САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ — ПРОЦЕССОВ УПРАВЛЕНИЯ
КАФЕДРА КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ И МНОГОПРОЦЕССОРНЫХ СИСТЕМ

Васильев Павел Сергеевич

Выпускная квалификационная работа бакалавра

Распознавание типов изображений документов

Направление 010400

Прикладная математика, фундаментальная информатика и основы программирования

Научный руководитель,
кандидат технических наук,
доцент
Гришкин В. М.

Санкт-Петербург
2016

Содержание

1 Введение

В данной работе рассматривается задача кластеризации данных применительно к изображениям. Целью работы является извлечение пригодных для кластеризации признаков изображения и построение на их основе хеш-множества, определяющего кластеры; полученные данные предполагается использовать как тестовое множество для обучения более сложных нейронных сетей в случаях, когда невозможно это сделать вручную в силу большого количества данных. Под изображением в данной работе понимается чёрно-белое изображение, содержащее текст, таблицы, либо другие элементы, присущие различного рода документам. Применение рассмотренных методов извлечения признаков ограничивается описанным выше содержимым изображений, в то время как рассмотренный метод кластеризации является независимым и может применяться в различных задачах, в том числе не имеющим отношения к изображениям.

В первой половине работы описывается постановка задачи, используемые для её решения понятия, определения, операции и методы. На их основе выводится алгоритм решения. Во второй половине описано решение задачи и результаты.

2 Постановка задачи

Дано множество изображений. Известно, что изображения принадлежат k типам документов, при этом нет изображений, заведомо не принадлежащих ни одному типу, так же как изображений, формирующих тип, состоящий только лишь из одного изображения. Подразумевается, что нет данных, позволяющих программно определить, относятся ли два произвольным образом взятых из выборки изображения к одному типу.

Необходимо построить алгоритм, позволяющий кластеризовать изображения за достижимое время. Воздействие человека при этом допускается только при выборе начальных точек кластеризации, однако задача должна решаться, пусть с уменьшением надёжности, и без такого воздействия.

3 Основные понятия и термины

Математическая морфология Серра — парадигма анализа изображений, основанная на нелинейных операторах, подробно описанная в [1].

Пусть множество X — некоторое конечное множество в дискретном пространстве размерности n с определённой евклидовой метрикой $\mathbb{D}^{\mathrm{n}}$.

Определение 1. Структурирующий элемент (структурный элемент) — некоторое множество $B \subset \mathbb{D}^{\mathrm{n}}$.

В данной работе рассматриваются и используются только выпуклые и односвязные структурирующие элементы.

Определение 2. Операция переноса множества Z — Z_t — задаётся следующим образом:

$$Z_t = \{z + t \mid z \in Z\}.$$

Определение 3. Операция дилатации (расширения) множества X со структурирующим элементом B — $X \oplus B$ — задаётся следующим образом:

$$X \oplus B = \bigcup_{b \in B} X_b.$$

Определение 4. Операция эрозии (сужения) множества X со структурирующим элементом B — $X \ominus B$ —задаётся следующим образом:

$$X \ominus B = \{x \in X \mid B_x \subseteq X\}.$$

Определение 5. Размер множества X — наибольшее число r, для которого выполнено равенство:

$$X \ominus B = \emptyset,\ B = \{x \mid \|x\| \leq r\}.$$

Определение 6. Операция замыкания (закрытия) множества X со структу-

рирующим элементом B — $X \bullet B$ — определяется следующим выражением:

$$X \bullet B = (X \oplus B) \ominus B.$$

Практический смысл операции замыкания в том, что она позволяет устранить пустоты в множестве X, если их размер не превышает размер B. Иными словами, если $X = Y \setminus H$, где $Y \subseteq M$ — односвязная область, $H \subset Y$, и $H \ominus B = \emptyset$, то после применения операции замыкания двусвязная область X станет односвязной, при этом размер остальных множеств рассматриваемого пространства останется неизменным. Результат выполнения операции закрытия проиллюстрирован на рис. 1.

Рис. 1: Результат применения операции «замыкание», [2]

Определение 7. Операция размыкания (открытия) множества X со структурирующим элементом B — $X \circ B$ — определяется следующим выражением:

$$X \circ B = (X \ominus B) \oplus B.$$

Практический смысл операции размыкания в том, что она позволяет устранить все $X \subset M$, размер которых не превышает размер B. Иными словами, если $X \subset M$ и $X \ominus B = \emptyset$, то после применения операции замыкания такое множество будет исключено из рассматриваемого пространства, при этом размер остальных множеств останется неизменным. Результат выполнения операции открытия проиллюстрирован на рис. 2.

