



<http://dx.doi.org/10.35596/1729-7648-2023-21-5-13-19>

Оригинальная статья
Original paper

УДК 519.2

ОПТИМАЛЬНЫЙ ГИСТОГРАММНЫЙ ФИЛЬТР

А. В. ОВСЯННИКОВ¹, В. М. КОЗЕЛ²

¹Белорусский государственный университет (г. Минск, Республика Беларусь)

²Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)

Поступила в редакцию 23.03.2023

© Белорусский государственный университет информатики и радиоэлектроники, 2023
Belarusian State University of Informatics and Radioelectronics, 2023

Аннотация. Рассматривается методика построения оптимального гистограммного фильтра и его модификаций с учетом априорной информации о предполагаемой плотности распределения вероятностей. Основная идея построения гистограммного фильтра заключается в применении специального преобразования, отображающего профиль участка любого закона распределения в эквивалентный ему постоянный уровень характеристических чисел – информационных весов. Это преобразование позволяет определить коэффициенты гистограммного фильтра. Оценка значения числа данных конкретного интервала гистограммы формируется характеристической функцией фильтра, содержащей реальные данные и соответствующей характеристическому числу. Показана сходимость оценок, полученных гистограммным фильтром, к истинным значениям вероятностей интервалов. Рассмотрены модификации оптимального гистограммного фильтра, требующие меньше вычислительных затрат на их реализацию. Получены верхние границы качественных характеристик фильтров. Установлено, что оптимальный гистограммный фильтр независимо от вида закона распределения обеспечивает в три раза лучшее качество идентификации (распознавания) в сравнении со стандартной гистограммной оценкой. Эффективность гистограммного фильтра подтверждается моделированием. Гистограммный фильтр является простым для реализации инструментом, который легко может быть встроен в любой открытый алгоритм идентификации (распознавания) закона распределения.

Ключевые слова: гистограммный фильтр, идентификация, характеристическая функция, интервал группирования, плотность распределения вероятности.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Овсянников, А. В. Оптимальный гистограммный фильтр / А. В. Овсянников, В. М. Козел // Доклады БГУИР. 2023. Т. 21, № 5. С. 13–19. <http://dx.doi.org/10.35596/1729-7648-2023-21-5-13-19>.

OPTIMAL HISTOGRAM FILTER

ANDREI V. AUSIANNIKAU¹, VICTOR M. KOZEL²

¹Belarusian State University (Minsk, Republic of Belarus)

²Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 23.03.2023

Abstract. The article discusses a technique for constructing an optimal histogram filter and its modifications, taking into account a priori information about the expected probability distribution density. The main idea of constructing a histogram filter is to apply a special transformation that displays the profile of a section of any distribution law into a constant level of characteristic numbers equivalent to it. This transformation allows to determine

the coefficients of the histogram filter. An estimate of the value of the number of data of a particular interval of the histogram is formed by the characteristic function of the filter containing real data and equivalent to the characteristic number. The convergence of the estimates obtained by the histogram filter to the true values of the interval probabilities is shown. Modifications of the optimal histogram filter that require less computational costs for their implementation are considered. The upper bounds of the qualitative characteristics of filters are obtained. It has been established that the optimal histogram filter, regardless of the type of distribution law, provides three times the best quality of identification (recognition) in comparison with the standard histogram estimate. The efficiency of the histogram filter is confirmed by simulations. The histogram filter is an easy-to-implement tool that can be easily integrated into any open distribution law identification (recognition) algorithm.

Keywords: histogram filter, identification, characteristic function, grouping interval, probability density distribution.

Conflict of interests. The authors declare no conflict of interests.

For citation. Ausiannikau A. V., Kozel V. M. (2023) Optimal Histogram Filter. *Doklady BGUIR*. 21 (5), 13–19. <http://dx.doi.org/10.35596/1729-7648-2023-21-5-13-19> (in Russian).

