

УДК 004.932.72'1

С.Г. Антошук, д-р техн. наук, проф.,
Н.А. Годовиченко, магістр,
Одес. нац. політехн. ун-т

АНАЛИЗ ТОЧЕЧНЫХ ОСОБЕННОСТЕЙ ИЗОБРАЖЕНИЯ В СИСТЕМЕ «МОБИЛЬНЫЙ ВИРТУАЛЬНЫЙ ГИД»

С.Г. Антошук, М.А. Годовиченко. Аналіз точкових особливостей зображення у системі «Мобільний віртуальний гід». Розроблено підсистему розпізнавання музейних експонатів для системи «Мобільний віртуальний гід». Модернізовано алгоритм детектування точкових особливостей SIFT, що використовується для розпізнавання музейних експонатів, з метою підвищення швидкодії алгоритму. Проведено тестування підсистеми на тестовому наборі зображень.

Ключові слова: точкові особливості, розпізнавання, SIFT, мобільний віртуальний гід.

С.Г. Антошук, Н.А. Годовиченко. Анализ точечных особенностей изображения в системе «Мобильный виртуальный гид». Разработана подсистема распознавания музейных экспонатов для системы «Мобильный виртуальный гид». Модернизирован алгоритм детектирования точечных особенностей SIFT, который используется для распознавания музейных экспонатов, с целью увеличения быстродействия алгоритма. Проведено тестирование подсистемы на тестовом наборе изображений.

Ключевые слова: точечные особенности, распознавание, SIFT, мобильный виртуальный гид.

S.G. Antoschuk, N.A. Godovichenko. Image local features analysis for “Mobile virtual guide” system. A subsystem for museum exhibits recognition for “Mobile virtual guide” system was developed. The SIFT algorithm of local feature detection, used for recognition of museum exhibits, was modernized to improve its performance. The subsystem was tested using a test set of images.

Keywords: local features, recognition, SIFT, mobile virtual guide.

Повсеместное распространение мобильных технологий привело к тому, что в настоящее время мобильные устройства используются во многих сферах жизни человека. Одним из возможных применений мобильных устройств является создание мобильных виртуальных «гидов» для музеев и галерей. Использование мобильных устройств позволяет предоставлять посетителям разнообразную информацию об экспонатах в виде полноценных мультимедийных презентаций.

Рассмотренные автором аналоги системы «Мобильный виртуальный гид» для музеев [1...4] работают по следующему принципу:

- на мобильное устройство предварительно устанавливается специальное приложение;
- мобильное устройство считывает информацию об экспонате;
- установленное приложение обрабатывает считанные данные и распознает экспонат;
- данные об экспонате извлекаются из базы экспонатов и предоставляются пользователю.

Получение информации об экспонате с использованием мобильных устройств может быть осуществлено несколькими способами. Некоторые из них предполагают размещение на экспонате специальных радиометок и штрих-кодов, что в большинстве случаев, является нежелательным, так как они занимают место, визуалью портят интерьер, и посетитель не всегда может подойти близко к экспонату, чтобы считать метку.

Поэтому предпочтительным является визуальное распознавание экспоната, которое не требует установки меток и кодов и не влияет на сам экспонат. При реализации системы «Мобильный виртуальный гид» на основе визуального распознавания решаются две основные задачи: идентификация фотоизображения экспоната по характерным признакам и поиск по этим признакам подобного изображения в фотобазе системы.

При этом повышаются требования к надежности системы распознавания, так как идентифицировать экспонат по специальной метке или штрих-коду легче, чем распознать его по сделанному фотоснимку.

Исходя из специфики предметной области, для достижения большей надежности распознавания, в разрабатываемой системе «Мобильный виртуальный гид», введены следующие требования:

— фотографируемый экспонат должен занимать более 30 % фотоснимка, при этом большинство экспонатов выставляется на однородном фоне без посторонних объектов, которые могли бы помешать процессу распознавания;

— фотоизображение экспоната не содержит существенного шума и размытия, угол съемки — не более 45 градусов;

— обработка фотоизображения происходит на мобильном устройстве.

С учетом предъявленных требований, наиболее перспективным является метод распознавания экспонатов с использованием точечных особенностей изображений, который позволяет осуществлять сравнение изображений по характеристикам выделенных ключевых точек [5].

Структура системы «Мобильный виртуальный гид», с учетом визуального распознавания экспонатов с помощью выделения точечных особенностей изображения, представлена на рис. 1.

