

Метод поиска несоответствий границ объектов между результатом 2D-3D конвертации и используемыми картами глубины

С.В. Долганов, Д.С. Ватолин

Аннотация—При создании полнометражных и любительских трехмерных фильмов с помощью конвертации отснятого в 2D материала достаточно часто возникают дефекты, связанные с качеством используемых карт глубины. Такого рода артефакты, даже если они находятся вне салиентных регионов, могут заметно ухудшить самочувствие зрителя, и, более того, вызвать головную боль. В данной статье предлагается метод поиска объектов переднего плана, границы которых в стереосцене не совпадают с действительностью, то есть объектов, «слитых» с фоном. Предлагаемый метод извлекает информацию о движении в сцене и находит несоответствия результата конвертации между извлеченным движением и глубиной сцены. Метод был применен на 39 полнометражных фильмах, что позволило найти 125 сцен с заметными несоответствиями границ объектов между картой движения и картой глубины, использованной при конвертации.

Ключевые слова—стерео видео, оценка качества, сопоставление стерео, 2D-3D конвертация, оценка глубины.

I. ВВЕДЕНИЕ

A. Стереоскопическое видео и конвертация

В наше время практически в каждом кинотеатре мира обязательно найдется сеанс в 3D. Такой интерес порождает повышенный спрос на стереоконтент, однако развитие современных технологий производства 3D фильмов не способно в полной мере ему отвечать. Стандартный процесс создания контента происходит тремя возможными путями: съемка с помощью дорогостоящих стереоскопических систем камер, конвертация заранее отснятого двухмерного, использование компьютерной графики.

Статья получена 18 декабря 2017.

Станислав Викторович Долганов, МГУ им. М.В. Ломоносова, (email: stanislav.dolganov@graphics.cs.msu.ru).

Дмитрий Сергеевич Ватолин, МГУ им. М.В. Ломоносова, (email: dmitriy@graphics.cs.msu.ru).

Данное научное исследование проводится при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере по программе «УМНИК» по теме «Разработка программного инструмента для контроля качества работ по конвертации видео в стерео формат» в рамках договора №09143Г/2015 от 24.12.2015г.

Последний подход относится к специфичной области мультипликационных фильмов, а также позволяет создавать различные спецэффекты на этапе постобработки. Съемка видео с помощью стереокамер требует постоянной калибровки цветowych и геометрических параметров камер системы, что является нетривиальным процессом и потому зачастую приводит к появлению некачественного контента. Учитывая последние разработки в области стереоконвертации [1], [2], которые заметно улучшили визуальное качество результата, а также сравнительно меньшие затраты на производство стереофильма с помощью конвертации, легко заметить, что многие киностудии предпочитают конвертацию остальным подходам. В данном утверждении можно убедиться, если посмотреть на ежегодные соотношения конвертированных и отснятых фильмов за последние 60 лет [3].

B. Проблемы конвертации

Хотя существующие программные инструменты на порядок облегчают различные этапы конвертации видео, сам процесс остается достаточно трудоемким и далеким от автоматизации. Майк Сеймор в своем обзоре [4] описывает основные проблемы генерации стереовидео по исходному 2D материалу: эффекты параллакса, кулисности, отжимающего действия рамки, агрессивная величина параллакса, некачественная карта глубины. Общее качество карты глубины, которая является необходимой частью процесса конвертации, прямым образом влияет на результат, поэтому очень важно контролировать соответствие карты действительности, что касается как верного выбора расположения объектов по глубине в сцене, так и точных границ объектов.

Эффект кулисности и остальные дефекты карт глубины, которые возникают в процессе конвертации из 2D в 3D часто ухудшают общее восприятие зрителя от просмотра фильма. Наиболее значимые проблемы создают объекты, находящиеся в салиентных областях сцены, границы по движению которых не совпадают с границами по глубине, то есть таких объектов, которые

