

Глухова Н.В.

Національний технічний університет «Дніпровська політехніка»

ВИЯВЛЕННЯ ІНФОРМАТИВНИХ ОЗНАК ЗОБРАЖЕНЬ НА БАЗІ АНАЛІЗУ ГІСТОГРАМ ЯСКРАВОСТІ

У статті розглянуто проблеми виявлення інформативних ознак зображень. Проаналізовано актуальність та переваги методів вимірювань та контролю, які ґрунтуються на реєстрації, обробці та аналізі зображень. Описано напрями застосування засобів вимірювань та контролю на базі реєстрації рентгенівських знімків.

Розглянуто стандартну методику побудови гістограми яскравості півтонового зображення для 256 стовпчиків, які відповідають окремим рівням яскравості. Наведено результати побудови гістограм для знімків газорозрядного випромінювання різних типів води на рентгенівській плівці. Проаналізовано способи вибору кількості інтервалів гістограми з точки зору метрологічної практики. Усі відомі на тепер способи розрахунку кількості інтервалів гістограм носять емпіричний характер.

Запропоновано методику скорочення кількості стовпчиків гістограми, яка дає змогу виділити інформативні ознаки газорозрядного світіння, придатні для подальшого експертного оцінювання та побудови систем програмної класифікації об'єктів за характеристиками зображень їх газорозрядного випромінювання.

Спосіб виділення інформативних ознак на базі побудови гістограм яскравості зображень для 12 інтервалів знайшов практичне впровадження у галузі біології та медицини. Гістограма яскравості зображення зі скороченою кількістю інтервалів використана у разі експертного аналізу біологічних властивостей води. Інформативні ознаки зображення у вигляді 12 ординат стовпчиків гістограми яскравості використано для створення програмної класифікації води різних типів на основі розрахунку евклідової відстані. Інформативні ознаки, вилучені з гістограми зі скороченою кількістю інтервалів, також застосовано у галузі медичної діагностики під час побудови системи програмної класифікації на базі нечіткого кластерного аналізу параметрів випромінювання зразків крові.

Ключові слова: *аналіз даних, цифрова обробка зображень, гістограма, газорозрядне випромінювання.*

Постановка проблеми. Більшість сучасних наукових досліджень у різних галузях як первинні дані використовує результати вимірювань. При цьому спектр можливих методів вимірювань постійно розширюється, у тому числі завдяки швидкій еволюції технічних засобів для збору, передачі, обробки та зберігання даних.

Забезпечення автоматизованого збору та обробки великих обсягів інформації дало змогу останніми роками суттєво розширити сфери застосування методів вимірювань, які ґрунтуються на реєстрації та подальшому аналізі зображень. Незважаючи на те, що зображення містять значну кількість інформації про досліджувані об'єкти або процес, їх використання у метрологічній практиці потребує синтезу відповідних інформаційних технологій для їх обробки, аналізу та інтерпретації.

Аналіз останніх досліджень і публікацій. Різноманіття об'єктів, дослідження яких ґрунтується на отриманні візуальних даних, не дозволяє нині запропонувати універсальні способи обробки

зображень та вилучення їх інформативних ознак. Тому під час розв'язання багатьох прикладних завдань використовуються візуальні оцінки характеристик зображень. Зрозуміло, що такі оцінки мають суб'єктивний характер і обмежені можливості щодо отримання кількісних значень інформативних ознак. Часто алгоритми аналізу зображень будують як автоматизований програмний аналог експертних оцінок.

Особливу роль серед усього спектра методів вимірювань та контролю, які використовують як первинні дані зображення, відіграють методи, що базуються на реєстрації рентгенівських знімків. Незважаючи на те, що вибираючи об'єктом досліджень у цій роботі тільки рентгенівські знімки, значно звужується коло об'єктів аналізу, все ж таки воно залишається дуже широким та охоплює багато практичних завдань.

Наприклад, процедура отримання рентгенівських зображень застосовується у складі інспекційно-доглядових комплексів у митних органах. При цьому отримані зображення розпізнаються

співробітниками, які проходять відповідну спеціальну підготовку [1].

Рентгенівські зображення використовуються також для дослідження поверхонь та внутрішньої структури різноманітних речовин. Єдиного підходу для аналізу рентгенівських зображень у цій галузі також не запропоновано. Дослідники розробляють різні способи залежно від поставленого практичного завдання. Наприклад, у роботі [2] запропонований підхід до аналізу рентгенівських зображень у разі деформації кристалів за розподілами енергетичного спектра Фур'є.

