

УДК 004.932.72

## ИТЕРАТИВНЫЙ АЛГОРИТМ ОБНАРУЖЕНИЯ ОБЪЕКТОВ В ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ НА ОСНОВЕ ГИСТОГРАММНЫХ ПРОСТРАНСТВЕННЫХ МЕТРИК

И.А. БОРИСКЕВИЧ

*Белорусский государственный университет информатики и радиоэлектроники  
П. Бровки, 6, Минск, 220013, Беларусь*

*Поступила в редакцию 15 октября 2012*

Предложен итеративный алгоритм обнаружения объектов в видеопоследовательности, основанный на свойствах локальных гистограмм и адаптации стандартной оптимизации множества частиц к особенностям изображений. Результаты моделирования показывают, что алгоритм может быть использован для решения задач в реальном масштабе времени.

*Ключевые слова:* видеопоследовательность, глобальная оптимизация, гистограммные метрики, обнаружение объектов.

### Введение

Для обнаружения динамических объектов в видеопоследовательности широко используются методы вычитания фона, сегментации изображений и контролируемой классификации [1]. Недостатком данных технологий является высокая вычислительная сложность, что затрудняет их использование для решения задач обнаружения в реальном масштабе времени.

Одним из перспективных решений данной проблемы является использование технологии модифицированной стандартной оптимизации множества частиц (Standard Particle Swarm Optimization, SPSO) [2, 3]. SPSO – итеративный алгоритм, включающий инициализацию ряда векторов (называемых частицами) по заданному закону распределения в пределах пространства поиска целевой функции. Каждая частица является потенциальным решением проблемы, представленной в виде целевой функции. На каждом шаге выполнения алгоритма производится расчет значения целевой функции для каждой частицы, где в качестве входного параметра используется ее позиция. Сравнение значений целевых функций различных частиц показывает, позиция которой из них является наилучшей. В процессе выполнения алгоритма частицы перемещаются по пространству поиска под воздействием как их предыдущих наилучших позиций, так и наилучшей позиции для всего множества. Однако в настоящее время не определены условия выбора параметров основных модулей (инициализации, оптимизации, останова) алгоритма SPSO к задаче обнаружения объектов в видеопоследовательности.

В связи с этим целью работы является разработка быстрого итеративного алгоритма обнаружения объектов, основанного на свойствах локальных гистограмм [4] и адаптации стандартной оптимизации множества частиц к особенностям изображений [5].

### Итеративный алгоритм обнаружения объектов в видеопоследовательности на основе гистограммных пространственных метрик

Алгоритм обнаружения объектов включает следующие шаги.

1. Инициализация начальных параметров алгоритма.
  - 1.1. Выбор исходной видеопоследовательности с объектом наблюдения.

Исходный кадр  $I^f$  размером  $M \times N$  видеопоследовательности  $I = \{I^f \mid f = \overline{1, F}\}$  можно представить в виде:  $I^f = (I_{m,n}^f)_{M \times N}$ , где  $m = \overline{1, M}$  и  $n = \overline{1, N}$  – пространственные индексы исходного кадра;  $f = \overline{1, F}$ ;  $F$  – номер и число кадров в видеопоследовательности соответственно.

1.2. Выбор размера окрестности частицы  $I_p^f$  множества и искомого объекта  $I^r$ .

Окрестность частицы  $I_p^f$  множества (или искомого объекта  $I^r$ ) размером  $U \times V$  имеет следующий вид:

$$I_p^f = (I_p^f(k, l)) = \begin{pmatrix} I_p^f(m-k, n-l) & \cdots & I_p^f(m-k, n) & \cdots & I_p^f(m-k, n+l) \\ \vdots & & \vdots & & \vdots \\ I_p^f(m, n-l) & \cdots & I_p^f(m, n) & \cdots & I_p^f(m, n+l) \\ \vdots & & \vdots & & \vdots \\ I_p^f(m+k, n-l) & \cdots & I_p^f(m+k, n) & \cdots & I_p^f(m+k, n+l) \end{pmatrix},$$

где  $k$  и  $l$  – индексы окрестности центрального пиксела с координатами  $(m, n)$ , равными координатам наилучших позиций частицы (индивидуальной  $p_{jd}^t$  или глобальной  $p_{gd}^t$ );  $U = 2k + 1$  и  $V = 2l + 1$  – линейные размеры изображения частицы и объекта. Размеры объекта и окрестности частицы зависят от качества видеопоследовательности и сложности объекта.

