

- Метрики для оценки качества.

Одним из ключевых моментов является возможность добавления собственных алгоритмов классификации и обучения. Это достигается с помощью обобщения алгоритмов в группы с одинаковыми интерфейсами и возможностями языка Python.

После получения результатов тестирования их можно сравнить между собой, сохранить для последующего использования, а также экспортировать для печати.

Таким образом, была разработана система оценки качества алгоритмов классификации. Она имеет возможность выбирать различные алгоритмы классификации и разные метрики для оценки. Задачей дальнейших исследований является реализация функционала для алгоритмов кластеризации.

Список использованных источников:

1. <http://archive.ics.uci.edu/ml/> - Электронный ресурс.
2. Гуров, С.И. Как оценить надежность алгоритма классификации // Таврический вестник информатики и математики №1. – Симферополь, 2002.
3. Ramdass, D. Document Classification for Newspaper Articles / D. Ramdass, S. Seshasai // 6.863 Final Project – MIT, Spring 2009.
4. Michie, D. Machine Learning, Neural and Statistical Classification / D. Michie, D.J. Spiegelhalter, C.C. Taylor // Overseas Press, 2009 edition (August 28, 2009) – 290 p.

АЛГОРИТМ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ ТОПОЛОГИИ МИКРОСХЕМ НА МИКРОСКОПНЫХ ИЗОБРАЖЕНИЯХ СЛОЕВ ТОПОЛОГИИ СБИС

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Питкин А.Б.

Воронов А. А. – канд. техн. наук, доцент

В процессе производства сверхбольших интегральных микросхем (СБИС), возможно появление дефектов на топологических слоях. Такие дефекты возникают например в процессе переноса изображения или в процессе фотолитографии. Автоматический контроль качества позволит в кратчайшие сроки выявлять различные дефекты. Идея заключается в поиске заданных шаблонов известных дефектов, либо шаблонов с изображением эталонов элементов на значительно большем изображении представляющем собой снимок слоя готовой микросхемы.

Простейшим способом детектирования на изображении является поиск по шаблону (template matching) [2]. Предварительно изображение обрабатывается с использованием фильтра дистантного преобразования [3], что хорошо выделяет контуры объектов на изображении. Далее происходит постепенный параллельный перенос изображения шаблона в окне над исходным изображением слоя и вычисление корреляции между ними по одной из формул [4]. На основании полученного значения корреляции изображений делается вывод о совпадении шаблона с фрагментом под ним.

Однако элементы на слое интегральной микросхемы могут располагаться с вращением кратным 90° в плоскости изображения, что увеличивает число проходов по изображению в четыре раза. Также изображение элемента может незначительно отличаться от эталонного изображения в связи с особенностями технологического процесса производства микросхемы.

Целью алгоритма является задача разработать метод поиска изображения устойчивый к поворотам изображения и масштабированию и позволяющий с высокой степенью вероятности найти все повторения заданного шаблона на изображении.

Универсальными методами поиска изображений являются методы, основанные на выделении ключевых точек на изображении. Поиск на изображении определенных характерных примитивов позволит перейти от описания изображения в виде массива цветных пикселей к массиву параметров описывающих эти примитивы. Таким образом, сокращается количество параметров, по которым необходимо производить дальнейшее сравнение изображения. Поиск на изображении при использовании методов основанных на ключевых точках состоит из трех этапов:

1. Этап детектирования ключевых точек. На данном этапе происходит сканирование изображения и нахождение на нем характерных точек интереса.
2. Этап создания дескрипторов для ключевых точек, найденных на предыдущем шаге. Таким образом, на данном этапе формируется математическое описание найденных ключевых точек в виде вектора значений.
3. Сопоставление наборов значений дескрипторов полученных в процессе обработки двух изображений и нахождение близких друг к другу значений.

Наиболее эффективными типами дескрипторов в данный момент являются дескрипторы SIFT (SIFT - Scale Invariant Feature Transform) [5] и SURF (Speeded Up Robust Features) являющийся улучшением

предыдущего [6]. SURF производит поиск особых точек с помощью матрицы Гессе. Детерминант матрицы Гессе (так называемый гессиан) достигает экстремума в точках максимального изменения градиента яркости. Он хорошо детектирует пятна, углы и края линий. Технически детектирование основано на вычислении перепадов яркости изображения.

Точки, найденные с помощью дескриптора SIFT устойчивы к повороту изображения, растяжению. В результате сопоставления дескрипторов на изображении было обнаружено большое число ложных срабатываний дескриптора. Данная проблема связана с малым размером шаблона, по которому выполняется поиск на изображении и как следствие малым количеством информации, которую несет в себе этот шаблон. Это приводит к тому, что дескрипторы шаблону случайным образом совпадают с некоторыми точками на изображении.

Перед применением алгоритма поиска ключевых точек необходимо произвести предварительную обработку изображения слоя топологии интегральной микросхемы. Одной из задач тут является удаление шумов на изображении, а также сглаживание изображения.

Рассмотрим два детектора ключевых точек при поиске заданных шаблонов на изображении. Важным параметром при работе детекторов на последнем этапе является максимально допустимое расстояние между дескрипторами на изображении и на шаблоне. Слишком малое значение отсечет часть верных результатов, слишком высокое значение будет давать слишком большое число ложных срабатываний.

В качестве тестовых изображений использовалось изображение слоя топологии, а также пять шаблонов изображений. Также были созданы копии изображений шаблонов, повернутые на 90 градусов. В процессе тестирования использовались алгоритмы, включенные в библиотеку OpenCV.

