

УДК 004

Е.Е. ФЕДОРОВ<sup>1</sup>, Э. СЛЕСОРАЙТИТЕ<sup>2</sup><sup>1</sup> Донецкая академия автомобильного транспорта, Украина<sup>2</sup> Вильнюсский университет, Литва

## РАЗРАБОТКА ГИБРИДНОГО МЕТОДА РАСПОЗНАВАНИЯ ЛИЦА ЧЕЛОВЕКА

Для создания интеллектуальной компьютерной системы распознавания лица человека в статье был предложен гибридный метод распознавания на основе антропометрических признаков лица, искусственной нейронной сети, нечеткой логики и генетического алгоритма. Были предложены архитектура нечеткой искусственной нейронной сети и эффективные варианты операторов генетического алгоритма (репродукции, кроссинговера, мутации и редукции) на основе имитации отжига, учитывающие этапы генетического алгоритма. Разработанный метод был исследован на стандартной базе ORL и сравнен с аналогичными методами распознавания.

**Ключевые слова:** распознавание лица человека, антропометрические признаки лица, нейронная сеть, нечеткая логика, генетический алгоритм, имитация отжига, гибридная интеллектуальная компьютерная система.

## Введение

**Постановка проблемы.** В настоящее время актуальной является разработка систем, предназначенных для распознавания зрительных образов (например, в авиационной промышленности), причем наиболее важной задачей является распознавание лиц людей.

**Анализ исследований.** Современные методы распознавания [1-7] или обладают низкой вероятностью распознавания (например, сопоставление эластичных графов; сопоставление с эталоном, представленным PCA-коэффициентами (выделены на основе анализа главных компонент), ICA-коэффициентами (выделены на основе анализа независимых компонент) и LDA-коэффициентами (выделены на основе линейного дискриминантного анализа)) или требуют длительного времени обучения и большого количества обучающих данных (например, псевдо двумерные скрытые марковские модели (СММ); сверточная нейронная сеть (CNN), неоконитрон, нейронная сеть на основе вероятностного принятия решений (PDBNN), многослойный перцептрон (MLP), нейронная сеть радиально-базисных функций (RBFNN), бинарное дерево машин опорных векторов (SVM)), причем процесс определения структуры сети для таких методов не автоматизирован и требует длительного исследования.

**Цель и задачи исследования.** Для повышения вероятности распознавания лиц людей, увеличения скорости обучения и автоматизации процесса определения структуры модели необходимо разработать гибридный метод распознавания, базирующийся на

подходах искусственного интеллекта.

В статье:

- определяются и формализуются признаки лица человека;
- создается нейросетевая нечеткая система распознавания;
- осуществляется построение генетического алгоритма;
- проводится численное исследование.

## 1. Формализация признаков изображения

На рис.1 представлены антропометрические точки лица человека и расстояния между ними.

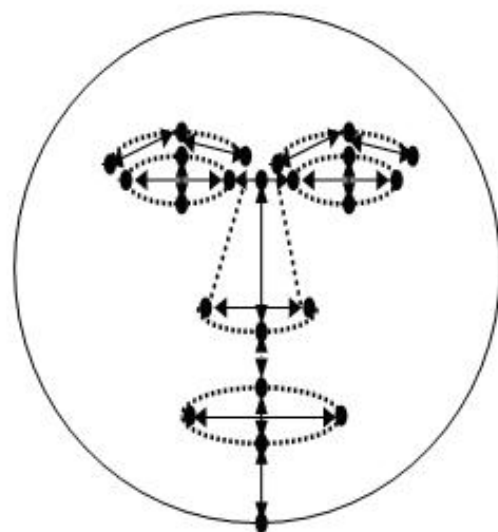


Рис. 1. Антропометрические точки лица и расстояния между ними

Антропометрические точки лица (базовые, опорные точки): 3 точки для левой брови (головка, точка излома, хвост брови); 3 точки для правой брови (головка, точка излома, хвост брови); 4 точки для левого глаза (левая, правая, верхняя, нижняя); 4 точки для правого глаза (левая, правая, верхняя, нижняя); 4 точки для носа (переносица, основание носа, левое и правое крыло носа); 4 точки для рта (левая, правая, верхняя, нижняя); 1 точка для подбородка;