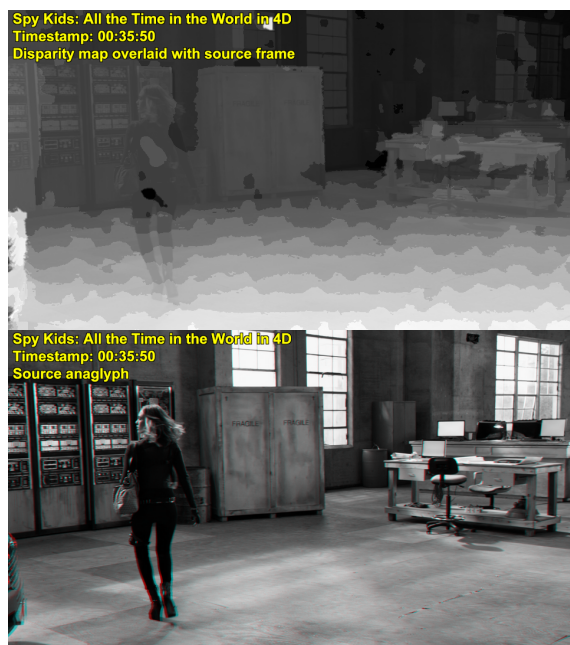


Рис. 1. Пример найденной с помощью предлагаемого метода сцены, содержащей дефект: объект переднего плана оказался слитым с фоном. Тело актрисы не было нарисовано на карте глубины, поэтому сцена демонстрирует невозможную ситуацию и может вызвать визуальный дискомфорт. На верхнем изображении показана карта диспаратности поверх исходного ракурса, на нижнем — исходная сцена в анаглифе.

движутся по пространству сцены не равномерно или деформируются, что невозможно в реальной жизни. Исследования в этой области [5], [6] подтверждают ухудшение состояния группы людей, которые смотрят такого рода стереовидео.

С. Полуавтоматический контроль качества конвертации

Основной целью данной работы является создание программного инструмента, который позволит автоматически оценить качество используемых карт глубины, что в свою очередь позволит киностудиям создавать более качественный видеоконтент. Частой практикой киностудии является разделение фильма на крупные части, которые затем отправляются на конвертацию в специализированные компании. Такой подход ухудшает качество, так как компании не знают о результатах работы друг друга. Основной выигрывш для киностудии — это уменьшение времени, которое понадобится для конвертации, однако из-за сжатых сроков, а так же ручного тестирования качества результата, часто возникают такие сцены, как на Рис. 1. Предлагаемый программный инструмент позволит проводить проверку в полуавтоматическом режиме, что несомненно повысит как качество, так и время производства. В качестве входных данных предлагаемому алгоритму требуются стереовидеопоследовательность и опционально карты глубины, используемые во время конвертации в 3D формат данной видеопоследовательности.

В случаях, когда информация об используемых картах глубины недоступна, метод оценивает карту диспаратности с помощью [7], [8]. Он находит дефекты, сравнивая для каждого отдельного кадра границы карты глубины/диспаратности с границами карты амплитуды движения. Такой подход позволяет находить движущиеся объекты, которые частично или полностью отсутствуют на оцениваемой карте глубины. Так как не существует общепринятого быстрого и качественного алгоритма оценки векторов движения, было принято решение использовать в качестве такого алгоритма [7], который применяется к последовательности левых ракурсов. Затем полученная последовательность оцененных карт амплитуды движения итеративно улучшается с помощью временных и пространственных фильтров [9], [10], [11]. Результирующая оценка для каждого кадра высчитывается, как величина несоответствия глубины и движения. Оценка для сцены высчитывается, как средневзвешенная оценка соответствующих кадров, где веса высчитываются на основании доверия к оцененным векторам движения.

II. ОБЗОР ПРЕДМЕТНОЙ ОБЛАСТИ

А. Методы оценки качества стереоскопического видео

Основными направлениями исследований в области оценки качества стереоскопического видео являются множественные несоответствия между ракурсами и временная стабильность динамического диапазона глубины. Voronov в соавторстве [12] предложил способ нахождения разницы между ракурсами по цвету, резкости, а также различных геометрических искажений. В статье [13] авторами был предложен метод нахождения перепутанных ракурсов в стереовидео.