1.3. Выбор метрики сходства между частицей и объектом.

Для эффективного обнаружения объектов используются метрики на основе локальных гистограмм интенсивности, которые обладают инвариантностью к сдвигу, повороту и искажению и имеют низкую вычислительную сложность. В качестве целевых функций оптимизации поведения частиц множества могут быть использованы следующие метрики.

– Расстояние Бхаттачария:

$$F(p_{(j/g)d}^t) = \frac{1}{\sqrt{1 - BC(H(I_p^f), H(I^r))}},$$

где  $BC(H(I_p^f), H(I^r)) = \sum_{c \in C} \sqrt{H_p^f(c) \times H^r(c)}$  – расстояние Бхаттачария между двумя локальными гистограммами  $H(I_p^f)$  и  $H(I^r)$ ;  $H(I_p^f) = (H_p^f(c) \mid c \in [0, 255])$  – локальная гистограмма интенсивности окрестности размером  $U \times V$  с координатами центрального пиксела  $(m, n)$ , равными координатам наилучшей глобальной позиции частицы множества  $p_{gd}^t$  или индивидуальной позиции частицы  $p_{jd}^t$ ;  $H(I^r) = (H^r(c) \mid c \in C = [0, 255])$  – локальная гистограмма интенсивности изображения искомого эталонного объекта размером;  $H_p^f(c)$ ,  $H^r(c)$  – значения локальных гистограмм, равные количеству пикселей с интенсивностью  $c$  изображений  $I_p^f$  и  $I^r$ .

– Расстояние Лорентциана:

$$LO(H(I_p^f), H(I^r)) = \sum_{c \in C} \log(1 + |H_p^f(c) - H^r(c)|).$$

– Расстояние Хаусдорфа:

$$HA = \max \left[ \max_m \left( \min_n [H_p^f(m) - H^r(n)] \right), \max_n \left( \min_m [H_p^f(m) - H^r(n)] \right) \right].$$

– Пересечение последовательностей:

$$SI(H(I_p^f), H(I^r)) = 1 - \frac{\sum_{m \in C} \min(H_p^f(m), H^r(m))}{\min\left(\sum_{m \in C} H_p^f(m), \sum_{n \in C} H^r(n)\right)}.$$

– Расстояние Минковского:

$$MA(H(I_p^f), H(I^r)) = \left( \sum_{c \in C} |H_p^f(c) - H^r(c)|^a \right)^{1/a},$$

где  $a \geq 1$  – порядок метрики, увеличение значения которого соответствует большей чувствительности к искажениям.

Частным случаем расстояния Минковского является Евклидово расстояние ( $a = 2$ ):

$$EU(H(I_p^f), H(I^r)) = \sqrt{\sum_{c \in C} (H_p^f(c) - H^r(c))^2}.$$

Для уменьшения требований к качеству и размеру эталонного изображения целесообразно использовать методы обработки гистограмм изображений.

1.4. Выбор параметров множества частиц для кадра  $I^f$  размером  $M \times N$  видеопоследовательности.

Начальные параметры  $(p_j^0, v_j^0)$  для  $j$ -й частицы задаются в виде  $\{(p_j^0, v_j^0) \mid j = \overline{1, J}\}$ , где  $p_j^0 = \{p_{j1}^0, \dots, p_{jd}^0, \dots, p_{jD}^0\}$  – начальная позиция  $j$ -й частицы, заданная случайным образом в  $D$ -мерном пространстве поиска;  $v_j^0 = \{v_{j1}^0, \dots, v_{jd}^0, \dots, v_{jD}^0\}$  – начальное смещение  $j$ -й частицы, заданное случайным образом из диапазона  $v_{jx/y}^0 = [v_{\min}; v_{\max}]$ ;  $d = 1, 2, \dots, D$  – произвольная координата  $D$ -мерного пространства поиска;  $v_{\min}$  и  $v_{\max}$  – минимальное и максимальное допустимые смещения частиц, последнее устанавливается с целью предотвращения преждевременной сходимости частиц;  $J$  – количество частиц множества. Для изображения  $D = 2$ .

1.5. Выбор параметров оптимизации динамики поведения частиц на изображении.

