицируются неверно. В связи с этим рассматривалась возможность дополнительного использования в классификационной процедуре параметров:

$$P5 = (Y_2, |Y_3|, Z_{dr}, AN, DPS);$$

$P6 = (Y_2, |Y_3|, I_1, TAN); P7 = (TAN, Z_h, Z_{dr})$ , обеспечивающих вероятности распознавания сильной грозы 0,49; 0,53 и 0,86, а сухой крупы – 0,48; 0,27 и 0,19, соответственно. Дальнейшее увеличение размерности векторных ПП нецелесообразно, поскольку достоверность распознавания остается практически на прежнем уровне (табл. 8), а снижение числа отсчетов, необходимых для принятия решения, ограничивается условием  $n_{\min} \geq 3$ .

Таблица 3

Состав двухкомпонентных смесей как результат распознавания  $\gamma$  класса  $a$  по признаку  $I_1$

$a$	$M$	$a$	$M$	$a$	$M$
дождь	701	гроза	1016	крупа	698
$\gamma$	$m$	$\gamma$	$m$	$\gamma$	$m$
Смесь	69	Смесь	39	Смесь	69
снег	Cb 20	Cb-St	CN 17	Cb	CN 6
гроза	Cb 19	Cb-St	CP 1	Cb	CP 2
гроза	крупа 18	Cb-St	снег 4	Cb	крупа 2
гроза	Cb-St 6	CP	снег 11	Cb	снег 20
гроза	CN 6	CP	дождь 3	Cb	Cb-St 2
		CP	CN 3	Cb	гроза 30
				Cb-St	гроза 7

Так как последовательная процедура с отбрасыванием классов требует наименьших затрат времени на принятие решения, то она наиболее предпочтительна при разработке алгоритмов классификации ГМО по  $N$ -мерным ПП. Вместе с тем, ни по одному из сформированных векторных признаков в отдельности не удается достичь высокой достоверности распознавания всех радиолокационных классов ГМО. Таким образом, существует необходимость разработки более эффективных алгоритмов определения специфических типов ГМО по поляризационным данным.

### Синтез алгоритма распознавания

В качестве базы для синтеза классификационной процедуры был выбран последовательный алгоритм с отбрасыванием классов. На основании результатов работы этого алгоритма можно составить сеть переходов от пространства исходов  $\gamma_i$  (принятых на основании ПП решений) к пространству состояний  $a_k$  (классам, предъявленным для распознавания),  $\gamma_i \in \{1 \dots K\}$ ; при этом из рассмотрения исключаются исходы, вероятность которых пренебрежимо мала.

Переходы для P1 приведены в табл. 6, из которой видно, что только для классов 1, 4, 5, 8 наблюдается четкое соответствие между исходом  $\gamma$  (либо его номером  $j$ ) и классом  $a$  (либо его номером  $k$ );

Таблица 4

Оценки вероятностей  $\hat{P}(\gamma_j|a_k)$  принятия решения в пользу  $j$ -го класса при условии, что для распознавания предъявлен класс  $a_k$  для алгоритма с накоплением информации по  $N$ -мерным векторным ПП

