

УДК 519.6

Определение оптимальных уровней восстановления и квантования плотности нормального распределения в метрике квадратичного отклонения для алгоритмов сжатия данных

Пастухов Ю.Ф., канд. ф. -м.н., доцент

Пастухов Д.Ф., канд. ф. -м.н., доцент

Богуш Р.П., канд. т.н., доцент

Полоцкий государственный университет, Республика Беларусь

Пастухов А.Ю.

В настоящее время для решения многих практических задач актуальным является сжатие информации из-за большого объема передаваемых или хранимых данных, например, в системах дистанционного зондирования Земли на основе радиолокаторов с синтезированной апертурой антенны передаваемый поток достигает 4 Гб/с [29]. В большинстве случаев данные обладают избыточностью, благодаря чему возможно применение алгоритмов сжатия, причем, его эффективность будет значительно выше, если учитывать особенности обрабатываемых данных. Известно, что значения компонент принимаемого сигнала радиолокатором с синтезированной апертурой имеют нормальное распределение с нулевым средним, поэтому для подобного типа сигналов используются алгоритмы типа блочного адаптивного квантования с коэффициентами для нормального типа распределения. Определение пороговых уровней в существующих алгоритмах в значительной степени зависит от начальных данных и, соответственно, не обеспечивает их эффективность для максимально точного восстановления отсчетов сжатого сигнала на приемной стороне. Поэтому новым в данной работе является алгоритм квантования функции плотности нормального распределения в пространстве ступенчатых функций на заданном интервале, который свободен от указанного недостатка [1-29].

Одним из самых распространенных алгоритмов сжатия является алгоритм Ллойда, имеющий описанные выше недостатки, результат сходимости в нем существенно зависит от начальных данных. В поставленной в данной работе задаче предлагается метод и алгоритм, отличный от алгоритма Ллойда, данные работы которого известны для максимального числа ступеней равного 4 на одну сторону (всего 8 – на обе стороны от оси ординат) и дающий существенно лучший, чем у Ллойда результат (отношение сигнал/шум). Для черно-белых 8-битовых файлов превышение показателя сигнал/шум по сравнению с алгоритмом Ллойда составило более 10%.

Определение. Пусть $n \in \mathbb{N}$. Функция $f_n : [a; b] \rightarrow \mathfrak{R}$ ($a < b$) называется n -кусочно-постоянной (или n -ступенчатой) на $[a, b]$, если $\exists x_1 < x_2 < \dots < x_{n-1}$ такие что:

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < b = x_n,$$

$$f_n(x) = c_i = \text{const} \quad \forall x \in (x_{i-1}, x_i),$$

$$f_n(x_i) = c_i \text{ или } f_n(x_i) = c_{i+1}, \quad c_i \neq c_{i+1}, \quad \forall i = \overline{1, n-1}.$$

Множество n -ступенчатых функций $f_n : [a; b] \rightarrow \mathfrak{R}$ ($a < b$) обозначим как $S_n[a, b]$.

Пусть $f : [a, b] \rightarrow \mathfrak{R}$, $f \in C^2[a, b]$, $f'(x) < 0 \quad \forall x \in [a, b]$. $n \in \mathbb{N}$. Для минимизации ошибки квантования требуется в пространстве n -ступенчатых функций найти

наилучшее приближение $h_n : [a; b] \rightarrow \mathfrak{R}$ функции $f : [a, b] \rightarrow \mathfrak{R}$ в метрике квадратичного отклонения, такое что $\text{dist} = \|f - h_n\|_{C^2[a, b]} = \min_{f_n \in S_n[a, b]} \|f - f_n\|_{C^2[a, b]}$. С учетом этого, расстояние оценивается как:

$$\text{dist} = \|f - h_n\|_{C^2[a, b]} = \sqrt{\int_a^b (f(x) - h_n(x))^2 dx} = \min_{f_n \in S_n[a, b]} \sqrt{\int_a^b (f(x) - f_n(x))^2 dx} = \min_{f_n \in S_n[a, b]} \|f - f_n\|_{C^2[a, b]}$$

Пусть $h_n(x) = c_k$ при $x \in (B_{k-1}, B_k)$ $k = \overline{1, n}$, функция $G(B_1, \dots, B_{n-1}, C_1, \dots, C_n) = \sum_{k=1}^n \int_{B_{k-1}}^{B_k} (f(x) - c_k)^2 dx$

