

## К ВОПРОСУ ФОРМАЛИЗАЦИИ ФУНКЦИОНИРОВАНИЯ СЛОЖНЫХ СИСТЕМ РАСПОЗНАВАНИЯ ОБРАЗОВ

И. А. Глазкова, М. А. Ивановский✉, Б. Х. М. Эль Эиссави

*Кафедра «Информационные системы и защита информации»,  
ivanovskiy\_62@mail.ru; ФГБОУ ВО «ТГТУ», Тамбов, Россия*

**Ключевые слова:** декомпозиция алгоритмов распознавания; информационная система; лингвистическая лотерея; система распознавания образов; сложная система; тезаурус; формализация функционирования.

**Аннотация:** Для одноуровневых систем распознавания образов предложено рассматривать поиск вектора признаков как модель распределенной информационной системы на тезаурусе. Для многоуровневых сложных систем распознавания предложен способ декомпозиции, который связан с включением в процесс распознавания специалиста-аналитика данных. При формальном описании системы распознавания на этапе синтеза образов учтены возможные неточности значений показателей из-за вероятностного характера величин, входящих в показатели, использована лингвистическая лотерея, аддитивная свертка показателей при нечеткой исходной информации.

---

### Введение

В настоящее время имеется тенденция к созданию универсальных многопрофильных междисциплинарных центров распознавания образов, функционирующих как в экспресс-режиме, так и режиме углубленного анализа, что связано с экономией аппаратных и программных ресурсов, поскольку современные информационные системы и технологии включают в себя большое количество типовых процедур, моделирующих или поддерживающих процесс интеллектуального анализа данных.

К простейшим процедурам такого типа относится любая классификация количественных данных по заданным пользователем критериям. Более сложные обеспечивают анализ сцен, процессов, явлений в целях выделения объектов с заданными характеристиками или свойствами, присутствуют не только в задачах анализа изображений, но и при обработке сигналов в технических системах, медицинской диагностике, биологии, социологии и других областях человеческой деятельности.

### Декомпозиция системы распознавания образов

В многоуровневых сложных системах распознавания важен способ декомпозиции, который обычно связан с моментами включения в процесс распознавания специалиста-аналитика данных, то есть с интерактивным режимом обработки из-за неполноты исходной информации или трудностей ее формального описания на этапе синтеза образов [1, 2].

В одноуровневых системах распознавание осуществляется на основе одного словаря признаков одним алгоритмом распознавания, что фактически приводит к поиску на тезаурусе, результатом может служить документ с вектором признаков.

Рассмотрим модель распределенной информационной системы на тезаурусе.

Тезаурус – словарь, отражающий смысловые связи между словами и иными смысловыми элементами и, следовательно, предназначенный для поиска слов по их смыслу. Тезаурусом называется непустое множество  $V$  слов  $v$ , отвечающих условиям:

- имеется непустое подмножество  $V_0 \subset V$ , называемое множеством дескрипторов;
- имеется симметричное транзитивное рефлексивное отношение  $R \subset V \times V$ , такое, что:

$$v_1 \neq v_2 \cap v_1 R v_2 \Rightarrow (v_1 \in V \setminus V_0) \cup (v_2 \in V \setminus V_0);$$

$$v_1 \in V \setminus V_0 \Rightarrow (\exists v \in V_0)(v R v_1);$$

- имеется транзитивное и несимметричное отношение  $K \subset V_0 \times V_0$ , называемое обобщающим отношением.

Отношение  $R$  называется *синонимическим отношением*, слова  $v_1$  и  $v_2$ , отвечающие этому отношению, называются *синонимическими дескрипторами*.

В случае, если два дескриптора  $v_1$  и  $v_2$  удовлетворяют отношению  $v_1 K v_2$ , полагается, что дескриптор  $v_1$  более общий, чем дескриптор  $v_2$ .

Элементы множества  $V \setminus V_0$  называются *множеством аскрипторов*.

Информационно-поисковый тезаурус позволяет находить нужные дескрипторы для адекватного выражения информационной потребности.

Дескриптор служит для описания основного содержания документа или формулировки информационного запроса при поиске документов в информационно-поисковой системе. Дескриптор ставится в однозначное соответствие группе ключевых слов естественного языка, отобранных из текста определенной области знания для построения дескрипторного языка [3].

*Информационной системой с тезаурусом* называется четверка  $(V, D, M, \delta)$ , где  $V$  – тезаурус с дескриптивным множеством  $V_0$ ;  $D$  – коллекция документов;  $M$  – множество вопросов;  $\delta: M \rightarrow 2D$  – отображение, сопоставляющее каждому вопросу множество документов.

Пусть описание любого документа  $d \in D$  может быть представлено в виде  $v(d) = \{v_1, v_2, \dots, v_k\}$  и удовлетворяет условию: никакие два дескриптора не встречаются в одном  $v(d)$ , если они удовлетворяют отношению  $K$ .