На основе антропометрических точек формируется вектор  $X = (x_1, \dots, x_N)$ , состоящий из следующих признаков: расстояние между головкой и точкой излома левой брови  $x_1$ ; расстояние между хвостом и точкой излома левой брови  $x_2$ ; расстояние между головкой и точкой излома правой брови  $x_3$ ; расстояние между хвостом и точкой излома правой брови  $x_4$ ; длина глазной щели левого глаза  $x_5$ ; высота глазной щели левого глаза  $x_6$ ; длина глазной щели правого глаза  $x_7$ ; высота глазной щели правого глаза  $x_8$ ; расстояние между глазами  $x_9$ ; ширина носа (расстояние между левым и правым крылом)  $x_{10}$ ; высота носа (расстояние между переносицей и основанием носа)  $x_{11}$ ; высота верхней губы (расстояние между основанием носа и верхней точкой рта)  $x_{12}$ ; длина рта  $x_{13}$ ; высота рта  $x_{14}$ ; высота подбородка (расстояние между нижней точкой рта и подбородком)  $x_{15}$ .

Предполагается, что оператор компьютерной системы сам выделяет указанные выше антропометрические точки.

## 2. Структура гибридной интеллектуальной системы распознавания лица человека

Авторская гибридная интеллектуальная система распознавания лица человека представляет собой нейросетевую нечеткую систему, которая обладает следующими достоинствами:

- возможность использования априорной информации (знаний экспертов).
- представление знаний в виде правил, легко доступных для понимания человеком;
- возможность быстрого обучения и адаптации;
- параллельная обработка информации, которая повышает вычислительную мощность;
- отсутствуют трудности с определением структуры сети.

Для распознавания лица человека используем рекуррентную пятислойную сеть (рис. 2).

Входной (нулевой) слой содержит  $N^{(0)} = N$  нейронов (соответствует количеству признаков).

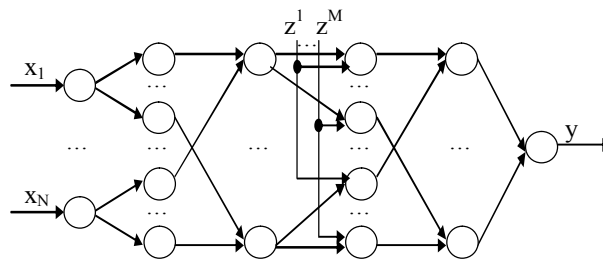


Рис. 2. Структура гибридной интеллектуальной системы

Первый скрытый слой реализует фаззификацию и содержит  $N^{(1)} = MN$  нейронов (соответствует количеству значений лингвистических переменных). Второй скрытый слой реализует агрегирование подусловий и содержит  $N^{(2)} = M$  нейронов (соответствует количеству классов лиц  $M$ ). Третий скрытый слой реализует активизацию заключений и содержит  $N^{(3)} = M^2$  нейронов. Четвертый скрытый слой реализует агрегирование заключений и содержит  $N^{(4)} = M$  нейронов. Выходной слой реализует дефаззификацию и содержит  $N^{(5)} = 1$  нейрон.

На вход нейросети подается вектор признаков  $X = (x_1, \dots, x_N)$ , сформированный на основе антропометрических точек лица человека. На нейроны третьего слоя подается вектор  $Z = (z_1, \dots, z_M)$ , значение каждой компоненты которого соответствует номеру класса лица. Переменная  $u$  содержит номер класса лица.

Все весовые коэффициенты нейронной сети равны 1.

## 3. Этапы создания гибридной интеллектуальной системы распознавания лица человека

Создание нейросетевой нечеткой системы распознавания предусматривает выполнение следующих этапов:

- формирование базы нечетких правил;
- фаззификацию;
- агрегирование подусловий;
- активизацию заключений;
- агрегирование заключений;
- дефаззификацию.

На этапе формирование базы нечетких правил  $j$ -е нечеткое правило представим в виде

$$R^j : \text{ЕСЛИ } \tilde{x}_1 \text{ есть } \tilde{\alpha}_1^j \text{ И } \dots \text{ И } \tilde{x}_N \text{ есть } \tilde{\alpha}_N^j,$$

$$\text{ТО } \tilde{y} \text{ есть } \tilde{\beta}^j, \quad j \in \overline{1, M}, \quad i \in \overline{1, N},$$

где  $\tilde{x}_i$  – имя входной лингвистической переменной;  $\tilde{y}$  – имя выходной лингвистической переменной;

ной;

$\tilde{\alpha}_i^j$  – нечеткая переменная (значение лингвистической переменной  $\tilde{x}_i$ );

$\tilde{\beta}^j$  – нечеткая переменная (значение лингвистической переменной  $\tilde{y}$ ).