описывает квадрат отклонения (ошибки) функции $h_n : [a; b] \rightarrow \mathfrak{R}$ от функции $f : [a, b] \rightarrow \mathfrak{R}$ отклонения. Необходимое условие экстремума функции $G(B_1, \dots, B_{n-1}, C_1, \dots, C_n)$ описывается системой уравнений:

$$G'_{B_i} = 0, \quad i = \overline{1, n-1}, \quad G'_{C_j} = 0, \quad j = \overline{1, n},$$

из которой следует:

$$\begin{cases} f(B_i) = \frac{1}{2}(C_i + C_{i+1}), i = \overline{1, n-1} \\ \int_{B_{j-1}}^{B_j} f(x)dx = C_j(B_j - B_{j-1}), j = \overline{1, n} \end{cases} \quad (1)$$

Система (1) имеет $2n-1$ уравнений и $2n-1$ неизвестных $B_1, B_2, \dots, B_{n-1}, C_1, C_2, \dots, C_n$.

При квантовании необходимо приблизить функцию плотности нормального распределения $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

ступенчатыми функциями, таким образом, чтобы ошибка этого приближения была минимально возможной. Условия $f \in C^2[a, b], f'(x) < 0 \forall x \in [a, b]$ гарантируют существование и единственность решения системы (1) для произвольного натурального n , а также наличие локального минимума для решения (1).

Учитывая экспоненциальное убывание плотности нормального распределения к нулю, в качестве правого края отрезка установлено значение 3,0 в качестве левого края отрезка принималось нулевое значение.

Авторами был разработан алгоритм решения системы (1), в предположении, что последняя ступенька равна 0-лю, (то есть n ненулевых ступенек и одна – последняя равна 0-лю, всего $n+1$ ступенька), то есть для системы ($2n$ переменных и $2n$ уравнений):

$$\begin{cases} f(B_i) = \frac{1}{2}(C_i + C_{i+1}), i = \overline{1, n-1} \\ f(B_n) = \frac{1}{2}C_n \\ \int_{B_{j-1}}^{B_j} f(x)dx = C_j(B_j - B_{j-1}), j = \overline{1, n} \end{cases} \quad (2)$$

Для $n+1$ ненулевой ступеньки система (1) имеет вид :

$$\begin{cases} f(B_i) = \frac{1}{2}(C_i + C_{i+1}), i = \overline{1, n} \\ \int_{B_{j-1}}^{B_j} f(x)dx = C_j(B_j - B_{j-1}), j = \overline{1, n+1} \end{cases} \quad (3)$$

И содержит $2n+1$ уравнений и $2n+1$ неизвестных.

Для исследования рассмотренного алгоритма разработано программное обеспечение, которое позволяет получать пороговые уровни и ошибку приближения для различного числа n (таблица 1).

Таблица 1. Результаты исследований ошибки квантования для различного количества уровней

n	8	16	32	64	128
G	$5,13 * 10^4$	$1,53 * 10^4$	$2,11 * 10^5$	$6,64 * 10^6$	$1,17 * 10^9$
$dist$	$2,27 * 10^2$	$1,24 * 10^2$	$4,59 * 10^3$	$2,58 * 10^3$	$3,42 * 10^5$

На рисунке 1 представлен пример квантования функции плотности нормального распределения на основе предложенного подхода в метрике квадратичного отклонения для числа ступеней $m=10$ и $m=20$ (на одну сторону)

Полученные значения пороговых уровней для $n=20$ ($m=10$ ступеней): 0; 0,350; 0,502; 0,644; 0,769; 0,892; 1,011; 1,130; 1,253; 1,379; 1,515; 1,658; 1,826; 2,006; 2,257; 2,552