Артефакты, влияющие на качество конвертированного стереоскопического видео, более разнообразны и менее изучены в научных кругах. Vokov в соавторстве [14] предложил алгоритм поиска трех основных проблем, возникающих при конвертации: резкость границ объектов, эффект кулисности, плоские сцены. Ожидается, данный алгоритм подходит для обнаружения проблем только для тех объектов, которые присутствуют на карте глубины. В данной статье предлагается дополнительный метод, расширяющий возможности оценки качества конвертированного стереовидео, позволяющий находить объекты переднего плана, частично или полностью отсутствующие на используемых картах глубины.

В. Обзор базовых алгоритмов

Предлагаемый метод является композицией двух крупных областей исследований: извлечение структуры сцены на основании информации о движении, а также алгоритмов сопоставления краев и

контуров между двумя картами границ. Обе области достаточно хорошо изучены.

Многие методы извлечения структуры сцены из движения предполагают минимизацию функции энергии с членом, отвечающим амплитуде движения, извлеченной из сцены. Яркий представитель [15] дополнительно учитывает ограничение, обеспечивающее временную стабильность, а также уточняет положение объектов с помощью ротации в пространстве плоскостей, соответствующих цветовым сегментам кадра, в сторону уменьшения энергии. Такой подход позволяет извлекать карты глубины высокого качества, обладающие временной стабильностью. Однако временная сложность алгоритма не подходит для решения прикладных задач, в частности для анализа полнометражного стереоскопического фильма. Недавно предложенный метод фильтрации [11], сохраняющий края объектов, позволил положить в основу разрабатываемого метода менее вычислительно сложный подход: изначально извлекается сырая карта амплитуды движения с помощью блочного алгоритма оценки векторов движения [7], а затем улучшается путем свертки фильтром, сохраняющим границы объектов.

Для решения задачи сопоставления контуров, было рассмотрено несколько современных точных алгоритмов сопоставления замкнутых контуров. Например, Хи в соавторстве [16] предложил метод, основанный на дескрипторе гибкости контуров. Расширение на случай не замкнутых контуров возможно, однако потребует использования не самого точного алгоритма водоразделов [17], что повлечет значительное увеличение ошибок второго рода. Другой подход [18] заключается в использовании особой структуры — иерархического дерева форм. К сожалению, данный метод не расширяется на разрывные контура, что критично в нашем случае, так как оценка карт амплитуды движения не является идеальной и содержит в том числе большое количество контуров, не являющихся непрерывными и замкнутыми. Таким образом, в ходе анализа предметной области было решено использовать менее точный, но более устойчивый к качеству анализируемых контуров метод, схожий с distance transform [19].

III. ПРЕДЛОЖЕННЫЙ МЕТОД

A. Краткое описание метода

На Рис. 2 изображена схема предлагаемого метода поиска несоответствий между движением объекта и картой глубины, используемой во время конвертации видео из 2D в 3D. Входными данными метода являются стереоскопическое видео, конвертированное из 2D, и, опционально, карта глубины, участвующая в процессе конвертации. Результатом работы алгоритма, реализующего метод, является покадровая оценка, которую можно интерпретировать как периметр

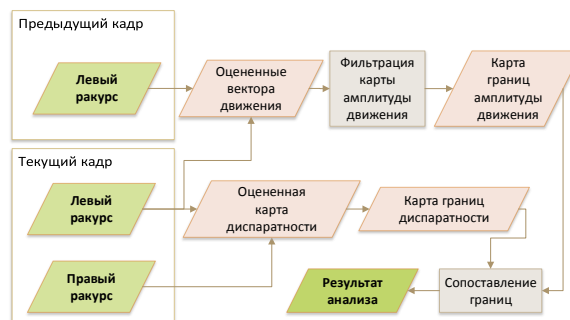


Рис. 2. Схема предлагаемого метода. Шаг оценки карты диспаратности может быть опущен в случае, когда доступна оригинальная карта глубины.