Нечеткое множество  $\tilde{A}_i^j$  является областью значений нечеткой переменной  $\tilde{\alpha}_i^j$ , нечеткое множество  $\tilde{B}^j$  является областью значений нечеткой переменной  $\tilde{\beta}^j$ .

На *этапе фаззификации* определим степень истинности  $i$ -го подусловия, т.е. установим соответствие между входными переменными  $x_i$   $j$ -го правила и значениями функции принадлежности  $\mu_{\tilde{A}_i^j}(x_i)$ . Поскольку ряд методов, связанных с распознаванием использует функцию Гаусса, то выберем эту функцию в качестве  $\mu_{\tilde{A}_i^j}(x_i)$ , т.е.

$$\mu_{\tilde{A}_i^j}(x_i) = \exp \left[ -\frac{1}{2} \left( \frac{x_i - m_i^j}{\sigma_i^j} \right)^2 \right],$$

где  $m_i^j$  – математическое ожидание,

$\sigma_i^j$  – среднее квадратичное отклонение.

На *этапе агрегирования подусловий* функция принадлежности условия для  $j$ -го правила определяется в виде

$$\mu_{\tilde{A}_j}(\bar{x}) = \mu_{\tilde{A}_1^j}(x_1) \dots \mu_{\tilde{A}_n^j}(x_n), \quad j \in \overline{1, M}.$$

На *этапе активизации заключений* функция принадлежности заключения для  $j$ -го правила определяется в виде

$$\mu_{\tilde{C}^j}(y) = \mu_{\tilde{A}_j}(\bar{x}) \mu_{\tilde{B}^j}(y), \quad j \in \overline{1, M},$$

$$\mu_{\tilde{B}^j}(y) = \begin{cases} 0, & y \leq j-0,5; \\ \frac{y-(j-0,5)}{0,5}, & j-0,5 \leq y \leq j; \\ \frac{(j+0,5)-y}{0,5}, & j \leq y \leq j+0,5; \\ 0, & y \geq j+0,5, \end{cases}$$

где  $\mu_{\tilde{B}^j}(y)$  – треугольная функция принадлежности.

На *этапе агрегирования заключений* функция принадлежности итогового заключения определяется как

$$\mu_{\tilde{C}}(y) = \max(\mu_{\tilde{C}^1}(y), \dots, \mu_{\tilde{C}^M}(y)).$$

На *этапе дефаззификации* для получения номера класса используется метод максимума функции принадлежности

$$y = \arg \max_{z^j} \mu_{\tilde{C}}(z^j); \quad j \in \overline{1, M},$$

где  $z^j$  – центр нечеткого множества  $\tilde{C}^j$ .

Таким образом, математическую модель гибридной интеллектуальной системы распознавания лица человека (рис. 2) можно представить в виде

$$y = \arg \max_{z^k} \max_{j \in \overline{1, M}} \mu_{\tilde{B}^j}(z^k) \prod_{i=1}^N \mu_{\tilde{A}_i^j}(x_i), \quad k \in \overline{1, M}.$$

#### 4. Этапы построения генетического алгоритма

Построение генетического алгоритма предусматривает выполнение следующих этапов [8]:

- создание исходной популяции;
- определение фитнес-функции;
- задание оператора репродукции (селекции);
- задание оператора кроссинговера;
- задание оператора мутации;
- задание оператора редукции;
- определение условия останова.

На *этапе создания исходной популяции* выбраны вещественные гены в силу следующих причин:

- возможность поиска в больших пространствах, что трудно делать в случае двоичных генов, когда увеличение пространства поиска сокращает точность решения при неизменной длине хромосомы;
- способность к локальной настройке решений;
- отсутствие операций кодирования/декодирования, которые необходимы для двоичных генов, повышает скорость работы алгоритма;
- близость к постановке большинства прикладных задач (каждый вещественный ген отвечает за одну переменную или параметр, что невозможно в случае двоичных генов).