Можно также считать, что каждый вопрос  $m \in M$  представлен в форме, аналогичной описанию документов. Множество описаний вопросов и документов частично упорядочено отношением включения  $\leq$  следующим образом  $\forall d_1, d_2 \in D$ :

$$v(d_1) < v(d_2) \Leftrightarrow (\forall v' \in v(d_1)) \times (\exists v'' \in v(d_2)) (v' K v'') \vee (d' = d''),$$

то есть каждый дескриптор из  $v(d_1)$  представляет собой обобщение дескриптора из  $v(d_2)$  или идентичен дескриптору из  $v(d_2)$ . Отношение  $\leq$  позволяет сформулировать ответ  $Q$  на вопрос  $m \in M$  в виде  $Q = \delta(m) = \{d : d \in D \wedge m \leq x(d)\}$ .

Пусть  $S_1, S_2, \dots, S_n$  – локальные информационные системы, где  $S_j = (V_j, D_j, M_j, \delta_j)$ ,  $j = 1, n$ . Соединим системы  $S_1, S_2, \dots, S_n$  в одну систему  $S$ , которую будем называть *распределенной информационной системой*, базирующейся на глобальном тезаурусе  $V = \bigcup V_j$ .

Пусть имеется распределенная информационная система  $S = (V, D, M, \delta)$  с синонимическим отношением  $R_n$ , обобщающим отношением  $K$ . Тогда последова-

тельность информационных систем можно определить следующим образом:  $S_j = (V_j, D_j, M_j, \delta_j)$ , где  $V_j \subset V$ ;  $D_j \subset D$ ;  $M_j \subset M$ ;  $\delta_j$  – сужение  $\delta$  на  $M_j$ ;  $j = \overline{1, n}$ .

Кроме того, отношения  $R_j = R \cap (V_j \times V_j)$  и  $K_j = K \cap (V_{0j} \times V_{0j})$ ;  $V_{0j}$  – множество дескрипторов системы  $S_j$ ; отношение предпочтения  $j \leq (M_j \times M_j)$ ;  $\forall m \in M_j \delta(m) = \{d : d \in D_j \wedge m \leq j_v(d)\}$ .

Определим понятие подсистемы, которое позволит формализовать отношение включения одной системы в другую. Пусть  $S_1 = (V_1, D_1, M_1, \delta_1)$  и  $S_2 = (V_2, D_2, M_2, \delta_2)$  – информационные системы. Система  $S_1$  является подсистемой  $S_2$ ,  $S_1 \subset S_2$ , если:

$$\begin{aligned} & (V_1 \subset V_2) \wedge (K_1 = K_2 \cap (V_{01} \times V_{02})) \wedge (R_1 = R_2 \cap (V_{01} \times V_{02})); \\ & D_1 \subset D_2; \quad (V_1 \subset V_2) \wedge \leq 2 \cap (M_j \times M_j); \\ & \delta_1(m) = \delta_2(m_2) \cap D_1; \quad m \in M_1. \end{aligned}$$

Локальные информационные системы – подсистемы распределенной информационной системы.

Ввиду того что множество документов распределенной системы – это объединение множеств документов локальных систем, можно выразить ответ на вопрос к распределенной системе как результат некоторых операций над ответами от локальных систем. Пусть  $\bar{m} = \{m\}$  – вопрос, содержащий один дескриптор,  $\delta(\bar{m})$  – ответ на вопрос  $m$ . В подсистеме  $S_i$  ответ на локальный вопрос  $\bar{m} = \{m\}$  описывается в следующей форме

$$\delta_j(m) = \{d : d \in D_j \wedge j_v(d)\}.$$

В свою очередь, в распределенной системе  $S$  глобальный ответ на вопрос  $\bar{m} = \{m\}$  будет объединением локальных ответов  $\delta(\bar{m}) = \bigcup_j \delta_j(\bar{m})$ .

Ответ на произвольный вопрос  $\bar{m} = \{m_1, \dots, m_k\}$  в распределенной системе выражается формулой  $\delta(\bar{m}) = \bigcap_{i=1}^k \delta(\bar{m}_i)$ ,  $\bar{m}_i = \{m_i\}$ .

Распределенная информационная система обладает свойством включения, если множество вопросов частично упорядочено отношением  $\leq$  и выполняется условие

$$\{\bar{m}_1, \bar{m}_2\} \subset M \wedge \bar{m}_1 \leq \bar{m}_2 \Rightarrow \delta(\bar{m}_1) \supset \delta(\bar{m}_2).$$

Свойства включения позволяют формулировать цепь ответов на цепь вопросов, уточняя ответы более специализированными вопросами. В некоторых случаях при выполнении свойства аддитивности операции над ответами локальных систем могут быть упрощены.