ПП	$a_k \backslash \gamma_j$	$\hat{P}(\gamma_j a_k)$										$\bar{n}$
		CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	
P1	CD	0,999	7,75e-4	0	0	0	0	0	0	0	0	3
	Cb	3,74e-3	0,992	0	0	0	1,56e-4	0	0	0	4,06e-3	3
	Cb-St	0	0	0,982	0	2,33e-3	1,36e-3	1,13e-2	1,36e-3	0	1,94e-3	4
	CN	1,57e-3	5,66e-3	0	0,991	0	0	1,57e-4	0	0	1,42e-3	3
	CP	0	0	1,33e-2	0	0,975	8,41e-3	0	0	0	3,19e-3	3
	дождь (5 мм/ч)	0	0,181	0	0	1,56e-3	0,782	0	0	0	3,64e-2	4
	дождь (30 мм/ч)	0	0	8,62e-3	0	3,08e-4	0,801	9,23e-3	3,08e-4	1,23e-3	0,179	6
	гроза (15 мм/ч)	0	0	0,544	0	0	3,59e-2	0,321	2,77e-2	0	7,21e-2	6
	гроза (55 мм/ч)	0	0	3,34e-2	0	0	0	1,83e-4	1,83e-4	0,966	0	4
	крупа (сухая)	0	0	2,30e-2	0	8,98e-4	0,673	9,58e-3	0	0	0,293	6
	крупа (мокрая)	0	0	4,73e-2	0	0	0	0	0,857	9,53e-2	2,98e-4	6
	град (сухой)	0	0	3,11e-4	0	0	0	1,56e-4	1,40e-3	0,995	2,64e-3	3
	град (мокрый)	0	0	4,75e-3	1,98e-4	0	0	1,98e-4	2,18e-2	0,973	5,93e-4	4
	снег (сухой)	0	0,364	0	0	0	0	0	0	0	0,636	4
снег (мокрый)	0	0	2,61e-3	0	2,90e-4	0,384	1,16e-3	0		0,611	6	
P2	CD	0,994	5,29e-3	0	0	0	0	0	0	4,67e-4	1,56e-4	3
	Cb	5,17e-3	0,986	0	1,57e-4	1,57e-4	2,51e-3	0	0	1,57e-4	6,11e-3	3
	Cb-St	0	0	0,979	0	2,63e-3	1,58e-3	1,14e-2	4,38e-3	0	5,26e-4	3
	CN	1,57e-4	1,13e-2	6,29e-4	0,982	0	3,15e-4	3,15e-3	4,72e-4	0	2,36e-3	3
	CP	0	3,39e-4	5,46e-2	0	0,933	1,15e-2	1,70e-4	1,70e-4	0	1,70e-4	3
	дождь (5 мм/ч)	0	0,172	1,28e-3	0	8,61e-3	0,806	0	0	0	1,15e-2	4
	дождь (30 мм/ч)	0	0	3,79e-2	0	0	0,91	3,71e-2	2,64e-3	2,83e-3	9,61e-3	4
	гроза (15 мм/ч)	0	0	0,525	0	2,31e-4	3,84e-2	0,389	4,76e-2	6,94e-4	0	5
	гроза (55 мм/ч)	0	0	0,831	0	1,74e-4	0	1,74e-4	1,57e-3	0,167	1,74e-4	3
	крупа (сухая)	0	0	7,23e-2	0	1,96e-4	0,857	4,23e-2	3,92e-3	3,72e-3	2,08e-2	4
	крупа (мокрая)	0	0	0,273	0	0	0	2,67e-3	0,689	3,49e-2	4,45e-4	4
	град (сухой)	0	1,58e-4	1,78e-2	0	1,58e-4	1,58e-4	7,89e-4	2,37e-3	0,975	3,94e-3	3
	град (мокрый)	0	0	8,43e-2	1,95e-4	0	0	7,80e-4	8,23e-2	0,832	7,80e-4	4
	снег (сухой)	0	0,339	0	0	0	9,10e-4	0	0	3,64e-4	0,659	4
снег (мокрый)	0	0	6,01e-3	0	3,88e-4	0,788	7,75e-3	3,30e-3	4,71e-2	0,147	4	
P3	CD	0,999	9,30e-4	0	0	0	0	1,55e-4	0	1,55e-4	1,55e-4	3
	Cb	3,45e-3	0,991	0	7,84e-4	0	0	1,57e-4	0	0	4,71e-3	3
	Cb-St	0	0	0,97	0	0	4,15e-4	2,26e-2	2,08e-3	1,04e-3	3,53e-3	4
	CN	1,26e-3	1,03e-2	0	0,986	0	0	3,16e-4	0	0	1,74e-3	3
	CP	0	0	2,00e-2	0	0,968	6,99e-3	0	0	0	4,83e-3	3
	дождь (5 мм/ч)	0	0,157	1,97e-4	0	1,08e-2	0,78	0	0	0	5,14e-2	4
	дождь (30 мм/ч)	0	0	1,80e-2	0	0	0,216	2,61e-2	0	1,76e-2	0,723	4
	гроза (15 мм/ч)	0	0	0,6	0	0	8,16e-3	0,291	2,86e-2	1,80e-2	5,43e-2	8
	гроза (55 мм/ч)	0	0	5,66e-4	0	1,89e-4	0	0,427	3,77e-4	0,572	0	4
	крупа (сухая)	0	0	2,89e-2	0	4,51e-4	0,407	1,53e-2	0	6,54e-3	0,542	4
	крупа (мокрая)	0	0	1,34e-2	0	0	0	6,67e-2	0,84	7,71e-2	2,98e-3	7
	град (сухой)	0	1,69e-4	1,69e-4	5,06e-4	0	5,06e-4	1,70e-2	4,72e-3	0,964	1,33e-2	3
	град (мокрый)	0	2,42e-4	0	2,42e-4	0	0	4,42e-2	2,25e-2	0,931	1,93e-3	5
	снег (сухой)	0	0,498	0	0	0	0	0	0	0	0,502	4
снег (мокрый)	0	0	1,73e-3	0	1,92e-4	0,121	1,92e-3	1,92e-4	1,79e-2	0,858	4	



Таблица 5

Оценки вероятностей  $\hat{P}(\gamma_j|a_k)$  для последовательного алгоритма с отбрасыванием классов по  $N$ -мерным векторным ПП