В качестве хромосомы, которая представляет  $i$ -ю особь популяции  $H = \{h_i\}$ , выступает упорядоченный вектор параметров (математических ожиданий и среднее квадратичных отклонений)

$$h_i = (lx_1^1 + i * \Delta m_1^1, lx_1^2 + i * \Delta m_1^2, \dots,$$

$$lx_n^1 + i * \Delta m_n^1, lx_n^2 + i * \Delta m_n^2,$$

$$lx_1^1 + i * \Delta \sigma_1^1, lx_1^2 + i * \Delta \sigma_1^2, \dots,$$

$$lx_n^1 + i * \Delta \sigma_n^1, lx_n^2 + i * \Delta \sigma_n^2), \quad i \in \overline{1, |H|},$$

$$\Delta m_k^j = \frac{rx_k^j - lx_k^j}{|H|}, \quad \Delta \sigma_k^j = \frac{rx_k^j - lx_k^j}{|H|}, \quad j \in \overline{1, M},$$

где  $|H|$  – мощность популяции,

$lx_k^j, rx_k^j$  – левая и правая границы значений  $k$ -го признака, вычисленные экспериментально.

На *этапе определения фитнес-функции* пред-

ложена следующая фитнес-функция, соответствующая вероятности правильного распознавания

$$F = \frac{1}{P} \sum_{p=1}^P I(y_p - d_p) \rightarrow \max_{m_i^j, \sigma_i^j},$$

$$I(a) = \begin{cases} 1, & a = 0 \\ 0, & \text{иначе} \end{cases},$$

где  $d_p$  – отклик, полученный с объекта (человека),

$y_p$  – отклик, полученный по нейросети,

$P$  – количество тестовых объектов в виде изображений.

На *этапе задания оператора репродукции (селекции)* для отбора векторов параметров для скрещивания и мутации в качестве оператора репродукции в статье предложена комбинация случайного и линейно упорядоченного отбора с имитацией отжига

$$P(h_i) = \frac{1}{|H|} \exp(-1/g(t)) +$$

$$+ \frac{1}{|H|} \left( a - (2a-2) \frac{i-1}{|H|-1} \right) (1 - \exp(-1/g(t))),$$

$$g(t) = \beta g(t-1), \quad g(0) = T_0,$$

где  $\beta$  – параметр,  $0 < \beta < 1$ ,

$T_0$  – начальная температура отжига,  $T_0 > 0$ ,

$t$  – номер итерации.

Таким образом, на ранних стадиях работы генетического алгоритма используется равновероятный отбор, обеспечивающий исследование всего пространства поиска (случайный выбор хромосом), а на заключительных стадиях используется линейно упорядоченный отбор, делающий поиск направленным (текущие лучшие хромосомы сохраняются). Эта комбинация не требует масштабирования и может использоваться при минимизации фитнес-функции.

На *этапе задания оператора кроссинговера* для комбинирования двух вариантов вектора параметров отобранных оператором репродукции, в качестве оператора кроссинговера используется равновероятный кроссинговер. Для выбора родителей в статье предложена следующая комбинация – на ранних стадиях работы генетического алгоритма используется аутбридинг (первый родитель выбирается случайно, а второй – как максимально далекий от первого либо как находящийся от первого на расстоянии большем, чем заданное  $\varepsilon$ , в качестве расстояния между родителями используется расстояние Манхэттена), обеспечивающий исследование всего пространства поиска, а на заключительных стадиях используется инбридинг (первый родитель выбирается случайно, а второй – как максимально близкий к первому либо как находящийся от первого на рас-

стоянии меньшем, чем заданное  $\varepsilon$ , в качестве расстояния между родителями используется расстояние Манхэттена), делающий поиск направленным. Эта комбинация не требует масштабирования и может использоваться при минимизации фитнес-функции. После выбора родителей осуществляется скрещивание, и производятся два потомка.

Вероятность кроссинговера над выбранными родителями на основе аутбридинга определена посредством имитации отжига в виде

$$P_c = P_0 \exp(-1/g(t)),$$

$$g(t) = \beta g(t-1), \quad g(0) = T_0,$$

где  $P_0$  – начальная вероятность кроссинговера.

Вероятность кроссинговера над выбранными родителями на основе инбридинга определена посредством имитации отжига в виде

$$P_c = P_0 (1 - \exp(-1/g(t))),$$

$$g(t) = \beta g(t-1), \quad g(0) = T_0.$$

Для глобального поиска оптимального вектора параметров необходимо повысить разнообразие вариантов.

На *этапе задания оператора мутации* для обеспечения разнообразия вариантов вектора параметров после кроссинговера используется неоднородная мутация.