Если множества документов локальных информационных систем не пересекаются, то глобальный ответ распределенной системы представляется как объединение локальных ответов

$$\left( \bigvee_{1 \leq k \neq n} D_k \bigcap_{1 \leq l \neq n} D_l = 0 \right) \wedge (\bar{m} = \{m_1, \dots, m_s\}) \Rightarrow \delta(m) = \bigcup_{j=1}^n \delta_j(\bar{m}).$$

Если все локальные информационные системы заданы на одном тезаурусе  $S_j = (V_j, D_j, M_j, \delta_j)$ ,  $j = \overline{1, n}$ , а множество документов разнесены, то ответ на глобальный вопрос будет объединением локальных ответов, то есть если

$$\bar{m} = \{m_1, \dots, m_k\} \text{ – вопрос } \forall_{1 \leq i \leq n} V_i = V, \text{ тогда } \delta(\bar{m}) = \bigcup_{j=1}^n \delta_j(\bar{m}).$$

Введем веса, которые будут описывать распределение информации на каждом слове. Информацию, заложенную в описании документов, будем рассматривать в виде единицы, следовательно, вес  $w_i$  соответствует объему информации, падающей на дескриптор  $v_i$ , то есть  $v(d) = \{ \langle v_1, w_1 \rangle, \langle v_2, w_2 \rangle, \dots, \langle v_k, w_k \rangle \}$  и выполняются условия:

$$v_i \neq v_j \wedge v_i K v_j \Rightarrow (\forall w_i \in [0, 1]) (v_i, w_i) \in v \left( \forall w_j \in (v_j, w_j) \in v(d) \right), \sum_{i=1}^k w_i = 1.$$

Описание документов удовлетворяют свойству включения. Дадим определение понятию подобия вопросов и описаний документов.

Пусть  $v(d_1), v(d_2)$  – описания документов:

$$v(d_1) = \{ \langle v_{11}, w_{11} \rangle, \langle v_{12}, w_{12} \rangle, \dots, \langle v_{1k_1}, w_{1k_1} \rangle \};$$

$$v(d_2) = \{ \langle v_{21}, w_{21} \rangle, \langle v_{22}, w_{22} \rangle, \dots, \langle v_{2k_2}, w_{2k_2} \rangle \}.$$

Описание  $v(d_1)$  подобно описанию  $v(d_2)$  с точностью  $\rho$  ( $0 \leq \rho \leq 1$ ), если  $v(d_1) \leq v(d_2) \wedge w_{1i}, \rho \leq \sum w_{2j}, \langle v_{11}, w_{11} \rangle \in v(d_1) j : [v_{2j} \in J(v_{1j})]$ , где  $J(v_{1j}) = \{v_{2j} : v_{1j} K v_{2j} \wedge (1 \leq j \leq K_2) \text{ при } 1 \leq i \leq K_1\}$ .

Если  $v(d_1)$  и  $v(d_2)$  удовлетворяют условию, то пишется  $v(d_1) \triangleleft_{\rho} v(d_2)$ .

Представляет интерес и другая величина, характеризующая коэффициент подобия или меру корреляции пар векторов, удовлетворяющих отношению  $\leq$ .

Для векторов

$$\mu(v(d_1), v(d_2)) = \sum_{i(v_{1i}, w_{1i}) \subset v(d_1)} \min(w_{1i}, \sum w_{2j}), \quad j : [v_{2j} \in J(v_{1i})],$$

где  $J(v_{1i}) = \{v_{2s} : v_{1i} K v_{2s} \wedge 1 \leq s \leq k_2\}$  при  $1 \leq i \leq k_1$ . Мера  $\mu$ , по сути дела, оценивает информацию, заключенную в описании документов  $v(d_1), v(d_2)$ .

*Информационной системой на тезаурусе с весами* называется четверка  $(V, D, M, \delta)$ , где  $V$  – тезаурус с дескрипторным множеством  $V_0 \subset V$ ;  $D$  – корреляция документов;  $M$  – множество вопросов;  $\delta$  – отображение  $\delta : M \times [0, 1] \rightarrow \rightarrow 2D \times [0, 1]$ , сопоставляющее каждой паре (вопрос, точность подобия) множество пар (документ мера корреляции).

Ответом системы на вопрос  $\bar{m}$  с требуемой точностью  $\rho = c$  называется множество

$$Q = \delta(\bar{m}, c) = \{(d, \alpha) : d \in D \wedge \bar{m} c v(d) \wedge \alpha = \mu(\bar{m}, v(d))\} \subseteq D \times [0, 1].$$

В определении ответа  $N = \delta(\bar{m}, c)$  отношение подобия осуществляет выбор документов, точность подобия которых не менее  $c$ . Мера корреляции  $\mu(\bar{m}, v(d))$  показывает, какая часть информации в документе соответствует ответу на вопрос.

Одно из фундаментальных свойств ответа связано с точностью  $\rho$  и мерой корреляции  $\mu$  вопроса к описанию документа, включенного в ответ.

Если документ  $d \in D$  с мерой корреляции  $\alpha$  включен в ответ, то есть  $(d, \alpha) \in \delta(\bar{m}, c)$ , то верно неравенство  $\alpha \geq c$ . Обратное утверждение неверно.

Если  $\rho_1 \geq \rho_2$ , то  $\delta(\bar{m}, \rho_1) \subset \delta(\bar{m}, \rho_2)$ .

В многоуровневых системах распознавания важной функцией аналитика данных является оценка качества распознавания на текущем этапе и возникающей из-за этого потери информации на следующем.

При разработке алгоритмов формализации в условиях неопределенности по характеристикам объекта аналитик может использовать модели, критерии и показатели  $R_1, \dots, R_i$ .