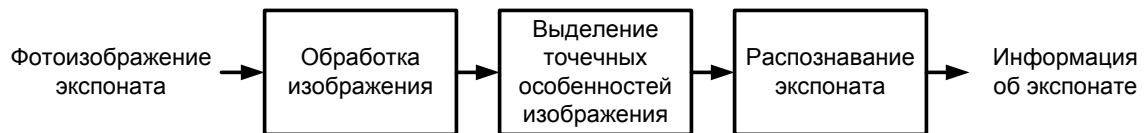


Рис. 1. Структура системы «Мобильный виртуальный гид»

Среди различных алгоритмов обнаружения и описания точечных особенностей можно выделить алгоритм SIFT [6], который обеспечивает нахождение ключевых точек, инвариантных к линейному масштабированию, изменению ориентации и частичную инвариантность к аффинным преобразованиям и изменению освещенности. К недостаткам этого алгоритма можно отнести то, что он генерирует большое количество дескрипторов, по которым идет поиск информации в фотобазе экспонатов системы.

Целью данной работы является повышение быстродействия алгоритма SIFT за счет снижения количества генерируемых дескрипторов.

Алгоритм выделения точечных особенностей SIFT основан на нахождении ключевых точек в масштабируемом пространстве изображения и реализуется несколькими этапами.

Этап 1. Поиск локальных экстремумов в разностях гауссиан.

Разностью гауссиан $D(x, y, \sigma)$ называют изображение, которое было получено с помощью вычитания гауссианы исходного изображения из гауссианы с другим радиусом размытия:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma), \quad (1)$$

где $G(x, y, \sigma)$ — гауссово ядро;

σ — радиус размытия;

* — операция свертки;

$I(x, y)$ — исходное изображение;

$L(x, y, \sigma)$ — масштабированное изображение;

k — шаг масштабирования.

Кандидатами в ключевые точки являются локальные экстремумы, найденные в разности гауссиан двух соседних масштабов.

Этап 2. Локализация ключевых точек.

Координаты кандидатов определяются с субпиксельной точностью с помощью аппроксимирования функции разности гауссиан многочленом Тейлора второго порядка, взятого в точке вычисленного экстремума:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X, \quad (2)$$

где D — функция разности гауссиан;

$X = (x, y, \sigma)^T$ — вектор смещения относительно точки разложения;

$\frac{\partial D^T}{\partial X}$ — градиент;

$\frac{\partial^2 D}{\partial X^2} X$ — матрица Гессе.

Этап 3. Нахождение ориентации ключевых точек.

Направление точки изображения $L(x, y)$ вычисляется исходя из направлений градиентов точек, соседних с ключевой. Вычисление градиентов осуществляется в пирамиде гауссиан с масштабом, наиболее близким к масштабу ключевой точки. Величина градиента $m(x, y)$ и его направление $\theta(x, y)$ вычисляется по следующим формулам:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}, \quad (3)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}. \quad (4)$$

Гистограмма направлений состоит из 36 компонент, которые равномерно покрывают область в 360 градусов. Каждая точка окна с координатами (x, y) вносит вклад, равный $mG(x, y, \sigma)$, в ту компоненту гистограммы, которая покрывает промежуток, содержащий направление градиента $\theta(x, y)$.

Направление ключевой точки лежит в промежутке, покрываемом максимальной компонентой гистограммы. Значения максимальной компоненты и двух соседних с ней интерполируются параболой и точка максимума этой параболы берется в качестве направления ключевой точки. Если в гистограмме есть еще компоненты с величинами не меньше $0,8 \cdot \max$, генерируются дополнительные ключевые точки.

Этап 4. Построение дескриптора.

В алгоритме SIFT дескриптором является вектор. Дескриптор вычисляется на гауссиане, ближайшей по масштабу к ключевой точке, и, исходя из градиентов, в некотором окне ключевой точки. Перед вычислением дескриптора окно поворачивается относительно направления ключевой точки, чем и достигается инвариантность относительно поворота. Затем происходит свертка окна с гауссианой со значением σ равным 1,5 от масштаба ключевой точки. Дескрипторы ключевой точки, как правило, состоят из 128 компонент, каждая из которых охватывает окно 4×4 пикселя.

С целью сокращения вычислительных затрат, авторами данной работы была предложена модификация алгоритма SIFT.

Специфика работы системы «Мобильный виртуальный гид» состоит в том, что фотографирование экспонатов производится в помещении с небольшого расстояния. Фотографии подобного рода представляют собой, как правило, изображение экспоната на фоне стены или стен, то есть, с преобладанием вертикальных плоскостей на изображении.

Если предположить, что ракурс изображения будет стабилен относительно оси наблюдения, то ориентация дескрипторов ключевых точек на вертикальных плоскостях также

будет стабильна, и не будет поворачиваться. Из этого следует, что для таких ключевых точек нет необходимости обеспечивать инвариантность к повороту.