Определение 8. Морфологический спектр — характеристика X, полученная

Рис. 2: Результат применения операции «размыкание», [2]

в результате морфологического преобразования $X \bullet B$ или $X \circ B$ и зависящая от размера B.

Определение морфологического спектра строится на основе терминов математической морфологии Ж. Серра [1], приведённых выше, по аналогии с частотным спектром, строящимся через преобразование Фурье. Преобразование Фурье заключается в одномерном умножении значения преобразуемой функции $f(x)$ на $e^{-ix\omega}$ и последующем вычислении площади под полученным сигналом $f(x)e^{-ix\omega}$. Будем говорить, что $e^{-ix\omega}$ — структурный элемент преобразования Фурье с параметром ω. При замене операндов на дискретные множества, описанные в предыдущем пункте, а оператора — на замыкание либо размыкание, получаем морфологическое преобразование — $X \diamond B$. Морфологическим спектром для этого преобразования будет, соответственно, являться площадь нового сигнала, а параметром — размер B.

Определение 9. Хеширование (англ. hashing) — операция преобразования входного массива данных произвольной длины в массив данных фиксированной длины по определённому алгоритму. Значения, получаемые в результате хеширования называются «хеш-сумма», «хеш-код» либо просто «хеш».

Пусть входным массивом является вектор, а выходной битовой строкой — целое число, не превосходящее ω (это условие можно считать эквивалентным условию фиксированности длины выходной строки в битах). Введём функцию H, такую, что для любого n-мерного вектора $v = (v_1\ v_2 \cdots v_n) \in \mathbb{R}^n$ выполнено $H(v) = \alpha$, где α — вещественное число, удовлетворяющее вышеприведённым ограничениям. Такую функцию будем называть хеш-функцией, определённой в пространстве $\mathbb{R}^n$.

Определение 10. LSH (Locality-Sensitive Hashing, локально-чувствительное хеширование) — хеширование, использующее локально-чувствительное семейство хеш-функций.

Пусть задано семейство хеш-функций $\mathcal{H} = \{H_1, \cdots, H_m\}$. $\mathcal{H}$ будем называть локально-чувствительным (обладающим свойством локальной чувствительности), если для любых $q,\ v$ функция $p(d) = \Pr_{\mathcal{H}}\left[h(q) = h(v) \mid \|q - v\| = d\right]$ является строго убывающей по d, то есть вероятность совпадения хешей двух векторов строго убывает с увеличением расстояния между ними.

Задача 1. Задача о нахождении k R-ближайших соседей.

Пусть задано пространство вещественных чисел $\mathbb{R}^d$ с определённой на нём евклидовой метрикой, множество $\mathcal{P} \subset \mathbb{R}^d$ и радиус $R > 0$. R-ближайшим соседом для точки x будем называть такую точку ξ множества $\mathcal{P}$, что $\|\xi - x\| \leq R$. Задача состоит в том, чтобы по заданному параметру R и точке x определить до k точек множества $\mathcal{P}$ включительно, являющихся R-ближайшими соседями данной точки.

Задача 2. Задача о нахождении k (R, c)-ближайших соседей.

Пусть задано пространство вещественных чисел $\mathbb{R}^d$ с определённой на нём евклидовой метрикой, множество $\mathcal{P} \subset \mathbb{R}^d$ и радиус $R > 0$ и параметр $c > 0$. (R, c)-ближайшим соседом для точки x будем называть такую точку ξ множества $\mathcal{P}$, для которой $\|\xi - x\| \leq cR$ при условии, что существует точка $\eta \in \mathcal{P}$, такая, что $\|\eta - x\| \leq R$. Задача состоит в том, чтобы найти множество $\mathcal{S}$, содержащее до k различных точек множества $\mathcal{P}$ включительно, удовлетворяющих условию задачи.

Задача 3. Задача о нахождении k $(R, 1-\delta)$-ближайших соседей (вероятностное нахождение k R-ближайших соседей).

