

Отзыв официального оппонента д. т. н. О. С. Ушмаева

на диссертацию Р. В. Шаповалова «Методы структурного обучения в задачах совместной разметки», представленную на соискание ученой степени кандидата физико-математических наук по специальности 01.01.09 — дискретная математика и математическая кибернетика.

Актуальность. В диссертационной работе Р. В. Шаповалова определен класс задач совместной разметки, и разработаны математические методы для решения этих задач. Практическая применимость методов проиллюстрирована на примере задач семантической сегментации в компьютерном зрении и категоризации текстовых документов. В отличие от стандартных задач распознавания образов, таких как классификация и регрессия, предсказываемая случайная величина предполагается многомерной, при этом не делаются предположения о независимости отдельных компонент этого случайного вектора. Для её решения формулируются вероятностные модели так называемых случайных полей, и рассматриваются их факторизации. В то время как задача вероятностного вывода в подобных моделях изучена достаточно хорошо, методология настройки их параметров слабо разработана как с теоретической, так и с практической точек зрения. Существующие методы, как правило, используют недостаточно гибкую линейную параметризацию, требовательны к вычислительным ресурсам и нуждаются в полностью размеченной обучающей выборке. Диссертационная работа направлена на решение этих проблем.

Содержание. Глава 1 содержит обзор ненаправленных графических моделей (случайных полей). Описана формулировка модели условного случайного поля, рассмотрено понятие факторизации. Проведен обзор методов вывода максимальной апостериорной оценки факторизованного распределения, включающий методы на основе целочисленного программирования, а также специализированные методы, такие как передача сообщений на фактор-графе, методы на основе разрезов графов и двойственного разложения минимизируемого функционала. Также дан обзор методов настройки параметров ненаправленных графических моделей по обучающей выборке, включающий максимизацию правдоподобия и её приближения и структурный метод опорных векторов. Обзор показывает, что для обучения большинства представленных в научно-технической литературе моделей требуются большие обучающие выборки, для которых задана полная разметка, получение которой часто не представляется возможным.

Эта проблема рассмотрена в главе 2. В ней сформулировано обобщение структурного метода опорных векторов на случай ограничений с латентными переменными, подобное рассмотренному в работе [Yu and Joachims, ICML 2009]. Такие ограничения позволяют учитывать слабую аннотацию экземпляров обучающей выборки, представляющую собой

некоторую статистику полной разметки. Проведен теоретический анализ модели. Показана связь алгоритма настройки параметров в такой модели с ЕМ-алгоритмом, хорошо зарекомендовавшим себя при обучении других вероятностных моделей с латентными переменными, например, при восстановлении смесей распределений (mixture models). В этой главе в качестве приложения рассмотрена задача семантической сегментации изображений. Традиционно для обучения в этой задаче требуется попиксельная разметка изображения. Вместо этого в работе рассматриваются следующие виды слабой аннотации:

- 1) аннотация метками изображений, в которой задаются присутствующие на изображении категории объектов, но не их локации,
- 2) аннотация плотными рамками, в которой объекты задаются описывающими прямоугольниками, то есть, фактически, двумя точками в пространстве изображения,
- 3) аннотация зёрнами объектов, в которой объекты задаются одной точкой, примерно описывающей локацию.

Использование подобных типов аннотации, однако, невозможно без специальной доработки алгоритма обучения: меняется формулировка подзадач дискретной оптимизации (так называемых оракулов), которые необходимо решать на итерациях алгоритма обучения. В работе описаны эффективные алгоритмы дискретной оптимизации, решающие эти задачи.

В главе 3 рассмотрены другие обобщения структурного метода опорных векторов. В частности, рассмотрена более гибкая параметризация апостериорной вероятности случайного поля. С одной стороны, снимается требование ассоциативности потенциальных функций, которое обычно используется при вероятностном выводе, что приводит к увеличению числа настраиваемых параметров. С другой стороны, описана нелинейная вариация параметризации на основе ядерного перехода, позволяющая настроить более гибкую модель, и тем самым повысить точность в случае достаточного количества обучающих данных. Подобные модификации, среди прочего, позволяют учитывать сложные контекстуальные зависимости в данных. Использование метода продемонстрировано на примере задачи сегментации данных лазерного сканирования, где примером таких зависимостей, которые могут быть «выучены» моделью, являются пространственные отношения вроде «здание находится выше земли».

Идеи использования пространственного контекста были развиты автором работы в главе 4. Модели на основе структурного метода опорных векторов требуют значительных вычислительных ресурсов, как на этапе настройки модели, так и при предсказании. Последнее может оказаться критичным в некоторых приложениях. В работе предложен новый метод совместной разметки на основе последовательной классификации, главным достоинством которой является высокая скорость работы. Алгоритм вывода разметки работает итеративно,

уточняя разметку на каждой итерации. На вход последующей итерации передаётся часть выхода предыдущей. Конкретный вид таких зависимостей задаётся вручную, исходя из априорных знаний о предметной области. Особенностью предложенного метода является двухэтапная классификация, при которой на каждой итерации сначала применяются классификаторы, учитывающие предыдущую разметку для одного из видов зависимостей, а затем выходы классификаторов агрегируются. Виды зависимостей формализованы с помощью понятия типов факторов. Метод опробован на задаче сегментации карт глубины, полученных датчиком Kinect.

