Vol.1, Chapter 8, the M.I.T. Press, Cambridge, MA 1986. P. 318–362.
5. Jang, J.-S. R. and C.-T. Sun, "Neuro-fuzzy modeling and control", Proceedings of the IEEE, March 1995.
6. R. Agrawal, R. Srikant. "Fast Discovery of Association Rules", In Proc. of the 20th International Conference on VLDB, Santiago, Chile, September 1994.
7. R. Agrawal, T. Imielinski, A. Swami. 1993. Mining Associations between Sets of Items in Massive Databases. In Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data, 207216.

Приведены результаты исследований по интеграции интеллектуальных методов и технологий для синтеза модели зависимости брака керамических изделий от физико-химических свойств керамической суспензии. Проведен анализ адекватности синтезированной модели и возможности ее применения в системах поддержки принятия управленческих решений в АСУТП керамического производства.

Ключевые слова: *нейро-нечеткие модели, керамическое производство, интеллектуальные методы*

The article describes the results of the research on intellectual methods and techniques integration for synthesis of dependency model of ceramic products' defects from physical and chemical properties of ceramic suspension. Adequacy analysis of the synthesized model and study of its implementation in Decision Support Systems in Process Control Systems of ceramic production was also carried out.

Keywords: *neuro-fuzzy models, ceramic production, intelligent techniques*

Рекомендовано до публікації д.т.н. Г.В. Кузнецовим. Дата надходження рукопису 06.10.10

УДК 681.5: 519.7

В.І. Корнієнко, Л.В. Будкова

ІДЕНТИФІКАЦІЯ ТА ПРОГНОЗУВАННЯ СТОХАСТИЧНИХ ПРИРОДНИХ ПРОЦЕСІВ ЗА ЧАСОВИМИ РЕАЛІЗАЦІЯМИ

V.I. Korniyenko, L.V. Budkova

IDENTIFICATION AND PREDICTION OF STOCHASTIC NATURAL PROCESSES BY TEMPORAL REALIZATIONS

Розроблено модифіковану методику ідентифікації складних природних процесів за часовими реалізаціями, що включає визначення режиму роботи породжуючої системи та його характеристик, синтез математичної моделі. На прикладі прогнозування витрат річок Дніпра, Неману та Сіверського Донця показана ефективність цієї методики.

Ключові слова: *ідентифікація, прогнозування, стохастичний процес, часова реалізація, нелінійна система*

Вступ. Такі природні процеси, як розливи річок під час повені і паводків викликають руйнування будівель, пошкодження гідротехнічних і дорожніх споруд, загибель посівів і лісів, а також приводять до людських жертв.

Тому розробка ефективних методів аналізу й прогнозу стоків річок є актуальною.

Постановка завдання. Природні процеси характеризуються нестаціонарністю, стохастичністю й нелінійністю (включаючи хаотичну динаміку й фрактальну розмірність).

Ідентифікація процесу, як динамічної системи, полягає в одержанні за експериментальними даними його математичної моделі.

Динамічна система зображується векторним рівнянням стану [1, 2]

$$\dot{x} = \Phi(x, \lambda), \tag{1}$$

де Φ – нелінійна функція розмірності d; x і λ – вектори координат і параметрів системи.

Процес у вигляді потоку (1) може бути також представлений дискретним відображенням Пуанкаре

$$x[k+1] = \Phi\{x[k], \lambda\} \ ; x[k] = \{x_1[k], ..., x_{d-1}[k]\}, \tag{2}$$

де k – такт часу $t = k \cdot T$; T – період дискретизації.

Залежно від значень параметрів порядку λ системи (1) – (2) мають чотири стійких рішення: стан рівноваги, коли після перехідного процесу система досягає стаціонарного стану; періодичне й квазіперіодичне рішення, а також хаос. Цим типам рішень відповідають атрактори системи у вигляді стійкої рівноваги, граничного циклу, квазіперіодичного атрактора й хаотичного (дивного) атрактора.

Відмінною рисою останнього є його чутливість до початкових умов і дробова розмірність – фрактальність (властивість самоподоби на різних масштабах).

Оскільки параметри порядку λ безпосередньо не спостерігаються, то актуальною є ідентифікація режимів і характеристик природних процесів за експериментальними часовими реалізаціями.

