

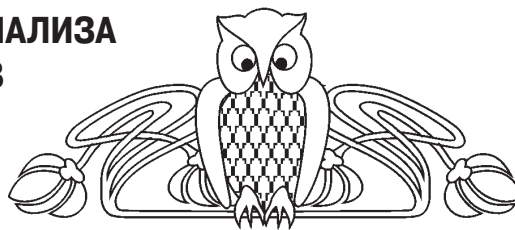


УДК 343.98.065

ПРИМЕНЕНИЕ МЕТОДОВ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ОЦЕНКИ СХОЖЕСТИ СЛЕДОВ БОЙКОВ

М. В. Корнилов

программист лаборатории микроанализа и моделирования результатов применения оружия, Образовательно-научный институт наноструктур и биосистем Саратовского государственного университета
E-mail: kornilovmv@gmail.ru



В. А. Федоренко

кандидат физико-математических наук, доцент, заведующий лабораторией микроанализа и моделирования результатов применения оружия, Образовательно-научный институт наноструктур и биосистем Саратовского государственного университета
E-mail: fed77@yandex.ru

Введение. Одной из важнейших задач судебно-баллистической экспертизы является идентификация оружия по следам бойка. Данная задача обычно связана с поиском совпадающих (парных) следов по большим базам данных. Для этого широко используются баллистические идентификационные системы, такие как ПОИСК, ТАИС, Арсенал и другие, в которых используются автоматические алгоритмы поиска парных следов, основанные на подсчете функции взаимной кросс-корреляции. В данной работе предложен иной подход, разработанный для изображений следов бойков, содержащих признаки в виде крупных областей неопределенной формы. К ним применялись методы кластерного анализа для формирования приоритетного списка. **Теоретическая часть.** В теоретической части дается алгоритм выделения и сравнения дескрипторов – характеристик, описывающих изображения следа бойка. После чего предложен способ формирования приоритетного списка, анализируя который, эксперт сможет сделать вывод о парности следов. **Экспериментальная часть.** Разработанный алгоритм был применен к базе данных, состоящей из более чем 100 изображений следов бойков 24 экземпляров оружия. **Заключение.** Разработанный алгоритм позволяет формировать приоритетный список из 20 изображений, в который входят парные следы (если такие имеются в электронной базе данных) с вероятностью 100%. Таким образом, предложенная методика позволяет существенно сократить время поиска следов. Данный алгоритм осуществляет предварительную фильтрацию и позволяет выделить список следов, к которым имеет смысл применять более сложные критерии, такие как сравнение контуров, особых точек изображений и т.п.

Ключевые слова: баллистические системы, сравнение следов бойка, идентификация оружия, кластерный анализ.

Введение

Одной из важных задач эксперта является идентификация оружия по следам бойка, которая связана с проверкой по большим массивам гильзотек. Для этого широко используются баллистические идентификационные системы, такие как ПОИСК, ТАИС, Арсенал и др. В данных системах применяются автоматизированные алгоритмы поиска совпадающих в криминалистическом

смысле следов (парных), которые формируют приоритетный список по степени их схожести. Однако вывод о принадлежности следов бойка одному и тому же экземпляру оружия всегда остается за экспертом. Зачастую алгоритмы поиска парных следов в баллистических системах основаны на подсчете функции взаимной корреляции. Такой подход требует больших временных затрат в случае поиска по большой базе данных. Кроме этого, оценка степени совпадения изображений следа бойка, содержащих признаки в виде крупных пятен неопределенной формы, методом корреляционного анализа неэффективна. Это обусловлено тем, что такие пятна занимают большую часть изображения следа бойка.

В данной статье предлагается новый подход к решению таких задач. В работе [1] было предложено разделять признаки следа бойка по основным морфологическим типам: крупные, мелкие (сравнимые с шумом поверхности), в виде контуров неопределенной формы, наслоенный или набора трасс. Таким образом, для каждого типа признаков целесообразно использовать адаптированные алгоритмы, которые позволяют их эффективно выделять и сравнивать. В данной работе исследуются возможности кластерного анализа для идентификации оружия по индивидуальным признакам в виде крупных пятен неопределенной формы.

Индивидуальные признаки огнестрельного оружия, отобразившиеся в следах бойка, можно описать с помощью различных дескрипторов – характеристик, определяемых однозначным образом. В данной работе использовались следующие дескрипторы, не зависящие от ориентации изображения: площадь, периметр, моменты инерции. Эти дескрипторы формируют вектор состояний признака. Кластерный анализ позволяет на основе сравнения таких векторов делать



выводы о схожести следов. При этом существенно сокращается количество вычислений при поиске парных следов боя, так как требуется оценивать вектор, содержащий 4 значения, а не рассчитывать функцию кросс-корреляции для полутоновых изображений.