ПП	$a_k \backslash \gamma_j$	$\hat{P}(\gamma_j a_k)$										$\bar{n}$
		CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	
P1	CD	0,999	7,74e-4	0	0	0	0	0	0	1,55e-4	0	3
	Cb	3,90e-3	0,992	1,56e-4	0	0	1,56e-4	0	0	0	4,21e-3	3
	Cb-St	0	0	0,949	0	3,46e-3	2,55e-3	3,77e-2	4,37e-3	1,82e-4	2,55e-3	4
	CN	1,72e-3	7,84e-3	0	0,989	0	1,57e-4	1,57e-4	0	0	1,25e-3	3
	CP	0	1,64e-4	1,62e-2	0	0,966	1,48e-2	0	0	0	2,95e-3	3
	дождь (5 мм/ч)	0	0,195	0	0	3,62e-3	0,757	0	0	0	4,41e-2	4
	дождь (30 мм/ч)	0	0	1,03e-2	0	0	0,744	2,87e-2	4,70e-4	7,04e-4	0,216	5
	гроза (15 мм/ч)	0	0	0,46	0	0	5,57e-2	0,358	3,97e-2	6,58e-4	8,54e-2	4
	гроза (55 мм/ч)	0	0	4,34e-2	0	0	0	1,74e-4	2,62e-3	0,954	0	3
	крупа (сухая)	0	0	2,87e-2	0	1,16e-3	0,63	2,48e-2	2,32e-4	4,63e-4	0,314	5
	крупа (мокрая)	0	0	6,57e-2	0	0	0	9,00e-4	0,779	0,154	2,25e-4	4
	град (сухой)	0	0	3,11e-4	0	0	1,55e-4	1,55e-4	1,87e-3	0,994	3,11e-3	3
	град (мокрый)	0	0	5,77e-3	1,80e-4	0	0	3,60e-4	5,97e-2	0,934	5,41e-4	4
	снег (сухой)	0	0,374	0	0	0	1,73e-4	0	0	0	0,626	3
снег (мокрый)	0	0	1,76e-3	0	0	0,417	8,80e-4	2,20e-4	2,86e-3	0,578	4	
P2	CD	0,994	5,44e-3	0	0	0	0	0	0	6,22e-4	1,55e-4	3
	Cb	6,41e-3	0,984	3,13e-4	1,56e-4	0	2,82e-3	0	0	3,13e-4	6,10e-3	3
	Cb-St	0	0	0,955	0	3,03e-3	2,19e-3	3,39e-2	5,22e-3	0	5,05e-4	3
	CN	0	1,18e-2	7,84e-4	0,98	6,28e-4	0	3,61e-3	4,71e-4	0	2,67e-3	3
	CP	0	3,30e-4	5,58e-2	0	0,928	1,54e-2	1,65e-4	4,95e-4	0	0	3
	дождь (5 мм/ч)	0	0,185	1,22e-3	0	9,78e-3	0,789	0	0	6,99e-4	1,40e-2	3
	дождь (30 мм/ч)	0	0	3,85e-2	0	0	0,883	5,77e-2	3,91e-3	5,33e-3	1,14e-2	3
	гроза (15 мм/ч)	0	0	0,471	0	0	5,72e-2	0,411	5,80e-2	2,52e-3	7,75e-4	4
	гроза (55 мм/ч)	0	1,70e-4	0,826	0	1,70e-4	0	3,40e-4	2,72e-3	0,17	1,70e-4	3
	крупа (сухая)	0	0	7,23e-2	0	1,29e-3	0,836	5,53e-2	4,77e-3	5,88e-3	2,50e-2	4
	крупа (мокрая)	0	0	0,252	0	0	0	3,31e-3	0,68	6,24e-2	2,34e-3	4
	град (сухой)	0	1,58e-4	1,76e-2	0	3,15e-4	3,15e-4	7,87e-4	2,83e-3	0,973	5,35e-3	3
	град (мокрый)	0	3,60e-4	7,95e-2	1,80e-4	0	0	8,99e-4	0,117	0,801	1,44e-3	4
	снег (сухой)	0	0,345	0	0	0	1,04e-3	0	0	5,21e-4	0,654	3
снег (мокрый)	0	0	6,19e-3	0	7,28e-4	0,767	8,73e-3	3,28e-3	5,48e-2	0,159	4	
P3	CD	0,998	1,55e-3	0	0	0	0	1,55e-4	0	1,55e-4	1,55e-4	3
	Cb	4,54e-3	0,989	0	7,82e-4	0	0	1,57e-4	0	0	5,16e-3	3
	Cb-St	0	0	0,943	0	3,83e-4	7,65e-4	4,32e-2	7,27e-3	1,34e-3	3,63e-3	4
	CN	1,26e-3	1,24e-2	0	0,984	0	0	3,15e-4	0	0	1,89e-3	3
	CP	0	0	2,12e-2	0	0,965	8,47e-3	0	0	0	5,05e-3	3
	дождь (5 мм/ч)	0	0,172	3,65e-4	0	1,39e-2	0,753	0	0	0	6,07e-2	4
	дождь (30 мм/ч)	0	0	2,02e-2	0	1,94e-4	0,236	3,36e-2	3,89e-4	2,76e-2	0,682	4
	гроза (15 мм/ч)	0	0	0,532	0	0	1,29e-2	0,304	6,49e-2	3,50e-2	5,09e-2	5
	гроза (55 мм/ч)	0	0	8,82e-4	0	1,76e-4	0	0,451	3,53e-4	0,547	1,76e-4	3
	крупа (сухая)	0	0	3,22e-2	0	5,88e-4	0,425	2,41e-2	0	1,02e-2	0,508	4
	крупа (мокрая)	0	0	1,78e-2	0	0	0	1,13e-1	0,729	1,37e-1	2,88e-3	5
	град (сухой)	0	3,32e-4	1,66e-4	4,97e-4	0	3,65e-3	2,29e-2	6,63e-3	0,949	1,71e-2	3
	град (мокрый)	0	4,07e-4	0	2,03e-4	0	0	9,13e-2	4,76e-2	0,859	1,63e-3	4
	снег (сухой)	0	0,502	0	0	0	0	0	0	0	0,498	3
снег (мокрый)	0	0	1,61e-3	0	1,79e-4	0,152	2,50e-3	1,79e-4	2,42e-2	0,819	3	

для остальных классов неоднозначность отношения  $\gamma \leftrightarrow a$  приводит к неустойчивости работы алгоритма и неправильной классификации ГМО. Так, гроза интенсивностью 55 мм/ч интерпретируется последовательным алгоритмом по P1 как град с вероятностью решения  $P(9|7) = 0,995$ , слабая гроза (15 мм/ч) – либо как гроза ( $P(7|7) = 0,32$ ), либо как грозовое облако ( $P(3|7) = 0,54$ ); если мокрая ледяная крупа распознается достаточно уверенно ( $P(8|8) = 0,86$ ), то сухую ледяную крупу процедура ошибочно относит, в основном, к классам "дождь" ( $P(6|8) = 0,67$ ) или "снег" ( $P(10|8) = 0,29$ ).

Для определения степени зависимости исходов между собой использовались статистические оценки вероятности согласованных решений по парам ПП из множества  $\{P1...P7\}$  для классов "дождь", "крупа", "снег" (табл. 7), что дает основание говорить, по крайней мере, о неэквивалентности исходов; более того, исходы, за исключением решений по группе P1...P4, слабо коррелированы. С учетом того, что существует достаточно высокая вероятность принятия неодинаковых решений по каждому отдельному ПП, для повышения достоверности распознавания предлагается принимать решение по нескольким ПП. При этом необходимо оценить как условные вероятности получения  $i$ -го исхода при наличии  $k$ -го класса, так и совместные вероятности соответствия  $j$ -й комбинации исходов  $k$ -му классу.

Использование последовательного алгоритма для векторного пространства ПП приводит к образованию неоднозначного пространства исходов (множества решений  $\{\bar{\gamma}_k\}$ , принятых на основании каждого вектора ПП). Для выбора определенного решения из множества  $\{\bar{\gamma}_k\}$  необходимо определить отношения между различными комбинациями исходов первой ступени процедуры распознавания и множеством классов  $A$ . Тогда решение возможно принимать в пользу того класса, для которого полученная комбинация исходов наиболее вероятна.