Введение

Разработка точных и быстрых методов, алгоритмов идентификации законов распределения, работающих, в том числе, на малых объемах данных, является востребованной и актуальной задачей для множества систем и устройств обработки данных общего и специального назначения [1–6]. Для облегчения встраивания программно-аппаратных реализаций алгоритмов идентификации в устройства первичной обработки данных эти алгоритмы должны разрабатываться на основе простых, универсальных методов, не требующих больших вычислительных затрат. В этих целях в статье рассматриваются модификации гистограммных оценок плотностей распределения вероятностей (ПРВ), эффективность которых можно значительно улучшить применением гистограммных фильтров (ГФ) [7, 8]. Использование ГФ значительно ослабляет зависимость гистограммных оценок от выбора числа интервалов группирования данных [4], устраняет сильную изрезанность гистограммы при относительно малом числе данных, повышает точность правильной идентификации ПРВ [7, 8].

Основная идея построения ГФ состоит в применении в нем функций принадлежности данных конкретному интервалу гистограммы вместо индикаторной функции включения. Однако эффективность таких ГФ с увеличением числа данных снижается вследствие отсутствия сходимости оценок ГФ к истинным при неограниченно возрастающем объеме данных. Устранить эту особенность можно специальным преобразованием с использованием априорной информации о предполагаемой ПРВ, в результате которого определяются коэффициенты ГФ, и сам ГФ становится оптимальным в смысле сходимости оценок частот интервалов к истинным значениям вероятностей. Это преобразование связано с преобразованием профиля участка предполагаемой ПРВ в постоянные характеристические числа участка, определяющие его информационный вес.

Цель исследований состояла в разработке алгоритмов реализации гистограммного фильтра на основе характеристических функций реальных данных с использованием специального преобразования для нахождения коэффициентов фильтра с учетом априорной информации об идентифицируемой ПРВ.

Методика построения оптимального гистограммного фильтра

Пусть имеется выборка случайных данных $\{x_i\}$, $i = \overline{1, n}$, и определено разбиение числовой прямой на m непересекающихся и примыкающих друг к другу интервалов A_j , $j = \overline{1, m}$, равной длины $\Delta_x = X_{j+1} - X_j = R/m$. $X_{m+1} = x_{\max} = \max_i x_i$, $X_1 = x_{\min} = \min_i x_i$, где X_j – граница интервалов; $R = x_{\max} - x_{\min} = m\Delta_x$ – размах диапазона данных. Заменяем обычную индикаторную функцию на диапазоне Δ_x , используемую при стандартном способе построения гистограммы, функцией принадлежности $\mu_j(x)$, $\mu_j(x) \in [0, 1]$, с областью определения $\Delta_\mu \leq 3\Delta_x$. Эта функция будет определять принадлежность данных интервалу группирования A_j [7, 8]. Введение в рассмотрение функций принадлежности предполагает обеспечение дополнительного условия $\sum_{j=1}^m \mu_j(x) = 1$. Для удобной практической реализации ГФ заменим непрерывную функцию $\mu_j(x)$ ступенчатой

с дополнительным индексом t – номером ее подынтервала. Тогда для внутренних и конечных интервалов можем записать:

$$\mu_{j,t}(x) = \begin{cases} \mu_{j,t}, & j = \overline{1, m}, \quad t = j, (j-m)(m-3)/(m-1) + (m-1); \\ \mu_{j,t}, & j = \overline{2, m-1}, \quad t = \overline{j-1, j+1}. \end{cases} \quad (1)$$

Исходя из условий нормировки, на ступенчатые функции (1) должны быть наложены ограничения

$$\sum_t \mu_{j,t} = 1. \quad (2)$$

Рассмотрим алгоритм ГФ с индексацией, соответствующей (1), в следующем виде:

$$u_j = \sum_t \varphi_{j,t} v_t, \quad (3)$$

где u_j – число данных (в общем случае не целое число), отнесенных ГФ к интервалу A_j ; v_t – число данных, попавших в t -й подынтервал j -й функции принадлежности; $\varphi_{j,t}$ – коэффициент ГФ.

