

ОТЗЫВ

официального оппонента д.ф.-м.н. **Ю. В. Визильтера** на диссертационную работу **Осокина Антона Александровича «Субмодулярная релаксация в задаче минимизации энергии марковского случайного поля»**, представленную на соискание ученой степени кандидата физико-математических наук по специальности 01.01.09 - Дискретная математика и математическая кибернетика

Скрытые марковские модели в настоящее время активно используются при решении задач анализа данных различных типов. В частности, в компьютерном зрении при помощи таких моделей решаются задачи сегментации двумерных и трехмерных изображений (например, томограмм), стерео отождествления при реконструкции трехмерных сцен по множеству двумерных изображений, выделения и прослеживания движущихся объектов на видеопоследовательностях, аппроксимации трехмерных данных кусочно-аналитическими моделями и т.п. Такие марковские модели предполагают, что состояние наблюдаемых переменных определяется значениями некоторых скрытых переменных, описывающих структуру анализируемых данных. Например, в задаче сегментации изображений такими скрытыми переменными являются метки принадлежности пикселей тем или иным областям разбиения кадра, для каждой из которых характерны одинаковые яркость, цвет или текстурные характеристики. Задачей анализа изображения при этом является оценка значений скрытых переменных на основе максимизации апостериорной вероятности гиббсовского распределения, описывающего соответствующее марковское поле. Эта задача обычно формулируется в терминах минимизации энергии марковского поля (логарифма апостериорной вероятности), имеющей две основные составляющие: штраф за несоответствие наблюдений и модели, и штраф за сложность (неоднородность) самой модели. Последний штраф, как правило, выступает в качестве регуляризатора, обеспечивающего предпочтение более простых вариантов сегментации, с разбиением кадра на меньшее число более крупных областей. Различные функции штрафа позволяют также учитывать различные специфические требования к модели сегментации, например, гладкость контуров областей, размеры областей и др. Начиная с пионерской работы (Geman & Geman, 1984), данный подход к сегментации изображений, основанный на минимизации энергии марковских моделей, является чрезвычайно привлекательным для специалистов по компьютерному зрению, поскольку это регулярный способ получения

среднеуровневых описаний изображений в терминах выделяемых на них областей или объектов, учитывающий не только локальную, но и глобальную информацию о свойствах изображения. Однако решение задачи минимизации энергии на изображениях за исключением простейших случаев приводит к дорогостоящим и неэффективным процедурам итеративной оптимизации.

Первой такой приближенной вычислительной процедурой был предложенный в (Geman & Geman, 1984) имитационный отжиг (imitation annealing). Дальнейший прогресс был связан с нахождением способов точного решения задачи минимизации энергий на графах: для ациклических графов (методы динамического программирования или распространения сообщений) и парно-сепарабельных субмодулярных энергий на графах произвольной конфигурации. Вычислительно эффективная реализация последней идеи методом разреза графов (graph cut) была предложена в работах Ю. Бойкова и В. Колмогорова (Boykov, Kolmogorov, 2004), породивших как ряд известных практических приложений, так и целую волну современных исследований в данном направлении. К сожалению, за исключением специальных случаев работы с полутоновыми изображениями, в которых метки областей могут быть упорядочены по яркости – с использованием техники уровневых срезов (level sets), как в работе (Darbon, 2005) – методы разреза графов не могут быть использованы для получения оптимальных решений в случае многих классов (более двух меток областей), функций штрафа высоких порядков (зависящих от сочетаний более двух скрытых переменных) или функций штрафа с дополнительными ограничениями. В этих случаях по-прежнему приходится использовать громоздкие процедуры итеративной оптимизации различного типа, что явно сдерживает практическое внедрение данного класса методов в системы компьютерного зрения. Таким образом, тема диссертации А. А. Осокина, связанная с поиском эффективного метода решения таких задач, является безусловно актуальной как в теоретическом, так и в практическом плане.

Для решения указанной проблемы в работе предложены:

1. Оригинальный подход к решению задачи минимизации энергии на основе субмодулярной релаксации.
2. Доказательства эквивалентности разработанного подхода ряду существующих аналогов в случае парно-сепарабельных энергий и энергий с потенциалами высоких порядков специального вида.
3. Алгоритм покоординатного подъема для максимизации нижней оценки в случае ассоциативных парно-сепарабельных энергий.