Таблица 6

Сеть переходов от пространства исходов  $\gamma_j$  к пространству состояний  $a_k$

$j$	1	2	3	4	5	6	7	8	9	10
$k$	1	2 6	3 7	4	5	6 8 10	7 8	8	9	6 8 10

Таблица 7

Оценки вероятности согласованных решений  $P_{k,k}$  по классам дождь, крупа и снег для 7 ПП

$P_{k,k}$	Дождь						
	P1	P2	P3	P4	P5	P6	P7
P1	1,00	0,71	0,42	0,32	0,43	0,51	0,22
P2	0,71	1,00	0,25	0,14	0,53	0,64	0,30
P3	0,42	0,25	1,00	0,80	0,14	0,17	0,23
P4	0,32	0,14	0,80	1,00	7,0e-2	9,2e-2	0,15
P5	0,43	0,53	0,14	7,0e-2	1,00	0,57	0,33
P6	0,51	0,64	0,17	9,2e-2	0,57	1,00	0,26
P7	0,22	0,30	0,23	0,15	0,33	0,26	1,00
$P_{k,k}$	Крупа						
	P1	P2	P3	P4	P5	P6	P7
P1	1,00	0,61	0,60	0,50	0,19	0,28	0,35
P2	0,61	1,00	0,41	0,37	0,26	0,40	0,51
P3	0,60	0,41	1,00	0,74	0,12	0,16	0,36
P4	0,50	0,37	0,74	1,00	8,5e-2	0,13	0,33
P5	0,19	0,26	0,12	8,5e-2	1,00	0,47	0,29
P6	0,28	0,40	0,16	0,13	0,47	1,00	0,28
P7	0,35	0,51	0,36	0,33	0,29	0,28	1,00
$P_{k,k}$	Снег						
	P1	P2	P3	P4	P5	P6	P7
P1	1,00	0,53	0,63	0,62	0,15	0,19	0,20
P2	0,53	1,00	0,30	0,26	0,24	0,32	0,25
P3	0,63	0,30	1,00	0,87	6,8e-2	8,5e-2	0,24
P4	0,62	0,26	0,87	1,00	4,4e-2	6,4e-2	0,22
P5	0,15	0,24	6,8e-2	4,4e-2	1,00	0,45	0,17
P6	0,19	0,32	8,5e-2	6,4e-2	0,45	1,00	0,11
P7	0,20	0,25	0,24	0,22	0,17	0,11	1,00

Если  $k$ -му классу соответствует некоторое подмножество исходов по одному ПП  $\{\gamma_i\}$ ,  $1 \leq i \leq K$ , с условными вероятностями  $P(\gamma_i|a_k)$ , то, согласно формуле полной вероятности, вероятность класса  $a_k$  при условии, что принято решение  $\gamma_i$ :

$$P(a_k|\gamma_i) = \frac{P(\gamma_i|a_k)P(a_k)}{\sum_j P(\gamma_i|a_j)P(a_j)}, \quad (9)$$

или, для равновероятного случая,

$$P(a_k|\gamma_i) = P(\gamma_i|a_k) / \sum_j P(\gamma_i|a_j). \quad (10)$$

Таблица 8

Оценка вероятности  $\hat{P}(\gamma_j|a_k)$  по вектору ПП ( $H_{dr}, Z_{dr}, I_1, |I_2|, ELV, ORH, HGM, DPS, TAN, AN$ )

$a_k \backslash \gamma_j$	$\hat{P}(\gamma_j a_k)$										$\bar{n}$
	CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	
CD	0,972	2,36e-02	0	0	0	0	0	1,89e-03	1,89e-3	1,10e-3	3
Cb	4,44e-3	0,963	1,74e-3	1,58e-3	3,17e-4	2,06e-3	0	5,86e-3	6,34e-4	2,00e-2	3
Cb-St	4,95e-4	1,65e-4	0,936	0	4,45e-3	1,65e-4	5,77e-2	9,89e-4	0	3,30e-4	3
CN	0	8,93e-3	2,87e-3	0,943	8,29e-3	0	3,24e-2	2,07e-3	9,57e-4	1,12e-3	3
CP	1,61e-4	0	8,15e-2	1,61e-4	0,913	2,58e-3	1,13e-3	9,68e-4	0	4,84e-4	3
дождь	3,61e-4	5,41e-4	5,54e-2	0	9,02e-4	0,71	0,183	4,27e-2	3,61e-4	6,67e-3	4
гроза	1,84e-4	1,84e-4	0,38	0	2,03e-3	1,79e-2	0,541	5,78e-2	0	7,37e-4	4
крупа	1,89e-4	3,78e-4	0,126	3,78e-4	6,43e-3	0,476	0,315	7,29e-2	3,78e-4	2,65e-3	4
град	5,37e-4	2,25e-2	6,40e-2	1,79e-3	1,07e-3	1,15e-2	8,21e-2	2,45e-2	0,785	7,15e-3	3
снег	0	0,319	3,46e-4	3,46e-4	0	5,20e-4	0	2,77e-3	1,73e-4	0,677	3

Таблица 9

Оценки вероятности соответствия  $k$ -го класса  $i$ -му исходу,  $\hat{P}(a_k|\gamma_i)$  и  $i$ -го исхода  $k$ -му классу,  $\hat{P}(\gamma_i|a_k)$

$\hat{P}(\gamma_i a_k)$													
Исход, $i$	Класс, $k$	P1			P2			P3			P4		
		8	6	10	8	6	10	8	6	10	8	6	10
		крупа	дождь	снег	крупа	дождь	снег	крупа	дождь	снег	крупа	дождь	снег
дождь	6	0,631	0,740	0,408	0,835	0,880	0,757	0,425	0,244	0,153	0,421	0,142	0,107
снег	10	0,308	0,217	0,579	0,027	0,013	0,164	0,510	0,675	0,812	0,548	0,813	0,874

$\hat{P}(a_k \gamma_i)$													
Исход, $i$	Класс, $k$	P1			P2			P3			P4		
		8	6	10	8	6	10	8	6	10	8	6	10
		крупа	дождь	снег	крупа	дождь	снег	крупа	дождь	снег	крупа	дождь	снег
дождь	6	0,355	0,416	0,229	0,338	0,356	0,306	0,517	0,297	0,186	0,628	0,212	0,160
снег	10	0,279	0,197	0,524	0,132	0,064	0,804	0,255	0,338	0,407	0,245	0,364	0,391