Для определения коэффициентов $\varphi_{j,t}$ уравнения (2) недостаточно. Требуются дополнительные сведения о связи коэффициентов с априорной информацией относительно предполагаемой ПРВ. Такую связь установим в два этапа. На первом введем специальное преобразование, которое будет отображать профиль участка любого закона распределения в эквивалентный ему постоянный уровень:

$$\rho_j = \mu_{j,t} p_t = \text{const}_j, \quad (4)$$

где $\rho_j > 0$ – характеристическое число, определяющее информационный вес участка A_j .

С учетом (2) эти числа ρ_j определяются простыми арифметическими выражениями по множеству $\{p_t\}$: $\rho_j = S_j^{-1} H_j$, где $H_j = H_j(\{p_t\})$ – среднегармоническое значение элементов множества $\{p_t\}$; $S_j = (\sum \{p_t\})$ – сумма элементов множества $\{p_t\}$; $S_j = 2$ для $j = \overline{1, m}$ и $S_j = 3$ для $j = \overline{2, m-1}$.

Идея преобразования (4), поясняющая его смысл, состоит в следующем. Во-первых, определенные преобразованием (4) коэффициенты $\mu_{j,t} = \rho_j / p_t$, примененные к набору частот $\{v_t / n\}$, должны обеспечить близкий к предполагаемому профиль участка гистограммы; во-вторых, результат (4) можно получить исходя из максимизации энтропии набора информационных весов $-\sum_t \mu_{j,t} p_t \ln(\mu_{j,t} p_t) \rightarrow \max$. Решение этой задачи известно и соответствует (4).

Далее на втором этапе на основе преобразования (4) можно записать $\rho_j = S_j^{-1} \sum_t \mu_{j,t} p_t$. По аналогии с этим сформируем характеристическую функцию по реальным данным $\rho_j^* = \mu_{j,j} u_j / n = S_j^{-1} \sum_t \mu_{j,t} v_t / n$, которую преобразуем в уравнение ГФ (3), где коэффициенты $\varphi_{j,t}$ будут определяться формулой

$$\varphi_{j,t} = S_j^{-1} (p_j / p_t) = S_j^{-1} (\mu_{j,t} / \mu_{j,j}). \quad (5)$$

Таким образом, связь коэффициентов ГФ с априорной информацией относительно предполагаемой ПРВ выражается уравнением (5). Тривиальным примером алгоритма ГФ является идентификация равномерной ПРВ. Алгоритм (3) принимает вид простого усреднения с постоянным весом $u_j = S_j^{-1} \sum_t v_t$. Это показывает неправомерность применения такого алгоритма гистограммного сглаживания для произвольных законов распределения.

На основании соотношений (4) и (5) можно оценить сходимость результатов, полученных применением ГФ (3), к истинным значениям вероятностей на каждом из m интервалов: $\lim_{n \rightarrow \infty} (u_j / n) = S_j^{-1} \lim_{n \rightarrow \infty} \sum_t (p_j / p_t) (v_t / n) = p_j$. В силу этой сходимости ГФ можно считать оптимальным для предполагаемой ПРВ.

Алгоритмы (3) и (5) могут быть охарактеризованы как алгоритмы обобщенного сглаживания (фильтрации) гистограммы. В отличие от множества известных алгоритмов сглаживания (равномерного, медианного и др.), алгоритмы (3) и (5) используют настраиваемые веса $\varphi_{j,t}$, определяющиеся априорной информацией о предполагаемой ПРВ.

Оценим эффективность ГФ (3). Наиболее просто это можно сделать на основе критерия χ^2 -квадрат для простой гипотезы. Подставляя в стандартное выражение критерия χ^2 -квадрат значения (3), получим $\chi_u^2 = \sum_{j=1}^m (n\rho_j^* / \mu_{j,j} - np_j)^2 / np_j = n \sum_{j=1}^m [(\rho_j^* - \rho_j) / \rho_j]^2 p_j$. Воспользовавшись неравенством Коши–Буняковского, после ряда громоздких преобразований, которые здесь опускаются, находим верхнюю границу оценки математического ожидания критерия

$$\mathbf{M}\chi_u^2 \leq \mathbf{M} \sum_{j=1}^m \left[S_j^{-1} \sum_{\{t\}} \phi_{j,t} (v_t - np_t)^2 / np_t \right] \leq 1 + \mathbf{M}\chi_v^2 / 3, \quad (6)$$

где $\mathbf{M}\chi_v^2 = m - 1$.