Учитывая вышесказанное, из алгоритма был исключен этап нахождения ориентации ключевых точек и был модифицирован этап построения дескрипторов ключевых точек, из которого была удалена процедура поворота окна дескриптора относительно направления ключевой точки.

Повышение быстродействия алгоритма SIFT было достигнуто за счет следующих факторов:

— ускорение процесса нахождения дескриптора для каждой ключевой точки изображения, за счет исключения этапа нахождения ориентации ключевых точек;

— сокращение итогового количества найденных ключевых точек, за счет исключения процедуры генерации дополнительных ключевых точек изображения на этапе нахождения ориентации ключевых точек;

— ускорение процесса распознавания экспоната за счет сокращения количества дескрипторов для описания каждого экспоната в базе данных.

Пример работы модифицированного алгоритма обнаружения точечных особенностей SIFT представлен на рис. 2.

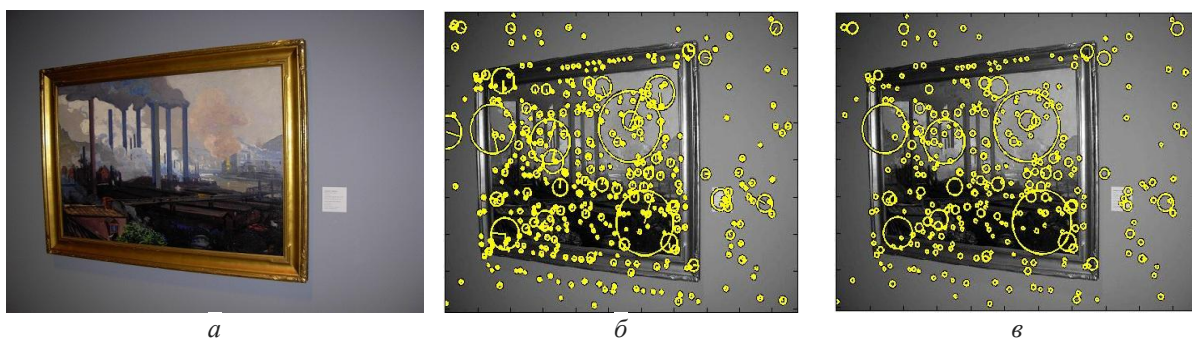


Рис. 2. Пример обнаружения точечных особенностей фотоизображения экспоната: а — исходное изображение; б — результат работы алгоритма SIFT; в — результат работы модифицированного алгоритма SIFT

В результате тестирования, было установлено, что, несмотря на исключение процедуры обеспечения инвариантности к повороту, модифицированный алгоритм SIFT робастен к незначительным поворотам — не более 20 градусов. Дальнейшее увеличение угла поворота приводит к существенному снижению качества распознавания.

Для тестирования модифицированного алгоритма была создана база дескрипторов SIFT для 65 экспонатов. Фотоизображение экспоната фиксировалось камерой мобильного устройства с различных ракурсов и расстояний. Идентификация экспонатов — обнаружение ключевых точек, на фотоизображениях — выполнялась с помощью модифицированного алгоритма SIFT. Распознавание экспоната происходило сопоставлением дескрипторов фотоизображения, с дескрипторами экспонатов хранимых в базе системы. Фотографирование экспонатов производилось с разных ракурсов. Процесс сравнения осуществлялся с помощью алгоритма Best-Bin-First [7], который обеспечивает приближенное решение задачи поиска ближайшего соседа с точностью 95 %. В качестве мобильного устройства был использован смартфон Samsung Galaxy Pocket S5300, который оснащен процессором Broadcom BCM21553, 832 МГц и 290 Мбайтами оперативной памяти, версия ОС Android 2.3.

Экспериментальная проверка модифицированного алгоритма SIFT показала:

— время нахождения ключевых точек и генерации дескрипторов было сокращено, в среднем, на 30 %;

— время распознавания экспоната было сокращено на 14 %;

— количество найденных ключевых точек уменьшилось, в среднем, на 16 %.

— процент успешного распознавания экспоната не изменился, при фронтальном фотографировании экспоната, и несколько снизился, в среднем на 3...4 %, при фотографировании экспоната под разными углами.

Результаты тестирования работы системы «Мобильный виртуальный гид» при разных условиях фотосъемки приведены в таблице.