Для виявлення певних деталей або фрагментів зображень натеper розроблено низку способів зафарбовування, підкреслення меж, а також алгоритми для нелінійного підсилення яскравості [3], але вони також можуть застосовуватися для обмеженого кола завдань.

Одним з найпоширеніших методів аналізу зображень (не тільки на рентгенівській плівці) є побудова гістограми зображення [4]. Гістограма яскравості зображення являє собою дискретну функцію, яка для рівня яскравості надає кількість пікселів, що характеризуються відповідною яскравістю.

У цій роботі предметом досліджень виступають зображення, зафіксовані на рентгенівській плівці, тому є півтоновими, тобто містять відтінки сірого кольору. Півтонові зображення можна охарактеризувати на основі побудови однієї гістограми, яка охоплює всі пікселі зображення.

Натеper у більшості прикладних програм для обробки й аналізу зображень цифрові растрові півтонові зображення представляються 256 різними відтінками яскравості, що стандартно кодуються цілими числами від 0 (чорний колір) до 255 (білий). Усі проміжні значення між граничними 0 та 255 являють собою відтінки сірого кольору. Натомість кольорове зображення неможливо представити на основі використання однієї гістограми. Для аналізу кольорових зображень зазвичай будують три гістограми, що відповідають палітрі RGB (тобто три гістограми розподілу яскравості для червоного, зеленого та блакитного кольорів) [5–6]. Окрім гістограм яскравості, також застосовуються гістограми контрасту [7], які являють собою дискретну функцію, що вказує на кількість пікселів з певним рівнем контрасту.

Перевагами застосування гістограм для обробки й аналізу цифрових зображень є те, що вони забезпечують інтегральну оцінку параметрів зображення, враховуючи яскравість (або контраст) кожного пікселя.

Гістограми є універсальним інструментом для аналізу, який може бути застосований фактично

для будь-якого цифрового зображення. Гістограми застосовуються на різних етапах опрацювання візуальних даних: як на етапах попередньої обробки [7], так і на етапі виділення інформативних ознак та інтерпретації зображень [8]. Однак, оскільки гістограма охоплює всі пікселі зображення, вона містить великий обсяг параметрів, а саме 256 значень кількості пікселів з певною яскравістю. Це призводить до ускладнення можливості аналізу людиною-експертом гістограм та виділення інформативних ознак візуальних даних на етапі інтерпретації зображення.

Наприклад, під час аналізу зображень газорозрядного випромінювання зразків води, зафіксованих на рентгенівській плівці, враховуються експертні оцінки як самого зображення, так і отриманої гістограми. Оскільки 256 кількісних параметрів гістограми (ординати висот її стовпчиків) не можуть усі аналізуватися експертом, то необхідно скоротити обсяг інформації. У роботі [8] зменшення обсягу кількісної інформації виконано шляхом розбиття гістограми на окремі ділянки (зони). Недоліком такого способу є суб'єктивний та неточний характер розподілу на зони.

Другим прикладом, який демонструє необхідність зменшення обсягів кількісної інформації про інформативні ознаки зображень, може слугувати завдання класифікації (розбиття на окремі класи або групи зображень для об'єктів зі схожими властивостями). У такому разі кількісна інформація про 256 ординат пікселів відповідної яскравості, отримана під час побудови гістограми, може виявитися надлишковою і для побудови автоматизованої програмної системи класифікації зображень буде досить значно меншої кількості числових інформативних ознак [9–11].

Постановка завдання. З метою забезпечення зручності інтерпретації гістограм зображень людиною-експертом і у разі розв'язання завдань автоматизованої програмної класифікації зображень актуальним є завдання розробки обґрунтованої методики скорочення кількості числових інформативних ознак зображень, які виділяються на основі побудови їх гістограм.

Виклад основного матеріалу дослідження.

Під час розробки методики виділення числових інформативних ознак зображень було проаналізовано декілька вибірок зображень, отриманих на основі методу газорозрядного випромінювання [11]. Вибірki зображень формувались у експериментальному дослідженні властивостей різних типів води, зокрема дистильованої, природної, водопровідної, питної бутильованої тощо.