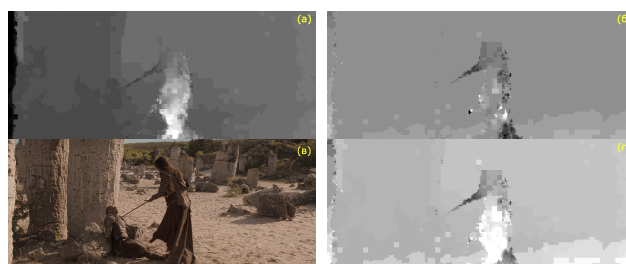


Рис. 3. Пример карт горизонтального (а) и вертикального (б) движения, извлеченные с помощью блочного алгоритма оценки движения [7] для левого ракурса исходного кадра (в). Результаты объединяются в одну общую карту амплитуды движения (г).

объектов, отсутствующих на карте глубины. Метод состоит из следующих шагов:

- 1) Оценить вектора движения левого ракурса между предыдущим и текущим кадром.
- 2) Оценить карту диспаратности между левым и правым ракурсом (в случае недоступности карты глубины). Использование карты диспаратности вместо карты глубины допустимо, так методу требуется знание только о границах объектов по глубине, которые совпадают с границами на карте диспаратности.
- 3) Рассчитать карты амплитуды движения, используя поле векторов движения, улучшенное с помощью временной и пространственной фильтрации для достижения временной целостности.
- 4) Провести сопоставление границ между картами амплитуды движения и глубины.
- 5) Рассчитать площадь границ карты амплитуды движения, которые отсутствуют на карте глубины. Эта площадь и будет являться финальной оценкой.

B. Карта амплитуды движения

Для извлечения как карты движения, так и карты диспаратности, применяется блочный подход, описанный в [7]. Для оценки значения доверия к получившимся картам используется ограничение left/right-consistency (LRC) [20] (которое, в случае



Рис. 4. Конвейер пространственно-временных фильтров. Такой подход позволяет значительно улучшить грубую карту интенсивности движения, которую выдает блочный алгоритм оценки векторов движения.

оценки движения, означает разницу между исходным пикселем и пикселем после перехода по вектору движения вперед и назад по вектору, построенному из прошлого в текущий кадр.

Элементы карты амплитуды движения M высчитываются, как величины в поле векторов движения (см. Рис. 3 (а) и (б)). Затем, для увеличения временной стабильности полученной карты интенсивности, применяются последовательно временные и пространственные фильтрации, составляющие конвейер алгоритмов, изображенный на Рис. 4.

Процесс улучшения начинается с применения линейного преобразования к карте амплитуды движения; данное преобразование выставляет минимальное значение элементов в карте равным 0, а максимальное — 1 (также известное, как автоматическая тоновая коррекция). Рис. 3 (г) демонстрирует промежуточный результат для этого шага. Из-за непостоянства скоростей движения камеры и объектов в сцене, а также из-за ошибок в поле векторов движения, результирующие карты интенсивности заметно мерцают. Для сокращения влияния данного эффекта, происходит выравнивание среднего и дисперсии интенсивности каждого кадра на величину, равную соответствующим значениям, рассчитанным за последние n кадров (в экспериментах использовалось $n = 10$):

$$M_i'^k = \frac{\text{Std}_{k-n}^{k-1}}{\text{Std}[M^k]} (M_i^k - E[M^k]) + E_{k-n}^{k-1} \quad (1)$$

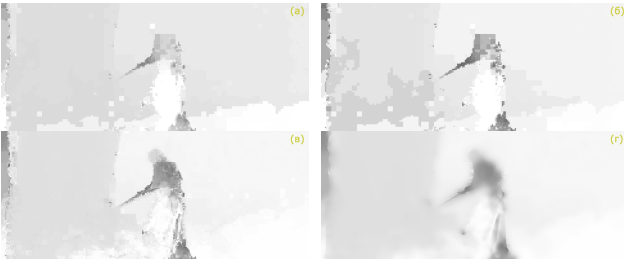


Рис. 5. Примеры, иллюстрирующие каждый промежуточный этап конвейера пространственно-временных фильтров из Рис. 4. Изображены результаты после стабилизации среднего и дисперсии (а), сопоставления гистограмм (б), временной фильтрации (в) и применения Guided filter (г).