Дополним условие задачи 1. Пусть также задано число $\delta \in (0, 1)$ (можно условно назвать это число вероятностью ошибки второго рода, то есть вероятностью не отнести точку ξ к ближайшим соседям, в то время как на деле она таковой является) и функция $F(x) \xrightarrow[\delta \to 0]{} \left\{ \bigcup\limits_{i \leq k} \xi_i \mid \|\xi_i - x\| \leq R \right\}$, значением

которой с вероятностью не менее $1 - \delta$ будет являться множество из не более, чем k точек, удовлетворяющих задаче 1. Необходимо составить алгоритм (функцию $F(x)$ для составления такого множества.

Замечание. Вариантом данной задачи является задача о нахождении всех $(R, 1-\delta)$-ближайших соседей. В таком случае значение k принимается равным бесконечности, либо мощности конечного множества, содержащего все точки.

4 Метод извлечения признаков изображения и их кластеризации

На практике вычисление морфологического спектра зачастую является трудновыполнимой задачей из-за существенных различий в получаемых характеристиках в зависимости от вида структурного элемента. Точные алгоритмы являются трудоёмкими, в то время как приближённые — узкоспециализированными. Наиболее эффективный из известных точных алгоритмов решения этой задачи был предложен и подробно описан в работе [3]. На рис. 3-4 показан пример вычисления морфологического спектра тестового изображения по такому алгоритму.

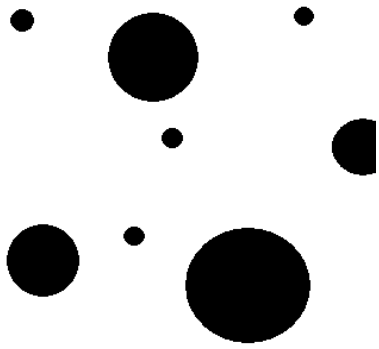


Рис. 3: Фрагмент входного изображения

В некоторых случаях допустимо использовать не морфологический спектр в строгом смысле определения 8, а его приближённые аналоги [3, 4, 5]. В частности, в работе [3] были выдвинуты предположения о возможности использования в качестве основной операции по вычислению морфологического спектра не открытия/закрытия, а эрозии/дилатации с последующим вычислением изменения закрашенной площади. Именно такой подход использует данная работа для выделения признаков изображения. Такой спектр не сохраняет форму и, в гораздо большей степени, размер объектов, что делает невозможным его применение для сегментации изображения, но даёт пред-

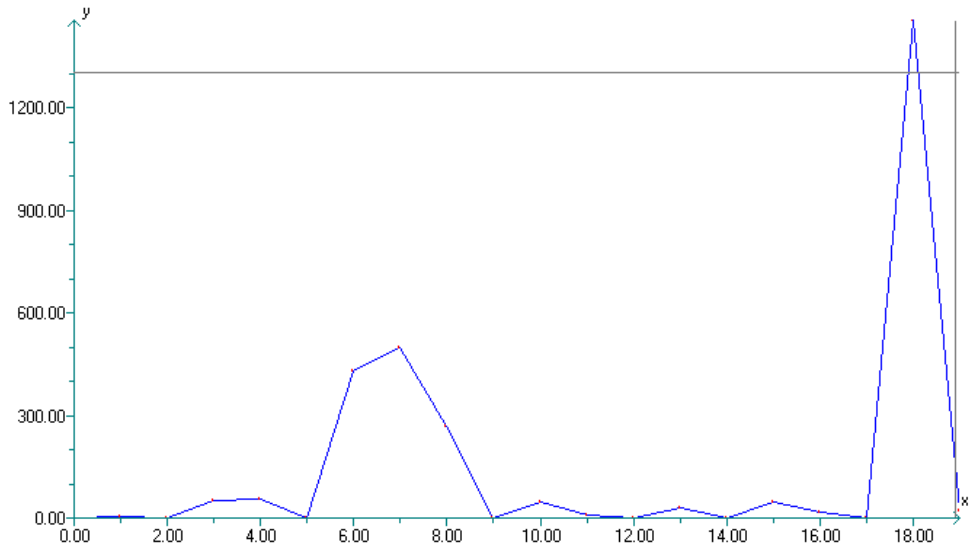


Рис. 4: Морфологический спектр фрагмента со сферическим структурным элементом. Видны резкие изменения площади при использовании круглого структурного элемента с радиусом, совпадающим с размером изображённых фигур.

ставление о пустотах значительных размеров, тем самым давая представление о заполненности плоского изображения текстом, что позволяет использовать его для извлечения признаков текстовых изображений.

