

В.Э. Подольский, К.М. Гречишев,  
Е.К. Пугачёв

## **ОШИБОЧНОЕ РАСПОЗНАВАНИЕ ОБРАЗОВ ИСКУССТВЕННЫМИ НЕЙРОСЕТЯМИ И МЕТОДЫ ТЕОРИИ ХАОСА КАК ОСНОВА НОВОГО ПОДХОДА К МАШИННОМУ ТВОРЧЕСТВУ**

*Приведена идея авторского подхода к реализации механизма машинного творчества на основе внесения в исходный образ шума и применения средств теории хаоса. Решающая роль отводится ошибкам распознавания искусственными нейронными сетями и методам хаотической динамики. В предлагаемом подходе прослеживается аналогия с механизмами спонтанного творчества человека.*

**E-mail: v.e.podolskiy@gmail.com,  
KMGrechishchev@ya.ru, ekpugachev@rambler.ru**

**Ключевые слова:** *машинное творчество, искусственный интеллект, теория хаоса, ошибки распознавания, искусственные нейронные сети.*

Творчество человека малоизучено, несмотря на то, что специалисты многих направлений современной науки – биологии, психологии, кибернетики, педагогики, социологии и даже математики – предпринимают попытки для осознания этого механизма и поиска его источников. В частности, классическая кибернетика тяготеет к перебору вариантов [1] и механизмам ассоциативной памяти [2]. Одна из областей кибернетики и информатики – искусственный интеллект (ИИ) – ставит перед собой в качестве ключевой задачи создание механизма машинного творчества. Причем результаты работы этого механизма должны в максимальной степени быть похожи на результаты работы человека-творца (художника, скульптора, архитектора и т. д.). В настоящее время ни один из предложенных подходов к машинному творчеству не дал удовлетворительного результата. Поэтому в данной работе предлагается новый подход к реализации механизма машинного творчества при помощи учета ошибок распознавания и использования средств современной математической теории динамического хаоса.

**Анализ работ по тематике машинного творчества.** Для понимания машинного творчества определимся с понятием творчества как такового. Творчество – деятельность, порождающая нечто качественно новое и отличающееся неповторимостью, оригинальностью и общественно-исторической уникальностью [3]. Однако согласно

тому же источнику «творчество специфично для человека». Поэтому в применении к ИИ и вычислительным машинам говорят не о творчестве, а о машинном творчестве и artificial creativity [4].

Машинное творчество – творческий процесс, который может быть осуществлен полностью средствами вычислительной техники, причем он может использовать как механизмы, заложенные в человеческом мозге, так и совершенно новые средства.

Основные требования к машинному творчеству были сформулированы в работе [5]:

- результат, представленный творческой программой, должен быть новым и полезным;
- результат требует отказа от ранее принятых идей;
- результат получен под воздействием высокой мотивации и настойчивости;
- результат получен от прояснения ранее расплывчатого вопроса.

В настоящее время систем и подходов, полностью удовлетворяющих представленным критериям, практически не существует.

Работы в области машинного творчества начаты давно, с момента зарождения ИИ. Интересные детерминистские и стохастические подходы к синтезу музыкальных произведений на вычислительной машине представлены Зариповым [6]. Ньюэллом и Саймоном была создана система GPS (General Problem Solver – общий решатель задач) [7], которая удовлетворяла перечисленным выше требованиям. Некоторые ученые попытались связать машинное творчество с математикой. Построение изображений с помощью самоподобных структур – фракталов – считается еще одной ветвью в машинном творчестве [8]. Недавно Шмидтом и Липсоном была предпринята попытка формализации научного творчества: ими создана программа выявления общих закономерностей (формул) физических и иных естественных процессов на базе генетических алгоритмов [9]. В Российской ассоциации искусственного интеллекта проблемой машинного творчества занимаются В.Л. Стефанюк, В.Э. Карпов, И.Б. Фоминых и другие ученые [10]. В нашей стране в области машинного творчества перспективной считается реализация аппарата теории решения изобретательских задач (ТРИЗ) на вычислительной машине [11]. Исследования в этом направлении ведутся, но результаты [12] довольно сомнительны с позиции перечисленных критериев.