Теоретическая часть

В работе анализировались изображения, полученные с помощью баллистической системы «POISC», которые представляют собой черно-белые фотографии следа боя. Белым цветом выделены области с максимальным отражением света. В качестве объекта исследования рассматривались изображения, содержащие крупные признаки (рис. 1), которые представляют собой пятна неопределенной формы. Для

этих признаков выделялись такие параметры, как площадь, периметр, максимальные и минимальные моменты инерции. Вычисление дескрипторов по полутоновым изображениям не представляется возможным. Поэтому на первом этапе проводилось преобразование изображения в бинарное (в данной работе белым цветом обозначаются индивидуальные признаки следа боя, а черным – фон). Для этого использовался алгоритм адаптивной бинаризации Ниблэка [2, р. 115–116], который позволяет свести к минимуму влияние неоднородности освещения. Далее вычислялись общая площадь (количество белых точек изображения), периметр (длина границы признака в бинарном изображении), максимальный и минимальный моменты инерции индивидуальных признаков.

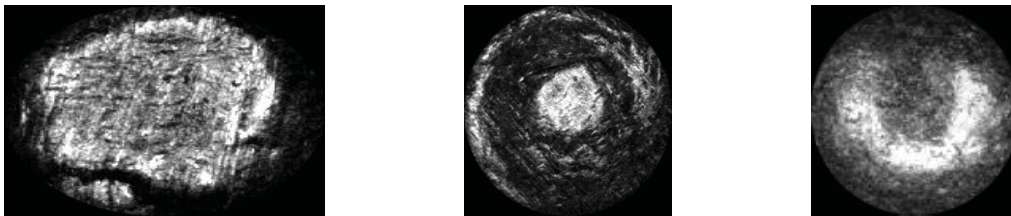


Рис. 1. Изображения следов различных бойков

Выбранные дескрипторы формируют так называемое пространство признаков. В этом случае каждый объект характеризуется точкой в пространстве признаков или вектором признаков. Чтобы сравнивать векторы признаков, рассчитывалось Евклидово расстояние от исследуемого объекта до каждого тестового:

$$d = \sqrt{\sum_{i=1}^N (A_i - B_i)^2}, \tag{1}$$

где A и B – тестовый и исследуемый объекты, i – i -я компонента вектора признаков (один из признаков: площадь, периметр, и т.д.).

Выбранные дескрипторы имеют разную размерность (м, м², б/р), а также большой разброс значений (значение периметра может отличаться на порядок от значения площади и моментов инерции). Поэтому, чтобы уравнивать их вклад в подсчет расстояния (1), разброс значений дескрипторов необходимо нормировать – преобразовать их значения в интервал (0, 1). Для этого были определены максимальные величины, до которых могут изменяться значения дескрипторов. После чего они были линейно нормированы на эти величины.

Для лучшей иллюстративности (рис. 2) проводилась перенормировка расстояний по формуле

$$\varphi = \ln(1 + d). \tag{2}$$

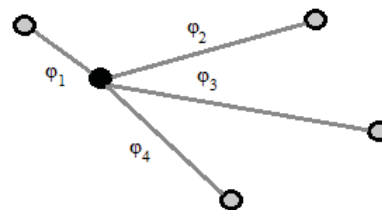


Рис. 2. Оценка расстояний до тестовых объектов в исследуемой точке векторного пространства, φ_i – перенормированные расстояния до тестовых объектов

В соответствии с этой формулой чем ближе значение величины φ к единице, тем ближе дескрипторы тестового объекта к дескрипторам исследуемого следа. В соответствии с этой закономерностью в дальнейшем и строился приоритетный список.

Приоритетный список формировался по возрастанию значения φ : чем больше φ , тем дальше от начала приоритетного списка располагается след. В работе изучалось попадание следов, парных к исследуемому, в список, состоящий из 5, 10 и 20 элементов. Результат считался успешным, если в приоритетный список попадал хотя бы один парный след.

Экспериментальная часть

В эксперименте использовались следы бойков 24 экземпляров оружия. Тестовый массив



составлял 110 цифровых изображений следов бойков, на основании которых была сформирована электронная база данных. В качестве исследуемых следов бойков использовались следы, парные к различным группам (кластерам) тестового массива.

Все следы тестового массива были разбиты на 24 кластера, в каждый из которых входили следы бойка одного экземпляра оружия. Предварительно была оценена эффективность отнесения исследуемого следа к одной из групп, близость к которой определялась по среднему значению его потенциальной функции:

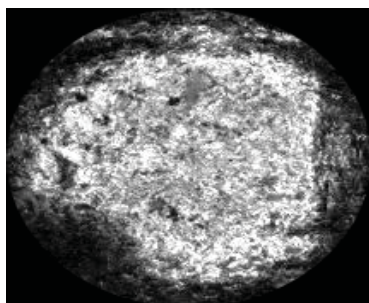
$$\varphi_j = \sum_{i=1}^N \frac{1}{1 + d_{ij}}, \quad (3)$$

где j – номер кластера, N – количество элементов кластера, d_i – расстояние от тестового объекта до i -го элемента кластера. Однако расчеты показали низкую эффективность данного подхода. Обусловлено это достаточно большой вариативностью признаков, вследствие чего многие кластеры пересекаются друг с другом.