Имеем в виду, что значения показателей получаем на основании алгоритмов оценивания эффективности распознавания. Предположим аддитивность критериев [4, 5]. Учитывая возможные неточности значений показателей из-за вероятностного характера величин, входящих в показатели, наличия значений показателей, представляется целесообразным использование лингвистического подхода. Проранжируем рассматриваемые варианты на основе аддитивной свертки показателей при нечеткой исходной информации [6, 7].

Нечеткие оценки вариантов и коэффициенты важности показателей зададим функциями принадлежности в треугольной форме. Причем коэффициенты важности определяются исходя из предпочтений исследователя на основе аппарата теории одномерной полезности (лингвистического подхода) [8].

Определим методику получения коэффициентов важности показателей оценивания эффективности. Задано множество показателей важности  $R = \{R_1, \dots, R_i\}$ . Каждый из показателей может принимать конкретные значения  $x_1, x_2, \dots, x_m$ , которые наступают с нечеткими (лингвистическим) вероятностями  $p_1, p_2, \dots, p_n$ .

Предпочитаемая величина значений показателей (в дальнейшем полезность) определяется в ходе диалога с исследователем и задается в виде нечеткой функции принадлежности  $w_1, w_2, w_3, \dots, w_n$ . Определим нечеткую ожидаемую важность (полезность) каждого из показателей. Смысл диалога с исследователем по определению предпочитаемых значений показателей оценивания заключается в следующем. В ходе такого диалога исследуется поведение в условиях риска. Для этого исследователь:

- в соответствии со своими предпочтениями упорядочивает значения показателей:  $x_1 \prec x_2 \prec \dots \prec x_n$ , где  $x_1$  и  $x_n$  – соответственно наименее и наиболее предпочтительные значения;

- обозначая предпочитаемую величину показателя (полезность)  $w_1$  для наихудшего значения,  $w_n$  для наилучшего, для каждого значения показателя  $x_j$  ( $j = 1, 2, \dots, n$ ) назначают вероятность  $p_j$  такую, что ему будут безразличны следующие ситуации:

- получить значение  $x_j$ ;
- получить значение  $x_n$  с вероятностью  $p_j$ ;
- получить значение  $x_1$  с вероятностью  $1 - p_j$ .

Получаем тройку  $(x_n; p_j; x_1)$ , то есть лотерею, полезность (важность) которой находится с помощью выражения  $w_j = w_n p_j + w_1 (1 - p_j)$ . Для вычисления нечеткой ожидаемой важности каждого показателя будем использовать  $\alpha$ -сечение нечеткого множества  $R$  (множество уровня  $\alpha$ ) [4]:  $R_\alpha = \{x \in X; \mu_k(x) \geq \alpha\}$ , где  $X$  – базовое множество,  $\alpha \in [0; 1]$ ,  $\mu_k(x)$  – функция принадлежности,  $\alpha$  – сечение нечеткой вероятности и нечеткой важности значения показателя оценивания  $x_j$ .

Представим в форме треугольных функций принадлежности  $\mu(p_j)$  и  $\mu(w_j)$  (рис. 1, 2). Если функции принадлежности нечеткой вероятности и нечеткой важности выпуклые и непрерывные, то  $\alpha$ -сечение есть замкнутый интервал. Функцию принадлежности каждого показателя будем строить для тех же  $\alpha$ , и  $\alpha$ -сечения важности каждого показателя есть замкнутый интервал с границами [9, 10]:

$$\inf \omega = \sum_{j=1}^{k-1} M_j w_j + \left( 1 - \sum_{j=1}^{k-1} M_j - \sum_{j=k+1}^n m_j \right) w_k + \sum_{j=k+1}^n m_j w_j$$

при  $k = k^-$  и  $w_j = a_j$ ;

$$\sup \omega = \sum_{j=1}^{k+1} m_j w_j + \left( 1 - \sum_{j=1}^{k+1} m_j - \sum_{j=k+1}^n M_j \right) w_k + \sum_{j=k+1}^n M_j w_j$$

при  $k = k^+$  и  $w_j = b_j$ .

Индексы  $k^-$  и  $k^+$  находим из условий:

$$\left\{ 1 - \sum_{j=1}^{k^- - 1} M_j - \sum_{j=k^- + 1}^n m_j \right\} \in [m_{k^-}; M_{k^-}];$$

$$\left\{ 1 - \sum_{j=1}^{k^+ + 1} m_j + \sum_{j=k^+ + 1}^n M_j \right\} \in [m_{k^+}; M_{k^+}].$$

Формулы имеют место для данного показателя  $i$  и данного уровня  $\alpha$ , поэтому индексы  $i$  и  $\alpha$  в них отсутствуют (для удобства их восприятия). Так как значения показателей  $x_1, x_2, \dots, x_n$ , а также их важности (полезности) необходимо предварительно упорядочить, то необходимо упорядочить и границы  $a_j; b_j: a_1 < a_2 < \dots < a_n; b_1 < b_2 < \dots < b_n$ .