**Суть подхода на основе ошибок распознавания.** Ни один из представленных выше подходов не приближает ученых к решению проблемы машинного творчества. Редко предлагаются методы машинного творчества, включающие одновременно несколько принципов. В настоящей статье представлен новый механизм машинного творчества: антропо-математический принцип моделирования оши-

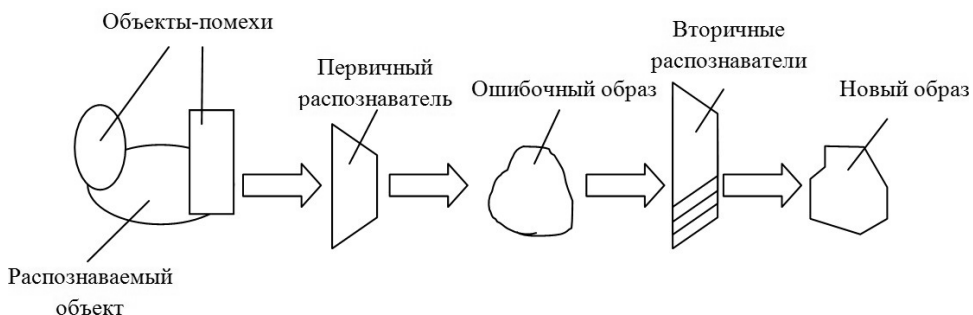
бок распознавания и построения на их основе новых образов, а также оценки новизны таких образов при помощи современного математического аппарата.

Ввиду явной несостоятельности существующих подходов, авторами настоящей статьи предлагается в некоторой степени универсальный (в рамках ИИ) подход к машинному творчеству. Чтобы полнее объяснить идею подхода, рассмотрим пример.

Как известно, система визуального распознавания образов человека (глаза – зрительный нерв – затылочная часть коры головного мозга) иногда дает сбой. Если человек находится в ярко освещенной естественным и искусственным светом комнате и изолирует ее полностью от естественного освещения, то с высокой долей вероятности в процессе адаптации глаз испытуемого к смене светового режима у человека возникнут короткие по продолжительности галлюцинации, т. е. он будет видеть то, чего на самом деле нет. Это происходит вследствие того, что некоторые палочки и колбочки (визуальные рецепторы на сетчатке глаза) сохранили предыдущее световое воздействие [13]. Из-за различий в скорости реакции разных рецепторов зрения перед глазами возникают световые пятна. Часто мозг старается достроить эти световые пятна до некоторых образов, виденных ранее. Если такое световое пятно не похоже в точности ни на один из виденных ранее образов, то происходит комбинирование разных образов в один. В результате такого процесса человек может видеть то, чего никогда в жизни не видел, но что напоминает некоторые виденные им ранее образы. Этот случай можно отнести к ошибке системы распознавания образов человека.

Основная идея статьи – формализация и построение творческого процесса в виде обработки ошибки распознавания образа и доведения ее до некоторого другого цельного образа с последующей оценкой его оригинальности и осуществимости.

Возможная схема реализации механизма машинного творчества в обобщенном виде приведена на рис. 1.



**Рис. 1. Абстрактная схема реализации механизма машинного творчества на основе ошибочного распознавания**

В соответствии с представленным рисунком объект распознается первичным распознавателем с некоторой ошибкой. В результате этого формируется ошибочный образ (его наличие обуславливается неполнотой информации об объекте или несовершенством распознавателя). Ошибочный образ далее интерпретируется как самостоятельный объект реального мира и при помощи комбинации вторичных распознавателей формирует на выходе новый образ, полученный комбинацией образов от каждого вторичного распознавателя, причем эти образы создаются на основе распознавания образа на выходе первичного распознавателя. Таким образом первичный образ дополняется до некоторого цельного образа, который может одновременно принадлежать нескольким группам (это относится к задачам нечеткой кластеризации [14]).

