

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

УДК 004.93

А.А. Дмитрук

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ
ВИДОВ ДЕСКРИПТОРОВ В ЗАДАЧЕ ПОИСКА
МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ**

Приводятся результаты сравнительного анализа эффективности различных видов дескрипторов в контексте задачи поиска медицинских изображений по содержанию. Поиск осуществляется по 11 тестовым базам с использованием 137 видов дескрипторов. Наилучшие результаты показывают дескрипторы на основе матриц совместной встречаемости, а также локальных бинарных шаблонов.

Введение

Коллекции медицинских изображений являются ценным источником информации, играющим важную роль в образовании, исследованиях по медицинской тематике и поддержке принятия решений в клинической практике. При этом большую проблему представляет постоянный рост размера цифровых архивов медицинских учреждений в связи с большой доступностью техники, способной производить цифровые снимки. К такой технике относятся рентгеновские аппараты, томографы, ультразвуковые сканеры и микроскопы. Среднестатистическое отделение радиологии в настоящее время производит несколько терабайт данных в год [1]. Указанные причины побуждают исследователей к разработке инструментов для быстрого и эффективного доступа к данным в архивах медицинских изображений. Индексация и каталогизация архивов изображений традиционно выполняются вручную, в связи с чем являются довольно трудоемкими процедурами, которые нередко подвержены ошибкам [2]. Таким образом, существует практическая необходимость в автоматизации индексирования коллекций медицинских изображений для улучшения скорости и качества поиска по ним.

Системы поиска медицинских изображений обычно позволяют осуществлять текстовый поиск по аннотациям или описаниям к изображениям, которые были заранее внесены вручную. В последние десятилетия в дополнение к традиционному текстовому поиску возник поиск изображений по их содержанию (Content-Based Image Retrieval, CBIR). Его отличительной особенностью является отсутствие необходимости в ручной аннотации изображений. Вместо этого содержание изображений автоматически описывается с помощью визуальных признаков (например, характеристик цвета, формы или текстуры). Поиск похожих изображений осуществляется на основе сравнения визуальных признаков (дескрипторов) изображения-образца и остальных изображений коллекции с целью нахождения наиболее близких в пространстве признаков.

В настоящее время существует огромное число видов дескрипторов изображений, применяющихся для широкого круга задач. Однако дескрипторы, положительно показавшие себя для одного типа изображений и решаемых задач, часто оказываются неэффективными для других [3]. Поскольку эффективность поиска зависит от многих причин, включающих тип используемой метрики сравнения дескрипторов, тип применяемого классификатора, способ оценки качества поиска, объективное сравнение дескрипторов возможно лишь при неизменности указанных условий. Существующая проблема объективного сравнения дескрипторов отражена в литературе. В частности, в статье [4] на примере известной базы изображений COREL показано, как, манипулируя процедурой оценки качества поиска, а также самими данными, можно улучшать характеристики системы поиска по содержанию. В итоге делается вывод о том, что объективное сравнение таких систем практически невозможно, поскольку во многих статьях подробное описание процедуры их оценки не приводится. В качестве возможного решения данной проблемы авторы указывают унификацию не только тестовых баз медицинских изображений, но и процедур оценки качества поиска. Эти условия выполняются на международных мероприятиях, например ImageCLEF [5] или PASCAL VOC [6], по сравнению различных алгоритмов на одинаковых дан-

ных и одним и тем же способом. Таким образом, настоящее исследование ставит целью сравнительный анализ различных видов дескрипторов применительно к задаче поиска медицинских изображений по заданному образцу.

1. Тестовые базы изображений

Для тестирования были выбраны общедоступные базы преимущественно медицинских изображений. Краткая характеристика 11 тестовых баз, включающая размер баз и число исковых классов, приведена в табл. 1.