Таблица 10

Данные о соответствии  $j$ -й комбинации исходов ПП  $k$ -му классу

Исходы				дождь				крупа				снег				$\gamma_k$
P1	P2	P3	P4	$\hat{P}(\bar{\gamma} a_k)$	$\hat{P}(\bar{\gamma} \bar{x})$	$m_{\bar{\gamma}}$	$\hat{P}(\bar{\gamma})$	$\hat{P}(\bar{\gamma} a_k)$	$\hat{P}(\bar{\gamma} \bar{x})$	$m_{\bar{\gamma}}$	$\hat{P}(\bar{\gamma})$	$\hat{P}(\bar{\gamma} a_k)$	$\hat{P}(\bar{\gamma} \bar{x})$	$m_{\bar{\gamma}}$	$\hat{P}(\bar{\gamma})$	
6	6	6	6	0,19	0,26	343	0,092	0,77	0,58	771	0,208	0,04	0,16	229	0,058	8
6	6	6	10	0,44	0,45	366	0,099	0,42	0,31	253	0,068	0,14	0,24	205	0,052	6 8?
6	6	10	6	0,31	0,20	29	7,8e-3	0,56	0,59	85	0,023	0,13	0,21	32	8,1e-3	8
6	6	10	10	0,49	0,48	1631	0,440	0,20	0,22	759	0,205	0,30	0,29	1069	0,269	6?
6	10	6	6	0,07	0,16	2	5,4e-4	0,68	0,40	5	1,4e-3	0,24	0,44	6	1,5e-3	8 10?
6	10	6	10	0,13	0,00	0	0	0,27	0,62	3	8,1e-4	0,60	0,38	2	5,0e-4	8 10?
6	10	10	6	0,09	0,00	0	0	0,35	0,00	0	0	0,56	1,00	3	7,6e-4	10
6	10	10	10	0,09	0,10	6	1,6e-3	0,08	0,13	8	2,2e-3	0,83	0,78	52	0,013	10
10	6	6	6	0,11	0,15	26	7,0e-3	0,77	0,56	98	0,026	0,12	0,29	55	0,014	8
10	6	6	10	0,24	0,22	44	0,012	0,38	0,39	78	0,021	0,37	0,39	85	0,021	8 10?
10	6	10	6	0,16	0,13	37	0,010	0,49	0,69	192	0,052	0,34	0,18	54	0,014	8
10	6	10	10	0,21	0,25	560	0,151	0,15	0,23	511	0,138	0,64	0,52	1243	0,313	10
10	10	6	6	0,03	0,15	1	2,7e-4	0,47	0,44	3	8,1e-4	0,50	0,41	3	7,6e-4	8 10?
10	10	6	10	0,04	0,07	1	2,7e-4	0,13	0,26	4	1,1e-3	0,84	0,67	11	2,8e-3	10
10	10	10	6	0,03	0,12	5	1,4e-3	0,17	0,51	21	5,7e-3	0,80	0,37	16	4,0e-3	8 10?
10	10	10	10	0,02	0,05	32	0,009	0,03	0,09	53	0,014	0,95	0,86	537	0,135	10
M						3711				3711				3972		

Поскольку на данном этапе синтеза алгоритма интерес представляет не столько количественная, сколько качественная оценка соответствия вероятности принятого решения  $\gamma_k$  действительному состоянию контрольной выборки  $a_k$ , то для простоты анализа можно считать исходы векторов ПП независимыми случайными величинами; тогда условная вероятность совместного появления ( $j$ -й комбинации) исходов векторного пространства ПП по отношению к  $k$ -му классу определяется как произведение вероятностей отдельных исходов при условии, что для распознавания был предъявлен класс  $a_k$ . Согласно (10), вероятность получить  $j$ -ю комбинацию из  $p$  исходов  $\gamma_p^{\{j\}}$  при наличии класса  $a_k$ :

$$P(\gamma_p^{\{j\}}|a_k) = \prod_p P(a_k|\gamma_p^{\{j\}}) / \sum_l \prod_p P(a_l|\gamma_p^{\{j\}}), \quad (11)$$

где  $p$  – число векторных ПП, образующих признаковое пространство;  $j \in \{1, 2, \dots, C\}$ ;  $C = \prod_p c_p!$  – общее число перестановок исходов;  $c_p$  – количество исходов по  $p$ -му векторному ПП, условная вероятность которых для класса  $a_k$  не менее 0,5%.

Для вычисления  $P(\gamma_p^{\{j\}}|a_k)$  использовались оценки вероятности соответствия  $k$ -го класса  $i$ -му исходу,  $\hat{P}(a_k|\gamma_i)$  (см. табл. 9; данные  $\hat{P}(\gamma_i|a_k)$  были получены при уравнивании необходимых для принятия решения объемов контрольных выборок). В табл. 10 для сравнения приведены теоретическая (определенная по (11)),  $\hat{P}(\bar{\gamma}|a_k)$ , и статистическая,  $\hat{P}(\bar{\gamma}|\bar{x})$ , оценки вероятности соответствия  $j$ -й комбинации исходов ПП  $k$ -му классу; указаны  $\gamma_k$  – соответствующее решение о наличии класса  $a_k$ ,  $m_{\bar{\gamma}}$  – число появления в имитационном эксперименте  $j$ -й комбинации исходов для 4-х ПП,  $\hat{P}(\bar{\gamma})$  – экспериментальная оценка вероятности  $j$ -й комбинации исходов,  $M$  – количество циклов принятия решения при длине контрольной выборки 19500 отсчетов.