Результаты сведены в таблицу 1.

Таблица 1 – Результаты работы детекторов

Шаблон/ детектор	SURF			SIFT		
	Ключевых точек на шаблоне	Совпадений ключевых точек	Итог	Ключевых точек на шаблоне	Совпадений ключевых точек	Итог
Шаблон 1	17	7	+	17	2	+
Шаблон 1 повернутый	17	7	+	19	0	-
Шаблон 2	14	4	+	19	4	+
Шаблон 2 перевернутый	14	4	+	17	0	-
Шаблон 3	2968	1803	+	3066	2772	+
Шаблон 3 перевернутый	2967	1803	+	3054	702	+
Шаблон 4	117	49	+	59	12	+
Шаблон 4 перевернутый	113	56	+	54	1	+
Шаблон 5	6	1	+	15	1	+
Шаблон 5 перевернутый	6	1	+	19	0	-

Время поиска ключевых точек и вычисления дескрипторов в тестовой программе составило 33 секунды для метода SIFT и 23 секунды для метода SURF, при этом использовалось изображение 7000 на 3000 пикселей. В дальнейшем этап поиска ключевых точек и вычисления дескрипторов можно опустить, сохранив найденные значения в кэш. Время сопоставления ключевых точек составляет менее 100 миллисекунд.

Таким образом, в процессе тестирования было определено, что детектор SURF более качественно находит ключевые точки даже на перевернутых изображениях и справился со всеми десятью тестами. Дескриптор SIFT не всегда находил ключевые точки на повернутых изображениях и число совпадений дескрипторов меньше. Оба дескриптора показали хорошие результаты уникальности ключевых точек и отсутствие случайных совпадений отдельных точек с фрагментами на изображении. По сравнению с методами, основанными на поиске по шаблону, данные методы дают значительно меньшее число ложных срабатываний. Это объясняется тем, что происходит сравнение не яркостных характеристик шаблона и изображения, а сравниваются их морфологические особенности.

Были исследованы методы поиска на изображении с использованием ключевых точек. Методы обеспечивают нахождение совпадений в случае поворота изображения и незначительных изменений их размеров.

Список использованных источников:

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва, 2005. – 1072с
2. Template matching [Электронный ресурс]. – Режим доступа: https://en.wikipedia.org/wiki/Template_matching
3. Distance Transform [Электронный ресурс]. – Режим доступа: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/distance.htm>
4. Which are the matching methods available in OpenCV [Электронный ресурс]. – Режим доступа: http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html#which-are-the-matching-methods-available-in-opencv
5. Scale-invariant feature transform [Электронный ресурс]. – Режим доступа: https://en.wikipedia.org/wiki/Scale-invariant_feature_transform

МОБИЛЬНОЕ ПРИЛОЖЕНИЕ ДЛЯ ЭКСПРЕСС-АНАЛИЗА СОСТОЯНИЯ РАСТИТЕЛЬНОСТИ ПО ИЗОБРАЖЕНИЯМ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Романовский К. В.

Дудкин А. А. – д-р. техн. наук, профессор

Описана и решена проблема быстрой обработки большого количества данных в системе распознавания сельскохозяйственной растительности путем ее разделения на основную обработку (сервер) и экспресс-анализ (мобильная платформа).

При принятии решений в процессе мониторинга сельскохозяйственных полей основным является распознавание пораженных заболеваниями участков. Для решения данной задачи предложено использование программно-аппаратного комплекса, мощности которого представлены мобильной системой и сервером. Серверная часть предполагает использование мощного компьютера. Она представлена подсистемами параллельной обработки изображений и принятия решений. Первая из них включает в себя наиболее часто используемые методы обработки фотоаэроснимков, вторая – отвечает за выбор способа построения информационных признаков на основе данных дистанционного зондирования Земли (ДЗЗ). В результате эксперт получает возможность детального изучения обработанных фотоаэроснимков различными методами, сэкономив при этом время на обработку за счет распараллеливания на современных вычислительных системах.

Мобильное приложение (мобильная подсистема) – это программные средства, предоставляющие возможность отображать информацию о состоянии сельскохозяйственной растительности на карте, которая является набором упорядоченных по географическим координатам цифровых фотографий. Данная подсистема оснащена возможностью экспресс-оценки состояния растительного покрова и определения зон (сегментов) с нарушением фотосинтеза, позволяет получать наглядную информацию о состоянии растительности в рамках целого поля или о его отдельных участках. Скорость обработки данных обуславливается нахождением мобильной платформы непосредственно у изучаемого участка – цифровые фотографии попадают в систему сразу после окончания процесса съемки, без их длительной транспортировки на сервер. В процессе обработки эксперт формирует выборки необходимых для дальнейшей обработки фотографий, уменьшая количество поступающих на вход сервера данных.

Источником данных был выбран беспилотный летательный аппарат (БПЛА), но не исключено использование современной сельскохозяйственной техники и средств ДЗЗ. Главным требованием к источнику данных является возможность использовать информацию со спутников глобального позиционирования и добавлять ее к изображениям.

В докладе предложена структура мобильного приложения и алгоритм экспресс-анализа состояния сельскохозяйственной растительности. Использование мобильной платформы позволило ускорить принятие решений при мониторинге состояния растительности данных за счет более ранней обработки и экспресс-анализа, проводимых уже на этапе получения информации, а также за счет сокращения объема данных, посылаемых для обработки на сервер.

Разработанное мобильное приложение включает пять основных модулей (рис. 1): приемки и передачи информации, обработки изображений, построения карты, хранения данных, управления БПЛА.