Ідентифікація таких складних процесів традиційними способами вимагає великих витрат на експериментальні дослідження. Методи ж нелінійної динаміки дозволяють з єдиних позицій визначати (класифікувати) й досліджувати природні процеси за окремими часовими реалізаціями із синтезом їх адекватних моделей. При цьому для зниження витрат доцільно використовувати методи систем штучного інтелекту, зокрема, нейронні мережі та системи нечіткої логіки, що легко навчаються і є універсальними й ефективними апроксиматорами [3].

Крім того, відомі методи (наприклад, [4]), орієнтовані на ідентифікацію об'єктів керування з метою їх автоматичного керування, що вимагає корегування цих методів при розв'язанні задач прогнозування природних процесів.

Таким чином, невирішеною задачею є розробка маловитратних і ефективних методів ідентифікації та прогнозування стохастичних природних процесів.

Мета статті. Розробка модифікованої методики ідентифікації складних природних процесів за часовими реалізаціями, що включає визначення режиму роботи породжуючої системи, його характеристик і синтез математичної моделі, а також оцінку ефективності прогнозування процесів за цією методикою.

Методика ідентифікації складних природних процесів за часовими реалізаціями включає наступні етапи.

1. Визначення режиму функціонування системи:

– якісний аналіз вигляду часового сигналу, його спектра, кореляційної функції і часо-частотного перетворення;

– побудова фазового портрета атрактора по дискретному відображенню часової реалізації;

– обчислення кореляційної ентропії K_C;

– обчислення показника Херста H.

2. Визначення розмірності (порядку) системи:

– обчислення кореляційного інтервалу передбачуваності (глибини прогнозу) процесу T_C;

– обчислення кореляційної розмірності атрактора D_C;

– визначення розмірності вкладення атрактора d (розмірності фазового простору) системи.

3. Реконструкція моделі режиму системи:

– вибір базисних функцій Φ (структури);

– визначення параметрів λ моделі.

Експериментальні реалізації містять інформацію про породжуючі системи.

Для періодичних рухів спектр сигналу містить дискретні лінії, тоді як хаос через аперіодичність зображається смугою на низьких частотах.

Кореляційна функція для регулярних рухів постійна чи осцилює, а в хаотичному режимі експоненційно спадає. Крім того, якісними ознаками хаотичності руху системи є самоподібна структура його часо-частотного (вейвлет) перетворення.

Найважливішою характеристикою руху у фазовому просторі довільної розмірності є ентропія Колмогорова K, що описує динамічне поводження на атракторі [1, 2]. Ентропія K пропорційна швидкості втрати інформації про стан динамічної системи з часом і показує наскільки динамічна система хаотична: K-ентропія дорівнює нулю для регулярного руху, нескінченна для випадкових систем, позитивна й обмежена для систем з детермінованим хаосом.

Крім того, K-ентропія дозволяє визначити середній час, на який можна передбачити стан системи з динамічним хаосом. Точне прогнозування стану цієї системи можливе тільки на інтервалі часу T_d, такому, що $\varepsilon \cdot e^{KT_d} = 1$ [2], відкіля

$$T_d = \frac{1}{K}\ln\left(\frac{1}{\varepsilon}\right), \tag{3}$$

де ε – розмір осередку фазового простору.

Оцінкою (знизу) колмогорівської ентропії K є кореляційна ентропія $K_C \le K$, тоді оцінка (зверху) інтервалу передбачуваності виконується аналогічно виразу (3)

$$T_C = \frac{1}{K_C}\ln\left(\frac{1}{\varepsilon}\right) \ge T_d .$$

За часи, більші T_C, можливе тільки статистичне прогнозування, інтервал (глибина) якого залежить від кореляційної функції процесу [5].

Більшість природних процесів, включаючи річкові стоки, температуру, опади, сонячні плями, є „зміщеними випадковими блуканнями“ – трендами з шумом [2, 5]. Відношення тренда (детермінованого чинника) до рівня шуму (випадкового чинника) характеризує показник Херста H [2]. Його визначення дозволяє класифікувати часові ряди, відрізняючи випадковий ряд від невипадкового.

Відстань між найближчими точками атрактора до і після біфуркацій знаходиться в універсальному відношенні. Самоподоба такого явища описується за допомогою фрактальної розмірності, чисельне визначення якої виконується за допомогою кореляційної розмірності D_C [1, 2].