Здесь, E и Std — математическое ожидание и стандартное отклонение, соответственно; M_i^k — пиксель i карты амплитуды движения для кадра k ; E_{k-n}^{k-1} и Std_{k-n}^{k-1} — математическое ожидание и стандартное отклонение во времени, соответственно. Эти значения получаются по следующим формулам:

$$E_{k-n}^{k-1} = \frac{1}{\#} \sum_i \left(\frac{1}{n} \sum_{p=k-n}^{k-1} M_i^p \right) \quad (2)$$

$$\text{Std}_{k-n}^{k-1} = \sqrt{\frac{1}{\#-1} \sum_i \left(\frac{1}{n} \sum_{p=k-n}^{k-1} M_i^p - E_{k-n}^{k-1} \right)^2} \quad (3)$$

В равенствах (2) и (3) $\#$ обозначает число пикселей в кадре. На Рис. 5 (а) изображен промежуточный результат после этого шага. В качестве глобального преобразования применяется сопоставление гистограмм [9] к последовательности карт амплитуды движения. Данное преобразование сдвигает гистограмму для каждого кадра так, что она соответствует средней гистограмме для предыдущих n кадров. Пример можно увидеть на Рис. 5 (б).

Для того, чтобы убрать локальные мерцания, возникающие после глобальных преобразований, используется метод временной стабилизации, описанный в [10] (см. пример на Рис. 5 (в)).

На последнем шаге для согласования границ карты амплитуды движения с границами исходного ракурса кадра, а также для заполнения областей с низким доверием, используется guided filter [11], на вход которому в качестве направляющего изображения передается исходный ракурс. Рис. 5 (г) демонстрирует пример финальной карты амплитуды движения.

С. Карта диспаратности

Существуют ситуации, когда нет возможности оценить используемые карты глубины напрямую. В таких случаях оценивается карта диспаратности с помощью блочного алгоритма сопоставления. Для увеличения качества карты используется быстрая версия фильтрации взвешенной медианы, описанной в [8]. Данная фильтрация улучшает контуры объектов и устраняет посторонние границы в местах, которые отвечают плавным переходам по глубине.

Пример оцененной карты перед и после применения фильтрации можно найти на Рис. 6 (а) и (б).

Д. Сопоставление границ

Для извлечения границ применяется оператор Щарра [21] к карте амплитуды движения после фильтрации и к карте глубины. Результирующие карты границ изображены на Рис. 7 (а) и (б). Так как карта интенсивности движения содержит размытые границы, являющиеся следствием применения guided filter, используется информация с исходного RGB кадра для их уточнения. Однако простое

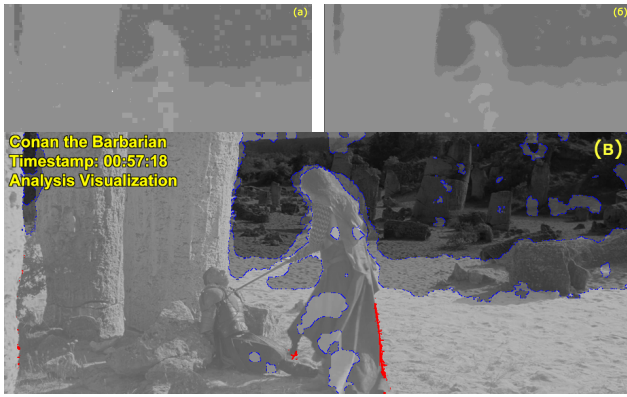


Рис. 6. Примеры карт диспаратности перед (а) и после (б) фильтрации взвешенной медианы. Итоговая визуализация (в), которая предлагается в качестве одного из выходов алгоритма, содержит карту диспаратности поверх исходного кадра. Синий цвет отмечает границы диспаратности, а красный — несоответствия между границами карт амплитуды движения и диспаратности.

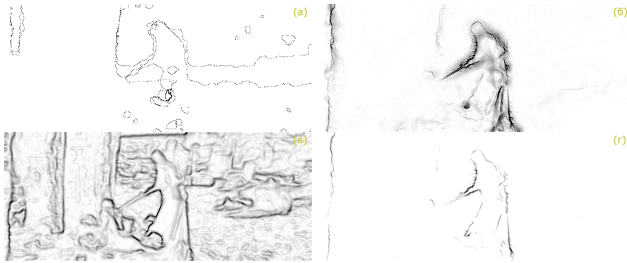


Рис. 7. Границы, извлеченные из карты диспаратности (а), улучшенной карты амплитуды движения (б), и исходного кадра после применения фильтрации, сохраняющей контуры объектов (в). Пересечение границ карты интенсивности движения и исходного кадра изображено на (г).