Построение морфологического спектра выполнено следующим образом. Изображение в градациях серого приводится к чёрно-белому путём пороговой фильтрации: всем пикселям с яркостью ниже заданного порога присваивается наименьшее значение яркости, пикселям выше порога — наибольшее. Далее происходит выравнивание изображения по следующему алгоритму.

1. Выполняется инверсия изображения.

2. По алгоритму Рурка [6] либо его ускоренной (аппроксимирующей) версии [7] находится прямоугольник наименьшей площади, описывающий текстовое содержимое (на практике, учитывая, что в документах все строки, заголовки и аннотации параллельны, достаточно найти прямоугольник наименьшей площади, описывающий всё содержимое отсканированного изображения; здесь полагается, что помехи, возникшие при сканировании, были отсеяны на этапе пороговой фильтрации). *Под содержимым понимаются ненулевые пиксели; соответственно, пло-*

щадь содержимого — это количество ненулевых пикселей.

3. Выполняется поворот изображения на угол поворота полученного прямоугольника в противоположную сторону.

К полученному исправленному изображению применяется следующий алгоритм (построения приближённого морфологического спектра):

1. Подсчитывается площадь содержимого (количество ненулевых пикселей) S. На данном этапе размер структурирующего элемента r_B равен 0, и никакие морфологические операции не выполняются.

2. r_B увеличивается на некоторую фиксированную величину. К исходному изображению применяется операция дилатации с структурирующим элементом размера r_B.

3. Снова подсчитывается площадь содержимого $\bar{S}$, после чего в качестве величины морфологического спектра берётся $\frac{S-\bar{S}}{S}$.

Этапы 2-3 выполняются некоторое количество раз для получения величин морфологического спектра для структурирующего элемента разных размеров. При практическом подсчёте используется сферический структурирующий элемент, размер которого увеличивается на 1 на каждой итерации.

Данный алгоритм можно оптимизировать. Последовательное увеличение размера структурирующего элемента на фиксированную величину, в случае со сферическим структурирующим элементом и операции эрозии, эквивалентно последовательному применению операции эрозии со структурирующим элементом, размер которого равен этой величине. Это порождает новую версию алгоритма, в котором на этапе 2 вместо эрозии *исходного* изображения увеличивающимся структурирующим элементом будет происходить эрозия изображения, полученного на предыдущем этапе, элементом фиксированной величины. Данный алгоритм, в свою очередь, можно свести к последовательному применению операции пороговой фильтрации и вовсе избавиться

от трудоёмких операций математической морфологии Серра, сохраняя эквивалентность преобразований. Для этого нужно вычислить преобразование дистанций исходного изображения, и к полученным данным последовательно применять пороговую фильтрацию.

Запишем обновлённый алгоритм:

1. Подсчитывается площадь содержимого S. Размер структурирующего элемента r_B равен 0. Понятие размера структурирующего элемента в контексте данного алгоритма имеет смысл только в качестве метрики преобразования, так как сам структурирующий элемент здесь уже не используется.

2. Находится преобразование дистанций. Результат содержит нули в тех точках, где исходное изображение было ненулевым, и расстояние до ближайшей ненулевой точки исходного изображения в противном случае.

3. r_B увеличивается на некоторую фиксированную величину. Применяется пороговая фильтрация с порогом r_B.

4. Снова подсчитывается площадь содержимого $\bar{S}$, после чего в качестве величины морфологического спектра берётся $\frac{S-\bar{S}}{S}$.

На шаге 3 выполняется операция, эквивалентная эрозии инвертированного оригинала изображения, которая, в свою очередь, эквивалентна дилатации оригинального изображения. Преимуществом данного алгоритма является также то, что время, за которое выполняются шаги 3-4, в значительной степени остаётся неизменным.

Для большей чувствительности к форме объектов в реализованном алгоритме дополнительно к построению морфологического спектра всего изображения была применена сегментация входного изображения по горизонтали и вертикали с последующим вычислением значения площади морфологического преобразования на каждом сегменте с фиксированным, заранее выбранным,