Розмірність фазового простору d визначається за залежністю $D_C(d)$, для якої D_C перестає змінюватися, що є мінімальною розмірністю вкладення атрактора, тобто найменша ціла розмірність фазового простору, що містить весь атрактор.

Разом з тим, із теореми про вкладення [2] випливає, що оцінка розмірності фазового простору d визначається через оцінку розмірності атрактора D_C реальної динамічної системи (формула Мане)

$$d \ge 2D_C + 1 . \tag{4}$$

Реконструкція моделі динамічної системи на основі аналізу часових реалізацій полягає у виборі базисних функцій (структури) і їх коефіцієнтів (параметрів) моделі, а також визначенні значень параметрів моделі, оптимальним чином відповідних часовій реалізації.

Для розв'язання задачі реконструкції моделі (1), (2) формується d-мірне відображення виду

$$x_{1,i+1} = \Phi_1\{x_{1,i}, x_{2,i}, \ldots, x_{d,i} | \lambda\};$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_{d,i+1} = \Phi_d\{x_{1,i}, x_{2,i}, \ldots, x_{d,i} | \lambda\},$$

де $x_{j,i}$ – координати вектора стану, розглянуті в моменти часу $i \cdot T$; $j = \overline{1,d}$; $i = \overline{d, N-1}$.

Далі еволюційні функції $\Phi_j\{x_i\}$ відображаються у вигляді розкладання по деякому базисі із шуканими коефіцієнтами (параметрами) λ.

Традиційно для апроксимації функцій використовуються поліноми Лежандра чи інші. Коефіцієнти цих поліномів утворюють невідомі параметри λ, значення яких обираються так, щоб щонайкраще відповідати часовим реалізаціям, що спостерігаються, наприклад, за критерієм мінімуму похибки

$$E_x^2 = \sum_{j=1}^{d} \sum_{i=d}^{N-1} \left[x_{j,i+1} - \Phi_j\{x_i\} \right]^2 \to \min. \qquad (5)$$

Більш продуктивним для апроксимації функцій $\Phi_j\{x_i\}$ є використання нейронних мереж чи гібридних нейронних мереж з нечіткою логікою, що є універсальними й ефективними апроксиматорами. Їх параметрами λ є ваги нейронів, а також коефіцієнти їх функцій активації і належності (для гібридних мереж). Навчання мереж здійснюється, наприклад, методом зворотного поширення похибки і полягає у визначенні значень параметрів λ, що оптимальним чином (відповідно до критерію (5)) відповідають часовим реалізаціям, що спостерігаються.