На рис. 1 представлено зображення (негатив) для зразка стічної води. У разі візуального аналізу людиною-експертом виявляються такі характерні для зображень зразків забрудненої води ознаки, як: наявність більш темного, ніж колір фону рентгенівської плівки, нерівномірного забарвлення внутрішнього кола зображення (внутрішнє коло відповідає зоні безпосереднього контакту зразка води з поверхнею рентгенівської плівки); нерівномірно розташовані навколо зовнішнього контуру випромінювання краплі рідини газоподібної розряди.

Характерною особливістю гістограм для зображень газорозрядного випромінювання забрудненої води є наявність чітко вираженого екстремуму в ділянці яскравостей, що умовно кодуються для цифрових зображень значеннями в діапазоні від

50 до 100. Наявність екстремуму огинаючої гістограми в цьому діапазоні відповідає забарвленню внутрішньої ділянки світіння краплі в зоні контакту з плівкою.

Для порівняння розглянемо зображення газорозрядного випромінювання на рентгенівській плівці для зразків дистильованої і природної вод (відповідно рис. 2 і 3) та відповідні гістограми яскравості. Випромінювання крапель дистильованої води характеризується світлим фоном внутрішнього кола в зоні контакту зразка з поверхнею плівки. Фактично відтінок внутрішнього кола збігається з фоном плівки поза межами зображення, що відображається наявністю декількох екстремумів у діапазоні відтінків яскравості 120–150.