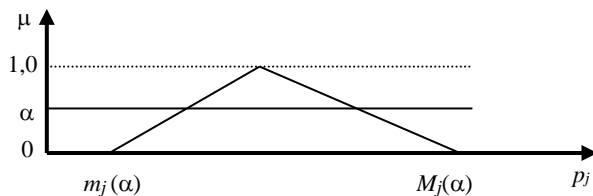


Рис. 1. Функция принадлежности

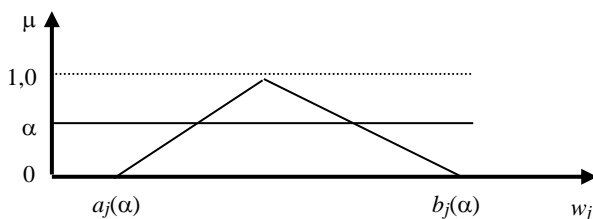


Рис. 2.  $\alpha$ -сечение функции принадлежности

Полученные функции принадлежности нечеткой ожидаемой важности показателей изображены графически на рис. 3. По треугольным функциям определяем значения коэффициентов важности показателей

$$p_{j\alpha} = [m_j(\alpha); M_j(\alpha)], \quad w_{j\alpha} = [a_j(\alpha); b_j(\alpha)]; \quad \forall \alpha[0;1].$$

Так как функции принадлежности нечетких вероятностей и нечетких значений показателей алгоритмов формализации в условиях неопределенности элементов моделей объектов распознавания известны, то задаваясь различными значениями  $\alpha$ , получаем тот материал, обработав который и найдем функции принадлежности важности показателей алгоритмов формализации в условиях неопределенности элементов моделей (рис. 4).

Проведем аддитивную свертку показателей. Варианты  $S = \{S_1, S_2, \dots, S_n\}$ , оцениваемые по  $n = i$  показателям, имеющим коэффициенты важности  $W_1, W_2, \dots, W_n$ , необходимо упорядочить. Через  $R_{ij}$  обозначим оценку  $i$ -го варианта ( $S_i$ ) по  $j$ -у показателю ( $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ ) и через  $R_i$  – оценку  $i$ -го варианта по всем показателям (взвешенную оценку). Оценки  $R_{ij}$  по показателям и коэффициенты относительной важности показателей  $W_{ij}$  задаем функциями принадлежности, соответственно  $\mu_{R_{ij}}(r_{ij})$  и  $\mu_{W_j}(\omega_j)$ . Взвешенная оценка  $i$ -го варианта вычисляется по формуле  $R_i = \sum_{j=1}^n W_j R_{ij}$  при условии, что оценки – нормированы. Будем

пользоваться треугольным представлением нечетких оценок и коэффициентов относительной важности показателей. Если функции принадлежности двух нечетких множеств  $X$  и  $Y$  имеют треугольное представление (см. рис. 1 – 4), то нечетное множество  $Z = X * Y$  также определяется функцией принадлежности треугольного вида, а границы и вершина находятся следующим образом:  $z' = x' * y'$ ,  $z'' = x'' * y''$ ,  $z = x * y$ , где  $x', y', x'', y'', x, y$  – соответственно левая, правая границы и вершины функций принадлежности нечетких множеств  $X$  и  $Y$ ,  $*$  – бинарная операция над функциями принадлежности нечетких множеств  $X$  и  $Y$  [11].