Неравенство (6) получено в предположении выполнения условия $\mathbf{M}\chi_v^2 \leq 9((I_f \Delta_x^2)^{-1} - 1) / 2$, где I_f – информация Фишера о предполагаемой ПРВ. Анализ этого условия в свою очередь показывает, что число интервалов группирования данных должно быть не меньше величины $m^* = \delta^{-2}(1 + \sqrt{1 + 63\delta^2}) / 9$, где $\delta^2 = (I_f R^2)^{-1}$ – информационное отношение. Можно показать, что теоретически параметр δ^2 является функцией вероятности $\gamma = \int_R f(x) dx$, где $f(x)$ – предполагаемая ПРВ. Например, для гауссовской ПРВ $\delta^2 = (2\sqrt{2} \text{Erf}^{-1}(\gamma))^{-2}$, (Erf^{-1} – обратная функция ошибок), для лапласовской и экспоненциальной ПРВ $\delta^2 = (\ln(1 - \gamma))^{-2}$. Поскольку величина γ задается заранее, всегда можно определить число интервалов группирования данных $m \geq m^*$, при котором справедливо неравенство (6). В табл. 1 приведены значения нижних границ числа интервалов в зависимости от доверительной вероятности для некоторых распределений.

Таблица 1. Нижние границы числа интервалов
Table 1. Lower limits of intervals

Доверительная вероятность	0,9	0,95	0,98	0,99
Гауссовская ПРВ	5	6	8	9
Лапласовская и экспоненциальная ПРВ	3	4	6	7

Неравенство (6) показывает независимость верхней границы оценки математического ожидания критерия χ^2 -квадрат от конкретного вида предполагаемой ПРВ. Это позволяет сделать важный вывод о верхней границе не для конкретной предполагаемой ПРВ, а для самого метода, который оказывается практически в три раза эффективнее стандартной гистограммной оценки ПРВ.

Модифицированный гистограммный фильтр на основе k -окна

Можно уменьшить вычислительные затраты на расчет коэффициентов фильтра (5), применяя преобразование (4) сразу на все m интервалов. Тогда в соответствии с критерием максимизации энтропии получим $\rho = \mu_j p_j = \text{const}$, $j = \overline{1, m}$, где $\sum_{j=1}^m \mu_j = 1$. Эти уравнения позволяют алгебраическим способом определить сразу все m весовых коэффициентов ГФ через среднегармоническое значение $\rho = H / m$, $H = H(\mathbf{p})$, $\mathbf{p} = (p_1, \dots, p_m)$, $\mu_j = \rho / p_j$. Введем обозначение для такого фильтра – ГФ(m). Число ρ в этом случае является характеристическим числом всего закона распределения. Далее, поскольку $\rho = \sum_{j=1}^m \mu_j p_j / m$ и можно сформировать характеристическую функцию $\rho_{(m)}^* = \sum_{j=1}^m \mu_j (v_j / n) / m$, уравнением ГФ(m) будет следующее:

$$u_{(m)j} = n\rho_{(m)}^* / \mu_j = m^{-1} \sum_{t=1}^m (\mu_t / \mu_j) v_j. \quad (7)$$

Критерий χ^2 -квадрат для такого ГФ(m) примет вид

$$\chi_u^2 = \sum_{j=1}^m (n\rho_{(m)}^* / \mu_j - np_j)^2 / np_j = n [(\rho_{(m)}^* - \rho) / \rho]^2 = (nH^2)^{-1} \left[\sum_{j=1}^m \mu_j (v_j - np_j) \right]^2.$$