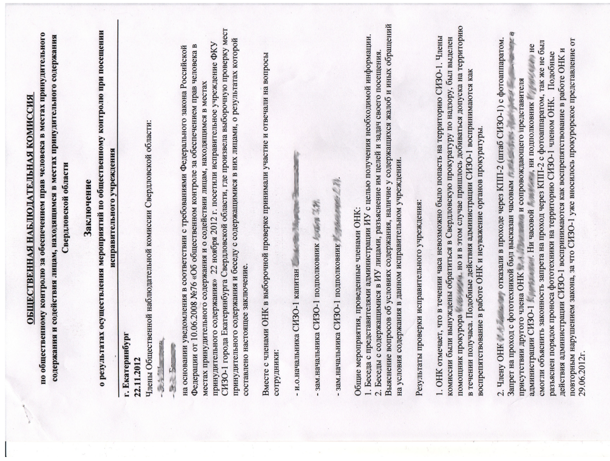
размером структурирующего элемента. Таким образом, для каждого изображения построены три характеристики — морфологический спектр, морфологическое преобразование по горизонтали и морфологическое преобразование по вертикали. Условимся, что разбиения по горизонтали и вертикали производятся с одинаковым количеством секторов — это позволит избежать неоднозначных результатов в случае альбомной ориентации содержимого на входных изображениях, так как содержимое заранее неизвестно. На выходе данный алгоритм даёт, соответственно, три вектора: $M = \{m_1, \cdots, m_k\}$, $H = \{h_1, \cdots, h_p\}$, $V = \{v_1, \cdots, v_q\}$. Необходимо учитывать, что размерность векторов M — k — ограничивается только размером входных изображений $k < \min\limits_{I \in \mathcal{I}} \{x(I), y(I)\}$, а размерность H и V — q — задаётся произвольно в пределах $\left(2, \frac{\min\limits_{I \in \mathcal{I}}\{x(I), y(I)\}}{2}\right)$, где $\mathcal{I}$ — множество всех входных изображений, однако значения размерности, близкие к 2, имеют малую практическую применимость, а значения, превышающие 30 представляют высокую вычислительную сложность, в большинстве случаев не сопоставимую с полученным увеличением точности, и, к тому же, порождают высокочастотные помехи, ограничивающие достижимую точность, поэтому на практике для p применяются значения в пределах $[5, 30]$, что для листа формата A4 соответствует сегментам с размерами от 7×9.9 миллиметра до 42×59.4 миллиметра при условии, что изображение не повёрнуто и обрезано по краю бумаги. На практике из-за этих факторов размеры сегментов несколько больше.

Полученные данные необходимо кластеризовать. Для этого в данной работе применяется метод k средних. В нашем случае известны и размерность, и число кластеров, поэтому такая задача имеет решение, однако в общем случае время решения задачи составляет $O\left(n^{dk+1} \log n\right)$, где k — количество кластеров, d — размерность пространства, n — количество кластеризуемых элементов, что было доказано в работе [8]. Затраты на решение могут быть уменьшены с использованием LSH для определения всех $(R, 1-\delta)$-ближайших соседей в евклидовом пространстве; к примеру, алгоритм, описанный в работе [9], работает за время $dn^{O\left(\frac{1}{c^2}\right)}$, где c — константа аппроксимации (см. задачу 2).

В такой модификации процесса нахождения кластеров на каждой итерации метода k средних вместо перебора всех точек, подсчёта для них расстояния и выбора ближайших используется одна макрооперация поиска по хешу, которая, как показано выше, является гораздо менее трудоёмкой.

5 Результаты

Описанный выше алгоритм извлечения и кластеризации признаков был протестирован на большой выборке изображений отсканированных документов (74 различных типа документов без титульных листов и приложений, значительно отличающихся по структуре от содержательной части документа; в среднем по 25 изображений каждого типа). В данной выборке, в отличие от первоначальной задачи, заранее известен тип каждого из документов — это сделано для того, чтобы оценить погрешность алгоритма. Для первоначального выбора центров кластеров используется ввод пользователем идентификатора кластера для случайно выбранного изображения из данного кластера. *Выбор изображения производится независимо от выполнения самого алгоритма и компенсирует отсутствие адаптивного алгоритма выбора центра кластера. В контексте решаемой задачи данное действие допускается как первоначальное задание параметров человеком для большого набора изображений.* В результате за сравнительно небольшое время при низкой сложности алгоритма были получены спектры входных изображений, содержащие данные о форме их содержимого. На рис. 5-7 проиллюстрирован результат выполнения алгоритма получения значений приближённого морфологического спектра для структурирующего элемента размером от 1 до 10 пикселей, а также результат подсчёта приближённого морфологического преобразования с использованием разбиения на 10 сегментов и различных размеров структурирующего элемента размером от 2 до 10 пикселей (трёхмерный график).