Шаг мутации определен в виде

$$\Delta = \begin{cases} (\text{Max}_j - h_{ij}) r \left(1 - \frac{t}{T}\right)^b, & r < 0,5; \\ (h_{ij} - \text{Min}_j) r \left(1 - \frac{t}{T}\right)^b, & r \geq 0,5, \end{cases}$$

где  $\text{Max}_j, \text{Min}_j$  – максимальное и минимальное значение  $j$ -го гена,

$t$  – номер итерации,

$T$  – максимальное количество итераций,

$r$  – случайное число,  $r \in [0,1]$ ,

$b$  – параметр, управляющий скоростью уменьшения шага,  $b > 0$ .

Для имитации отжига вероятность мутации определена в виде

$$P_m = P_0 \exp(-1/g(t)),$$

$$g(t) = \beta g(t-1), \quad g(0) = T_0,$$

где  $P_0$  – начальная вероятность мутации.

Таким образом, на ранних стадиях работы генетического алгоритма с высокой вероятностью происходит мутация с большим шагом, что обеспечивает исследование всего пространства поиска, а на заключительных стадиях вероятность мутации и ее шаг стремятся к нулю, что делает поиск направленным. Для формирования новой популяции на основе предыдущей популяции и векторов парамет-

ров, полученных путем кроссинговера и мутации, используется оператор редукции.

На *этапе задания оператора редукции* в качестве оператора редукции в статье предложена следующая комбинация – на ранних стадиях работы генетического алгоритма используется схема на основе равновероятного отбора (случайный выбор хромосом), обеспечивающая исследование всего пространства поиска, а на заключительных стадиях используется схема  $(\mu + \lambda)$  (особи предыдущей популяции численностью  $\mu$  и потомки численностью  $\lambda$  объединяются и упорядочиваются по значению фитнес-функции, и в новую популяцию отбирается  $\mu$  первых лучших особей), делающая поиск направленным (текущие лучшие хромосомы сохраняются). Эта комбинация не требует масштабирования и может использоваться при минимизации фитнес-функции.

Вероятность выбора схемы на основе равновероятного отбора определена посредством имитации отжига в виде

$$P_r = P_0 \exp(-1/g(t)), \\ g(t) = \beta g(t-1), g(0) = T_0,$$

где  $P_0$  – начальная вероятность редукции.

Вероятность выбора схемы  $(\mu + \lambda)$  определена посредством имитации отжига в виде

$$P_r = P_0 (1 - \exp(-1/g(t))), \\ g(t) = \beta g(t-1), g(0) = T_0.$$

На *этапе определения условия останова* предлагается следующее условие

$$1 - \max_i F(h_i) < \varepsilon \vee t \geq T.$$

Значение  $\varepsilon$  и  $T$  вычисляются экспериментально.

## 5. Численное исследование методов распознавания лица человека

В табл. 1 приведено сравнение предложенного гибридного метода и существующих методов на основе базы данных ORL. Исследование позволяет сделать вывод, что авторский метод обеспечивает вероятность распознавания на уровне лучших методов (псевдо двумерные СММ с ДКП-коэффициентами), причем превосходит их по скорости обучения (обучение СММ может занимать несколько месяцев).

### Заключение

*Новизна.* В статье предложен метод распознавания лиц людей, базирующийся на методах цифровой обработки сигнала, нейросетях, нечеткой логике

Таблица 1

Оценка методов распознавания

Методы распознавания	Ошибка распознавания (%)
Одномерная СММ с ДКП-коэффициентами	16
Псевдо двумерная СММ с ДКП-коэффициентами	0,5
Сопоставление с эталоном, представленным PCA-коэффициентами	9,5
Сопоставление с эталоном, представленным ICA-коэффициентами	15
Сопоставление с эталоном, представленным LDA-коэффициентами	10
Сопоставление эластичных графов	20
Байеса	7
Непрерывный n-элементный классификатор	3
Стандартный n-элементный классификатор	14
CNN	4
Неокогнитрон	3
MLP с ДКП-коэффициентами	6
Бинарное дерево SVM с PCA-коэффициентами	3
RBFNN с ДКП-коэффициентами	2,5
PDBNN с сжатым изображением	4
Сопоставление с опорным изображением, на основе расстояния Манхэттена	3
Гибридный	0,5

и генетическом алгоритме. Этот метод был численно исследован на базе ORL. Разработанная гибридная интеллектуальная система распознавания лица человека обладает следующими достоинствами: возможность использования априорной информации (знаний экспертов); представление знаний в виде правил, легкодоступных для понимания человеком; возможность быстрого обучения и адаптации; параллельная обработка информации; отсутствуют трудности с определением структурой системы; высокая вероятность распознавания. Предложены варианты операторов генетического алгоритма на основе имитации отжига, повышающие эффективность обучения системы распознавания.