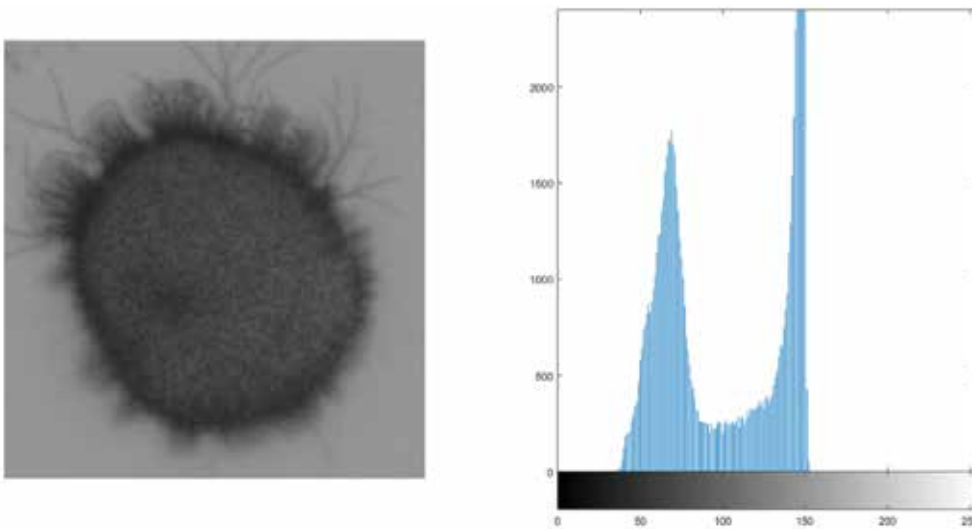


Рис. 1. Знімок (негатив) для стічної води та гістограма

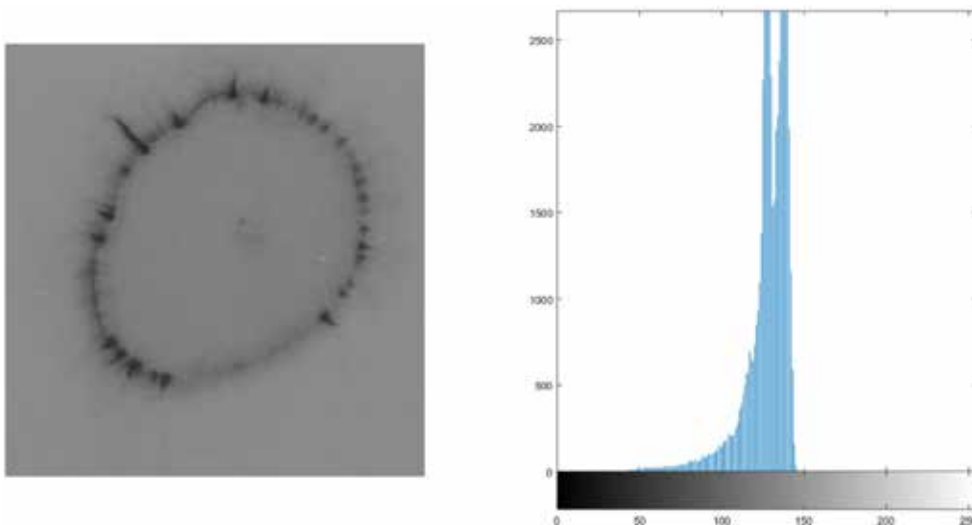


Рис. 2. Знімок (негатив) для дистильованої води та гістограма

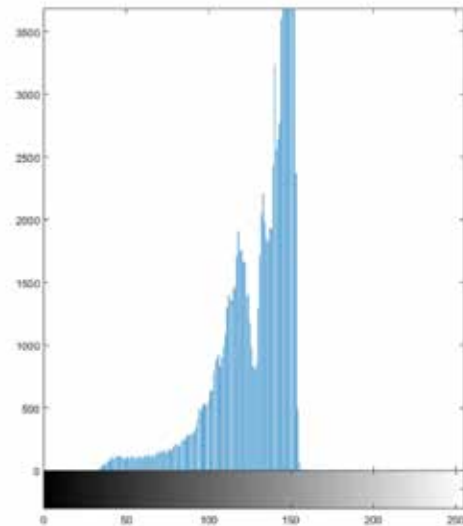
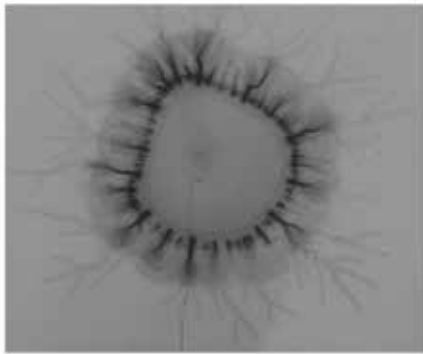


Рис. 3. Знімок (негатив) для природної води та гістограма

Внаслідок слабого та нерівномірного формування газорозрядних треків навколо зразка дистильованої води на гістограмі міститься дуже мало пікселів з яскравістю менше 100.

Для зразків питної, у тому числі природної води, максимум огинаючої гістограми яскравості також розташований у діапазоні 120–150. Порівняно зі зразком дистильованої випромінювання природної води характеризується рівномірною широкою газорозрядною «коронаю», що формується в результаті поширення окремих газових розрядів навколо зразка. Завдяки цьому на гістограмі виявляється екстремум огинаючої в діапазоні 100–120, що відповідає відтінку забарвлення «корони» випромінювання.

На рис. 1–3 (праворуч) представлено гістограми яскравості, побудовані стандартним способом, який використовується у разі отримання гістограм яскравості зображень у прикладних програмних пакетах, тобто гістограма містить 256 стовпчиків.

Однак, як було зазначено вище, під час розв'язання певних прикладних завдань аналізу зображень кількість інтервалів доцільно скоротити. Розглянемо рекомендації щодо побудови гістограм, які прийняті в прикладній статистиці та застосовуються в метрологічній практиці. Виявляється, що чітких рекомендацій та розрахункових формул натепер не розроблено і кількість так званих інтервалів групування у разі побудови гістограм може коливатися у значних межах [12]. Запропоновані в теорії ймовірностей та математичній статистиці методи вибору кількості інтервалів гістограми мають емпіричний характер [13].

Серед рекомендацій наводяться такі: кількість інтервалів не слід вибирати великою, оскільки графік стає невиразним і в ньому проявляються нерівномірні коливання висот; у разі занадто малої кількості інтервалів може спостерігатися розмивання особливостей розподілу й огинаюча гістограма відповідає отриманим експериментальним даним дуже грубо.

Найбільш поширеною в метрологічній практиці емпіричною залежністю для визначення кількості інтервалів розбиття є формула Стерджесса:

$$k = 3,32 \cdot \lg N + 1,$$

де N – кількість даних у вибірці.

Також часто використовують таке просте співвідношення:

$$k = \sqrt{N}.$$

Для виконання розрахунків кількості інтервалів гістограми застосовують формулу Брукса та Каррузера:

$$k = 5 \cdot \lg N.$$

Однак під час розв'язання завдання інтерпретації картин газорозрядного випромінювання на рентгенівській плівці (надалі відсканованих) стає очевидним, що для сучасних цифрових зображень навіть із середньою роздільною здатністю всі наведені вище формули дають кількість інтервалів гістограми 300 або більше. Але, як зазначалося вище, така кількість інтервалів не задовольняє вимогам щодо можливості експертного аналізу

отриманих гістограм зображень та раціонального розв'язання на подальших етапах обробки зображень завдань їх класифікації.