Применяя неравенство Коши–Буняковского, получим верхнюю границу оценки математического ожидания критерия χ^2 -квадрат

$$\mathbf{M}\chi_u^2 \leq (m^2 H)^{-1} \left[\sum_{j=1}^m \mu_j (v_j - np_j)^2 / np_j \right] \leq \mathbf{M}\chi_v^2 / (m^2 H) < (mH)^{-1}. \quad (8)$$

Анализ (8) позволяет сделать выводы об эффективности $\Gamma\Phi(m)$. Целесообразность применения $\Gamma\Phi(m)$ определяется неравенством $m^2 H > 1$. Поскольку для среднегармонического значения по множеству $\{p\}$ выполняется неравенство $p_{\min} = \min_j p_j \leq H \leq 1/m$, эффективность $\Gamma\Phi(m)$ достигается при $m^2 p_{\min} > 1$. Анализ этого неравенства может быть использован для оценки нижней границы числа интервалов группирования данных в зависимости от вида предполагаемой ПРВ.

В уравнении (7) использовалась характеристическая функция $\rho_{(m)}^*$, являющаяся слабо чувствительной к колебаниям отдельных частот ν_j/n , поэтому целесообразно ее заменить модифицированной укороченной характеристической функцией шириной не более k интервалов (k -окна). Для упрощения анализа этого варианта построения $\Gamma\Phi$ предположим, что множество m можно разбить некоторым алгоритмом на N_k непересекающихся множеств $\{T_i\}$, $i = \overline{1, N_k}$, приблизительно одинаковой мощности k (отличающихся по мощности не более чем на единицу). Тогда характеристическая функция для множества $\{T_i\}$ будет иметь вид $\rho_{j \in T_i}^* = k^{-1} \sum_{t \in \{T_i\}} \mu_t(\nu_t/n)$. Алгоритм работы $\Gamma\Phi(k)$ в этом случае запишется следующим выражением:

$$u_{j \in \{T_i\}} = n \rho_{j \in \{T_i\}}^* / \mu_j = k^{-1} \sum_{t \in \{T_i\}} \Phi_{j \in \{T_i\}, t} \nu_t, \quad (9)$$

где $\Phi_{j \in \{T_i\}, t} = k^{-1}(\mu_t / \mu_j)$.

Критерий χ^2 -квадрат для $\Gamma\Phi(k)$ определяется формулой $\chi_u^2 = n \sum_{j=1}^m [(\rho_{j \in T_i}^* - \rho_j) / \rho_j]^2 p_j$, $\rho_j = \mu_j p_j$. Выполняя преобразования с использованием неравенства Коши–Буняковского, получаем верхнюю границу оценки математического ожидания критерия χ^2 -квадрат

$$\mathbf{M} \chi_u^2 \leq \mathbf{M} \left(N_k \sum_{j \in \{T_i\}} k^{-1} \sum_{t \in \{T_i\}} \Phi_{j \in \{T_i\}, t} (\nu_t - n p_t)^2 / n p_t \right) < (k-1) \left(k^2 \min_i \{H_i\} \right)^{-1}, \quad (10)$$

где $H_i = H(\{p_{T_i}\})$, $\min_i \{H_i\} = p_{\min}$ – минимальное среднегармоническое значение по множеству $\{H_i\}$.

Для эффективной работы $\Gamma\Phi(k)$, как следует из (10), должно быть выполнено условие $k^2 p_{\min} (m-1) - k + 1 > 0$ с необходимым условием положительности дискриминанта квадратного уравнения $1 - 4 p_{\min} (m-1) > 0$. Из чего получаем неравенство, определяющее оптимальный выбор ширины окна $(k-1)/k^2 \leq p_{\min} (m-1) \leq 1/4$. Если мощность множества $\{T_i\}$ равна m , то неравенство (10) совпадает с (8). Из тех же соображений, по которым область определения функции принадлежности данных ограничена тремя интервалами группирования, и в целях обеспечения баланса между чувствительностью характеристической функцией $\rho_{j \in T_i}^*$ и минимизацией вычислительных затрат ширину окна целесообразно выбрать равной трем.