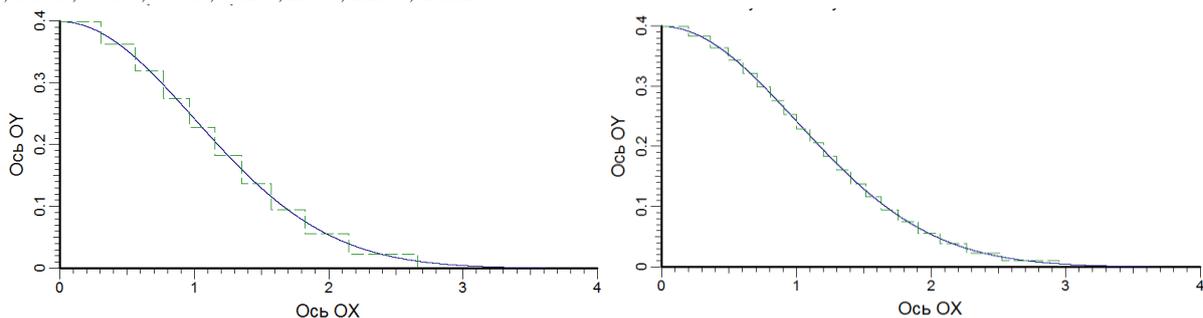


Рисунок 1. Результат квантования: а) для $m=10$; б) для $m=20$

Восстановление — процесс, обратный к сжатию данных, это значит, что поиск оптимальных уровней восстановления сводится к поиску оптимальных уровней квантования для функции, обратной к плотности нормального распределения.

Полученные значения пороговых уровней для $n=20$: 0.161; 0.330; 0.456; 0.582; 0.696; 0.811; 0.922; 1.033; 1.144; 1.256; 1.371; 1.486; 1.609; 1.732; 1.868; 2.004; 2.161; 2.319; 2.516; 2.712.

На рисунке 2 представлен пример восстановления функции плотности нормального распределения на основе предложенного подхода в метрике квадратичного отклонения для числа ступеней $m=10$ (уровней $n=20$) и $m=20$ (уровней $n=40$).

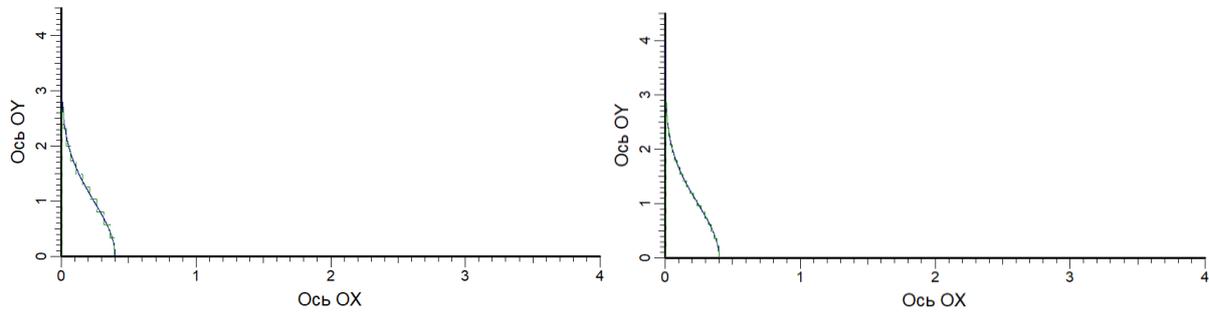


Рисунок 2. Результат восстановления: а) для $m=10$; б) для $m=20$

На рисунке 3,4 приведен результат работы программы для поиска оптимальных уровней квантования для числа ступеней $m=10$

На рисунке 5,6 приведен результат работы программы для поиска оптимальных уровней квантования для числа ступеней $m=10$

```

3,0
Введите число ступенек
10
Введите количество отрезков разбиения для расчета интеграла
По умолчанию количество отрезков разбиения для расчета интеграла=1000
Теперь кол-во отрезков разбиения для расчета интеграла можно вводить и 10000 (с
лед. параметр-количество отрезков подразбиения=100) - время расчета - где-то ми
нут 10-20-30
2000
Введите количество отрезков подразбиения для решения интегрального уравнения
По умолчанию количество отрезков подразбиения для решения интегрального уравнен
ия=100
10
Введите количество узлов (кратно 10) для расчета интеграла
20
Начало работы программы:
Время : год:2018 мес:11 дней:18 час:17 мин:28 сек:21 .
Прогресс: 100.0000 % Осталось: сут: 0 Час: 0 мин: 0 секунд: 0 .
Время : год:2018 мес:11 дней:18 час:17 мин:29 сек:19 .
Оценка приближения к решению = 3.485734061613373E-004
Вывод уровней квантования :
X( 1 )= 0.00000000000000E+000
X( 2 )= 0.306000000000000
X( 3 )= 0.438032075449999
X( 4 )= 0.560099999999972
X( 5 )= 0.665926937511052
X( 6 )= 0.768749999999949
X( 7 )= 0.866168501569582
X( 8 )= 0.962699999999928
X( 9 )= 1.05050868894321
X( 10 )= 1.154700000000002
X( 11 )= 1.253649496859061
X( 12 )= 1.354200000000015
X( 13 )= 1.46142082442052
X( 14 )= 1.571850000000028
X( 15 )= 1.69507733886316
X( 16 )= 1.824300000000044
X( 17 )= 1.97031191938808
X( 18 )= 2.147250000000065
X( 19 )= 2.38326328822217
X( 20 )= 2.662500000000097
Ошибка(квадрат расстояния)-интеграл квадрата разности f(x) и ступенчатой функции
= 3.485734061613373E-004
Расстояние = 1.867012067881023E-002
    
```

