

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
на диссертацию Кузьмина А.И.
«Методы обучаемой регуляризации в задачах сопоставления
изображений», представленную на соискание ученой степени
кандидата технических наук по специальности 05.13.17 –
теоретические основы информатики.

Актуальность

В диссертации рассматривается актуальная задача: разработка алгоритма плотного сопоставления изображений, обладающего высокой точностью и допускающего обработку изображений с частотой 25+ кадров в секунду на современных вычислительных устройствах. Актуальность задачи подтверждается многочисленными публикациями в мире на эту тему.

Достоверность результатов

Достоверность представленных в диссертации А.И. Кузьмина результатов обусловлена корректностью примененного автором математического аппарата и выполненными численными экспериментами. Воспроизводимость результатов обусловлена открытостью доступа в сети интернет к коллекции изображений, использованных в экспериментах.

Достоверность результатов по сопоставлению ультразвуковых медицинских изображений обусловлена постановкой экспериментов на серии специально изготовленных изображений (фантомов) известной геометрии, что позволило количественно оценить качество сопоставления. Полученные результаты соответствуют результатам сопоставления на данных, использованных другими авторами (Rivaz, 2014).

Основные результаты и их новизна

В рамках диссертационной работы А.И. Кузьминым получены следующие новые результаты:

1. Предложен новый метод сопоставления изображений в приложении к бинокулярной стерео-реконструкции, основанный на комбинации сверточной и рекуррентной нейросети. Разработанная модель имеет меньшую вычислительную трудоемкость по сравнению с моделями, предложенными в литературе. Обучаемая схема агрегирования тензора энергий позволяет избежать трудоемкого сравнения большого количества визуальных дескрипторов, который является основным этапом большинства прочих методов стерео-сопоставления, основанных на глубоком обучении машин.

2. Разработан новый нейросетевой метод для задачи сопоставления изображений в приложении к задаче вычисления оптического потока. Предложенный метод основан на представлении итераций оптимизационного алгоритма в виде слоев сверточной нейросети, что позволяет получить метод, имеющий более низкую вычислительную трудоемкость по сравнению с методами, предложенными в литературе. Основой предложенной модели является обучение оператора регуляризации, что позволяет улучшить качество по сравнению с использованием фиксированного оператора регуляризации, применяемого в прочих методах.
3. Предложен новый метод сопоставления серий из трех ультразвуковых снимков в приложении к задаче ультразвуковой эластографии. Метод основан на локальном сопоставлении фрагментов изображения с последующим применением адаптивной регуляризации. Такой подход позволяет получить метод, устойчивый к участкам неверного сопоставления, при этом имеющий низкую вычислительную сложность. Предложенный алгоритм регуляризации основан на обобщении функционала полной вариации с учетом качества локального сопоставления. Такой функционал является выпуклым и позволяет применять эффективные алгоритмы двойственной минимизации.

Содержание работы

Диссертация состоит из введения, четырех глав и заключения.

Во введении обосновывается актуальность проводимых исследований, научная и практическая ценность работы, сформулированы цели и задачи, а также сформулированы основные положения, выносимые на защиту.

Вторая глава «Задача плотного сопоставления изображений» содержит постановку задачи сопоставления изображений, обзор вариантов задачи для различных приложений, а также описание областей применения.

Третья глава «Бинокулярное стерео-сопоставление» содержит описание метода стерео-сопоставления, разработанного автором. Глава начинается с введения в стерео-реконструкцию: приводится описание эпиполярной геометрии двух видов, излагаются особенности стерео-сопоставления на реальных данных. Также содержится обзор существующих методов на основе постановки задачи оптимизации. Главу продолжает обзор методов фильтрации изображений, использующих нахождение границ объектов. Затем приводится описание методов глубокого обучения машин для стерео-реконструкции.

Основная часть главы содержит описание метода, основанного на комбинации сверточной и рекуррентной нейросетей, предложенного

автором, а также численные эксперименты с использованием изображений дорожных сцен из открытых коллекций.