**Постановка задачи и интерпретация ошибок.** Оригинальным будем называть тот образ  $I_{\text{нов}}$ , полученный в результате исполнения описанной выше процедуры, минимальное расстояние от которого  $d_{\min}$  до другого реально существующего образа  $I_i$  определенной группы (число образов в ней равно  $n$ ) в некотором метрическом пространстве превосходит на величину максимального расстояния  $d_{\max}$  между двумя образами  $I_i$  и  $I_j$ , принадлежащими одной группе. Итак, образ считается оригинальным, если:

$$d_{\min}(I_{\text{нов}}, I_i) > d_{\max}(I_j, I_k), \quad i = \overline{1:n}; j = \overline{1:n}; k = \overline{1:n}. \quad (1)$$

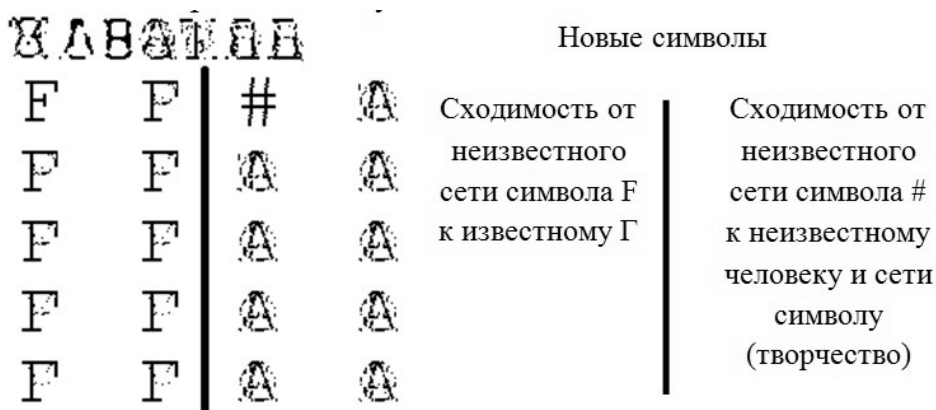
Очевидно, что формула (1) является геометрической интерпретацией задачи поиска выбросов. Задача поиска выбросов, заключающаяся в нахождении таких точек данных, которые больше всего отличаются от уже сгруппированных по некоторому признаку в классы и кластеры. Проблемной точкой этого простого подхода является определение функции  $d$  – степени отличия образов. В геометрической интерпретации функцию можно принять за расстояние. Но в целом эта функция зависит от рассматриваемой проблемы и механизма ее решения. Рассмотрим функцию  $d$  на примере нейросетей (НС). Также приведем основу нового подхода к машинному творчеству с применением математической теории хаоса [15].

Концепция искусственных нейронных сетей (ИНС) была разработана на основе принципов функционирования головного мозга человека (с позиции нейрофизиологии). К общим чертам ИНС и НС головного мозга человека можно отнести способность к обобщению, обучаемость, способность корректно обрабатывать информацию при ее высокой зашумленности и др. [16]. Обладающие подобными свойствами нейросетевые алгоритмы можно использовать в задачах, связанных с проблемой машинного творчества.

В работе [17] рассмотрена проблема распознавания символов многослойным перцептроном. Причинами ошибочной работы НС в процессе распознавания являются недостаточная или, напротив, чрезмерная обученность НС, слишком высокий уровень шума во входных данных или подача на входы НС образа, эталон которого отсутствовал в обучающей выборке. Там же предложен критерий, используя который можно выявить ошибку при распознавании или низкую степень уверенности НС.

Остановимся подробнее на проблеме обучения сети. При высокой степени обученности сети теряется одно из важнейших ее свойств: способность к обобщению. Так, сеть просто «запомнит» эталонные образы. В этом случае при незначительном отличии распознаваемого образа от оригинала сеть не будет уверена в своем решении. Если же сеть, наоборот, недостаточно обучена, то она может оказаться неспособной различить два полностью несовпадающих эталона.

Рассмотрим работу трехслойного перцептрона применительно к задаче машинного творчества. ИНС обучается повторять на своих выходах входной сигнал, причем на входы сети подаются изображения символов размером 32 на 32 пикселя. Число нейронов во входном и выходном слоях равно 1 024 (32×32). Эталонное множество состоит из символов русского алфавита. При подаче на входы ИНС эталонного образа на выходе получают образ, незначительно отличающийся от входного (рис. 2). Очень важно не «переобучить» сеть.