Рисунок 3. Уровни квантования: для $m=10$

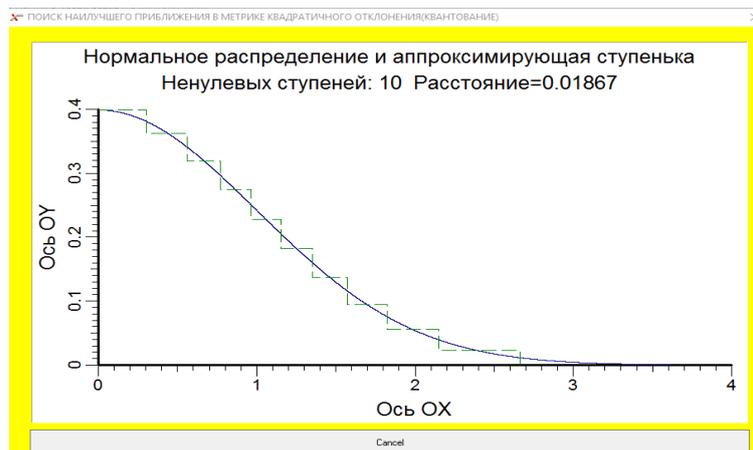


Рисунок 4. График уровней квантования: для $m=10$

```

E:\Lloyd_Result\Test3_1\Vosstanovlenie\Release\Kvant.exe"
Введите количество отрезков разбиения для расчета интеграла
По умолчанию количество отрезков разбиения для расчета интеграла=1000
Теперь кол-во отрезков разбиения для расчета интеграла можно вводить и 10000 (с
лед. параметр-количество отрезков подразбиения=100) - время расчета - где-то ми
нут 10-20-30
2000
Введите количество отрезков подразбиения для решения интегрального уравнения
По умолчанию количество отрезков подразбиения для решения интегрального уравне
ния=100
10
Введите количество узлов(кратно 10) для расчета интеграла
20
Начало работы программы:
Время : год:2018 мес:11 дней:18 час:17 мин:24 сек:47
Прогресс: 100.0000 % Осталось: сут: 0 час: 0 мин: 0 секунд: 0
Время : год:2018 мес:11 дней:18 час:17 мин:26 сек:7
Оценка приближения к решению = 7.943674477578655E-003
Вывод уровней восстановления:
X( 1 )= 0.160865976198399
X( 2 )= 0.329675626874377
X( 3 )= 0.455626544843146
X( 4 )= 0.581577462411915
X( 5 )= 0.696228853165765
X( 6 )= 0.810880243919615
X( 7 )= 0.921840102781452
X( 8 )= 1.03279996164329
X( 9 )= 1.14420897835909
X( 10 )= 1.25561799507489
X( 11 )= 1.37091357714337
X( 12 )= 1.48620915921184
X( 13 )= 1.60021530392186
X( 14 )= 1.73222144863187
X( 15 )= 1.86807270721979
X( 16 )= 2.00392396500771
X( 17 )= 2.16132200064482
X( 18 )= 2.31872003548193
X( 19 )= 2.51552636075299
X( 20 )= 2.71233268602406
Расстояние-корень интеграла квадрата разности f_1(x) и 0-лем =
0.865525405907619
Ошибка(квадрат расстояния)-интеграл квадрата разности f(x) и ступенчатой функции
= 2.283618672908710E-003
Расстояние = 4.778722290433615E-002
    
```