Четвертая глава содержит метод вычисления оптического потока, предложенный автором. В начале главы содержится обзор методов вычисления оптического потока, основанных на постановке оптимизационной задачи. Главу продолжает обзор методов, основанных на глубоком обучении машин. Приводится описание метода, предложенного автором, а также численные эксперименты.

Пятая глава содержит метод ультразвуковой эластографии, разработанный автором. Глава начинается с описания методов сопоставления ультразвуковых медицинских изображений. Далее приводится обзор методов эластографии по двум кадрам и многим кадрам. Описание метода, разработанного автором, сопровождается численными экспериментами с использованием данных, полученных с помощью численного моделирования, данных с использованием специально изготовленных фантомов известной геометрии, а также ультразвуковых снимков ткани человека. Главу завершает описание задачи сопоставления ультразвуковых сигналов, которая возникает в ультразвуковой отражательной томографии, а также описание метода реконструкции, предложенного автором.

Теоретическая значимость

Теоретическая значимость диссертации А.И. Кузьмина обусловлена предложенными в ней новыми архитектурами нейронных сетей с низкой вычислительной сложностью для решения задач сопоставления изображений.

Практическая значимость

Методы сопоставления изображений, предложенные А.И. Кузьминым, а также разработанный комплекс программ с применением графических ускорителей позволяют решать задачу сопоставления изображений в режиме реального времени для ряда практически важных задач, таких как стерео-реконструкция и анализ движения на дорожных сценах, обнаружение движения при видеонаблюдении, а также сопоставление медицинских ультразвуковых снимков в применении к эластографии.

Замечания

У диссертационной работы есть ряд недостатков, в основном связанных с некоторой небрежностью автора:

1. Текст диссертации требует корректуры. Неоднократно встречаются слова «афинное», «коэффициент», «порабола», «истинное»,

«предложная» (вместо «предложена»), «по сравннию», пропущены слова, некорректно используются деепричастные обороты. Повсеместно используются профессиональные жаргонизмы-англицизмы («сэмплы», «патчи»). Обозначения не всегда консистентны (эталонное поле диспаратета d_{gt} на стр. 34, но D_{gt} далее). Сиамская архитектура в работе почему-то пишется с заглавной буквы.

2. В некоторых местах диссертации текст также содержит и смысловые ошибки. На стр. 1 написано «аффинные преобразования, которые позволяют сопоставить изображения с учетом перспективных искажений», хотя перспективным искажениям соответствует другое подмножество проективных преобразований. На стр. 12 утверждается, что «Задача обнаружения движения ... называется вычислением оптического потока», хотя это две очень разные задачи. На стр. 110 «задана скорость света 1800 м/с», тогда как из контекста ясно, что речь идет о скорости звука.
3. Обязательные части диссертации оформлены с огрехами. В разделе «публикации» утверждается, что все пять печатных работ автора опубликованы «в изданиях, рекомендованных ВАК». Издания, рекомендуемые ВАК, являются журналами, а не сборниками конференций. Таких работ в списке – две из пяти. Цель работы следовало бы сформулировать конкретнее. В текущем виде для достижения цели достаточно запустить любой известный метод на суперкомпьютере. В самой работе раскрывается, что имеется в виду одна рабочая станция. Аналогично, везде, где ставится требование работы алгоритмов в реальном времени, не указывается класс вычислителя. Третье положение, выносимое на защиту, не фальсифицируемо по Попперу. Что такое эффективная параллелизация? Каков критерий?
4. В таблице 3.1 приводится оценка качества результатов работы предложенного метода стерео-сопоставления в чистом виде и с использованием пост-процессинга, в результате которого ошибка снижается. Далее автором предлагается метод апостериорной корректировки полученного поля диспаратета путем оценки присутствия заслоненных объектов. Все это интересным образом соотносится с утверждением в Заключение на стр. 114 об отсутствии необходимости постобработки. Что касается сложности