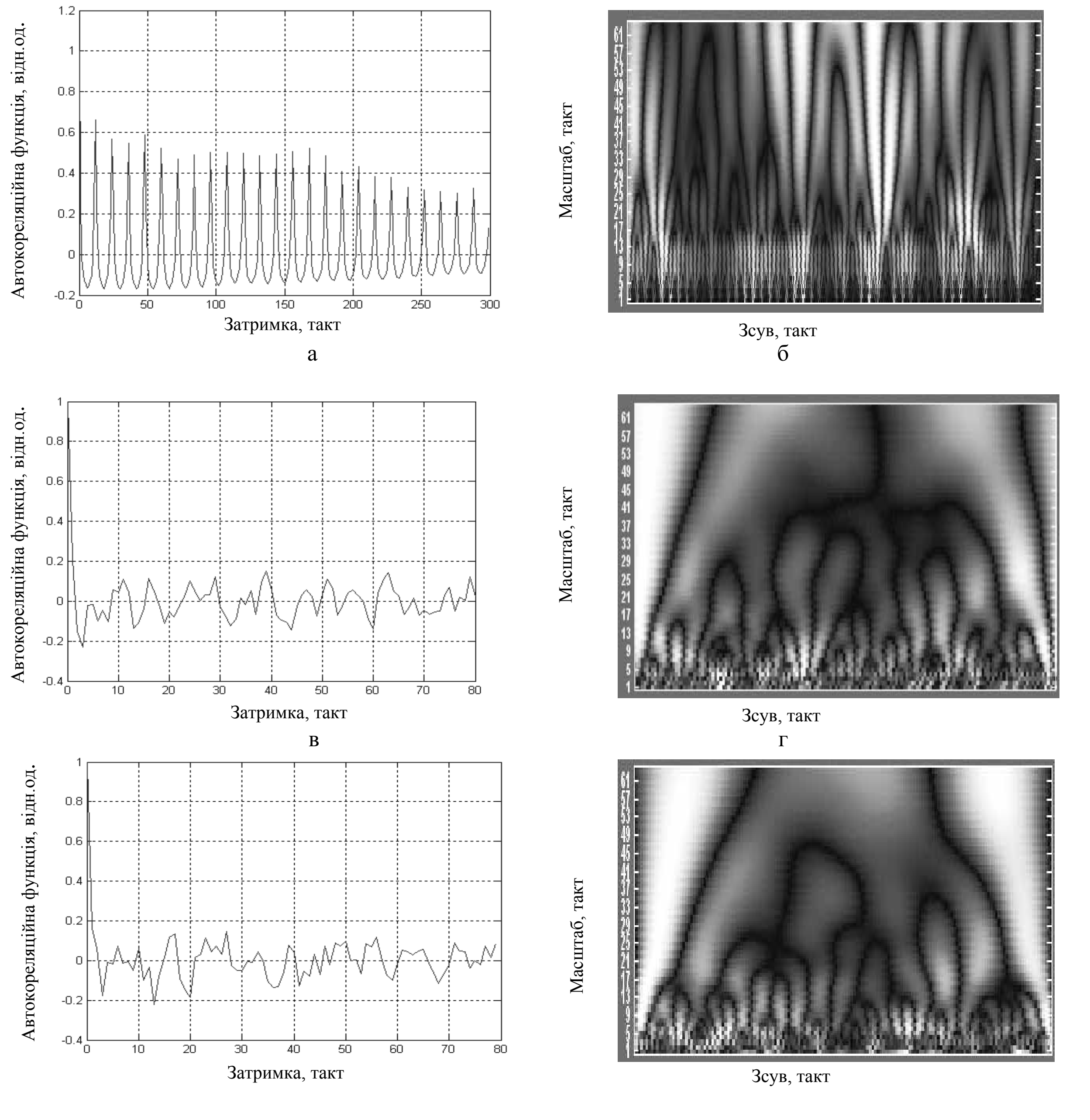

Рис. 1. Кореляційні функції та вейвлет перетворення витрат води Сіверського Донця (а, б), Дніпра (в, г) та Неману (д, е)

Прогнозування стохастичних природних процесів виконувалось відповідно до запропонованої методики за допомогою розроблених програм у середовищі Matlab.

Розглядалось прогнозування витрат річок за їх експериментальними часовими реалізаціями на прикладі Сіверського Донця, Дніпра та Неману [5]. При цьому інтервал дискретизації витрат для Сіверського Донця складав $T = 1$ місяць, а для Дніпра та Неману – $T = 1$ рік.

Кореляційні функції та вейвлет перетворення часових реалізацій витрат річок наведені на рис. 1.

За їх виглядом для Сіверського Донця можна сказати про регулярність породжуючого процесу, оскільки його кореляційна функція і вейвлет перетворення мають періодичний вигляд, характерний для полігармонійного авторегресійного процесу (рис. 1, *а* і 1, *б*)

Для Дніпра і Неману породжуючі процеси нерегулярні, оскільки їх кореляційні функції мають вид авторегресійного процесу (рис. 1, *в* і 1, *д*), а їх вейвлет перетворення мають самоподібний (фрактальний) характер (гілляста структура на рис. 1, *г* і 1, *е* зберігається на різних масштабах).

Побудова фазового портрету витрат води Сіверського Донця дозволила знайти зворотню періодичність руху породжуючої системи терміном у 6 місяців.

Для Дніпра і Неману фазові портрети витрат води при розмірності $d \le 3$ не мають закономірностей руху, що може бути викликано або великим рівнем шуму, або перебуванням породжуючих процесів у несталих режимах, або розмірність їх фазового простору більше 3.

При дослідженні витрат води Сіверського Донця значення показника Херста, обчислені за допомогою програми Fractan [6], склало $H_1 = 0{,}5307 \pm 0{,}1518 > 0{,}5$. Це персистентна реалізація, для якої найбільш ймовірним є збереження тенденції розвитку в майбутньому.