ОБЩЕСТВЕННАЯ НАБЛЮДАТЕЛЬНАЯ КОМИССИЯ
по общественному контролю за обеспечением прав человека в местах принудительного содержания и содействия лицам, находящимся в местах принудительного содержания
Свердловской области

Заключение

о результатах осуществления мероприятий по общественному контролю при посещении исправительного учреждения

г. Екатеринбург
22.11.2012

Члены Общественной наблюдательной комиссии Свердловской области:
- [illegible]
- [illegible]

на основании уведомления в соответствии с требованиями Федерального закона Российской Федерации от 10.06.2008 №76 «Об общественном контроле за обеспечением прав человека в местах принудительного содержания и о содействии лицам, находящимся в местах принудительного содержания» 22 ноября 2012 г. посетили исправительное учреждение ФКУ СИЗО-1 города Екатеринбурга Свердловской области, где произвела выборочную проверку мест принудительного содержания и беседу с содержащимися в них лицами, о результатах которой составлено настоящее заключение.

Вместе с членами ОНК в выборочной проверке принимали участие и отвечали на вопросы сотрудники:

- и.о.начальника СИЗО-1 капитан [illegible]

- зам.начальника СИЗО-1 подполковник [illegible]

- зам.начальника СИЗО-1 подполковник [illegible].

Общие мероприятия, проведенные членами ОНК:
1. Беседа с представителями администрации ИУ с целью получения необходимой информации.
2. Беседы с содержащимися в ИУ лицами, разъяснение им целей и задач своего посещения.
Выяснение вопросов об условиях содержания, наличие у содержащихся жалоб и иных обращений на условия содержания в данном исправительном учреждении.

Результаты проверки исправительного учреждения:

1. ОНК отмечает, что в течении часа невозможно было попасть на территорию СИЗО-1. Члены комиссии были вынуждены обратиться в Свердловскую прокуратуру по надзору, был выделен помощник прокурора [illegible], но и в этом случае пришлось добиваться допуска на территорию в течении получаса. Подобные действия администрации СИЗО-1 воспринимаются как воспрепятствование в работе ОНК и неуважение органов прокуратуры.

2. Члену ОНК [illegible] отказали в проходе через КПП-2 (штаб СИЗО-1) с фотоаппаратом. Запрет на проход с фототехникой был высказан часовым [illegible] в присутствии другого члена ОНК [illegible] и сопровождающего представителя администрации СИЗО-1 [illegible]. Ни часовой [illegible], ни подполковник [illegible] не смогли объяснить законность запрета на проход через КПП-2 с фотоаппаратом, так же не был разъяснен порядок проноса фототехники на территорию СИЗО-1 членом ОНК. Подобные действия администрации СИЗО-1 воспринимаются как воспрепятствование в работе ОНК и повторным нарушением закона, за что СИЗО-1 уже вносилось прокурорское представление от 29.06.2012г.