Для обґрунтованого скорочення кількості інтервалів гістограм було враховано рекомендації для вибірок великих розмірів, що надані в роботі [14]. Для вибірок обсягом від 500 до 1000 пропонується вибирати кількість інтервалів розбиття 10–16; для вибірок з кількістю значень 1000–10000 рекомендована кількість стовпчиків гістограми становить 12–22. При цьому відомо, що у разі вибору 10–11 інтервалів розбиття групувана вибірка зберігає 98% інформації.

Поряд з цим у результаті психологічних досліджень було встановлено, що у короткочасній пам'яті людини одночасно може знаходитися (та аналізуватися) від 5 до 9 градацій деякої властивості [15].

На підставі узагальнення рекомендацій щодо побудови гістограм для великих за обсягом вибірок даних та одночасно вимог до можливості експертного аналізу гістограм яскравості була розроблена методика опрацювання знімків газорозрядного випромінювання зразків рідини, заснована на побудові гістограм яскравості у разі розбиття на 12 інтервалів яскравості однакової ширини. Підкреслимо, що необхідність можливості залучення та врахування експертної думки

під час розв'язання завдань інтерпретації характеристик газорозрядного випромінювання зумовлена постановкою прикладних завдань у галузі медицини та біології.

На основі запропонованої методики розрахунку висот для 12 стовпчиків гістограми була побудована система класифікації типів води за властивостями її біологічної активності, що включала розрахунок евклідової відстані між висотами стовпчиків гістограми [16]. Розроблена методика побудови гістограми яскравості та вилучення інформативних ознак впроваджена у галузі медичної діагностики на базі використання особливостей газорозрядного світіння зразків крові [10] та в галузі біохімії для аналізу біологічних рідин [17; 18].

Висновки. У статті запропонована методика побудови гістограм яскравості зображень для 12 інтервалів розбиття. Скорочення кількості числових інформативних ознак зображення забезпечує можливість виконання експертного аналізу гістограми яскравості, а також дає можливість використання інформативних ознак гістограми для подальших етапів аналізу зображень, зокрема їх класифікації. Позитивний досвід практичного застосування методики під час розв'язання завдань у галузі біології та медицини підтвердив доцільність вибору 12 інтервалів групування для гістограми.

Список літератури:

1. Трофимчук А.М. Компьютерное моделирование рентгеновских изображений, полученных с помощью инспекционно-досмотровых комплексов. *Инженерный вестник Дона*. № 1. 2017. С. 5–21.
2. Баловсяк С.В. Анализ рентгеновских муаровых изображений деформированных кристаллов по радиальным распределениям энергетического спектра Фурье / С.В. Баловсяк, С.Н. Новиков, И.М. Фодчук, И.В. Яремчук. *Металлофізика, новітні технології*. 2019. Том 41. № 3. С. 389–402. URL: <https://doi.org/10.15407/mfint.41.03.0389>.
3. Гук А.П. Разработка методик автоматизированного дешифрирования аэрокосмических снимков. Дешифровочные признаки изображений объектов на многоспектральных космических снимках / Гук А.П., Евстратова Л.Г., Хлебникова Е.П., Алтынцев М.А., Арбузов С.А., Гордиенко А.С., Гук А.А., Симонов Д.П. *Геодезия и картография*. № 7. 2013. С. 31–40.
4. Obed Appiah. Fast Generation of Image's Histogram Using Approximation Technique for Image Processing Algorithms / Obed Appiah, J. Ben Hayfron-Acquah. *Image, Graphics and Signal Processing*. 2018. No. 3. Pp. 25–35. DOI: 10.5815/ijigsp.2018.03.04.
5. Глухова Н.В. Інформаційна технологія для аналізу кольорових зображень газорозрядного випромінювання / Н.В. Глухова, Л.А. Пісоцька. *Перспективні технології та прилади*. № 12. 2018. С. 48–52.
6. Глухова Н.В. Розробка методу аналізу кольорових зображень газорозрядного випромінювання / Н.В. Глухова, Л.А. Пісоцька. *Системи управління, навігації та зв'язку*. № 2(48). 2018. С. 59–62.
7. Хрящев Д.А. Об одном методе анализа цифрового изображения с применением гистограмм. *Вестник Астраханского гос. техн. ун-та*. 2010. № 1. С. 109–113.
8. Песоцкая Л.А. Анализ изображений кирлиановского свечения капель воды / Л.А. Песоцкая, Н.В. Глухова, В.Н. Лапицкий. *Науковий вісник Національного гірничого університету*. 2013. № 1. С. 91–96.
9. Глухова Н.В. Розробка методу експрес-оцінки біологічних властивостей води. *Східно-європейський журнал передових технологій*. 2014. № 6/5(72). С. 18–25.
10. Glukhova N.V. Cluster Analysis for Measurements of Gas-Discharge Emission Parameters for Blood Samples. *Meas Tech*. No. 58. Pp. 1066–1070. 2015. URL: <https://doi.org/10.1007/s11018-015-0843-7>.