$T$  – количество итераций;  $U \times V$  – высота и ширина окна поиска;  $c_1$  и  $c_2$  – коэффициенты ускорения соответственно индивидуального (когнитивного) и глобального поведения частиц множества.

Для повышения разнообразия распределения частиц на начальных итерациях и повышения сходимости алгоритма на его конечных итерациях используются коэффициенты ускорения, динамику изменения которых можно представить в виде соотношений:

$$c_1^t = c_{1in} - (c_{1in} - c_{1fin}) \frac{t}{T-1}, \quad c_2^t = c_{2in} - (c_{2in} - c_{2fin}) \frac{t}{T-1},$$

где  $c_{1in}^t$ ,  $c_{1fin}^t$ ,  $c_{2in}^t$  и  $c_{2fin}^t$  – начальные и конечные значения коэффициентов ускорения соответственно индивидуального (когнитивного) и глобального поведения частиц множества;  $t = \overline{1, T}$  – номер итерации алгоритма.

Для обеспечения оптимального соотношения между индивидуальным и глобальным воздействиями на частицу используется коэффициент инерции движения, динамику изменения которого можно представить в виде соотношения:  $w^t = w_{in} - (w_{in} - w_{fin}) \frac{t}{T-1}$ .

Начальное значение коэффициента задается как  $w_{in} = 1$ , конечное –  $w_{fin} = 0,5$ .

2. Распределение частиц множества на области поиска (исходном изображении).

Количество и характер распределения частиц зависят от размеров изображения и объекта поиска, условий наблюдения.

3. Вычисление метрики сходства  $j$ -й частицы и эталонного изображений на  $t+1$ -й итерации .

В общем случае вычисление метрики сходства  $F(p_{jd}^{t+1})$  между окрестностью  $j$ -й частицы и эталоном объекта поиска можно представить в виде:  $F(p_{(j/g)d}^t) = \varphi(H(I_p^f), H(I^r))$ .

4. Определение наилучшей индивидуальной позиции  $j$ -й частицы на  $t+1$ -й итерации.

В алгоритме *SPSO* смещение  $v_{jd}^{t+1}$   $j$ -й частицы на  $t+1$ -й итерации учитывает три компоненты (инерционную, индивидуальную и глобальную) и вычисляется с помощью соотношения:

$$v_{jd}^{t+1} = w \cdot v_{jd}^t + c_1 \cdot (\hat{p}_{jd}^t - p_{jd}^t) + c_2 \cdot (\hat{p}_{gd}^t - p_{jd}^t),$$

где  $p_{jd}^t$  – позиция  $j$ -й частицы на  $t$ -й итерации;  $\hat{p}_{jd}^t$  и  $\hat{p}_{gd}^t$  – наилучшие индивидуальная и глобальная позиции  $j$ -й частицы на  $t$ -й итерации.

Для  $t=1$  итерации положим, что начальная позиция  $j$ -й частицы  $p_j^0$  является ее наилучшей позицией  $\hat{p}_j^0$ .

Определение наилучшей индивидуальной позиции  $j$ -й частицы на  $(t+1)$ -й итерации производится с помощью следующего соотношения:

$$\hat{p}_{jd}^{t+1} = \begin{cases} p_{jd}^{t+1} & \text{при } F(p_{jd}^{t+1}) > F(\hat{p}_{jd}^t), \\ \hat{p}_{jd}^t & \text{при } F(p_{jd}^{t+1}) \leq F(\hat{p}_{jd}^t) \end{cases}$$

Определение значений наилучшей позиции множества частиц на  $(t+1)$ -й итерации производится с помощью следующего соотношения:

$$\hat{p}_{gd}^{t+1} = \begin{cases} \hat{p}_{jd}^{t+1} & \text{при } F(\hat{p}_{jd}^{t+1}) > F(\hat{p}_{gd}^t), \\ \hat{p}_{gd}^t & \text{при } F(\hat{p}_{jd}^{t+1}) \leq F(\hat{p}_{gd}^t) \end{cases}$$

В общем случае целевая функция (метрика сходства) вычисляется с помощью соотношения:  $\hat{p}_{jd}^t = \arg \max_{i \in \{1, J\}} (F(p_{id}^t))$ .

5. Проверка выполнения условия завершения алгоритма. Условие завершения алгоритма имеет следующий вид:  $F(p_g^t) / F(I^r) > Th$ , где  $F(I^r)$  – значение метрики для случая полного сходства объекта поиска и его эталона;  $Th$  – выбранное пороговое значение.