Если  $P(\gamma_p^{\{j\}}|a_k) = \max\{P(\gamma_p^{\{j\}}|A)\}$ , тогда  $j$ -й комбинации исходов можно поставить в соответствие класс  $a_k$ ,  $\gamma_p^{\{j\}} \rightarrow a_k$ . Данный подход позволяет существенно повысить качество распознавания слабо различимых классов. Так, в табл. 11 приведены результаты распознавания классов "дождь" (30 мм/ч), "крупя" (сухая), "снег" (мокрый) при совместном использовании 4-х ПП {P1...P4} в т.н. параллельном алгоритме, основанном на вероятностном анализе отношений пространства исходов отдельных ПП и пространства состояний контрольной выборки. Таким образом, на первой ступени процедуры различения в результате применения последовательного алгоритма с отбрасыванием классов к векторному пространству ПП по критерию ДКИ формируется множество решений. На второй ступени решение в пользу определенного класса принимается по критерию максимума апостериорной вероятности. Углубленная блок-схема алгоритма на векторном пространстве ПП {P1...P7} показана на рис. 1, а результаты распознавания классов ГМО приведены в табл. 12. Если отнести грозу и Сб-St к одному радиолокационному классу "грозоопасные объекты", то вероятности правильного решения  $P_{np}$  для облаков всех форм имеют величины порядка 0,9; для грозы и града – 0,8...0,9; для дождя – 0,6...0,7 (ошибочное решение  $\gamma_{oui} = 8$  (крупя)); для сухой крупы  $P_{np} > 0,4$  ( $\gamma_{oui} = \{6$  (дождь); 10 (снег)); для мокрой –  $P_{np} = 0,7...0,8$  ( $\gamma_{oui} = \{9$  (град); 7 (гроза)); для сухого снега  $P_{np} = 0,9$ ; для мокрого –  $P_{np} > 0,4$  ( $\gamma_{oui} = \{6$  (дождь); 8 (крупя)) (табл. 13).

Таблица 11

Оценки вероятности соответствия  $i$ -го исхода  $k$ -му классу для {P1...P4}

Исход, $i$	Класс, $k$			
	дождь 6	крупя 8	снег 10	
дождь	6	0,62	0,334	0,333
крупя	8	0,186	0,436	0,127
снег	10	0,183	0,208	0,535

Таблица 12

Оценки вероятностей  $\hat{P}(\gamma_j|a_k)$  для параллельного алгоритма по векторному пространству ПП  $\{P1...P7\}$   
 ( $R$  – интенсивность осадков, мм/ч;  $w$  – степень обводненности ледяных частиц (с – сухие, м – мокрые))

$a$	$R$	$w$	$\hat{P}(\gamma_j a_k)$											$\bar{n}$
			CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	смесь	
CD	-	-	0,997	1,37e-3	0	0	0	6,87e-4	0	0	0	6,87e-4	0	13
Cb	-	-	4,73e-3	0,895	0	0	0	4,73e-3	0	0	0	9,55e-2	0	8
Cb-St	-	-	0	0	0,787	0	3,12e-3	3,38e-3	0,201	5,19e-3	2,60e-4	0	0	5
CN	-	с	1,92e-3	7,99e-3	0	0,988	0	9,58e-4	0	0	0	9,58e-4	0	6
CP	-	с	0	0	1,36e-2	0	0,966	1,42e-2	4,53e-3	1,98e-3	0	0	0	6
дождь	5	-	0	5,16e-2	0	0	3,56e-3	0,668	0	0,204	0	7,26e-2	0	7
дождь	30	-	0	0	2,12e-3	0	7,05e-4	0,696	2,15e-2	0,209	3,53e-4	7,02e-2	0	7
гроза	15	-	0	0	9,10e-2	0	0	8,25e-2	0,735	8,51e-2	1,47e-3	4,79e-3	0	7
гроза	55	-	0	0	2,62e-4	0	0	0	0,92	4,99e-3	7,48e-2	0	0	5
крупа	1	с	0	0	5,89e-3	0	1,47e-3	0,426	4,78e-2	0,407	3,68e-4	0,111	0	7
крупа	2	с	0	0	2,61e-3	0	0	0,335	6,26e-2	0,438	1,45e-2	0,148	0	7
крупа	5	с	0	0	0	0	0	4,71e-3	1,58e-2	0,605	0,374	4,28e-4	0	8
крупа	1	м	0	0	4,17e-3	0	0	0	6,33e-2	0,799	0,133	0	0	7
крупа	2	м	0	0	2,82e-2	0	0	0	0,158	0,8	1,41e-2	0	0	6
крупа	5	м	0	0	5,47e-3	0	0	0	0,102	0,651	0,242	0	0	8
град	5	с	0	0	0	0	0	1,89e-3	2,77e-2	9,76e-3	0,96	3,15e-4	0	6
град	5	м	0	0	0	3,21e-4	0	0	0,109	5,88e-2	0,832	0	0	6
снег	5	с	0	0,129	0	0	0	5,05e-3	0	0	0	0,866	0	12
снег	5	м	0	0	1,44e-3	0	3,60e-4	0,414	3,60e-3	0,197	2,16e-3	0,382	0	7

Таблица 13

Оценки вероятности правильных  $P_{np}$  ( $P(a_k|a_k)$ ) и ошибочных решений  $P_{ош}$  ( $P(a_j|a_k)$ );  $j \neq k$ )

$\alpha$		CD	Cb	Cb-St	CN	CP	дождь		гроза	
$R, \text{мм/ч}$		-	-	-	-	-	5	30	15	55
$\gamma$	$P_{np}$	0,997	0,895	0,787	0,988	0,966	0,668	0,696	0,735	0,920
Cb-St	$P_{ош}$								9,1e-2	
дождь									8,3e-2	
гроза				0,201						
крупа							0,204	0,209	8,5e-2	
снег			9,6e-2							

$\alpha$		крупа					град		снег		
$w$		сухая		мокрая			сух.	мокр.	сух.	мокр.	
$R, \text{мм/ч}$		1	2	5	1	2	5	5	5	5	
$\gamma$	$P_{np}$	0,407	0,438	0,605	0,799	0,8	0,651	0,960	0,832	0,866	0,382
Cb	$P_{ош}$									0,129	
дождь		0,426	0,335								0,414
гроза						0,158	0,102		0,109		
крупа											0,197
град					0,374	0,133		0,242			
снег			0,111	0,148							