Отметим зависимость верхнего предела (8), (10) от вида идентифицируемой ПРВ, поскольку сам упрощенный алгоритм (7) и его модификация (9) используют информацию о всей предполагаемой ПРВ, в то время как алгоритмы (3) и (5) используют локальную информацию в пределах $\Delta_\mu \leq 3\Delta_x$.

Моделирование гистограммного фильтра

На рис. 1, 2 приведены результаты моделирования работы $\Gamma\Phi$ в виде вычисления средних для статистик $\chi_u^2 = \sum_{j=1}^m (u_j - n p_j)^2 / n p_j$ в сравнении со средними значениями $\chi_v^2 = \sum_{j=1}^m (\nu_j - n p_j)^2 / n p_j$. Результаты моделирования представлены в виде усредненных по количеству экспериментов (2000 статистических испытаний) критериев χ^2 -квадрат с изменяющимся объемом данных ($n = \overline{50, (50), 1000}$) и изменяющимся количеством интервалов их группирования ($m = \overline{5, 17}$) при простой гипотезе. Исследовались $\Gamma\Phi$ на примерах трех распределений с дисперсиями, равными единице: гауссовская, лапласовская и экспоненциальная ПРВ.

На рис. 1 представлены качественные характеристики работы $\Gamma\Phi$, заданного базовыми уравнениями (3), (5). Линия 1 (рис. 1) соответствует усредненной статистике χ_v^2 -квадрат, линия 2 характеризует верхнюю границу оценки математического ожидания критерия χ_u^2 -квадрат (5), маркер 3 (квадрат) – результат работы $\Gamma\Phi$ по гауссовской ПРВ с нулевым математическим ожиданием, маркер 4 (треугольник) – результат работы $\Gamma\Phi$ по лапласовской ПРВ с нулевым математическим ожиданием, маркер 5 (кружок) – результат работы $\Gamma\Phi$ по экспоненциальной ПРВ с единичным математическим ожиданием.

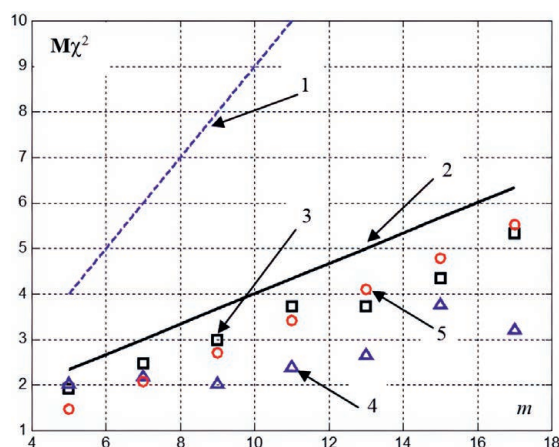


Рис. 1. Результаты работы гистограммных фильтров, заданных уравнениями (3), (5)
Fig. 1. The results of the histogram filters given by equations (3), (5)

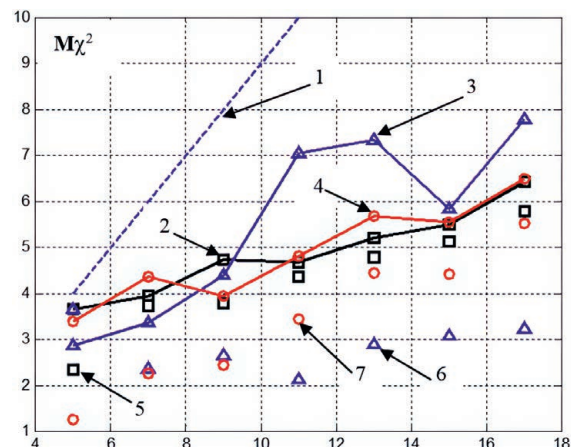


Рис. 2. Результаты работы гистограммных фильтров, заданных уравнениями (9), (10) ($k = 3$)
Fig. 2. The results of the histogram filters given by equations (9), (10) ($k = 3$)