Таблица 1

Тестовые базы изображений			
Тип изображений	Характеристика	Число изображений/ классов	Разрешение, пиксели
Текстуры	Натуральные текстуры из альбома Бродаса [7, 8]	999/111	213×213
Текстуры	База изображений отсегментированных предметов COIL-20 при разных положениях камеры относительно предмета [9]	1440/20	128×128
Биология	База снимков пыльцы различных видов растений Pollen [10]	630/7	25×25
Цитология	Флуоресцентные микроскопические снимки различных органоидов (постоянные специализированные структуры) культуры клеток HeLa, способных делиться бесконечное число раз [11]	862/10	382×382
Цитология	Флуоресцентные микроскопические снимки различных органоидов (постоянные специализированные структуры) культуры клеток яичников китайского хомячка CHO, которые широко используются в различных генетических исследованиях [11]	327/5	512×512
Гистология	Гистологические снимки яичника и щитовидной железы при 200-кратном увеличении (ткани обработаны специальным иммуногистохимическим маркером D2-40, подсвечивающим сеть лимфатических сосудов, окружающих опухоль)	50 000/2	512×512
Рентген	База японской ассоциации радиотехников JSRT, включающая рентгеновские изображения грудной клетки высокого разрешения с узловыми образованиями и без таковых [12]	247/2	2048×2048
Рентген	Снимки грудной клетки здоровых мужчин и женщин [13]	3000/2	950×1000
Компьютерная томография	КТ-изображения узловых образований (подвыборка базы из [14]) и остальных участков легкого	250/2	40×40
Лица	База изображений лиц AT&T с различным выражением эмоций [15]	400/40	92×112
Лица	База изображений лиц Yale Face Database с различным выражением эмоций, условиями освещенности, наличием закрывающих лицо деталей [16]	166/15	320×243

2. Типы дескрипторов изображений

В табл. 2 приведены виды преобразований и типы дескрипторов изображений, участвовавшие в сравнительном анализе. Основная часть дескрипторов подробно описана в [17, 18] и реализована в общедоступной утилите WND-CHARM. В указанной утилите признаки вычислялись как по исходным изображениям, так и по изображениям после их преобразований. Комбинация всех видов преобразований изображений и дескрипторов породила 124 различных типа дескрипторов. Кроме того, были рассчитаны дескрипторы на основе матриц совместной встречаемости, где в качестве характеристик отдельных пикселей элементарных структур рассматривались яркость I , градиент яркости G и анизотропия A (угол между векторами градиентов). Были сформированы две группы дескрипторов, включающие элементарные структуры в виде пар (ID, GGD, AAD) и триплетов (IID, GGGD, AAAD) пикселей. Длина стороны элементарных структур матриц была 1, 3 и 5 пикселей, а диапазон значений параметра пикселя (яркость, градиент, угол) разбивался на 16 интервалов. Подробное описание указанных дескрипторов приведено в работе [19]. Дополнительные дескрипторы включали обычные гистограммы, гистограммы ориентированного градиента и локальные бинарные шаблоны. В итоге в сравнительном анализе участвовало 137 типов дескрипторов. Все дескрипторы были редуцированы при помощи метода главных компонент с сохранением 95 % исходной вариации данных.

Таблица 2

Типы дескрипторов и виды преобразований изображений

Обозначение	Название
<i>Виды преобразований изображений в утилите WND-CHARM [17]</i>	
Chebyshev()	Преобразование на основе полиномов Чебышева, где порядок полинома соответствует размеру изображения
Edge()	Выделение углов с помощью оператора Прюитта
Fourier()	Преобразование Фурье
Wavelet()	Вейвлет-преобразование на основе симплета пятого порядка с уровнем разложения 1
<i>Типы дескрипторов в утилите WND-CHARM [17]</i>	
Chebyshev-Fourier Coefficients	Коэффициенты двухмерного преобразования Чебышева – Фурье
Chebyshev Coefficients	Коэффициенты полиномов Чебышева
Comb Moments	Первые четыре момента, рассчитанные для четырех направлений 0° , 45° , -45° и 90°
Edge Features	Различные статистики, рассчитанные по градиентам Прюитта
Fractal Features	Фрактальные признаки, которые вычислялись как среднее значение абсолютной разности яркостей пар пикселей по горизонтали и вертикали на различных расстояниях (масштабах)
Gabor Texture	Признаки, вычисленные по изображению после преобразования Габора
Gini Coefficient	Коэффициент Джини, который рассчитывался как среднее значение абсолютной разности яркостей всех пикселей изображения и его средней яркости
Haralick Textures	Четырнадцать признаков, вычисленные по классической матрице Харалика
Inverse-Otsu Object Features	Статистические показатели, вычисленные по инвертированному бинаризованному методом Оцу изображению
Multiscale Histograms	Четыре гистограммы яркостей пикселей с 3, 5, 7 и 9 интервалами
Otsu Object Features	Статистические показатели, вычисленные по бинаризованному методом Оцу изображению