Таблица 14

Результаты распознавания двухкомпонентных смесей для параллельного алгоритма

Компоненты смеси:						$\hat{P}(\gamma_j a_k)$										
1-ый	$R, \text{мм/ч}$	$p, \%$	2-ой	$R, \text{мм/ч}$	$p, \%$	CD	Cb	Cb-St	CN	CP	дождь	гроза	крупа	град	снег	смесь
дождь	10	50	крупа	2	50	0	0	1,16e-2	0	0	6,44e-4	0,113	0,836	3,83e-2	0	0
дождь	20	50	град	10	50	0	0	0	3,27e-4	1,31e-3	7,55e-2	8,92e-2	0,285	0,545	2,94e-3	0
дождь	20	80	град	10	20	0	0	0	0	5,58e-4	4,74e-3	7,20e-2	2,96e-2	0,893	2,79e-4	0

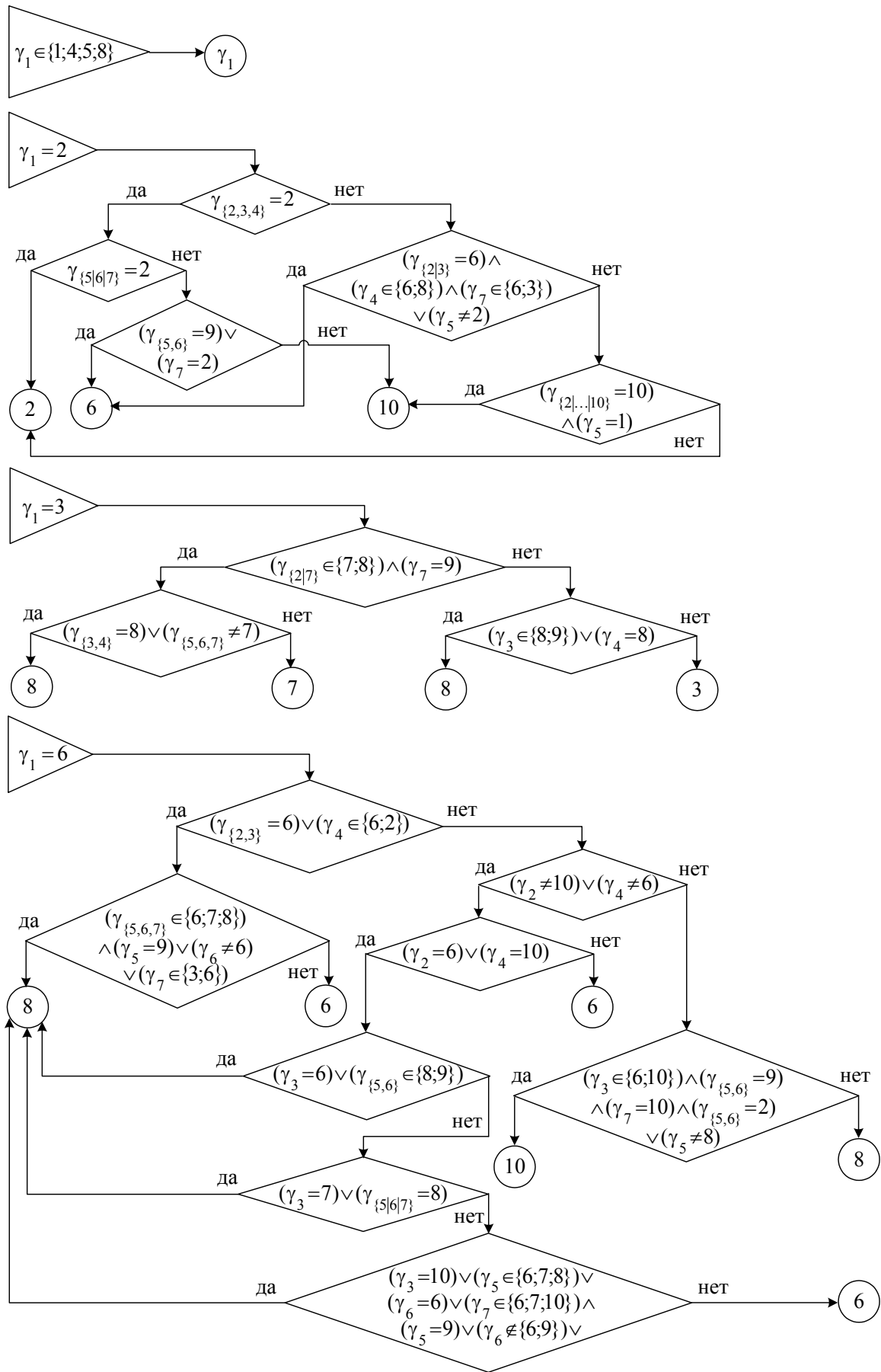


Рис. 1. Укрупненная блок-схема параллельного алгоритма по векторному пространству ПП  $\{P1...P7\}$

