

ОТЗЫВ

официального оппонента Лемпицкого Виктора Сергеевича
на **диссертацию** Шаповалова Романа Викторовича

«Методы структурного обучения в задачах совместной разметки»,

представленную на соискание учёной степени кандидата физико-математических наук по специальности 01.01.09 — дискретная математика и математическая кибернетика.

Актуальность темы диссертации

По мере развития методов распознавания образов и машинного обучения и проникновения этих методов в новые научные и технологические области, всё более актуальными становятся задачи предсказания со *структурированным результатом* (*structured output*), т.е. задачи предсказания, в которых каждому входному объекту присваивается выходной объект, обладающий нетривиальной структурой (примером служит задача попиксельной разметки изображений). Методы, решающие подобные задачи, обобщают стандартные методы распознавания образов с неструктурированным результатом (классификацию и регрессию), и являются востребованными в таких областях применений как компьютерное зрение, обработка естественного языка, распознавание речи, анализ интернет-данных, и т. п. Активная работа над подобными методами *структурного обучения* ведётся во многих исследовательских лабораториях в ведущих университетах, НИИ и индустрии.

Диссертация Шаповалова Р.В. посвящена развитию методов распознавания образов в задачах со структурированным результатом. Предложенные методы являются достаточно общими. В диссертации убедительно продемонстрирована применимость методов в ряде задач компьютерного зрения и в задаче, связанной с обработкой естественного языка. Актуальность темы диссертации, таким образом, не вызывает сомнения.

Структура работы

Работа включает в себя пять глав (в т. ч. вводную главу) и заключение. Вторая, третья и четвертая глава посвящены соответственно структурному обучению по слабоаннотированным тренировочным данным, структурному обучению неассоциативных марковских сетей и структурному обучению методом последовательной классификации. Список литературы включает 92 публикации.

Новизна научных положений и выводов диссертации

- В рамках второй главы проанализированы существующие методы структурного обучения по *слабоаннотированным* данным, в частности структурированный метод опорных векторов с латентными переменными (LV-SSVM). На основе анализа проведено обобщение LV-SSVM для наборов тренировочных данных, сочетающих полностью аннотированные и слабоаннотированные данные.

Метод LV-SSVM требует задания расширенных *алгоритмов вывода* в обучаемых моделях, позволяющих минимизировать функционалы *энергии*, дополненные функцией потерь, а также с ограничением на множество выходов, совместимых со слабой аннотацией. Основным научный вклад второй главы заключается в построении новых эффективных алгоритмов расширенного вывода для важных и широко распространенных классов слабых аннотаций, включающих (1) аннотацию изображений с помощью плотных и неплотных ограничивающих прямоугольников, (2) аннотацию изображений с помощью зерен, (3) аннотацию (*тегирование*) текстовых документов неполными наборами меток.

- В рамках третьей главы рассмотрены модификации LV-SSVM, включая *ядерную* версию этого метода, позволяющие обучать *неассоциативные марковские сети* для задач компьютерного зрения, в то время как подавляющее большинство предшествующих работ, связанных с обучением и применением марковских сетей в компьютерном зрении, накладывали требование ассоциативности. Немногие предшествующие работы, проводившие сравнение ассоциативных и неассоциативных марковских сетей в задачах разметки, приходили к выводу о неоправданности использования неассоциативных полей (поскольку не наблюдалось существенного прироста точности получаемой разметки или даже наблюдалось снижение точности, в то время как неассоциативность марковской сети влечет существенное осложнение задач вывода). Однако, автором показано, что при применении ядерной версии метода LV-SSVM и включении нелокальных (*дальнодействующих*) факторов неассоциативная марковская сеть достигает существенно большей точности разметки по сравнению с ассоциативной сетью, обученной одним из современных методов.

- В рамках четвертой главы предложен новый метод разметки данных трехмерного сканирования (*пространственная машина вывода*), существенно развивающий и улучшающий с т. з. получаемых результатов предшествующий метод (машины вывода). К главным новшествам метода относятся новый тип дальнодействующих факторов (т. н. *Д-факторы*), специализированных для задачи разметки трехмерных

данных, а также двухшаговый метод обучения. Предложенный метод относится к классу подходов, основанных на принципе последовательной классификации. Данные подходы характеризуются существенно упрощенными и ускоренными методами обучения и вывода (по сравнению с классическими моделями структурного предсказания, такими как марковские сети), и вместе с тем зачастую достигают большей точности по сравнению с классическими методами. Хотя практическая апробация метода проведена только для данных трехмерного сканирования, очевидна применимость аппарата пространственных машин вывода и к другим задачам разметки изображения (например, к задаче семантической разметке двумерных изображений).