пересечение не даст должной точности результата, так как исходный кадр содержит множество посторонних текстурных границ. Для устранения большей части «шумных» границ применяется фильтрация с учетом контуров объектов [22] к исходному кадру, а границы извлекаются из уже отфильтрованного изображения (см. Рис. 7 (г)).

В конце вычисляется пересечение извлеченных границ из карты амплитуды движения (M) и исходного кадра после фильтрации (I) следующим образом:

$$M_i'^k = M_i^k \max_{j \in B_r(i)} (I_j^k) \quad (4)$$

Здесь, $B_r(i)$ — это окрестность радиуса $r = 0$ вокруг пикселя i . Для улучшения границ карты диспаратности используется схожая процедура с $r = 5$.

Для получения итоговой оценки качества производится пересечение границ улучшенной карты амплитуды движения с границами карты диспаратности методом, основанным на алгоритме distance transform. Для каждого пикселя границы в карте интенсивности движения итеративно

вычисляется значение расстояния:

$$M_i^{k,0} = M_i^k, \quad M_i^{k,t} = M_i^{k,t-1} - \max_{j \in B_t(i)} w(i,j) \frac{(\vec{M}_i^{k,t-1}, \vec{D}_j)}{\|\vec{D}_j\|} \quad (5)$$

В этом выражении $w(i,j)$ — функция веса, зависящая от расстояния:

$$w(i,j) = \exp\left(-\frac{\|i-j\|^2}{2\sigma^2}\right) \quad (6)$$

Здесь, $\vec{M}_i^{k,t-1}$ и $\vec{D}_j$ — карты границ амплитуды движения и диспаратности, извлеченные с помощью оператора Щарра. Итоговая карта содержит только те границы, которые присутствуют на карте движения, но отсутствуют на карте глубины. Такие границы, вероятно, принадлежат объектам, частично или полностью отсутствующим на карте глубины. Пример итогового результат можно найти на Рис. 6 (в) (потерянные границы по мнению алгоритма, изображены красными линиями).

IV. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА

A. Валидация алгоритма

Фильм	Истинно положительные	Ложно положительные
Clash of the Titans	5	5
Conan the Barbarian	4	12
Star Trek Into Darkness	1	7
Harry Potter and the Deathly Hallows: Part 2	4	7
The Avengers	7	0

Таблица I

Количество истинно и ложно положительных срабатываний начальной версии алгоритма на валидационном множестве.

Алгоритм, реализующий метод, описанный в предыдущем разделе, был использован для определения отношения ложных и истинных срабатываний на множестве из 5 фильмов. Для создания валидационного множества, сцены, на которых алгоритм выдал высокие значения, были вручную классифицированы как истинно и ложно положительные. Итоговый набор сцен состоял из 21 истинно положительной сцены и 31 ложно положительной (см. Таблицу I).

Как показала классификация, основным источником ошибок второго рода оказались объекты, находящиеся перед монотонным фоном (например, небо). Так как блочный алгоритм оценки диспаратности выдает неточности в малотекстурных областях, предлагаемый метод не способен найти разницу между диспаратностью объекта переднего плана и фона, даже если объект присутствует на карте

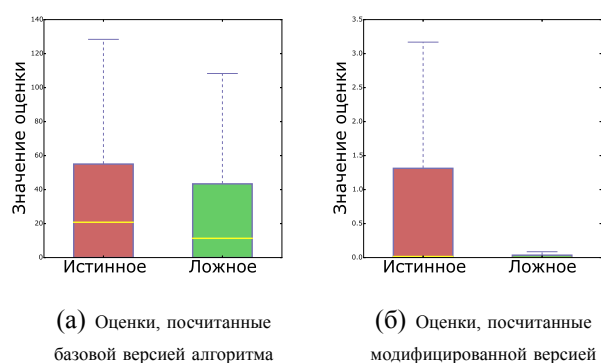


Рис. 8. Среднее оценок начальной версии алгоритма (а) для истинно и ложно положительных сцен, а также оценки посчитанные для тех же сцен с помощью модифицированной версии алгоритма, исключаяющей однотонные регионы (б).