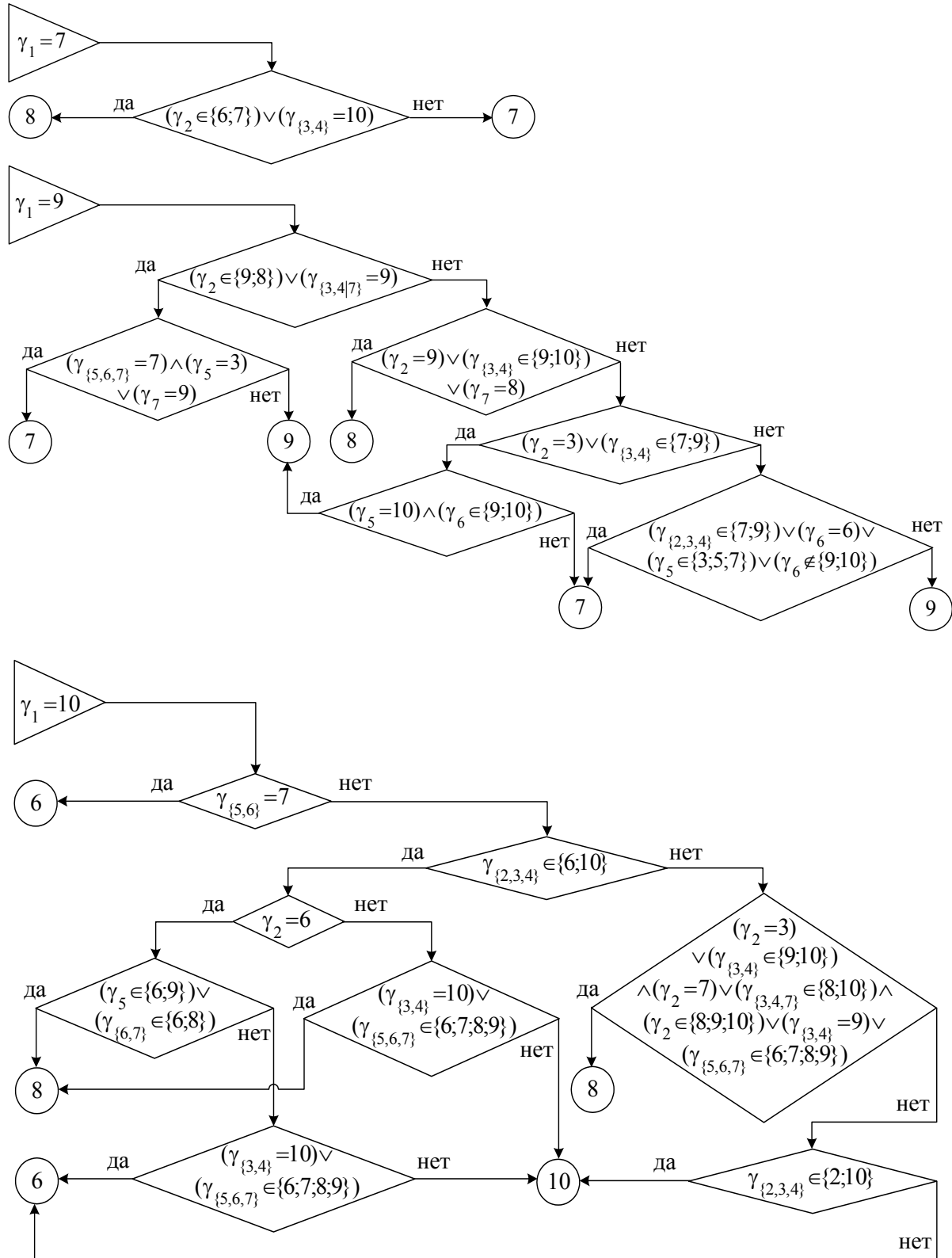


Рис. 1. Укрупненная блок-схема параллельного алгоритма по векторному пространству ПП {P1...P7} (окончание)

Результаты распознавания моделей смесей "дождь + крупа" и "дождь + град" (табл. 14) показали, что, независимо от соотношения компонентов смеси, алгоритм идентифицирует один класс, причем тот, который дает основной вклад в радиолокационную отражаемость (т.е., характеризующийся большим размером частиц).

### Заключение

На основе анализа известных методов принятия решения синтезирован двухступенчатый параллельный алгоритм распознавания классов ГМО по векторному пространству ПП. Использование последовательного метода с накоплением информации и отбрасыванием классов на первой ступени алгоритма позволяет значительно сократить длительность процедуры (среднее число отсчетов, необходимых для принятия предварительных решений, находится в пределах 5...13 в зависимости от класса, а максимальное по граничному условию не превышает 30). Выявленные статистические соотношения между пространством исходов первого этапа распознавания и пространством состояний контрольной выборки (номераами классов) легли в основу логической структуры второй ступени алгоритма, на которой окончательное решение принимается по критерию максимума апостериорной вероятности. Апробация алгоритма на модельных данных показала, что вероятность распознавания града и грозоопасных объектов составляет 0,8...0,9.

### Литература

1. Васильева И.К. Анализ информативности поляризационных признаков гидрометеорологических объектов // Радіоелектронні і комп'ютерні системи. – 2004. – № 4 (8). – С. 126 – 130.
2. Попов А.В., Васильева И.К. Моделирование поляризационных признаков гидрометеорологических образований // Радіоелектронні і комп'ютерні системи. – 2004. – № 1 (5). – С. 17 – 26.

3. Фомин Я.А., Тарловский Г.Р. Статистическая теория распознавания образов. – М.: Радио и связь, 1986. – 264 с.

4. Канарейкин А.Б., Потехин В.А., Шишкин И.Ф. Морская поляриметрия. – Л.: Судостроение, 1968. – 328 с.

5. Рыжков А.В. Поляризационные методы в метеорологической радиолокации // Зарубежная радиоэлектроника. – № 4. – 1993. – С. 18 – 28.

6. Straka J.M., Zrnicek D.S., Ryzhkov A.V. Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations // J. Appl. Meteor. – 2000. – Vol. 39. – P. 1341 – 1372.

7. Хан Г., Шапиро С. Статистические модели в инженерных задачах. – М.: Мир, 1969. – 369 с.

8. Краснов О.А., Попов А.В. Об оптимизационном подходе к построению математических моделей радиолокационных целей // Труды научно-технической конференции молодых ученых и специалистов. – X., 1986. – С. 127 – 132.

9. Бабаков М.Ф. Об одном способе аппроксимации распределений многомерных поляриметрических характеристик // Автоматизированные системы управления. – 1981. – Вып.3 – С. 166 – 167.

10. Пересада В.П. Автоматическое распознавание образов. – Л.: Энергия, 1970. – 90 с.

11. Кульбак С. Теория информации и статистика. – М.: Наука, 1968. – 302 с.

12. Попов А.В. Критерий информативности параметров сигналов для радиолокационного распознавания объектов // Авиационно-космическая техника и технология. – 1999. – Вып. 12. – С. 44 – 46.

13. Косенко Г.Г. Критерии информативности при различении сигналов. – М.: Радио и связь, 1982. – 216 с.

*Поступила в редакцию 27.09.2005*

**Рецензент:** д-р техн. наук, проф. Г.Я. Красовский, ГНПЦ «Природа», Харьков.