Розрахунки показника Херста H для Дніпра і Неману показали, що їх витрати води є ергодичними часовими рядами, оскільки $H_2 = 0{,}2197 \pm 0{,}1067$ і $H_3 = 0{,}3668 \pm 0{,}1244$, тобто $H < 0{,}5$. Такі часові реалізації мінливі і складаються з частих реверсів спад-підйом.

Виконані розрахунки також визначили значення кореляційних ентропії і розмірності атракторів часових реалізацій витрат води: Сіверського Донця – $K_{C1} = 0{,}4$ і $D_{C1} = 2{,}922$, Дніпра – $K_{C2} = 0{,}43$ і $D_{C2} = 3{,}08$ та Неману – $K_{C3} = 0{,}304$ і $D_{C3} = 3{,}178$. При цьому інтервали їх точної передбачуваності склали $T_{C1} = 4{,}02$, $T_{C2} = 3{,}74$ і $T_{C3} = 5{,}29$ такту (тривалістю T).

Для визначення розмірності фазового простору d за виразом (4) обчислювалась їх оцінка зверху: $d_1 \le 6{,}84$, $d_2 \le 7{,}16$ і $d_3 \le 7{,}4$, а для оцінки значення d знизу будувались залежності $D_C(d)$. Таким чином визначено, що $3 \le d_1 \le 7$, $4 \le d_2 \le 7$ і $5 \le d_3 \le 7$.

Для реконструкції моделей сигналів використовували адаптивну нейронну систему нечіткого висновку з дзвіноподібною функцією належності [3]. На її вхід подавали відповідні часові реалізації зі своїми розмірностями d (глибиною пам'яті $d-1$). Реалізації розбивалися на навчальну і перевірочну послідовності нарівно, а прогноз виконувався глибиною до 10 тактів. Як показник ефективності прогнозування використовували відносну середньоквадратичну похибку (рис. 2).

З аналізу отриманих результатів виходить, що похибка прогнозування витрат води Сіверського Донця після 6-го такту складає більше 10%, а для Дніпра та Неману – менше 8% при глибині прогнозу до 10 тактів. При цьому для інтервалів передбачуваності T_C похибка менше 2% для цих річок (для порівняння, відносна похибка прогнозування лінійними фільтрами складає більше 20%).

Статистична перевірка показала, що запропоноване прогнозування адекватне з імовірністю 0,01 розглянутим часовим реалізаціям при глибині прогнозу до 6–10 тактів.

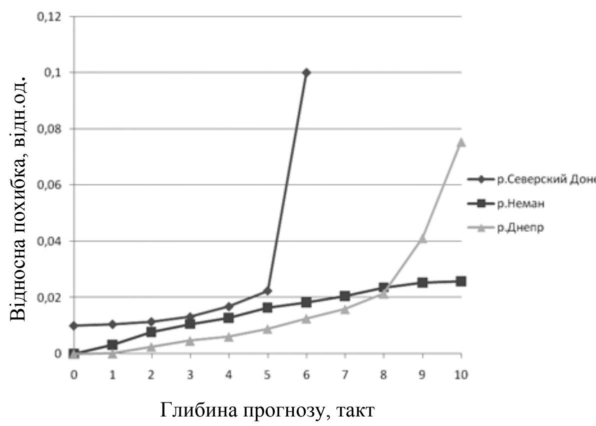

Рис. 2. Залежність відносної похибки від глибини прогнозу

Висновки. Розроблено модифіковану методику ідентифікації стохастичних природних процесів за часовими реалізаціями, що дозволяє визначати режим функціонування динамічної системи, її розмірність та реконструювати її модель з підвищеною точністю.

Ефективність запропонованої методики оцінена для часових реалізацій витрат води річок. Визначено, що реалізація цієї методики забезпечує зниження похибки прогнозування до 10% при глибині прогнозу в 6–10 тактів.

Подальші дослідження мають бути спрямовані на розробку методів оптимізації вибору базисних функцій моделей (структурно-параметричної ідентифікації) породжуючих природних процесів.