Основная идея предложенного подхода заключается в «релаксации» (смягчении) исходной задачи минимизации энергии за счет замены точных ограничений, накладываемых на группы скрытых переменных, на их «мягкие» аналоги с весами в виде множителей Лагранжа. Оригинальность данного подхода связана с тем, что среди переменных, подлежащих оптимизации, оказываются теперь как дискретные переменные (скрытые переменные марковской модели), так и непрерывные переменные (множители Лагранжа). Это позволяет разделить оптимизационную процедуру на ряд парных шагов, первый из которых является вычислительно эффективной процедурой дискретной оптимизации методом разреза графа, а второй – процедурой непрерывной оптимизации. Идеи подобного смягчения задачи минимизации энергии предлагались и ранее – например, в методе двойственной декомпозиции (dual decomposition) в работе (Komodakis et al, 2011). Однако лишь в работах А.А. Осокина данный прием разворачивается в последовательный и регулярный подход к решению целого класса задач, который к тому же всесторонне исследуется как теоретически, так и экспериментально.

При теоретическом исследовании данного подхода в диссертации был разрешен ряд вопросов, возникающих при появлении нового семейства нижних оценок, основанных на релаксации Лагранжа. Указан точный вид задачи линейного программирования, решение которой эквивалентно наилучшей нижней оценке. Разработаны алгоритмы поиска наилучшей оценки в рамках предложенного семейства нижних оценок.

Также автором проведено экспериментальное исследование предложенных процедур оптимизации, демонстрирующее их значительное превосходство над известными аналогами по экономии вычислительных затрат. Проведены эксперименты на реальных данных, позволяющие сравнить работу различных методов оптимизации нижней оценки. Выполнены эксперименты по применению предложенного подхода к задачам с потенциалами высоких порядков, с неассоциативными парными потенциалами и глобальными ограничениями на переменные.

В качестве достоинств данной работы необходимо отметить отличный стиль изложения, строгость определений и четкость доказательств. Следует также обратить внимание на отличный уровень публикаций (в частности, три статьи в трудах ведущей мировой конференции CVPR с индексом цитируемости SJR=2,245).

В качестве недочетов по содержанию и оформлению данной работы можно указать следующие:

1) Описание проведенных экспериментов с изображениями в автореферате является слишком лаконичным и не позволяет судить об их содержании. К сожалению, и в тексте диссертации описание экспериментов не является детальным и исчерпывающим. Не

приводятся конкретный вид функций энергии и ряд других важных подробностей, которые представляют существенный интерес.

2) Эксперименты с реальными изображениями носят чисто качественный и демонстративный характер. Неясно, каким практическим задачам они могли бы соответствовать. В работе нет никаких численных критериев качества решения задач, кроме значений функции энергии. Между тем, например, для задач стереотождествления (рис.1 реферата) имеется четкий численный критерий точности стерео реконструкции.

3) Пример сегментации изображения (рис.5.3 диссертации), где демонстрируется найденный глобальный минимум энергии, показывает, что данная энергия не обеспечивает сегментации изображения, которая соответствовала бы интуитивным представлениям человека-наблюдателя о присутствующих на изображении объектах. В связи с этим целесообразно было бы хотя бы кратко рассмотреть вопрос о способе формирования функции энергии, соответствующей некоторым конкретным типам практических задач. Это позволило бы более точно определить круг возможных практических приложений предложенных методов.

Отмеченные недостатки не влияют на общую высокую оценку работы. Работа производит целостное впечатление, написана на хорошем математическом уровне, содержит четкую постановку задачи, доказательство ряда теоретических утверждений и описание достаточного количества численных экспериментов. Оригинальные результаты работы не имеют аналогов в известных оппоненту публикациях, т.е. являются новыми научными результатами. Содержание диссертации в достаточной степени опубликовано в научной печати, в том числе – в трудах ведущих международных конференций и статьях в рецензируемых научных журналах из списка ВАК. Результаты работы прошли широкую апробацию на ведущих российских и международных конференциях. Научная достоверность и обоснованность полученных результатов определяются корректностью применения используемых в работе математических методов, а также результатами экспериментов на модельных и реальных данных.

Автореферат достаточно полно и правильно отражает содержание диссертации. Тематика и содержание диссертационного исследования соответствуют специальности 01.01.09 - Дискретная математика и математическая кибернетика.

В целом данная диссертационная работа представляет собой законченное научное исследование, выполненное на высоком научном уровне на актуальную тему, содержит новые научные результаты, имеющие существенное значение для рассматриваемой предметной области, и отвечает всем требованиям, предъявляемым ВАК к кандидатским диссертациям, а ее автор, Осокин Антон Александрович, заслуживает присуждения ему учёной степени кандидата физико-математических наук по специальности 01.01.09 «Дискретная математика и математическая кибернетика».

**Официальный оппонент,
начальник подразделения “Системы
интеллектуального анализа данных,
технического зрения, улучшенного и
синтезированного видения”
ФГУП «ГосНИИАС», д.ф.-м.н.**

Визильтер Ю. В.

**Подпись Визильтера Ю. В. заверяю
Ученый секретарь ФГУП «ГосНИИАС», д.т.н.**



Мужичек С. М.