Новизна научных положений и выводов. В работе рассмотрено обобщение структурного метода опорных векторов на случай обучения с латентными переменными. Проведен теоретический анализ этого обобщения. Приведен алгоритм обучения модели на основе выпукло-вогнутой процедуры, известной из методов оптимизации [Yuille and Rangarajan, NIPS 2002]. Доказана теорема о соответствии этого алгоритма EM-алгоритму обучения параметров условного случайного поля, модифицированному специальным образом. Предложена новая методика определения эмпирических функций потерь для слабо-аннотированных экземпляров на основе оценки матожидания функции потерь для полностью размеченных экземпляров. Введены функции потерь для разных типов слабой аннотации в задаче семантической сегментации, доказаны теоремы о том, что они являются соответствующими оценками.

Был рассмотрен метод обучения структурного метода опорных векторов по несбалансированной выборке, доказана лемма о минимизации среднеклассовой полноты на обучающей выборке.

Предложен метод семантической сегментации на основе идеи последовательной классификации. Новыми являются понятие типа факторов и двухэтапная модель принятия решения на каждой из итераций. При таком подходе сначала считается распределение на метки класса, возвращаемое каждым из типов факторов, затем их результаты агрегируются специальным образом.

Обоснованность выводов подтверждается соответствием теоретических и экспериментальных результатов. Эксперименты описаны достаточно подробно и допускают воспроизводимость. В качестве материала используются стандартные наборы данных, доступные в сети Интернет. Методы протестированы на данных из различных областей: двумерных фотографиях, данных лазерного сканирования, картах глубины, а также в принципиально другой области — категоризации юридических текстовых документов.

Полученные результаты развивают теорию распознавания образов. Сведение выпукло-вогнутой процедуры обучения структурного метода опорных векторов с латентными переменными к ЕМ-алгоритму для случайных полей позволяет использовать известные результаты теоретического анализа ЕМ-алгоритма для разработанного метода. Предложенная методология определения эмпирических функций потерь может использоваться для применения обучения по слабоаннотированным данным в широком круге **прикладных задач** совместной разметки: в компьютерном зрении, анализе изображений и видео, анализе транспортных потоков, обработке естественных текстов — в задачах, где полная аннотация данных излишне трудоемка. Предложенный метод обучения последовательной классификации может применяться для решения задач разметки, в которых присутствуют систематические зависимости, например, пространственный контекст в визуальных сценах. Особенно полезен этот подход будет при решении задач, в которых важна скорость работы предсказывающего алгоритма, например, в системах реального времени, так как метод не нуждается в итеративной оптимизации сложной целевой функции, а лишь фиксированное число раз применяет решающее правило.

Результаты работы докладывались на ведущих конференциях в области распознавания образов, опубликованы в семи статьях в ведущих научных изданиях, из которых четыре — в изданиях, рекомендованных ВАК. Автореферат соответствует содержанию диссертации.

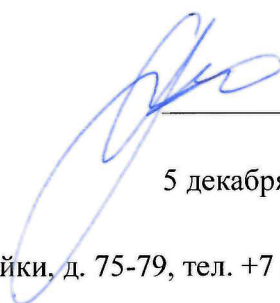
К **недостаткам** диссертации и автореферата можно отнести следующее.

1. Анализ результатов экспериментов использует средние значения характеристик точности на тестовой выборке, при этом мало внимания уделяется анализу конкретных экземпляров, на которых алгоритм допускает ошибки. В результатах главы 2 не приведен визуальный пример работы алгоритма на тестовых данных.
2. В работе сделан вывод о том, что при составлении обучающей выборки для решения задачи семантической сегментации изображений эффективнее всего с точки зрения трудозатрат использовать совместную аннотацию рамками объектов и метками изображений. При этом формально не определяются трудозатраты для каждого типа аннотации, и не приводится «стоимость» аннотации одного изображения.
3. В главе 4 предложена система семантической сегментации на основе последовательной классификации, в которой на каждой итерации классификатор применяется в две стадии: передачи сообщений и агрегации. При этом в работе не исследованы преимущества такой архитектуры по сравнению с алгоритмом «автоконтэкст» [Tu, CVPR 2008], который на каждой итерации использует один классификатор.

Отмеченные недостатки не снижают теоретической и практической ценности диссертационной работы.

Тема работы является актуальной и соответствующей паспорту специальности «Дискретная математика и математическая кибернетика». Работа соответствует критериям, установленным Положением ВАК о присуждении ученых степеней для диссертаций на соискание ученой степени кандидата наук, а Шаповалов Роман Викторович заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 01.01.09 — дискретная математика и математическая кибернетика.

Официальный оппонент,
Главный эксперт по концептуальному
инжинирингу — начальник департамента
геологии и разработки новых активов
ООО «Газпромнефть НТЦ», д.т.н.



О.С.Урмаев

5 декабря 2014 г.

Адрес: г. Санкт-Петербург, набережная реки Мойки, д. 75-79, тел. +7 (812) 313-69-24 (*3482).

Email: oushmaev@gmail.com

Подпись О.С. Урмаева заверяю.

Главный специалист Отдела кадрового администрирования
и развития персонала
ООО «Газпромнефть НТЦ»



Г. Н. Жемаева