**Рис. 2. Результаты использования натренированной на символах русского алфавита ИНС применительно к неизвестным ей символам латинского алфавита и спецсимволам**

После обучения на вход НС подается изображение неизвестного ей символа (например, цифры или символа латинского алфавита). Выполняя функцию обобщения, НС пытается классифицировать новое изображения к одному из известных эталонов. В первом случае,

если в эталонном множестве имелось изображение похожего символа (например, «E» для «F», или «I» для «N», «P» для «R»), то изображение на выходе НС будет соответствовать эталонному с некоторой степенью искажения. В противном случае (например, для изображений цифр) сеть не сможет соотнести изображение ни с одним из эталонов, и изображение на ее выходах будет состоять из элементов тех символов, которые более всего похожи на входной. Именно здесь и проявляется механизм машинного творчества на основе ошибок распознавания.

Проводя параллель со сказанным выше, следует отметить, что первый случай аналогичен попытке человеческого мозга достроить неизвестные световые пятна до некоторых образов, виденных ранее. Во втором случае сеть генерирует изображение, состоящее из элементов ранее изученных эталонов.

Интересный эффект наблюдается, если на  $(i+1)$ -й итерации подать на входы НС выходные изображения  $i$ -й итерации. При подаче на первой итерации неизвестного образа возможны два случая:

1) через  $N$  итераций на выходах сети будет получено изображение одного эталонного образца практически без искажений. При этом степень искажения выходного изображения будет уменьшаться с ростом  $N$ ;

2) сеть не сможет достроить новое изображение до эталонного. На выходах сети будет формироваться образ, состоящий из элементов эталонов. С ростом  $N$  отличия сигналов на выходах на  $(i+1)$ -й и  $i$ -й итерации становятся все менее значительными. Можно утверждать, что выходное изображение сходится к изображению некоторого не представленного ни в одном алфавите символа, т. е. символа, который человеку не известен.

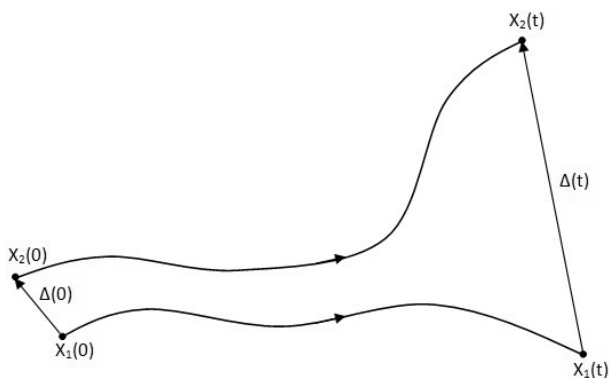
Как говорилось выше, хаотическая динамика также может быть применена в контексте рассматриваемого подхода.

**Динамический хаос в оценке оригинальности образа.** Предположим, что мы обладаем образом  $I$  (обобщая, считаем  $I$  многопризнаковым метрическим пространством образов) некоторого реально существующего объекта  $O$ , т.е. при помощи некоторого отображения образ был поставлен в соответствие реальному объекту:  $\psi: O \rightarrow I$  (это может быть сделано как человеком, так и при помощи некоторого алгоритма распознавания образов). Для наших целей также принимаем, что образ  $I$  воспроизводит объект  $O$  с необходимой точностью. После получения образа  $I$  внесем в него некоторый малый шум  $\varepsilon$ , в результате чего получим образ  $I' = I + \varepsilon$ , причем  $\varepsilon$  может быть настолько мало, что  $|\varepsilon| = |I' - I| \approx 10^{-20} \dots 10^{-10}$ . В общем случае значение шума зависит от того, насколько быстро должен дать результат процесс, который будет описан в дальнейшем. Теперь, если

будет найдена некоторая функция  $I_{t+1} = f(I_t)$ , обладающая существенной зависимостью от начальных условий, то можно будет сконструировать такой математический алгоритм, который способен оценить степень оригинальности образа  $I'$  и его производных во времени. Для наиболее полного описания предлагаемого подхода рассмотрим одну из теорем хаотической динамики.