При выполнении данного условия производится обнаружение объекта и алгоритм завершается.

6. Вычисление новой позиции  $j$ -й частицы на  $t+1$ -й итерации.

Новая позиция  $j$ -й частицы вычисляется с помощью соотношения:  $p_{jd}^{t+1} = p_{jd}^t + v_{jd}^{t+1}$ .

7. Проверка условия завершения алгоритма по критерию превышения максимального количества итераций  $t+1 > T$ .

При выполнении данного критерия алгоритм завершается. Для уменьшения вычислительной сложности алгоритма в качестве условия завершения может быть использован критерий превышения максимального количества вычисленных значений целевых функций:  $N_F > N_{F_{\max}}$ , где  $N_F$  – текущее значение количества вычисленных значений целевых функций.

8. Конечный результат работы алгоритма.

Выделенная область  $I_p^f$  размером  $U \times V$ , центральный пиксел которой имеет координаты  $\hat{p}_{jd}^t$ .

## Результаты моделирования

Для моделирования объектов поиска использованы тестовые эталоны (рис. 1).

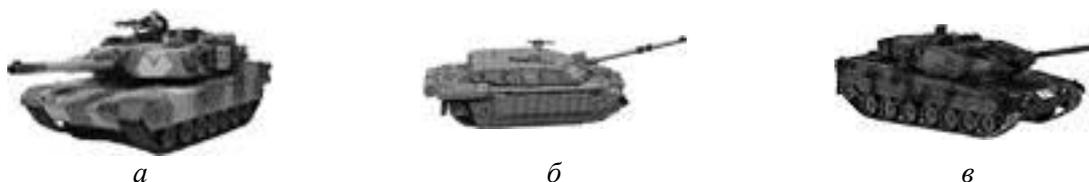


Рис. 1. Тестовые эталоны объектов поиска:

*a* – abrams.jpg (99×55 пикселей); *б* – challenger.jpg (99×43 пикселей); *в* – leopard.jpg (99×35 пикселей)

Рассмотрены следующие варианты расположения объектов на различных фоновых изображениях размером 640×480 пикселей (рис. 2): «1» – у левого края, «2» – у правого края, «3» – в левой части, «4» – в правой части, «5», «6», «7», «8» – в аналогичных позициях на большем удалении от места наблюдения. Выбор позиций на тестовом изображении обусловлен требованием к обнаружению объекта при его появлении в поле обзора камер.

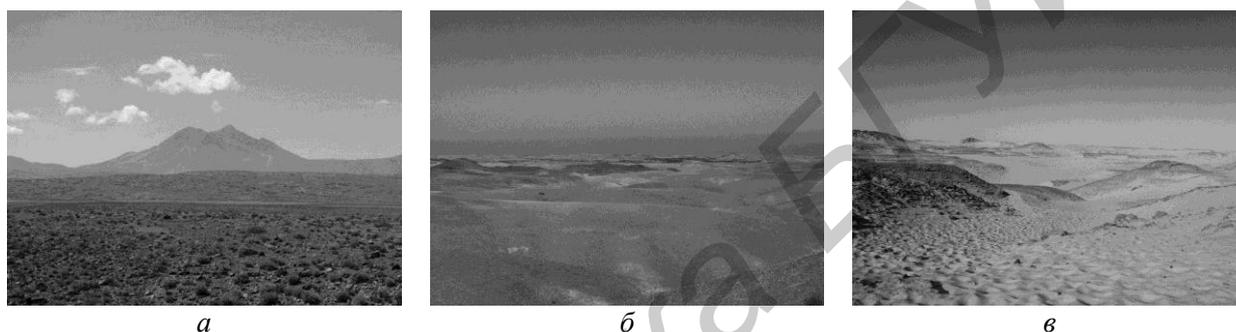


Рис. 2. Тестовые фоновые изображения: *a* – atakama.jpg; *б* – moroccan.jpg; *в* – sahara.jpg

Определены следующие начальные параметры алгоритма:  $N = 16$  частиц распределены равномерно вдоль правого, левого и верхнего краев изображения (в местах наибольшей вероятности появления объектов), что с учетом размеров области поиска соответствует примерно 9000 пикселям на одну частицу. Максимальное количество итераций  $T = 10$  обусловлено расстоянием между объектом и ближайшими частицами. Максимальное  $V_{\max} = 15$  и минимальное  $V_{\min} = 5$  смещения частиц выбраны исходя из критерия эффективности обнаружения. Начальные значения коэффициентов ускорения – глобальный  $C_1 = 1$ , индивидуальный  $C_2 = 0,5$ , коэффициента инерции –  $W = 0,8$ .