Окончание табл. 2

Обозначение	Название
Pixel Intensity Statistics	Пять признаков изображения, включающие минимальную, максимальную и среднюю яркость, а также медиану и стандартное отклонение.
Radon Coefficients	Коэффициенты преобразования Радона, которые вычислялись как гистограммы проекций яркостей пикселей на линии, проходящие через центр изображения под углом 0° , 45° , 90° и 135°
Tamura Textures	Шесть признаков текстур Тамуры, включающие контраст, направленность, гистограмму с тремя интервалами и сумму элементов параметра зернистости текстуры
Zernike Coefficients	Моменты Цернике
<i>Матрицы совместной встречаемости [19]</i>	
IID	Матрица совместной встречаемости пар пикселей на расстояниях 1, 3, 5 с яркостью I
IID	Матрица совместной встречаемости триплетов пикселей с яркостью I
GGD	Матрица совместной встречаемости пар пикселей с абсолютным значением градиента G
GGGD	Матрица совместной встречаемости триплетов пикселей с абсолютным значением градиента G
AAD	Матрица совместной встречаемости пар пикселей с углом между векторами градиентов A
AAAD	Матрица совместной встречаемости триплетов пикселей с углом между векторами градиентов A
<i>Дополнительные виды признаков изображений</i>	
Historam-16	Гистограмма яркости пикселей, 16 интервалов
HoG-16-1	Гистограммы ориентированного градиента (16 интервалов), рассчитанные по исходному изображению [20]
HoG-16-3	Гистограммы ориентированного градиента, рассчитанные по изображению, которое разбито на девять частей (мозаика 3×3)
HoG-16-5	Гистограммы ориентированного градиента, рассчитанные по изображению, которое разбито на 25 частей (мозаика 5×5)
LBP-u2	Дескрипторы на основе однородных локальных бинарных шаблонов в окрестности 1, 3 и 5 пикселей [21; 22, с. 13]
LBP-ri	Дескрипторы на основе вращательно-инвариантных локальных бинарных шаблонов в окрестности 1, 3 и 5 пикселей [21; 22, с. 18]
LBP-riu2	Дескрипторы на основе однородных вращательно-инвариантных локальных бинарных шаблонов в окрестности 1, 3 и 5 пикселей [21; 22, с. 16]

Следует отметить, что рассмотрены далеко не все существующие дескрипторы. Например, по результатам распознавания различных модальностей медицинских изображений на соревнованиях ImageCLEF 2011–2013 SIFT-подобные дескрипторы показали наилучшие результаты [23], однако процедура их построения более сложная, чем для исследуемых дескрипторов. Она зависит от способа вычисления, величины используемых параметров дескрипторов, а также от способа построения так называемого «мешка признаков» (bag-of-features), поэтому указанные дескрипторы не рассматриваются в данной работе.

3. Методика постановки экспериментов и оценки качества поиска изображений

Тестирование выполнялось согласно схеме поиска изображений по образцу. В качестве образца последовательно выступало каждое изображение из тестовой базы, а поиск осуществлялся по всем остальным. Результат поиска считался правильным, если найденное изображение принадлежало к тому же классу, что и образец для поиска.

В литературе при оценке систем поиска по содержанию изображений с заранее заданными классами часто используют следующую оценку [24]:

$$RR(q) = \frac{NF(\alpha, q)}{GT(q)} \times 100 \%,$$

где $RR(q)$ – качество поиска (retrieval rate) по q запросам, принимающее значения от 0 % (не найдено ни одного изображения) до 100 % (все найденные изображения искомого класса) для q запросов;

$GT(q)$ – размер базы изображений заранее заданных искомого классов (ground truth) для запроса q ;

$NF(\alpha, q)$ – число изображений искомого класса, найденных среди первых $\alpha \times GT(q)$ результатов для запроса q .

Таким образом, качество поиска определялось как процент правильно найденных изображений, т. е. изображений, принадлежащих к тому же классу, что и образец для поиска, от их общего числа вне зависимости, на каком месте они находились.

В рассматриваемом случае качество поиска рассчитывалось для первых $q = 5$ найденных изображений, наиболее близких к образцу. Степень близости вычислялась как расстояние L_1 (расстояние городских кварталов Манхэттена) между дескриптором изображения-запроса и дескрипторами остальных изображений.