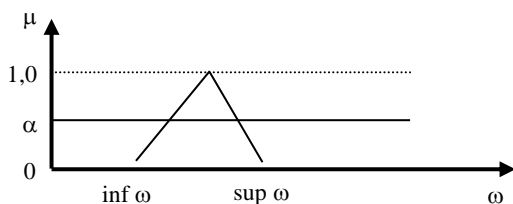


Рис. 3. Правые и левые границы функции принадлежности

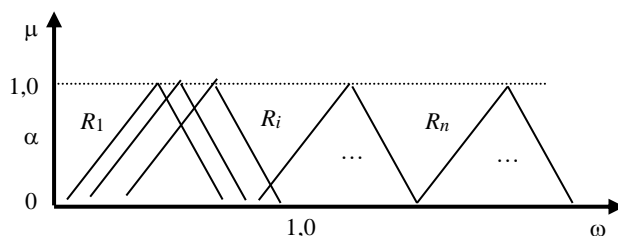
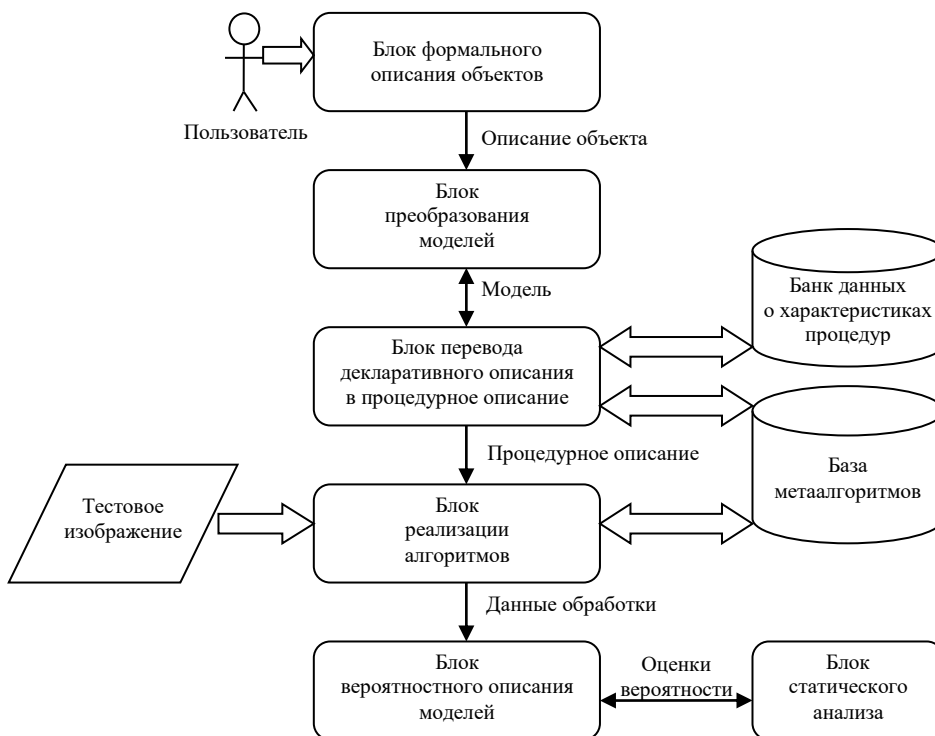


Рис. 4. Выбор наилучшего варианта. Бинарное сравнение



**Рис. 5. Структура системы выбора алгоритмов распознавания изображений**

После того как взвешенные оценки получены, необходимо сравнить варианты на их основе. Для этого вводится нечеткое множество  $I$ , заданное на множестве индексов вариантов  $\{1, 2, \dots, n\}$ , и значение соответствующей функции принадлежности интерпретируется как характеристика того, насколько вариант  $x_i$  является лучшим. Значение  $\mu(i)$  вычисляется по формуле  $\mu_I(i) = \sup \min_{j \in R_i / 1, n} \mu(r_j)$  [12, 13].

Графически значение  $\mu(i)$  равно ординате точки пересечения взвешенной оценки  $i$ -го варианта и взвешенной оценки наилучшего варианта (то есть того варианта, вершина функции принадлежности которого расположена справа от всех остальных). В простейшем случае для обнаружения контурных и полутоновых объектов на цветных и полутоновых изображениях система может иметь вид, показанный на рис. 5.

### Заключение

Таким образом, при проектировании и системном анализе процесса функционирования универсальных многопрофильных междисциплинарных центров распознавания образов, функционирующих как в экспресс-режиме, так и в режиме углубленного анализа, возможно рассматривать поиск вектора признаков как модель распределенной информационной системы на тезаурусе, методы лингвистической лотереи для свертки показателей при нечеткой исходной информации.

#### Список литературы

1. Алтуни, А. Е. Применение теории нечеткости для оценивания технологических параметров в АСУ ВПО «Тюменгазпром» / А. Е. Алтуни, С. Н. Чукаев, М. В. Семухин // Труды ЗапсибНИГНИ : Проблемы нефти и газа Тюмени. – Тюмень, 1983. – Вып. 58. – С. 57 – 59.