В табл. 1–3 представлены результаты обнаружения объектов на тестовых изображениях для различных метрик сходства между эталоном и окном поиска. Исследованы пять гистограммных метрик – расстояния Бхаттачария, Хаусдорфа, Лорентциана, Минковского и пересечение последовательностей. Сравнительный анализ проведен по критериям эффективности обнаружения и времени выполнения алгоритма. Объект считается обнаруженным, если он полностью или частично попал в окно поиска после превышения заданного порога обнаружения (в табл. 1–3 НО соответствует случаю, когда объект не обнаружен, ЛО – когда обнаружен ложный объект). Время обнаружения рассчитывается исходя из количества итераций, требуемых для правильного обнаружения. Среднее время вычисления одной итерации при заданных параметрах обнаружения приведено в табл. 4.

Таблица 1. Результаты обнаружения объекта, соответствующего эталону leopard.jpg, на тестовом изображении atakama.jpg

Метрика сходства	«1»	«2»	«3»	«4»	«5»	«6»	«7»	«8»
1. Бхаттачария	1	1	2	6	2	1	4	5
2. Минковского	ЛО	2	8	6	2	НО	5	НО
3. Лорентциана	1	2	4	6	2	НО	5	5
4. Хаусдорфа	5	НО	9	10	6	НО	9	НО
5. Пересечения	1	1	4	6	НО	НО	НО	НО

Таблица 2. Результаты обнаружения объекта, соответствующего эталону *abrams.jpg*, на тестовом изображении *mogoccam.jpg*

Метрика сходства	«1»	«2»	«3»	«4»	«5»	«6»	«7»	«8»
1. Бхаттачария	1	1	2	7	2	4	4	10
2. Минковского	2	НО	3	НО	2	НО	5	НО
3. Лорентциана	1	1	2	7	1	НО	2	НО
4. Хаусдорфа	1	НО	6	НО	3	НО	5	НО
5. Пересечения	1	1	2	8	1	НО	5	НО

Таблица 3. Результаты обнаружения объекта, соответствующего эталону *challenger.jpg*, на тестовом изображении *sahara.jpg*

Метрика сходства	«1»	«2»	«3»	«4»	«5»	«6»	«7»	«8»
1. Бхаттачария	8	ЛО	6	10	15	НО	5	6
2. Минковского	8	НО	5	НО	НО	НО	НО	НО
3. Лорентциана	6	2	4	8	НО	НО	6	НО
4. Хаусдорфа	1	НО	4	НО	ЛО	ЛО	3	ЛО
5. Пересечения	2	2	4	8	НО	НО	НО	НО

Таблица 4. Среднее время выполнения одной итерации для различных метрик сходства

Метрика сходства	Бхаттачария	Хаусдорфа	Лорентциана	Минковского	Пересечения
Время, с	0,048	0,586	0,057	0,066	0,049

На рис. 3 представлены начальное распределение частиц для тестового изображения *atakama.jpg* с эталоном *leopard.jpg* (объект поиска расположен в правой части и уменьшен в 1,3 раза по сравнению с эталоном), промежуточное распределение частиц с использованием метрики Бхаттачария после четвертой итерации и конечное распределение частиц при обнаружении объекта.