4. Результаты экспериментов

Для более компактного представления результатов экспериментов использовался суммарный ранг дескриптора, т. е. место, которое занимал определенный тип дескриптора относительно остальных. Поскольку исследовалось 137 типов дескрипторов на 11 тестовых базах, минимально возможный суммарный ранг мог равняться 11, а максимальный – 1507. Первый вариант соответствует ситуации, когда дескриптор во всех случаях занимал первое место (ранг), а второй – когда последнее. В результате была сформирована табл. 3, которая содержит суммарные ранги 30 дескрипторов с наилучшим качеством поиска по всем тестовым базам.

Таблица 3

Суммарный ранг наиболее эффективных дескрипторов

Тип дескриптора	Суммарный ранг
IID	129
LBP-u2	133
Zernike Coefficients (Fourier (Wavelet ()))	155
Historam-16	168
GGD	169
IID	195
Multiscale Histograms	202
Zernike Coefficients (Fourier (Edge ()))	211
Zernike Coefficients (Fourier ())	213
LBP-ri	219
HoG-16-3	263
Fractal Features (Wavelet ())	289
LBP-riu2	306
Fractal Features (Fourier (Chebyshev ()))	325
GGGD	357
Multiscale Histograms (Wavelet ())	367
AAAD	401
HoG-16-5	406
Edge Features	431

Окончание табл. 3

Тип дескриптора	Суммарный ранг
AAD	435
Zernike Coefficients (Edge ())	437
Fractal Features	446
Multiscale Histograms (Fourier (Chebyshev ()))	450
Comb Moments	452
Zernike Coefficients	455
Gabor Textures	457
Pixel Intensity Statistics	464
Radon Coefficients	465
Fractal Features (Chebyshev (Fourier ()))	485
Fractal Features (Fourier (Wavelet ()))	502

Наилучшие результаты показали дескрипторы на основе матриц совместной встречаемости как пар (IID, GGD, AAD), так и триплетов (IID, GGGD, AAAD) пикселей. Похожие результаты оказались у дескрипторов на основе локальных бинарных шаблонов (LBP-u2, LBP-gi, LBP-riu2). Процедура вычисления локальных бинарных шаблонов похожа на процедуру вычисления матриц совместной встречаемости: в ней так же используются пары пикселей в некоторой окрестности, которые затем определенным образом кодируются. Возможно, данный факт объясняет их похожие описательные способности. Высокие результаты показали моменты Цернике, вычисленные по изображениям после преобразования Фурье, а также классические гистограммы, гистограммы ориентированного градиента и фрактальные признаки.

Результаты сравнения дескрипторов, полученные для соответствующих семи тестовых баз изображений в [17, 18], в целом согласуются с результатами, представленными в настоящей статье. В указанных статьях точность поиска оценивалась с помощью перекрестной проверки с разбиением набора данных на 50 частей и с использованием заранее отобранных наиболее информативных признаков.

Заключение

В результате сравнительного анализа различных типов дескрипторов при решении задачи поиска медицинских изображений по образцу наилучшие результаты показали дескрипторы на основе матриц совместной встречаемости, а также на основе локальных бинарных шаблонов.

Следует отметить, что результаты сравнительного анализа дескрипторов справедливы лишь для рассмотренных типов дескрипторов и изображений и не могут быть распространены на все типы медицинских изображений в силу их большого разнообразия и широкого круга решаемых на их основе задач.

Список литературы

1. Partik, B. Digital (R) Evolution in Radiology / B. Partik, C. Schaefer-Prokop // Digital radiology in chest imaging. – Springer Vienna, 2001. – P. 189–203.
2. Quality of DICOM header information for image categorization / M.O. Guld [et al.] // Medical Imaging 2002. International Society for Optics and Photonics. – San Diego, 2002. – P. 280–287.
3. Deselaers, T. Features for image retrieval: an experimental comparison / T. Deselaers, D. Keysers, H. Ney // Information Retrieval. – 2008. – Vol. 11, № 2. – P. 77–107.
4. Müller, H. The truth about Corel-evaluation in image retrieval / H. Müller, S. Marchand-Maillet, T. Pun // Image and Video Retrieval. – Springer Berlin Heidelberg, 2002. – P. 38–49.
5. ImageCLEF – The CLEF Cross Language Image Retrieval Track [Electronic resource]. – 2015. – Mode of access : <http://www.imageclef.org>. – Date of access : 15.05.2015.
6. Pattern Analysis, Statistical Modelling and Computational Learning [Electronic resource]. – 2015. – Mode of access : <http://www.pascal-network.org>. – Date of access : 15.05.2015.