глубины. Более того, зрительная система человека терпимо относится к отсутствию различия по глубине в малотекстурированных областях. Таким образом, разумным подходом сокращения количества ложных срабатываний является исключение из рассмотрения алгоритма любых объектов, возникающих перед монотонным фоном. Это достигается с помощью добавления весового множителя к значениям итоговой карты пересечения, который будет равен цветовой дисперсии в некоторой окрестности для каждого пикселя исходного кадра. Рис. 8 демонстрирует сравнение оценок, вычисленных обычным подходом и модифицированным для сцен из валидационного множества. Модифицированной версии алгоритма удастся отличить сцены с артефактами от сцен, ошибочно оцененных простой версией.

В. Анализ полнометражных фильмов

После модификации алгоритм был запущен на 39 стереоскопических полнометражных фильмах с конвертированным стерео. В итоге было найдено 125 проблемных сцен. На Рис. 9 изображен график распределения количества сцен, найденных в каждом из анализируемых фильмов.

С. Программная реализация

Среднее время работы предложенного метода анализа при обработке видеопоследовательности с разрешением 960×540 около 3.5 секунд на кадр на машине с характеристиками: 2.67 GHz Intel Core i7, 24 GB of RAM.

У. ЗАКЛЮЧЕНИЕ

А. Результаты работы

В данной статье был предложен метод оценки качества карт глубины, используемых при конвертации стереоскопического видео. Метод находит объекты переднего плана, которые оказались «слиты» с фоном на карте глубины, что позволит создателям

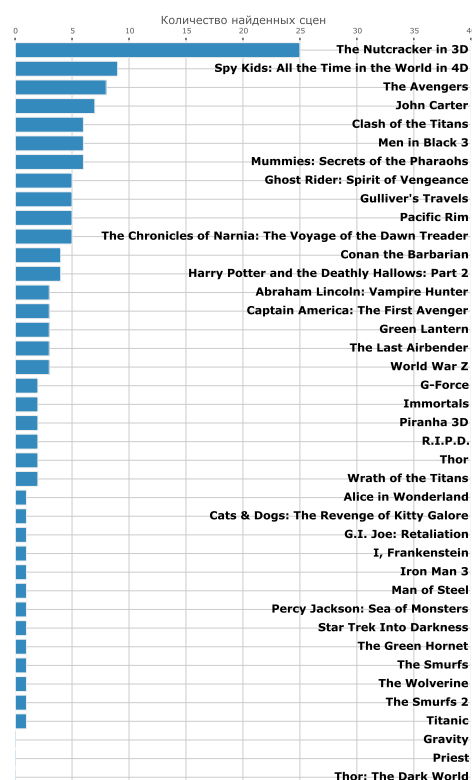


Рис. 9. Количество сцен с заметными артефактами, найденных в процессе анализа полнометражных фильмов.

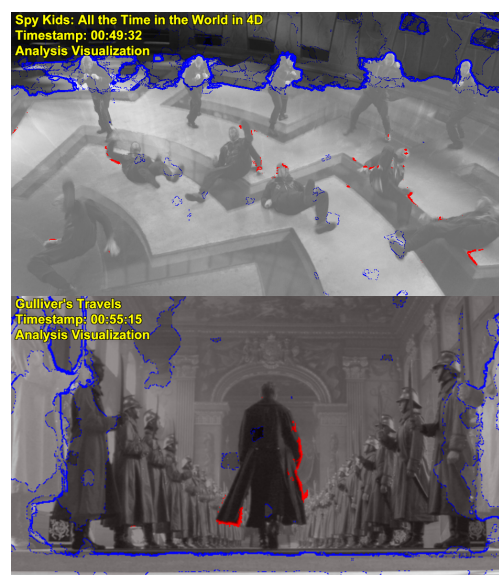


Рис. 10. Пример результатов срабатывания алгоритма. В данных сценах объекты переднего плана «слиты» с фоном. Для наглядности поверх исходного кадра изображены карта диспаратности, границы карты диспаратности (синие), и границы объектов, которые алгоритм посчитал потерянными (красные).