*Практическое значение.* Основные положения данной работы предназначены для реализации в

интеллектуальных компьютерных системах идентификации человека.

### Литература

1. Самаль, Д.И. Подходы и методы распознавания людей по фотопортретам [Текст] / Д.И. Самаль. – Минск: ИТК, 1998. – 54 с.

2. Brunelli, R. Face recognition: features versus templates [Text] / R. Brunelli, T. Poggio // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1993. – Vol. 15, № 10. – P. 235-241.

3. Face Recognition by Elastic Bunch Graph Matching [Text] / L. Wiskott, J.M. Fellous, N. Krueger, C. Malsburg // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1997. – Vol. 19. – P. 775-779.

4. Lucas, S.M. Face Recognition with the con-

tinuous *n*-tuple classifier [Text] / S.M. Lucas // *Proc. of the British Machine Vision Conference*. – Essex, 1997. – P. 222-231.

5. Wang, X. Bayesian Face Recognition Using Gabor Features [Text] / X. Wang, X. Tang // *Proceeding of the 2003 ACM SIGMM workshop on Biometrics methods and applications (WBMA '03)*. – Berkeley, 2003. – P. 70-73.

6. Samaria, F.S. HMM-based Architecture for Face Identification [Text] / F.S. Samaria, S. Young // *Image and Vision Computing*. – 1994. – Vol. 12, № 8. – P. 537-543.

7. Хайкин, С. Нейронные сети: полный курс [Текст] / С. Хайкин. – М.: Издательский дом «Вильямс», 2006. – 1104 с.

8. Скобцов, Ю.А. Основы эволюционных вычислений [Текст] / Ю.А. Скобцов. – Донецк: ДонНТУ, 2008. – 326 с.

Поступила в редакцию 1.04.2013, рассмотрена на редколлегии 24.04.2013

**Рецензент:** д-р техн. наук, проф., проф. каф. электронной техники Н.И. Чичикало, ГВУЗ «Донецкий Национальный технический университет», Донецк.

### РОЗРОБКА ГІБРИДНОГО МЕТОДУ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ ЛЮДИНИ

*Є.Є. Федоров, Е. Слесорайтите*

Для створення інтелектуальної комп'ютерної системи розпізнавання особи людини в статті було запропоновано гібридний метод розпізнавання на основі антропометричних ознак особи, штучної нейронної мережі, нечіткої логіки й генетичного алгоритму. Були запропоновані архітектура нечіткої штучної нейронної мережі й ефективні варіанти операторів генетичного алгоритму (репродукції, кросинговеру, мутації й редуцції) на основі імітації відпалу, що враховують етапи генетичного алгоритму. Розроблений метод було досліджено на стандартній базі ORL і зрівняно з аналогічними методами розпізнавання.

**Ключові слова:** розпізнавання особи людини, антропометричні ознаки особи, нейронна мережа, нечітка логіка, генетичний алгоритм, імітація відпалу, гібридна інтелектуальна комп'ютерна система.

### DEVELOPMENT OF A HYBRID METHOD FOR DETECTING HUMAN FACE

*E.E. Fedorov, E. Sliesoraityte*

For creation of intellectual computer system of recognition of the person face in article the hybrid method of recognition on the basis of anthropometrical features of the person, an artificial neural network, the fuzzy logic and genetic algorithm has been offered. The architecture of a fuzzy artificial neural network and effective variants of operators of genetic algorithm (have been offered a reproduction, crossingover, mutation and reduction) on the basis of imitation annealing, considering stages of genetic algorithm. The developed method has been investigated on standard base ORL and compared to similar methods of recognition.

**Key words:** recognition of the person of the person, anthropometrical signs of the person, neural network, the fuzzy logic, genetic algorithm, imitation annealing, hybrid intellectual computer system.

**Федоров Евгений Евгеньевич** – д-р техн. наук, доцент, заведующий кафедрой специализированных компьютерных систем, Донецкая академия автомобильного транспорта, Донецк, Украина, e-mail: fee75@mail.ru

**Слесорайтите Эгле** – старший преподаватель, Вильнюсский университет, Донецк, Украина