На рис. 2 представлены качественные характеристики работы ГФ, заданного уравнениями (9), (10). В качестве примера был выбран ГФ ($k = 3$), т. е. максимальная ширина окна не превышала трех интервалов. Линия 1 (рис. 2) соответствует усредненной статистике критерия χ^2 -квадрат, кривые 2–4 с маркерами характеризуют верхнюю границу оценки математического ожидания критерия χ^2 -квадрат (10) для разных распределений. Кривая 2 с маркером-квадратом – для гауссовской ПРВ, кривая 3 с маркером-треугольником – для лапласовской ПРВ, кривая 4 с маркером-кружком – для экспоненциальной ПРВ. Маркеры 5–7, заданные теми же символами, соответствуют результату работы ГФ ($k = 3$) для указанных распределений. Во всех приведенных на рис. 1, 2 случаях выполнялось соотношение $M\chi_u^2 < M\chi_v^2$.

Заключение

1. Уравнения гистограммных фильтров (3), (7), (9) структурно одинаковы, различие заключается в ширине окна фильтра и принципе его формирования.

2. Результаты моделирования наглядно подтверждают эффективность и целесообразность применения гистограммных фильтров вместо стандартной гистограммной оценки.

3. Оптимальный гистограммный фильтр (формулы (3), (5)) может быть эффективно применен в задачах идентификации (распознавания) плотностей распределения вероятностей для произвольного объема данных. Эффект применения гистограммного фильтра, согласно критерию χ^2 -квадрат, практически в три раза превосходит обычную гистограммную оценку (рис. 1). Поскольку такой эффект достигается при сопоставимых объемах данных, можно утверждать, что оптимальный гистограммный фильтр для достижения таких же качественных показателей, что и при стандартной гистограммной оценке, будет использовать в три раза меньший объем данных. Это, в свою очередь, уменьшает время идентификации плотностей распределения вероятностей.

4. Модификации оптимального гистограммного фильтра позволяют снизить вычислительные затраты на его реализацию при сохранении эффективности (рис. 2).

5. Гистограммный фильтр является простым для реализации инструментом, который легко может быть встроен в любой открытый алгоритм идентификации (распознавания) плотности распределения вероятности.

6. Окончательное принятие решения по распознаванию плотностей распределения вероятностей может осуществляться методами простого или взвешенного голосования. Гистограммный фильтр может быть использован в задачах обработки изображений, локализации мобильных робототехнических устройств и т. д.

Список литературы

1. Вожов, С. С. Параметрическая и непараметрическая идентификации закона распределения по интервальным данным / С. С. Вожов, Е. В. Чимитова // Метрология. 2019. № 1. С. 3–16.
2. Wang, Z. Nonparametric Density Estimation for High-Dimensional Data – Algorithms and Applications / Z. Wang, D. W. Scott // Wiley Interdisciplinary Reviews Computational Statistics. 2019. Vol. 11, No 4. P. 1461–1501.

3. Куликов, В. Б. Анализ методов идентификации законов распределения случайных величин и процессов / В. Б. Куликов // *Cloud of Science*. 2020. Т. 6, № 4. С. 565–589.
4. Орлов, Ю. Н. Оптимальное разбиение гистограммы для оценивания выборочной плотности функции распределения нестационарного временного ряда / Ю. Н. Орлов // *Препринты ИПМ имени М. В. Келдыша*. 2013. № 14. 26 с. <http://library.keldysh.ru/preprint.asp?id=2013-14>.
5. Gonzalez, R. *Digital Image Processing* / R. Gonzalez. New York: Pearson, 2018.
6. Бройнль, Т. Встраиваемые роботехнические системы: проектирование и применение мобильных роботов со встроенными системами управления / Т. Бройнль; под ред. В. Е. Павловского. Ижевск: Ин-т компьютер. исслед., 2012. 520 с.
7. Овсянников, А. В. Гистограммный фильтр с настройкой параметра сглаживания / А. В. Овсянников, В. М. Козел // *Доклады БГУИР*. 2022. Т. 20, № 8. С. 42–50. <http://dx.doi.org/10.35596/1729-7648-2022-20-8-42-50>.
8. Овсянников, А. В. Фильтрация гистограммной оценки плотности вероятности на основе нечеткой принадлежности данных интервалу группирования / А. В. Овсянников, В. М. Козел // *Доклады БГУИР*. 2021. Т. 19, № 4. С. 13–20. <http://dx.doi.org/10.35596/1729-7648-2021-19-4-13-20>.