Обоснованность и достоверность выводов

Обоснованность и достоверность выводов обусловлена грамотным использованием методов теории вероятностей, машинного обучения и распознавания образов, а также апробацией предложенных методов на реальных наборах данных разного типа, их экспериментальными сравнениями с ранее предложенными методами, решающими аналогичные задачи.

Практическая и теоретическая значимость полученных результатов

Полученные результаты имеют существенную практическую и теоретическую ценность. С практической точки зрения в диссертации убедительно продемонстрировано, что предложенные методы существенным образом повышают точность структурного предсказания в практически важных задачах разметки изображений и данных трехмерного сканирования, а также в задаче автоматической аннотации текстовых документов. Разработанные методы могут успешно применяться и в других областях, использующих методы распознавания со структурированным выходом.

Успешность использования слабоаннотированных данных для обучения структурированных моделей существенно повышает практическую привлекательность подобных моделей, поскольку получение детализированной (полной) разметки для наборов данных является чрезвычайно трудоемким.

Полученные результаты также свидетельствуют о существенном потенциале неассоциативных марковских сетей, ядерной версии структурного метода опорных векторов, а также методов структурного предсказания, основанных на последовательной классификации.

Полнота и уровень публикаций по теме диссертации

Результаты диссертации изложены в 7 публикациях, включающих российские журналы и рецензируемые труды международных конференций. Особенно хотелось бы отметить публикацию, излагающую результаты, связанные с пространственной машиной вывода, в трудах международной конференции по компьютерному зрению и распознаванию образов (CVPR) в 2013 г. Данная конференция характеризуется высочайшим уровнем (является самой «высокоимпактной» из всех конференций по инженерным и компьютерным наукам по данным системы Гугл Сколар), что обеспечивается строгим отбором поданных статей в условиях двойной анонимности. Работа Шаповалова Р. В. была отобрана для пленарного доклада, что для данной конференции происходит лишь для примерно 4% от работ, поданных на конференцию.

Недостатки работы

К недостаткам работы можно отнести:

1. Различие тестовых наборов данных, используемых для апробации методов, представленных в различных главах. В то же время, надо отметить, что подобное различие в большинстве случаев обуславливается различием протоколов и целей экспериментальных сравнений внутри соответствующих глав.
2. В главе 2 желательно более полное исследование зависимости точности разметки от применения различных схем инициализации оптимизационного процесса.
3. Поскольку существенная часть диссертации посвящена применениям разрабатываемых методов к данным трехмерного сканирования, обзор литературы следовало бы расширить с учетом работ, исследующих извлечение признаков из подобных данных (например, спиновых изображений/spin images).
4. В работе присутствует небольшое количество опечаток и грамматических несогласованностей, не влияющих существенным образом на восприятие (например, в определении 1.12 пропущена частица «не» в предложении после определяющей формулы, но сама формула задает верное определение).

В целом, отмеченные недостатки не снижают ценности и вклада диссертации.

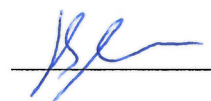
Заключение

Диссертация Шаповалова Р.В. выполнена автором самостоятельно и на высоком уровне. Полученные результаты представляют существенную практическую и теоретическую ценность, что доказывается успешной апробацией на реальных данных

и наличием публикаций высокого уровня. Автореферат соответствует содержанию диссертации. Тематика соответствует специальности 01.01.09 — дискретная математика и математическая кибернетика.

Диссертация является законченным исследованием и соответствует критериям, установленным Положением о присуждении учёных степеней, а соискатель заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 01.01.09 — дискретная математика и математическая кибернетика.

Официальный оппонент,
старший преподаватель Сколковского института
науки и технологий, к.ф.-м.н.
В.С. Лемпицкий
« 5 » декабря 2014 г.



143025, Московская область, Одинцовский район,
дер. Сколково, ул. Новая, д.100, тел. 8(495) 280-14-81 (доб. 3460).
Email: lempitsky@skoltech.ru

Подпись В.С. Лемпицкого заверяю,
Менеджер по административным
и кадровым вопросам
Л.Б. Коновалова