алгоритма, то автором утверждается, что его алгоритм требует $O(n * d_{max})$ операций, где n – число пикселей на изображении, d_{max} – максимальный диспаратет. Точнее было бы сказать, что сложность алгоритма – $O(n * d_{max} * k)$, где k – размер патча, используемого при вычислении *sensus transform*. Это важно, поскольку по утверждению автора «Трудоемкость наиболее быстрого метода, основанного на обучении машин, предложенного в литературе [9] оценивается как $O(n * d_{max} * k)$, где k – размерность глубоких дескрипторов» (т.е. оценка сложности та же), а одна из поставленных автором задач – «разработка новых методов обучения машин для сопоставления изображений, имеющих низкую вычислительную сложность на этапе исполнения». Кроме того, хотелось бы получить оценки сложности других этапов выполнения алгоритма для подтверждения утверждения о том, что вычисление тензора энергий является наиболее вычислительно сложным. Из таблицы 3.3 видно, что данный этап занимает порядка 8% от времени работы алгоритма, в то время как, например, рекурсивный фильтр ~ 56% и его сложность зависит от архитектуры сети.

5. Для ранжирования алгоритмов оценивания оптического потока помимо использованной автором евклидовой метрики обычно дополнительно используется величина угловой ошибки, более чувствительная к погрешностям алгоритма при малых величинах смещения. Но и для выбранной метрики автор приводит только качественные, примерные выводы: «Предложенный метод демонстрирует прирост производительности более чем в 2 раза по сравнению с моделью FlowNetS [3] при небольшом увеличении ошибки». Следовало бы привести точные цифры. Кроме того, окончательный вывод главы оказывается неожиданным для читателя после приведенного выше утверждения. Вывод гласит: «Такой подход позволяет улучшить качество оценки ... с незначительным увеличением вычислительной сложности».
6. В главе 5 отсутствует информация о кадровой частоте, достигнутой при реализации предложенных в ней методов.

Заключение

Работа вполне соответствует пунктам 5 и 7 паспорта специальности 05.13.17. Структура работы ясная, четко изложены практическая и научная значимость и новизна, присутствует адекватный обзор литературы. Работа

достаточно целостная: разработанный подход применен к решению трех различных задач (вычисление стереодиспаратности и оптического потока в анализе видеоизображений дорожных сцен и вычисление поля смещений в ультразвуковой эластографии). Поставленная задача успешно решена, что демонстрируется численным экспериментом, показавшим производительность порядка 29 кадров в секунду на изображениях разрешения 1242×375 пикселей. Часть результатов опубликована в авторитетном журнале *Computer Vision and Image Understanding*, апробация прошла на нескольких международных конференциях под патронажем IEEE.

Несмотря на приведенные выше замечания, высокий общий уровень работы не вызывает сомнения. Диссертация Кузьмина Андрея Игоревича «Методы обучаемой регуляризации в задачах плотного сопоставления изображений» является самостоятельной законченной научно-квалификационной работой. Диссертация обладает научно-теоретической и практической значимостью, ее результаты имеют существенное значение для развития направлений исследований в области теоретических основ информатики, а именно в исследовании методов машинного зрения и обработки изображений, в т.ч. ультразвуковых медицинских снимков.

Диссертация Кузьмина Андрея Игоревича «Методы обучаемой регуляризации в задачах плотного сопоставления изображений» отвечает всем требованиям пп. 9-14 «Положения о присуждении ученых степеней», предъявляемых к кандидатским диссертациям, а ее автор заслуживает присвоения ученой степени кандидата технических наук по специальности 05.13.17 – теоретические основы информатики.

Официальный оппонент

к.ф.-м.н. Николаев Дмитрий Петрович,
e-mail: dimonstr@iitp.ru,
тел: 8-916-146-24-87

Федеральное государственное бюджетное учреждение науки Институт проблем передачи информации им. А.А. Харкевича Российской академии наук (ИППИ РАН)

127051, г. Москва, Большой Каретный переулок, д.19 стр.1

И.о. заместителя директора ИППИ РАН по научной работе