Рис. 5: Тестовое изображение

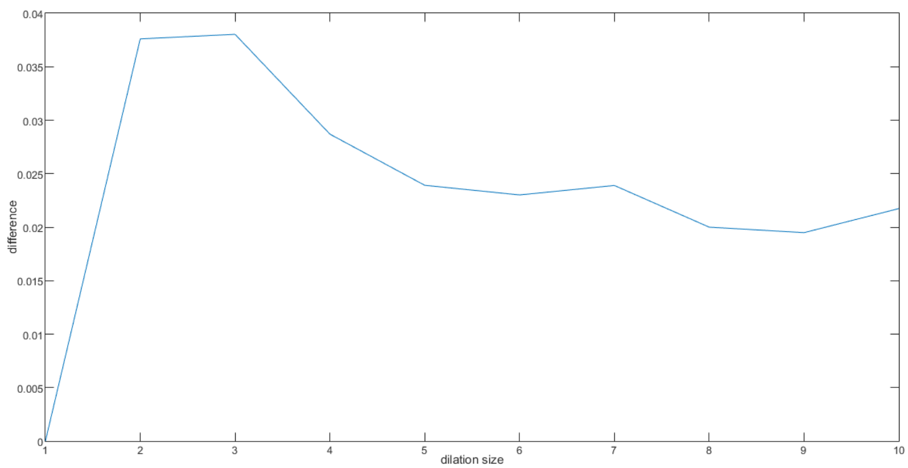


Рис. 6: Приближённый морфологический спектр тестового изображения

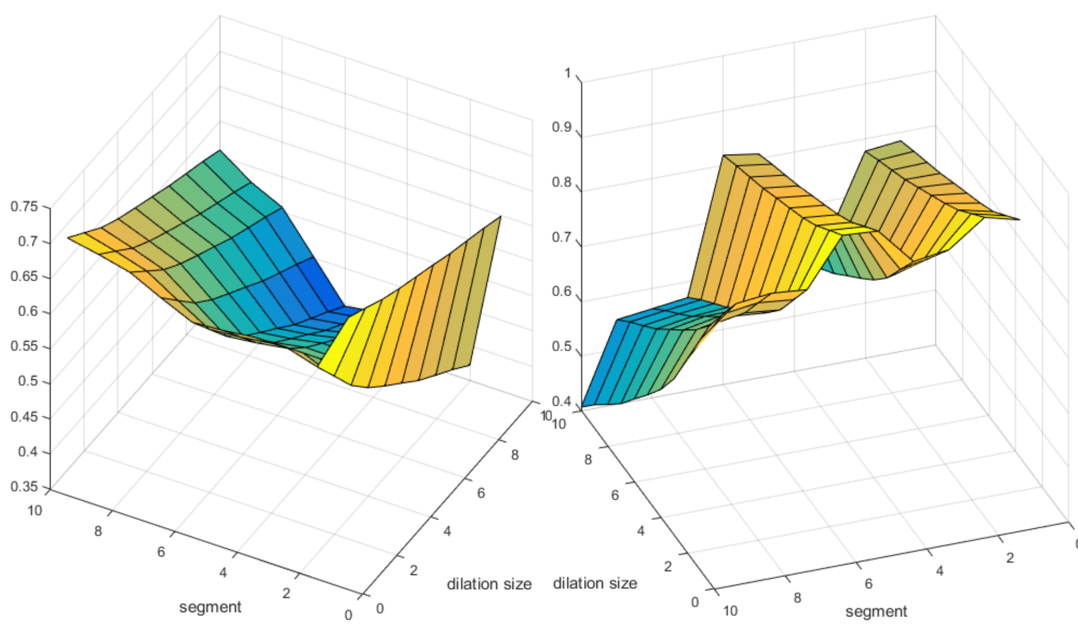


Рис. 7: Вертикальный и горизонтальный морфологические спектры сегментов тестового изображения, кол-во сегментов — 10×10

Для кластеризации полученных данных был применён описанный в предыдущей части модифицированный метод k средних. Из-за высокой размерности векторов, полученных при извлечении признаков из выборки изображений, иллюстрация процесса кластеризации вызывает затруднения, однако можно привести типичные ошибки алгоритма. На рис. 8- 9 показан пример неверного результата, когда страницы документов из разных классов обладают похожей структурой текста, из-за чего их отнесение к различным кластерам становится зависимым от всей выборки. Общая точность алгоритма в данном случае составила 76.8%.

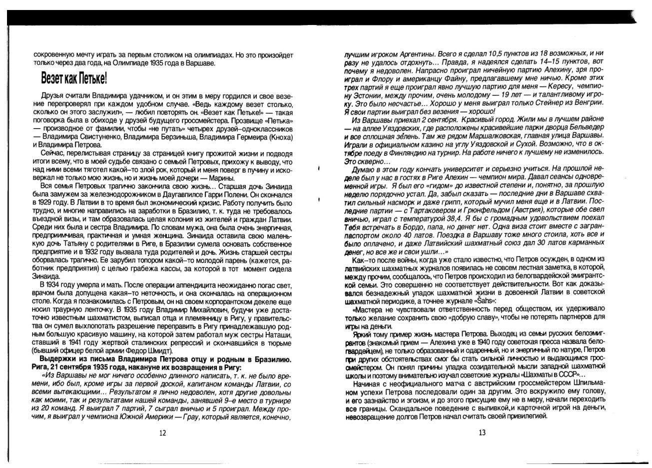


Рис. 8: Отсканированная книга «Звезда, погасшая до срока»