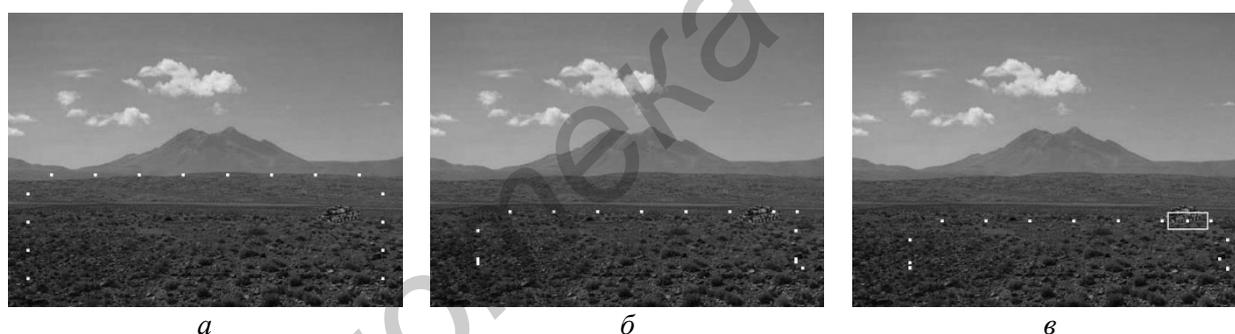


Рис. 3. Обнаружение объекта на фоновом изображении *atakama.jpg*:

*a* – начальное распределение частиц; *б* – промежуточное расположение частиц после четвертой итерации; *в* – конечное распределение частиц при обнаружении объекта

Из табл. 1–4 видно, что наилучшими показателями по критериям эффективности и времени обнаружения обладает расстояние Бхаттачария, несколько худший результат у метрики Лорентциана. Эффективное обнаружение масштабированных и зеркально отображенных объектов обусловлено инвариантностью гистограммных метрик. Размеры объектов поиска изменялись в пределах от  $99 \times 55$  до  $76 \times 27$  пикселей.

Предложенный алгоритм реализован в среде Matlab R2009A. Оценка эффективности алгоритма рассчитана на компьютере с процессором AMD Athlon (2,21 ГГц) и 1 ГБ ОЗУ.

### Заключение

Предложен итеративный алгоритм обнаружения объектов в видеопоследовательности, основанный на свойствах локальных гистограмм интенсивности и модифицированной стандартной оптимизации множества частиц. Он состоит из трех основных модулей: инициализации, оптимизации и останова. Первый модуль обеспечивает выбор следующих параметров: размеров области поиска и окрестности частиц множества, метрики сходства между частицами и объектом, динамики множества частиц для изображения (начальные позиция и смещение каждой частицы, минимальное и максимальное допустимые смещения частиц, количество ча-

стиц в множестве, коэффициенты ускорения индивидуального и глобального поведения частиц множества, пороговый локальный контраст между значениями, вычисленными для текущей и предыдущей наилучшей глобальной позиции, максимальное количество вычисленных значений целевой функции). Модуль оптимизации осуществляет распределение частиц множества на области поиска, вычисление метрики сходства, определение наилучшей индивидуальной и глобальной позиций, вычисление новых позиций частиц. Модуль останова алгоритма обеспечивает оптимальное соотношение между быстродействием и точностью распознавания за счет эффективного выбора условия завершения оптимизации.

Определены оптимальное количество частиц множества ( $N = 16$ ) и характер их распределения для изображений размером  $640 \times 480$  пикселей. Показано, что гистограммные метрики обладают свойством инвариантности к масштабу и положению объектов поиска на изображении. Установлено, что наилучшими характеристиками по критериям эффективности обнаружения и времени выполнения алгоритма обладает расстояние Бхаттачария. Результаты моделирования показывают, что данный алгоритм может быть использован для решения задач в реальном масштабе времени. Для уменьшения требований к качеству и размеру эталонного изображения целесообразно использовать методы обработки гистограмм изображений.

## **AN ITERATIVE ALGORITHM OF OBJECT DETECTION IN VIDEO SEQUENCE BASED ON HISTOGRAM SPATIAL MEASURES**

I.A. BARYSKIEVIC

### **Abstract**

A fast iterative algorithm of object detection in video sequence is proposed. It is based on local histogram properties and SPSO adaptation to image singularities. The parameters of algorithm are defined in terms of detection efficiency and computational complexity.

### **Список литературы**

1. *Yilmaz A., Javed O., Shah M.* // ACM Computing Surveys. Vol. 38, № 4. 2006. P. 45.
2. *Eberhart R., Kennedy J.* // Proceedings of the Sixth International Symposium on Micro Machine and Human Science. 1995. P. 39–43.
3. *Kennedy J., Eberhart R.* // IEEE International Conference on Neural Networks. 1995. P. 1942–1948.
4. *Борискевич И.А.* // Материалы X Междунар. научн. конф. «Ломоносов-2011». 2011. С. 252.
5. *Борискевич И.А.* // Россия – Беларусь – Сколково: тезисы Междунар. научн. конф. 2012. С. 202–203.