Рисунок 5- Уровни восстановления для $m=10$

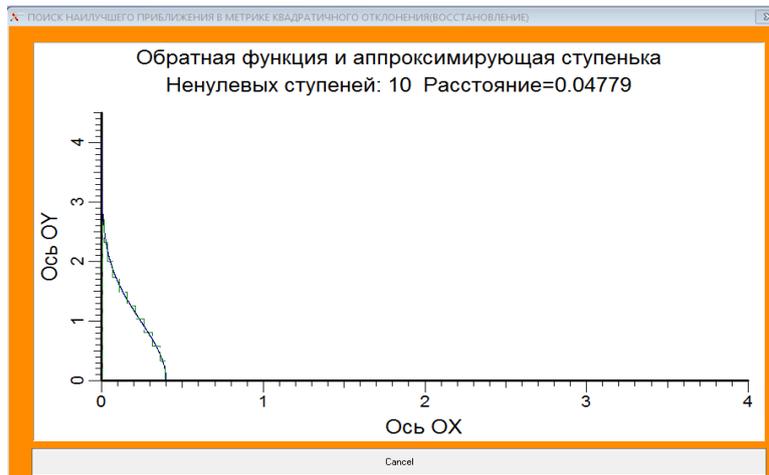


Рисунок 6. График уровней восстановления для $m=10$

Литература:

1. Пастухов, Д. Ф. Классификация профилей температуры в плюс - минус одно сантиметровом слое от поверхности раздела геотермального озера // Вестн. Моск. ун-та. Сер. 3. Физ. Астрон.. - 1995., Т.36. № 6 - С. 84.
2. Пастухов, Д. Ф. Метод эквивалентной струи для описания развитой турбулентной области температурного поля, порожденной взаимодействием нескольких конвективных струй. Деп. в ВИНТИ 20.04.94. № 934-В94. 13 с.
3. Пастухов, Д. Ф. Простая мультифрактальная модель источника плавучести эквивалентной турбулентной конвективной струи и вариационная задача минимизации скорости роста энтропии в струе.- Деп. в ВИНТИ 13.12.94 № 2887-В94. 1994, 15 с.
4. Анисимова, Е. П., Пастухов, Д. Ф. Расчет поля линий тока в вулканическом озере с геотермальными источниками на дне// Вулканология и сейсмология. №3. 1996. С. 94 - 100.
5. Анисимова, Е. П., Пастухов, Д. Ф., Сперанская, А. А., Фазлуллин, С. М. Особенности термогидродинамического режима вулканического озера// Вулканология и сейсмология. 1992. № 2. С. 71 - 76.
6. Анисимова, Е. П., Пастухов, Д. Ф., Сперанская, А. А. Температурный режим геотермального озера// Тезисы докладов 4-я конференция "Динамика и термика рек, водохранилищ, внутренних и окраинных морей". М., 1994, т.2, с.14-16.
7. Анисимова, Е. П., Пастухов, Д. Ф., Сперанская, А. А., Сперанская О.А. О роли аэрации в формировании термического режима геотермального озера// Изв. РАН, Физика атмосферы и океана, 1996. Т. 32. №2. С. 267-273.
8. Анисимова, Е. П., Пастухов, Д. Ф., Сперанская, А. А., Фазлуллин, С. М. Особенности термического режима геотермальных озер// Международная конференция "Геофизика и современный мир". Тезисы докладов. М., 1993, В 4, р.24.
9. Ю. Ф. Пастухов, "Необходимые условия в обратной вариационной задаче", Фундаментальная и прикладная математика, 7:1 (2001), 285–288.