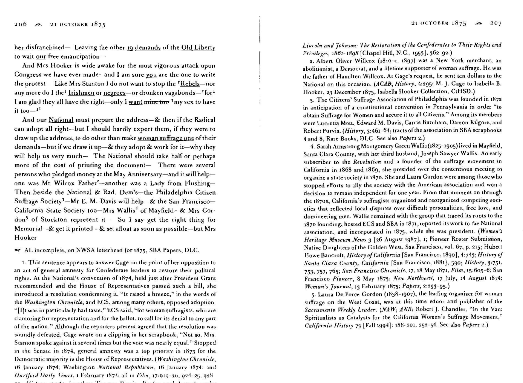


Рис. 9: Отсканированная книга «The Selected Papers of Elizabeth Cady Stanton & Susan B. Anthony»

В данной работе с целью уменьшения времени работы по рекомендации

авторов алгоритма [9, 10] в качестве радиуса поиска ближайших соседей выбрано медианное значение расстояния между кластерами. При практическом применении алгоритма кластеризации с вероятностным решением задачи о ближайших соседях часто возникают ситуации, когда одна точка многомерного пространства относится сразу к нескольким кластерам. В контексте применяемого алгоритма не существует [10] оптимального решения такого конфликта, за исключением уменьшения радиуса поиска, как следствие все подобные ситуации решаются прямым сравнением расстояний до предполагаемых кластеров и выбором кратчайшего. Учитывая, что для каждой «конфликтной» точки число предполагаемых кластеров значительно меньше общего числа кластеров, а также то, что самих «конфликтных» точек несравнимо меньше, чем всех точек множетсва, такой подход позволяет решить данную проблему без серьёзных временных затрат. Кроме того, все точки, не попавшие ни в один кластер по методу LSH, также относятся к ближайшему кластеру точным методом. К примеру, приведённый выше случай ложной кластеризации был разрешён за счёт такой поправки к алгоритму, и результирующая точность составила 81.6%.

Недостаток приведённого метода кластеризации текстовых изображений — относительно низкая точность, методы повышения которой предполагают вычислительные затраты, уменьшение которых является основной целью данной работы; достоинством подхода является простота реализации и вычислений и применимость в качестве инструмента для построения обучающего множества более сложных, адаптивных методов классификации изображений, таких как нейросети.

6 Заключение

Нами был рассмотрен метод быстрого извлечения признаков и последующей их кластеризации на основе полученных признаков, применимый к большим наборам данных. В сравнении с более сложными математически методами классификации с помощью нейросетей, данный алгоритм даёт меньшую точность, однако всё же достаточную для применения его при построении тестового множества нейросети

Среди направлений дальнейших исследований можно выделить улучшение алгоритма извлечения признаков с целью более точно выделять особенности структуры текста, а также алгоритма кластеризации с целью повышения надёжности. Кроме того, увеличить точность работы алгоритма можно с помощью адаптивной кластеризации, применения более сложных алгоритмов разбиения или исследования пространства признаков на возможность применения неевклидовых метрик.

Список литературы

1. J. P. F. Serra. Image analysis and mathematical morphology // Academic Press (1982).
2. Opencv documentation.
   URL `http://docs.opencv.org/2.4/doc/tutorials/imgproc/opening_closing_hats/opening_closing_hats.html`
3. E. R. Urbach, M. H. F. Wilkinson. Efficient 2-d grayscale morphological transformations with arbitrary flat structuring elements // IEEE Transactions on Image Processing (2008) стр. 1–8.
4. P. Soille, E. J. Breen, R. Jones. Recursive implementation of erosions and dilations along discrete lines at arbitrary angles // IEEE Transactions on Pattern Analysis and Machine Intelligence (1996) стр. 562–567.
5. Z. Yu-qian, G. Wei-hua, C. Zhen-cheng, T. Jing-tian, L. Ling-yun. Medical images edge detection based on mathematical morphology // Pattern Recognition Letters (2007) стр. 6492–6495.
6. J. O'Rourke. Finding minimal enclosing boxes // International Journal of Computer and Information Sciences (1985) стр. 183–199.
7. G. Barequet, S. Har-Peled. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions // Journal of Algorithms (38) (2001) стр. 91–109.
8. M. Inaba, N. Katoh, H. Imai. Applications of weighted voronoi diagrams and randomization to variance-based k-clustering // Proceedings of 10th ACM Symposium on Computational Geometry (1994) стр. 332–339.
9. A. Andoni, P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions // The IEEE Symposium on Foundations of Computer Science (2006) стр. 1–7,10.
10. M. Datar, N. Immorlica, P. Indyk, V. S. Mirrokni. Nearest-neighbor methods in learning and vision: Theory and practice // MIT Press (2006) стр. 221–222.