Пусть  $(X, d)$  – метрическое пространство, содержащее бесконечное множество точек. Если отображение  $f : X \rightarrow X$  непрерывно и транзитивно, а периодические точки  $f$  плотны в  $X$ , то  $f$  обладает существенной зависимостью от начальных условий [18]. Воспользуемся этой теоремой для составления алгоритма поиска функции  $f$ , обладающей существенной зависимостью от начальных условий.

Рассмотрим два близких начальных условия:  $x_1(0) = x_0$  и  $x_2(0) = x_0 + \Delta(0)$ . Представим, что со временем траектории  $x_1(t)$  и  $x_2(t)$  изменяются так, как представлено на рис. 3. К моменту времени  $t$



**Рис. 3. Пример сильной зависимости функции от начальных условий**

разница между двумя траекториями составляет  $\Delta(t) = x_2(t) - x_1(t)$ .

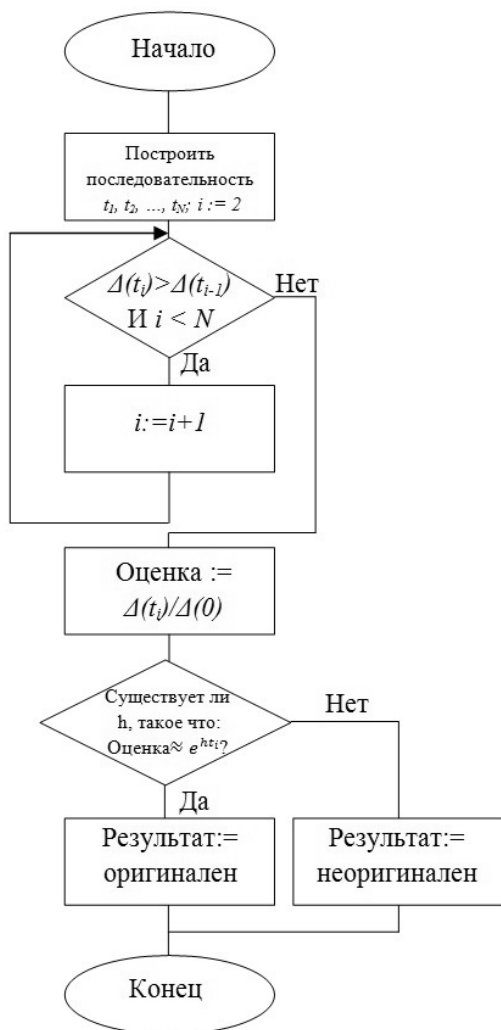
Если  $|\Delta(t)|$  растет экспоненциально для типичной ориентации вектора  $\Delta(0)$ , тогда говорят, что система выражает сильную зависимость от начальных условий. Само условие такой зависимости может быть записано в следующем виде:

$$\frac{|\Delta(t)|}{|\Delta(0)|} \sim e^{ht}, \quad h > 0.$$

Алгоритм, представленный на рис. 4, позволяет оценить степень оригинальности нового образа, полученного путем наложения шума на исходный образ. Оценка зависит от результатов анализа динамики разницы между траекториями представления образов (исходного и

ошибочного) при применении к ним некоторой функции. Алгоритм довольно прост. Конечно, в данной области необходимы дальнейшие исследования ввиду того, что проблема оценки оригинальности того или иного объекта, образа сама по себе является сложной и нерешенной задачей. По сути все машинное творчество может быть сведено к двум задачам:

- 1) синтез образа;
- 2) оценка оригинальности образа. Дополнительные критерии, представленные Ньэллом и другими учеными также могут быть задействованы по мере необходимости.



**Рис. 4. Вариант алгоритма оценки оригинальности образа**

Конечно, по окончании процесса получения нового образа и определения его оригинальности, нужно проверить полученный образ (его части) на существование в некоторой библиотеке образов.