References

1. Vozhov S. S., Chimitova E. V. (2019) Parametric and Nonparametric Identification of the Distribution Law by Interval Data. *Metrology*. (1), 3–16 (in Russian).
2. Wang Z., Scott D. W. (2019) Nonparametric Density Estimation for High-Dimensional Data – Algorithms and Applications. *Wiley Interdisciplinary Reviews Computational Statistics*. 11 (4), 1461–1501.
3. Kulikov V. B. (2020) Analysis of Methods for Identifying the Laws of Distribution of Random Variables and Processes. *Cloud of Science*. 6 (4), 565–589 (in Russian).
4. Orlov Yu. N. (2013) Optimal Histogram Partitioning for Estimating the Sample Density of the Distribution Function of an Unsteady Time Series. *Preprints of the Institute of Applied Mathematics name after M. V. Keldysh RAS*. (14), 26. <http://library.keldysh.ru/preprint.asp?id=2013-14> (in Russian).
5. Gonzalez R. (2018) *Digital Image Processing*. New York, Pearson Publ.
6. Brojnl' T. (2012) *Embedded Robotic Systems: Design and Application of Mobile Robots with Embedded Control Systems*. Izhevsk, Institute of Computer Science. 520 (in Russian).
7. Aysiannikau A. V., Kozel V. M. (2022) Histogram Filter with Smoothing Parameter Setting. *Doklady BGUIR*. 20 (8), 42–50. <http://dx.doi.org/10.35596/1729-7648-2022-20-8-42-50> (in Russian).
8. Aysiannikau A. V., Kozel V. M. (2021) Filtration of Histogram Evaluation of Probability Density Based on Fuzzy Data Accessibility to a Grouping Interval. *Doklady BGUIR*. 19 (4), 13–20. <http://dx.doi.org/10.35596/1729-7648-2021-19-4-13-20> (in Russian).

Вклад авторов

Овсянников А. В. разработал общую методику построения гистограммного фильтра с настройкой параметра сглаживания. Козел В. М. выполнил моделирование работы гистограммного фильтра.

Author's contribution

Aysiannikau A. V. developed a general technique for constructing a histogram filter with adjustment of the smoothing parameter. Kozel V. M. performed modeling of the histogram filter.

Сведения об авторах

Овсянников А. В., к. т. н., доцент, доцент кафедры информационных технологий Белорусского государственного университета

Козел В. М., к. т. н., доцент, доцент кафедры информационных радиотехнологий Белорусского государственного университета информатики и радиоэлектроники

Адрес для корреспонденции

220030, Республика Беларусь,
г. Минск, просп. Независимости, 4
Белорусский государственный университет
Тел.: +375 17 209-58-94
E-mail: andovs@mail.ru
Овсянников Андрей Витальевич

Information about the authors

Aysiannikau A. V., Cand. of Sci., Associate Professor, Associate Professor at the Department of Information Technologies of the Belarusian State University

Kozel V. M., Cand. of Sci., Associate Professor, Associate Professor at the Department of Information Radiotechnologies of the Belarusian State University of Informatics and Radioelectronics

Address for correspondence

220030, Republic of Belarus,
Minsk, Nezavisimosti Ave., 4
Belarusian State University
Tel.: +375 17 209-58-94
E-mail: andovs@mail.ru
Aysiannikau Andrei Vital'evich