11. Глухова Н.В. Метод оцінки біологічних та квантових властивостей води / Н.В. Глухова, Л.А. Пісоцька, Н.Г. Кучук. *Системи обробки інформації*. 2015. № 7(132). С. 195–200.

12. Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. Ленинград : Энергоатомиздат, 1991. 303 с.

13. Горошко А.В. Розрахунок допустимих значень параметрів об'єктів у випадку полімодальності їх ймовірнісних розподілів / А.В. Горошко, Ройзман В.П. *Вібрації в техніці та технологіях*. № 4(72). 2013. С. 19–27.

14. Бурдун Г.Д., Марков Б.Н. Основы метрологии. Москва : Изд-во стандартов, 1985. 120 с.

15. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова. Москва : Наука. Гл. ред. физ-мат. лит., 1986. 312 с.

16. Глухова Н.В. Розробка системи експрес-класифікації води на основі бази даних зображень газорозрядного випромінювання / Н.В. Глухова, Л.А. Пісоцька, Н.Г. Кучук. *Збірник наукових праць Харківського університету повітряних сил ім. І. Кожедуба*. 2015. Вип. 3(44). С. 112–118.

17. Povstyanyi V.A. Peculiarities of gas-discharge luminescence of biological fluid from the achilles tendon in the late postmortem period / V.A. Povstyanyi, L.A. Pesotskaya, N.V. Glukhova, N.M. Yevdokimenko, L.R. Nikogosyan, Ye.L. Koshelnik. *Journal of Education, Health and Sport. Licensee Open Journal Systems of Kazimierz Wielki University in Bydgoszcz, Poland*. 2017. No. 7(2), Pp. 498–508. DOI: <http://dx.doi.org/10.5281/zenodo.399324>.

18. Повстяный В.А. Особенности газоразрядного свечения жидкости из ахиллового сухожилия и изменений отдельных ее компонентов в позднем постмортальном периоде / В.А. Повстяный, Л.А. Песоцкая, Н.В. Глухова, Н.М. Евдокименко. *Український журнал медицини, біології та спорту*. 2017. № 2(4). С. 102–107.

Glukhova N.V. DETECTION OF INFORMATIVE SIGNS OF IMAGES ON THE BASIS OF ANALYSIS HISTOGRAMS OF BRIGHTNESS

The article considers the problems of detecting informative features of images. The relevance and advantages of measurement and control methods based on image registration, processing and analysis are analyzed. The directions of application of means of measurements and control on the basis of X-rays are described.

The standard technique of construction of the histogram of brightness of the halftone image for 256 columns which correspond to separate levels of brightness is considered. The results of construction of histograms for images of gas-discharge radiation of different types of water on X-ray film are given. Different ways of choosing the number of histogram intervals from the point of view of metrological practice are analyzed. All currently known methods of calculating the number of histogram intervals are empirical.

A method of reducing the number of columns of the histogram is proposed, which allows to identify informative features of gas-discharge glow, suitable for further expert evaluation and construction of software classification systems of objects according to the characteristics of images their gas-discharge radiation.

The method selection of informative features on the basis construction of histograms brightness of images for 12 intervals has found practical application in the field of biology and medicine. The histogram brightness of image with a reduced number of intervals was used in the expert analysis of biological properties of water. Informative features of the image in the form of 12 ordinates columns of the histogram brightness were used to create a software classification of water of different types based on the calculation of the Euclidean distance. Informative features extracted from the histogram with a reduced number of intervals were also used in the field of medical diagnostics in the construction of a software classification system based on fuzzy cluster analysis of radiation parameters of blood samples.

Key words: data analysis, digital image processing, histogram, gas discharge radiation.