Список літератури
1. Кузнецов С.П. Динамический хаос: монография / С.П. Кузнецов. – М.: Физматлит, 2002. – 296 с.
2. Шустер Г. Детерминированный хаос. Введение : пер. с нем. / Генрих Шустер. – М.: Мир, 1988. – 256 с.
3. Круглов В.В. Нечеткая логика и искусственные нейронные сети / В.В. Круглов, М.И. Дли, Р.Ю. Голунов. – М.: Солон, 1996. – 348 с.
4. Корнієнко В.І. Ідентифікація нелінійних процесів по часових реалізаціях / В.І. Корнієнко, Д.Ю. Скриль // Науковий вісник Національного гірничого університету – 2009. – №3. – С. 85–89.
5. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами / А.Г. Ивахненко. – К.: Техніка, 1975. – 312 с.
6. Сычев В. Фрактальный анализ. Программа Fractan 4.4 / Вячеслав Сычев. – http:// impb.ru/~sychyov/.

Разработана модифицированная методика идентификации сложных природных процессов по временным реализациям, которая включает определение режима работы порождающей системы и его характеристик, синтез математической модели. На примере прогнозирования расходов рек Днепра, Немана и Северского Донца показана эффективность этой методики.

Ключевые слова: *идентификация, прогнозирование, стохастический процесс, часовая реализация, нелинейная система*

It has been designed the modified methodology for identification of stochastic natural processes on temporal realizations, which includes determination of the mode of operation of the originative system, its descriptions and synthesis of mathematical model. The efficiency of this method is shown on the example of prediction of flow of the rivers Dniper, Neman and Siverskyi Donets.

Keywords: *identification, prediction, stochastic process, temporal realization, nonlinear system*

Рекомендовано до публікації д.т.н. Г.В. Кузнецовим. Дата надходження рукопису 15.09.10

УДК 681.3

Т.А. Рузова

УСОВЕРШЕНСТВОВАНИЕ АЛГОРИТМА ГРЭХЕМА ДЛЯ ПОСТРОЕНИЯ ВЫПУКЛЫХ ОБОЛОЧЕК ОБЪЕКТОВ РАСТРОВОЙ ГРАФИКИ

T.A. Ruzova

GRAHAM ALGORITHM DEVELOPMENT FOR CONSTRUCTING CONVEX HULLS OF BITMAPPED OBJECTS

Проведен анализ проблем, возникающих при использовании алгоритма Грэхема для построения выпуклых оболочек к объектам растровой графики. Предложено усовершенствование алгоритма для построения выпуклых оболочек объектов с целочисленными координатами. Модификация алгоритма позволяет избежать погрешностей, связанных с дискретным представлением контура, при построении оболочек объектов сложной конфигурации. Предложенный подход может быть использован при разработке компьютерных систем для решения широкого круга задач.

Ключевые слова: *алгоритм Грэхема, выпуклая оболочка, растровое изображение, граница объекта, полярный угол*

Введение. Понятие выпуклой оболочки широко используется при решении различных задач вычислительной геометрии [1, 2]. Таких, на первый взгляд, далеких от нее как задача о максимумах множества точек, а также для широкого круга задач статистики. Если многомерные статистические данные рассматривать как точки в евклидовом пространстве, то многие статистические задачи (такие как задача о робастном оценивании, монотонной регрессии и др.) можно рассматривать с точки зрения вычислительной геометрии. Во многих задачах построение выпуклой оболочки является определяющим моментом. Выпуклые оболочки находят применение и в задачах распознавания образов. Один из часто используемых способов описания формы объекта основан на построении его выпуклой оболочки и определении дефицита выпуклости [3, 4]. Понятие дефицита выпуклости используется также при построении деревьев вогнутости, в задачах о восстановлении формы объектов на изображении [5]. Выпуклые оболочки применяются также для сегментации агрегированных объектов [6]. Построение выпуклых оболочек является существенным этапом решения задач компоновки и раскроя материалов [7]. Поэтому методам построения выпуклых оболочек посвящено большое число публикаций [1, 8–12]. К числу наиболее известных относят алгоритмы Грэхема, Джарвиса, Чена.

Постановка задачи. Алгоритм Грэхема [11] является одним из наиболее простых и позволяет построить выпуклую оболочку N точек на плоскости за время *O(N log N)* при памяти O(N) с использованием только арифметических операций и сравнений [1]. Однако, применение этого алгоритма в дискретных системах, к которым относятся изображения растровой графики, сопряжено с рядом проблем, связанных с использованием полярных координат. При переходе из непрерывного пространства вещественных чисел к дискретным координатам объек-