Также возможна проверка непосредственно по применимости, если полученный образ не обнаружен в библиотеке (т.е. полностью соответствует второму критерию Ньюэлла, Шоу и Саймона).

В настоящей работе предпринята попытка рассмотреть механизм машинного творчества на основе ошибочных образов, а также интеграции в данную проблему математической теории хаоса. Если первое направление более тяготеет к этапу синтеза нового образа, то второе – к этапу оценки синтезированного образа (на новизну). Выполнение иных проверок по критериальной системе Ньюэлла – Шоу – Саймона на данном этапе работ представляется проблематичным ввиду крайне невысокой возможности их автоматизации. Особый интерес с научной точки зрения представляет дальнейшая проработка возможностей применения аппарата математической теории хаоса в динамических системах (какими и являются системы машинного творчества – творческие агенты) не только для оценки, но и непосредственно для синтеза новых образов. Например, требует рассмотрения идея новых (творческих) образов как неких аттракторов.

Применение машинного творчества (как такового, так и представленного в работе подхода) охватывает много направлений: синтез рифм, написание картин, создание образов новых персонажей для литераторов, синтез рассказов, творческое решение изобретательских задач и т. д.

Таким образом, в настоящей работе предложена идея нового подхода к актуальной проблеме машинного творчества. Сочетание новейших методов математики с классическими подходами искусственного интеллекта, используемыми при рассмотрении их стандартного функционала с новой позиции (например, ошибки классификации), может приблизить человечество еще на шаг к разгадке тайны машинного и человеческого творчества.

## СПИСОК ЛИТЕРАТУРЫ

1. Стобо Дж. Язык программирования Пролог: Пер. с англ. – М.: Радио и связь, 1993. – 368 с.
2. Кохонен Т. Ассоциативные запоминающие устройства. – М.: Мир, 1982. – 384 с.
3. Советский энциклопедический словарь / Гл. ред. А.М. Прохоров. – М.: Советская энциклопедия, 1987. – 1600 с.
4. Интернет-ресурс <http://www.thinkartificial.org>
5. Newell, A., Shaw, J. G., and Simon H. A. The process of creative thinking / H.E. Gruber, G. Terrell and M. Wertheimer (Eds.), *Contemporary Approaches to Creative Thinking*. – New York: Atherton, 1963. – P. 63–119.
6. Зарипов Р. Машинный поиск вариантов при моделировании творческого процесса. – М.: Наука. Гл. ред. физ.-мат. лит., 1983. – 232 с.
7. Newell A., Shaw J. C., Simon H. A. Report on a general problem-solving program / *Proceedings of the International Conference on Information Processing – USA*, 1959. – P. 256–264.

8. Федер Е. Фракталы: Пер. с англ. – М.: Мир, 1991. – 254 с.
9. Schmidt M., Lipson H. Distilling Free-Form Natural Laws from Experimental Data // Science. – 2009. – Vol. 324. – No. 5923. – P. 81–85.
10. Интернет-ресурс <http://raai.org>
11. Альтшуллер Г. Найти идею. Введение в ТРИЗ – теорию решения изобретательских задач. – М.: Альпина Паблишер, 2011. – 410 с.
12. Митрофанов В. В. Машина открытий. Компьютерная программа. Свидетельство об официальной регистрации программы для ЭВМ № 2000610103.
13. Чупров А., Кудрявцева Ю. Анатомия и физиология органа зрения. – Киров: КГМА, 2007. – 107 с.
14. Вятчинин Д. Нечеткие методы автоматической классификации. – Мн.: УП «Технопринт», 2004. – 219 с.
15. Ott E. Chaos in dynamical systems. – UK: Cambridge University Press, 1994. – P. 386.
16. Хайкин С. Нейронные сети: полный курс: Пер. с англ.– М.: Издательский дом «Вильямс», 2006. – 1104 с.
17. Гречищев К. М., Самарев Р. С. Особенности моделей представления изображений для нейросетевого распознавания символов. – М.: Эликс+, 2010. – С. 57–64.
18. Кроновер Р. Фракталы и хаос в динамических системах. Основы теории. – М.: Постмаркет, 2000. – 352 с.

Статья поступила в редакцию 14.05.2012