2. An access Control System Based on Multimodal Approach / V. V. Alekseev, D. V. Lakomov, G. Al Maamari, M. V. Moiseeva // AIP Conference Proceedings. – 2022. – Vol. 2467, No. 1. – P. 050001. doi: 10.1063/5.0093508
3. Алексеев, В. В. Анализ применимости алгоритма Кэнни для распознавания изображений в условиях неопределенности / В. В. Алексеев, П. И. Карасев, Д. В. Лакомов // Информатика: проблемы, методология, технологии : материалы XVI Междунар. науч.-метод. конф. (Воронеж, 11–12 февраля 2016 г.). – Воронеж, 2016. – С. 42 – 45.
4. Вагин, В. В. Интеллектуальный модуль определения изображений деструктивной направленности / В. В. Вагин, В. В. Алексеев, Е. Н. Емельянов ; под ред. А. А. Зацаринного, Д. Н. Борисова // Информатика: проблемы, методы, технологии : материалы XX Междунар. науч.-метод. конф. (Воронеж, 13–14 февраля 2020 г.). – Воронеж, 2020. – С. 363 – 367.
5. Анализ и синтез структур информационных целенаправленных систем : монография / Ю. Ю. Громов, М. А. Ивановский, О. Г. Иванова, А. В. Яковлев ; под ред. Ю. Ю. Громова. – Саарбрюкен (Германия) : LAP LAMBERT Academic Publishing. – 2015. – 164 с.
6. Моделирование информационных систем на основе использования дискретных и непрерывных законов распределения и их аппроксимации / Ю. Ю. Громов, Ю. В. Минин, О. Г. Иванова, А. П. Рыжков // Приборы и системы. Управление, контроль, диагностика. – 2019. – № 3. – С. 26 – 33.
7. Животников, Г. С. Теоретико-возможностные модели распознавания / Г. С. Животников, Ю. П. Пытьев // Математические методы распознавания образов : докл. 10-й Всерос. конф. – М., 2001. – С. 56–57.
8. Кобелев, К. С. Интеллектуальная система распознавания объектов на изображениях / К. С. Кобелев, Ю. В. Минин, С. В. Ковалев // Новые технологии в научных исследованиях, проектировании, управлении, производстве : тр. Междунар. науч.-техн. конф. (Воронеж, 09–10 ноября 2017 г.) – Воронеж, 2017. – Т. 1. – С. 295 – 299.
9. Лакомов, Д. В. Логико-лингвистическая модель определения контурных точек сосредоточенных и площадных объектов в условиях неопределенности / Д. В. Лакомов, Д. В. Поляков, В. В. Алексеев ; под ред. А. А. Зацаринного, Д. Н. Борисова // Информатика: проблемы, методы, технологии : материалы XX Междунар. науч.-метод. конференции (Воронеж, 13–14 февраля 2020 г.). – Воронеж, 2020. – С. 1484 – 1491.
10. Модель анализа изображений на основе метода нечёткой кластеризации / Д. В. Лакомов, В. В. Алексеев, Ю. В. Минин, Ю. В. Кулаков, Г. Н. Нурутдинов // Информация и безопасность. – 2017. – Т. 20, № 4(4). – С. 580 – 583.
11. Лакомов, Д. В. Модель анализа изображений на основе модернизированного алгоритма контурного анализа / Д. В. Лакомов, В. В. Алексеев // Нейрокомпьютеры и их применение : тезисы докл. XVII Всерос. науч. конф. (Москва, 19 марта 2019 г.). – М., 2019. – С. 259 – 261.
12. К вопросу об идентификации информационного процесса в системе поддержки принятия решений / Д. Х. Льюнг, А. С. Моисеев, Ю. В. Минин, В. В. Алексеев // Современные информационные технологии. – 2014. – № 19(19). – С. 28 – 31.
13. Перфильева, И. Приложения теории нечетких множеств / И. Перфильева // Итоги науки и техники. Серия Теория вероятностей. Математическая статистика. Теоретическая кибернетика. – 1990. – Т. 29. – С. 83 – 151.

## To the Problem of Formalizing the Operation of Complex Pattern Recognition Systems

I. A. Glazkova, M. A. Ivanovsky✉, Badr Khalil Mahmud El Eissawi

Department of Information Systems and Information Security, ivanovskiy\_62@mail.ru;  
TSTU, Tambov, Russia

**Keywords:** decomposition of recognition algorithms; information system; linguistic lottery; pattern recognition system; a complex system; thesaurus; formalization of functioning.

**Abstract:** For single-level pattern recognition systems, it is proposed to consider the search for a feature vector as a model of a distributed information system on a thesaurus. For multi-level complex recognition systems, a decomposition method is proposed, which is associated with the inclusion of a data analyst in the recognition process. When formally describing the recognition system at the stage of image synthesis, possible inaccuracies in the values of indicators are taken into account due to the probabilistic nature of the quantities included in the indicators, a linguistic lottery, and additive convolution of indicators with fuzzy initial information were used.

### References

1. Altunin A.E., Chukleev S.N., Semukhin M.V. [Problems of oil and gas of Tyumen], *Trudy ZapsibNIGNI: Problemy nefti i gaza Tyumeni* [Proceedings of ZapsibNIGNI], Tyumen, 1983, issue 58, pp. 57-59. (In Russ.)
2. Alekseev V.V., Lakomov D.V., Al Maamari G., Moiseeva M.V. An access Control System Based on Multimodal Approach, *AIP Conference Proceedings*, 2022, vol. 2467, no. 1, pp. 050001. doi: 10.1063/5.0093508
3. Alekseev V.V., Karasev P.I., Lakomov D.V. *Informatika: problemy, metodologiya, tekhnologii: materialy XVI Mezhdunar. nauch.-metod. konf* [Informatics: problems, methodology, technologies: Materials of the XVI International Scientific and Methodological Conference], Voronezh, February 11-12, 2016, Voronezh, 2016, pp. 42-45. (In Russ.)
4. Vagin V.V., Alekseev V.V., Emelyanov E.N.; Zatsarinny A.A., Borisov D.N. (Eds.). *Informatika: problemy, metody, tekhnologii: materialy XX Mezhdunar. nauch.-metod. konf.* [Informatics: problems, methods, technologies: materials of the XX International. scientific method. conf.], (Voronezh, February 13–14, 2020), Voronezh, 2020, pp. 363-367 (In Russ.)
5. Gromov Yu.Yu. (Ed.), Ivanovsky M.A., Ivanova O.G., Yakovlev A.V. [Analysis and synthesis of structures of information purposeful systems: monograph], Saarbrücken (Germany): LAP LAMBERT Academic Publishing, 2015, 164 p. (In Russ.)
6. Gromov Yu.Yu., Minin Yu.V., Ivanova O.G., Ryzhkov A.P. [Modeling of information systems based on the use of discrete and continuous distribution laws and their approximation], *Pribory i sistemy. Upravleniye, kontrol', diagnostika* [Devices and systems. Management, control, diagnostics], 2019, no. 3, pp. 26-33. (In Russ., abstract in Eng.)
7. Zhivotnikov G.S., Pyt'yev Yu.P. *Matematicheskiye metody raspoznavaniya obrazov : dokl. 10-y Vseros. konf.* [Mathematical methods of image recognition: reports. 10th All-Russian conf.], Moscow, 2001, pp. 56-57. (In Russ.)
8. Kobelev K.S., Minin Yu.V., Kovalev S.V. *Novyye tekhnologii v nauchnykh issledovaniyakh, proyektirovaniy, upravlenii, proizvodstve: tr. Mezhdunar. nauch.-tekhn. konf.* [New technologies in scientific research, design, management, production Proceedings of the International Scientific and Technical Conference], (Voronezh, November 09-10, 2017), Voronezh, 2017, vol. 1, pp. 295-299. (In Russ.)
9. Lakomov D.V., Polyakov D.V., Alekseev V.V.; Zatsarinny A.A., Borisov D.N. (Eds.). *Informatika: problemy, metody, tekhnologii: materialy XX Mezhdunar. nauch.-metod. konf.* [Informatics: problems, methods, technologies. Proceedings of the