7. Brodatz, P. Textures: A Photographic Album for Artists and Designers / P. Brodatz. – N.Y. : Dover Publications, 1966. – 112 p.
8. Brodatz texture database with CASIA V3 Iris database naming scheme [Electronic resource]. – 2012. – Mode of access : <http://staff.neu.edu.tr/~kkilic/prj/lac/brodatz/brodatz.html>. – Date of access : 11.01.2015.
9. Columbia University Image Library (COIL-20) [Electronic resource]. – 2013. – Mode of access : <http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php>. – Date of access : 01.06.2015.
10. Geometric features of Pollen grains [Electronic resource]. – 2013. – Mode of access : <http://ome.grc.nia.nih.gov/iicbu2008/pollen>. – Date of access : 01.06.2015.
11. Identifying Sub-cellular Organelles [Electronic resource]. – 2013. – Mode of access : <http://ome.grc.nia.nih.gov/iicbu2008/hela/index.html>. – Date of access : 01.06.2015.
12. JSRT Database [Electronic resource]. – 2013. – Mode of access : <http://www.jsrt.or.jp/jsrt-db/eng.php>. – Date of access : 01.06.2015.
13. Ковалев, В.А. Влияние мер близости в пространстве признаков на качество поиска медицинских изображений по содержанию / В.А. Ковалев, А.А. Дмитрук // Информатика. – 2011. – № 30. – С. 5–11.
14. Дмитрук, А.А. Обнаружение узловых образований в легком по данным компьютерной томографии / А.А. Дмитрук // Информатика. – 2011. – № 29. – С. 25–33.
15. The Database of Faces [Electronic resource]. – 2013. – Mode of access : <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. – Date of access : 01.06.2015.
16. Yale Face Database [Electronic resource]. – 2013. – Mode of access : <http://vision.ucsd.edu/content/yale-face-database>. – Date of access : 01.06.2015.
17. WND-CHARM: Multi-purpose image classification using compound image transforms / N. Orlov [et al.] // Pattern recognition letters. – 2008. – Vol. 29, № 11. – P. 1684–1693.
18. General pattern recognition in images using WND-CHARM [Electronic resource]. – 2013. – Mode of access : <http://ome.grc.nia.nih.gov/wnd-charm>. – Date of access : 01.06.2015.
19. Ковалев, В.А. Анализ структуры трехмерных медицинских изображений / В.А. Ковалев. – Минск : Белорус. наука, 2008. – 263 с.
20. Dalal, N. Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // Computer Vision and Pattern Recognition. CVPR 2005. IEEE Computer Society Conf. – Montbonnot-Saint-Martin, 2005. – Vol. 1. – P. 886–893.
21. A general Local Binary Pattern (LBP) implementation for Matlab [Electronic resource]. – 2013. – Mode of access : <http://www.cse.oulu.fi/CMV/Downloads/LBP Matlab>. – Date of access : 01.06.2015.
22. Pietikainen, M. Computer vision using local binary patterns / M. Pietikainen. – Springer, 2011. – Vol. 40. – 207 p.
23. Improved medical image modality classification using a combination of visual and textual features / I. Dimitrovski [et al.] // Computerized Medical Imaging and Graphics. – 2015. – Vol. 39. – P. 14–26.
24. Color and texture descriptors / B.S. Manjunath [et al.] // IEEE Transactions on Circuits and Systems for Video Technology. – 2001. – Vol. 11, № 6. – P. 703–715.

Поступила 25.06.2015

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: dmitruk@newman.bas-net.by*

A.A. Dmitruk

**COMPARATIVE STUDY OF DESCRIPTORS
FOR MEDICAL IMAGE RETRIEVAL**

This paper presents results of comparative analysis of image descriptors in the context of content-based medical image retrieval problem. Experiments are carried out on 11 test databases using 137 different types of descriptors. Image descriptors based on co-occurrence matrices and local binary patterns showed the best performance.