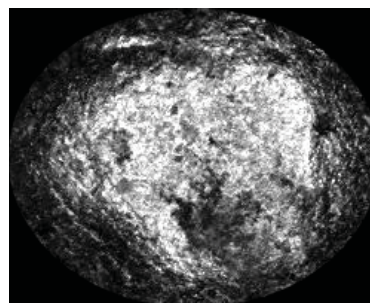
Более эффективным оказался поиск отдельных «парных» объектов по минимальному значению потенциальной функции следов те-

стового массива. Проведенные исследования показали, что «парные» к исследуемому объекту следы попадают в первую пятерку с вероятностью 90%, в десятку – 95%, в двадцатку – 100%. Следует отметить, что статистику сильно ухудшают следы, на изображении которых присутствуют артефакты в виде элементов коррозии, частичек лака и т.д. (рис. 3).

Таким образом, в приоритетный список, включающий 20% объектов из тестового массива, обязательно будет входить «парный» след, если такой будет присутствовать в базе данных. Получившийся список еще может быть достаточно большим для быстрого выбора экспертом «парного» следа. Например, для массива, включающего 2000 однотипных следов, 20% составят 400 объектов. Поэтому для дальнейшего сокращения круга поиска «парных» следов (например, до численности, равной 1–2% от общего числа объектов тестового массива) необходим дополнительный анализ следов из приоритетного списка по дескрипторам, описывающим форму объекта. К таким дескрипторам можно отнести контуры, построенные по внешней границе выделенных признаков. Применение контурного анализа впоследствии позволит более корректно сформировать «короткий» приоритетный список.



а



б

Рис. 3. Цифровое изображение следа бойка: а – хорошего качества; б – имеющее артефакты в виде темных пятен

Заключение

Предложенный алгоритм позволяет выделить следы бойков, близкие по совокупности параметров к исследуемому изображению. Проведенные исследования показали перспективность формирования приоритетного списка по дескрипторам, не зависящим от ориентации изображений и представляющим собой числовые значения.

Выбранные дескрипторы, такие как суммарная площадь выделенных признаков, их периметр, максимальный и минимальный момент

инерции, позволяют использовать быстрый алгоритм их сравнения и исключить из дальнейшего анализа 80% объектов тестового массива.

Дальнейшее сокращение приоритетного списка до 1–2% от тестового массива можно проводить методами, учитывающими пространственные особенности индивидуальных признаков и, соответственно, требующими больших временных затрат на сравнение изображений. Одним из перспективных методов решения данной задачи является применение контурного анализа.



Список литературы

1. Федоренко В. А., Корнилов М. В., Сидак Е. В. Повышение эффективности работы баллистических идентификационных систем за счет применения растровой электронной микроскопии и цифровых

технологий // Судебная экспертиза : научно-практический журнал. Волгоград : ВА МВД России, 2012. Вып. 3(31). С. 70–79.

2. Niblack W. An Introduction to Image Processing. N.J.: Prentice Hall, Englewood Cliffs, 1986. 626 p.

Use of the Cluster Analysis to Compare the Digital Images Traces of the Firing Pin

M. V. Kornilov

Education Research Institute of Nanostructures and Biosystems of the Saratov State University, 83, Astrakhanskaya, Saratov, 410012, Russia
E-mail: kornilovmv@gmail.ru

V. A. Fedorenko

Education Research Institute of Nanostructures and Biosystems of the Saratov State University, 83, Astrakhanskaya, Saratov, 410012, Russia
E-mail: fed77@yandex.ru

Introduction. One of the most important tasks of forensic ballistics examination is the identification of weapons from marks of the striker. This problem is usually associated with the searching for matching traces from cartridge case repository. Ballistic identification systems (such as POISC, TAIS, Condor, Arsenal and others) are used to help expert in such task. All of them use automatic algorithms for searching matching traces based on the calculation of the cross-correlation function. In this paper cluster analysis methods is used to form a priority list of striker traces images which contains individual attributes in the form of large irregularly shaped areas. **Theoretical part.** In the theoretical part the algorithm is described for extraction and comparison of fire pin marks. Then it is used for forming a priority list. **Experimental part.** The developed algorithm was applied to cartridge case repository consisting more than 100 images of the strikers traces of 24 weapon samples. **Conclusion.** The algorithm allows to form a priority list of 20 images, which includes paired tracks (if bullets repository contains it) with a probability of 100%. Thus it helps to reduce amount of images for comparison by 80%. The proposed method can significantly reduce the search time. To further reduce the priority list should be used more intensive methods, such as contour analysis or singular points comparing.

Key words: ballistic system, comparing firing pin traces, firearm identification, cluster analysis.

References

1. Fedorenko V. F., Kornilov M. V., Sidak E. V. Povyshenie jeffektivnosti raboty ballisticheskikh identifikacionnyh sistem za schet primenenija rastrovoj jelektronnoj mikroskopii i cifrovyyh tehnologij [Improvement of efficiency of the ballistic identification systems' functioning

due to the application of raster electron microscopy and digital technologies]. *Sudebnaja ekspertiza: nauchno-prakticheskij zhurnal* [Forensic examination: scientific-practical journal]. Volgograd, 2012, iss. 3(31). pp. 70–79.

2. Niblack W. An Introduction to Image Processing. N.J., Prentice Hall, Englewood Cliffs, 1986. 626 p.