XX International Scientific and Methodological Conference], (Voronezh, February 13-14, 2020), Voronezh, 2020, pp. 1484-1491. (In Russ.)

10. Lakomov D.V., Alekseev V.V., Minin Yu.V., Kulakov Yu.V., Nurutdinov G.N. [An image analysis model based on the fuzzy clustering method], *Informatsiya i bezopasnost'* [Information and security], 2017, vol. 20, no. 4 (4), pp. 580-583 (In Russ., abstract in Eng.)

11. Lakomov D.V., Alekseev V.V. *Neyrokomp'yutery i ikh primeneniye: tezisy dokl. XVII Vseros. nauch. konf.* [Neurocomputers and their application: abstracts of the XVII All-Russian Scientific Conference], (Moscow, March 19, 2019), Moscow, 2019, pp. 259-261 (In Russ.)

12. Luong H.D., Moiseev A.S., Minin Yu.V., Alekseev V.V. [On the issue of identification of the information process in the decision support system], *Sovremennyye informatsionnyye tekhnologii* [Modern information technologies], 2014, no.19(19), pp. 28-31 (In Russ., abstract in Eng.)

13. Perfil'eva I. [Applications of the theory of fuzzy sets], *Itogi nauki i tekhniki. Seriya Teoriya veroyatnostey. Matematicheskaya statistika. Teoreticheskaya kibernetika.* [Results of Science and Technology. Series Theory of Probability. Math statistics. Theoretical cybernetics], 1990, vol. 29, pp. 83-151. (In Russ., abstract in Eng.)

---

### **Zur Frage der Formalisierung des Funktionierens komplexer Mustererkennungssysteme**

**Zusammenfassung:** Für einstufige Mustererkennungssysteme ist vorgeschlagen, die Suche nach einem Merkmalsvektor als Modell eines verteilten Informationssystems in einem Thesaurus zu betrachten. Für mehrstufige komplexe Erkennungssysteme ist eine Zerlegungsmethode vorgeschlagen, die mit der Einbeziehung eines Datenanalysten in den Erkennungsprozess verbunden ist. Bei der formalen Beschreibung des Erkennungssystems im Stadium der Bildsynthese sind mögliche Ungenauigkeiten der Indikatorenwerte aufgrund des probabilistischen Charakters der darin enthaltenen Größen berücksichtigt. Bei den Indikatoren sind eine linguistische Lotterie, eine additive Faltung von Indikatoren mit unscharfen Ausgangsinformationen verwendet.

---

### **Vers une question de la formalisation du fonctionnement des systèmes complexes de la reconnaissance des images**

**Résumé:** Pour les systèmes de reconnaissance de images à un seul niveau, est proposé de considérer la recherche des vecteurs des caractéristiques comme un modèle de système d'information distribué sur le thésaurus. Pour les systèmes de reconnaissance complexes à plusieurs niveaux, est proposée une méthode de la décomposition qui implique l'inclusion d'un analyste de données spécialisé dans le processus de reconnaissance. Lors de la description formelle du système de reconnaissance au stade de la synthèse des images, sont pris en compte les inexactitudes possibles des valeurs des indicateurs en raison de la nature probabiliste des grandeurs entrant dans les indicateurs; sont utilisées la loterie linguistique et la convolution additive des indicateurs avec des informations brutes floues.

---

**Авторы:** *Глазкова Инга Александровна* – магистрант кафедры «Информационные системы и защита информации»; *Ивановский Михаил Андреевич* – кандидат технических наук, доцент, доцент кафедры «Информационные системы и защита информации»; *Эль Эиссави Бадр Халил Махмуд* – аспирант кафедры «Информационные системы и защита информации», ФГБОУ ВО «ТГТУ», Тамбов, Россия.