10. Пастухов, Ю.Ф. Задача построения поля линий тока по температурному разрезу / Ю.Ф. Пастухов, Д.Ф. Пастухов // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2015. – № 4. – С. 27–36.
11. Пастухов, Д.Ф. Аппроксимация уравнения Пуассона на прямоугольнике повышенной точности / Д.Ф. Пастухов, Ю.Ф. Пастухов // Вестник Полоцкого университета. Серия С, Фундаментальные науки. – 2017. – № 12. – С. 62–77.
12. Пастухов, Ю.Ф. Инварианты в расслоениях скоростей произвольного порядка / Ю. Ф. Пастухов, Д. Ф. Пастухов, Голубева О. В. // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2015. – № 12. – С. 117–123.
13. Пастухов, Ю.Ф. Тензор обобщенной энергии / Ю. Ф. Пастухов, Д. Ф. Пастухов, С. В. Чернов // Вестник Полоцкого государственного университета. Серия С, Фундаментальные науки. - 2017. - № 12. - С. 78-100.
14. Пастухов, Ю.Ф. Группы преобразований, сохраняющие вариационную задачу со старшими производными / Ю. Ф. Пастухов, Д. Ф. Пастухов // Вестник Полоцкого государственного университета. Серия С, Фундаментальные науки. - 2018. - № 4. - С. 194.
15. Пастухов, Д. Ф. К вопросу о редукции неоднородной краевой задачи Дирихле для волнового уравнения на отрезке / Д. Ф. Пастухов, Ю. Ф. Пастухов, Н. К. Волосова // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2018. – № 12. – С. 54–68.
16. Пастухов, Д. Ф. Оптимальный параметр аппроксимации разностной схемы волнового уравнения на отрезке / Пастухов Д. Ф., Пастухов Ю. Ф., Волосова Н. К. // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. - 2018. - № 4. С. 167-186.
17. Соловьёв, А. А., Сперанская, О. А., Пастухов, Д. Ф., Пастухов, Ю. Ф. Временная последовательность состояний системы в гидродинамических задачах.
В сборнике: Прикладная математика и информатика: современные исследования в области естественных и технических наук Сборник научных статей IV научно-практической международной конференции (школы-семинара) молодых ученых: в двух частях. 2018. С. 496-502.
19. Карабанов, Р. Ю., Пастухов, Д. Ф., Пастухов, Ю. Ф., Богуш, Р. П. наилучшее приближение монотонно убывающих функций кусочно – постоянными функциями в метрике квадратичного отклонения.
В сборнике: Прикладная математика и информатика: современные исследования в области естественных и технических наук Сборник научных статей IV научно-практической международной конференции (школы-семинара) молодых ученых: в двух частях. 2018. С. 48-53.
20. Пастухов Ю. Ф., Пастухов Д. Ф., Богуш Р. П. Квантование функции плотности нормального распределения в метрике квадратичного отклонения. В сборнике: Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ-2018) Электронный сборник статей I международной научно-практической конференции, посвященной 50-летию Полоцкого государственного университета. 2018. С. 92-95.
21. Сперанская, О. А., Волосова, Н. К., Пастухов, Д. Ф., Пастухов, Ю. Ф., Соловьёв, А. А., Голубева, О. В., Ехилевский, С. Г., Радкевич, Д. С. Моделирование быстропериодического прерывания потока жидкости под давлением в оболочке ANSYS Fluent . В сборнике: Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ-2018) Электронный сборник статей I международной научно-практической конференции, посвященной 50-летию Полоцкого государственного университета. 2018. С. 73-79.
22. Герец, А. Ю. Согласование порядков аппроксимации дифференциального и граничного операторов в краевой задаче для уравнений в частных производных / А. Ю. Герец, А. А. Зеленкевич, Н. А. Гурьева, Ю. Ф. Пастухов, Д. Ф. Пастухов // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. 2015. № 12. С. 102-109.
23. Пастухов, Ю. Ф. Закон преобразования обобщенного импульса / Ю. Ф. Пастухов, С. Г. Ехилевский, О. В. Голубева, Д.Ф. Пастухов // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. 2017. № 4. С. 85-99.
24. Пастухов, Ю. Ф. Обобщение теоремы Гамильтона – Остроградского в расслоениях скоростей произвольного порядка / Ю.Ф. Пастухов, Ехилевский С. Г., Голубева О.В., Пастухов Д.Ф. // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. 2016. № 12. С. 125-133.
25. Пастухов Ю.Ф., Пастухов Д.Ф. Сборник статей по дифференциальной геометрии. <http://elib.psu.by:8080/handle/123456789/22094>
26. Пастухов, Ю. Ф. Лагранжевы сечения / Ю. Ф. Пастухов, Д.Ф. Пастухов // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2018. – № 12. – С. 69–93.
27. Пастухов, Ю. Ф. Исследование решения обратной вариационной задачи произвольного порядка. – Деп. в ВИНТИ № 1328 – В – 96.
28. Пастухов, Ю. Ф. Исследование локального решения обратной вариационной задачи // Тезисы докладов научно-практической конференции г.Новокузнецк: НГПИ, 1995.
29. Богуш, Р. П. Моделирование сжатия радиолокационных данных дистанционного зондирования земли алгоритмом ЕСВАQ и представления восстановленных данных в формате CEOS / Р.П. Богуш, И.Ю. Захарова, Н.М. Наумович // Вестник ПГУ, серия С. Фундаментальные науки, 2017г. №12 – с. 2-9