Результаты тестирования модифицированного алгоритма SIFT

Название алгоритма	Среднее время поиска ключевых точек, с	Среднее время поиска экспоната по базе, с	Среднее количество ключевых точек	Процент успешного распознавания (%) фото-изображения, расположенного			
				фронтально		под разными углами	
				1 метр	4 метра	1 метр	4 метра
Алгоритм SIFT	6,25	3,63	562	96,5	92,2	88,7	86,5
Модифицированный алгоритм SIFT	4,02	3,12	470	96,5	92,0	87,4	85,4

В данной работе была разработана система «Мобильный виртуальный гид» для музеев и галерей. Для идентификации экспоната был использован подход визуального распознавания объекта на изображении с помощью выделения точечных особенностей.

В качестве алгоритма выделения точечных особенностей был использован алгоритм SIFT, который обеспечивает инвариантность ключевых точек к линейному масштабированию, изменению ориентации и частичную инвариантность к аффинным преобразованиям и изменению освещенности. Для повышения быстродействия работы системы была предложена модификация алгоритма SIFT, которая позволила снизить время выделения точечных особенностей и поиска по базе экспонатов без существенного ухудшения процента успешного распознавания экспонатов.

Предложенная система может быть использована при создании мобильных «гидов» для музеев, галерей, выставочных залов. Предложенная модификация алгоритма SIFT может быть использована при выделении точечных особенностей в условиях небольшой глубины изображения, а также при фотографировании объекта с небольшого расстояния.

Литература

1. Evaluating the usability of a mobile guide: The influence of location, participants and resources / J. Kjeldskov, C. Graham, S. Pedell [and others] // Behavior and Informational Technology. — 2005. — Vol. 4. — P. 51 — 65.
2. Proctor, N. Mobile Apps for Museums: The AAM Guide to Planning and Strategy [Electronic resource] / N. Proctor. — Washington, 2012. — <http://www.mobileappsformuseums.wordpress.com>. — 14.12.2012.
3. Damala, A. Interaction design and evaluation of mobile guides for the Museum Visit: the case study in multimedia and mobile augmented reality // A. Damala, P. Cubaud. — Conservatoire national des arts et métiers. — 2009. — 365 p.
4. Bruns, F. Enabling Mobile Phones To Support Large-Scale Museum Guidance / F. Bruns, B. Brombach, T. Zeidler, O. Bimber // Multimedia, IEEE. — 2007. — No 3. — P. 16 — 25.
5. Bernal, J. Feature Detectors and Feature Descriptors: Where We Are Now / J. Bernal, F. Villarino, J. Sanchez // CVC Technical Report # 154. — 2010. — P. 1 — 108.
6. Lowe, D. G. Distinctive Image Features From Scale-Invariant Keypoints / D.G. Lowe // International Journal of Computer Vision 60. — 2004. — Vol. 4. — P. 91 — 110.

-
7. Beis, J. Shape Indexing Using Approximate Nearest-Neighbour Search In High-Dimensional Spaces / J. Beis, D. Lowe // IEEE Computer Society Conference. — 1997. — Vol. 1. — P. 1000 — 1006.

References

1. Evaluating the usability of a mobile guide: The influence of location, participants and resources / J. Kjeldskov, C. Graham, S. Pedell [and others] // Behavior and Informational Technology. — 2005. — Vol. 4. — pp. 51 — 65.
2. Proctor, N. Mobile Apps for Museums: The AAM Guide to Planning and Strategy [Electronic resource] / N. Proctor. — Washington, 2012. — <http://www.mobileappsformuseums.wordpress.com>. — 14.12.2012.
3. Damala, A. Interaction design and evaluation of mobile guides for the Museum Visit: the case study in multimedia and mobile augmented reality // A. Damala, P. Cubaud. — Conservatoire national des arts et métiers. — 2009. — 365 p.
4. Bruns, F. Enabling Mobile Phones To Support Large-Scale Museum Guidance / F. Bruns, B. Brombach, T. Zeidler, O. Bimber // Multimedia, IEEE. — 2007. — # 3. — pp. 16 — 25.
5. Bernal, J. Feature Detectors and Feature Descriptors: Where We Are Now / J. Bernal, F. Villarino, J. Sanchez // CVC Technical Report # 154. — 2010. — pp. 1 — 108.
6. Lowe, D. G. Distinctive Image Features From Scale-Invariant Keypoints / D.G. Lowe // International Journal of Computer Vision 60. — 2004. — Vol. 4. — pp. 91 — 110.
7. Beis, J. Shape Indexing Using Approximate Nearest-Neighbour Search In High-Dimensional Spaces / J. Beis, D. Lowe // IEEE Computer Society Conference. — 1997. — Vol. 1. — pp. 1000 — 1006.

Рецензент д-р техн. наук, проф. Одес. нац. политехн. ун-та Положаенко С.А.

Поступила в редакцию 16 декабря 2012 г.