конвертированных фильмов найти и исправить данную проблему.

Качество метода было продемонстрировано с помощью валидационного множества сцен, состоящего из 5 фильмов. Количество ложных

срабатываний было уменьшено с помощью исключения из анализа малотекстурированных областей. Использование быстрого извлечения глубины из движения позволило проанализировать 39 полнометражных конвертированных стереоскопических фильмов и найти 125 сцен со значительными артефактами, которые могут вызвать визуальный дискомфорт зрителя. Результаты могут быть найдены по ссылке [23].

Список литературы

- [1] Depth image-based rendering with advanced texture synthesis for 3-d video / P. Ndjiki-Nya, M. Köppel, D. Doshkov et al. // *Multimedia, IEEE Transactions on*. — 2011. — Vol. 13, no. 3. — P. 453–465.
- [2] Tolstaya E., Pohl P., Rychagov M. Depth propagation for semi-automatic 2d to 3d conversion // *IS&T/SPIE Electronic Imaging / International Society for Optics and Photonics*. — 2015. — P. 939303–939303.
- [3] Is it Real or Fake 3D? — <http://www.realorfake3d.com>.
- [4] Seymour M. Art of stereo conversion: 2D to 3D — 2012. — <http://www.fxguide.com/featured/art-of-stereo-conversion-2d-to-3d-2012/>. — 2012.
- [5] Visual comfort assessment metric based on salient object motion information in stereoscopic video / Y. J. Jung, S. Lee, H. Sohn et al. // *Journal of Electronic Imaging*. — 2012. — Vol. 21, no. 1. — P. 011008–1.
- [6] Li J., Barkowsky M., Callet P. Le. Visual discomfort of stereoscopic 3d videos: Influence of 3d motion // *Displays*. — 2014. — Vol. 35, no. 1. — P. 49–57.
- [7] Fast video super-resolution via classification / K. Simonyan, S. Grishin, D. Vatolin, D. Popov // *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on / IEEE*. — 2008. — P. 349–352.
- [8] Zhang Q., Xu L., Jia J. 100+ times faster weighted median filter (wmf) // *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on / IEEE*. — 2014. — P. 2830–2837.
- [9] Fecker U., Barkowsky M., Kaup A. Time-constant histogram matching for colour compensation of multi-view video sequences // *Proc. 26th Picture Coding Symp.(PCS 2007)*. — 2007.
- [10] Temporal filtering for depth maps generated by kinect depth camera / S. Matyunin, D. Vatolin, Y. Berdnikov, M. Smirnov // *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011 / IEEE*. — 2011. — P. 1–4.
- [11] He K., Sun J., Tang X. Guided image filtering // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. — 2013. — Vol. 35, no. 6. — P. 1397–1409.
- [12] Towards automatic stereo-video quality assessment and detection of color and sharpness mismatch / A. Voronov, D. Vatolin, D. Sumin et al. // *3D Imaging (IC3D), 2012 International Conference on / IEEE*. — 2012. — P. 1–6.
- [13] Automatic left-right channel swap detection / D. Akimov, A. Shestov, A. Voronov, D. Vatolin // *3D Imaging (IC3D), 2012 International Conference on / IEEE*. — 2012. — P. 1–6.
- [14] Automatic detection of artifacts in converted s3d video / A. Bokov, D. Vatolin, A. Zachesov et al. // *IS&T/SPIE Electronic Imaging / International Society for Optics and Photonics*. — 2014. — P. 901112–901112.
- [15] Consistent depth maps recovery from a video sequence / G. Zhang, J. Jia, T. Wong, H. Bao // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. — 2009. — Vol. 31, no. 6. — P. 974–988.
- [16] Xu C., Liu J., Tang X. 2d shape matching by contour flexibility // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. — 2009. — Vol. 31, no. 1. — P. 180–186.
- [17] Roerdink Jos BTM, Meijster Arnold. The watershed transform: Definitions, algorithms and parallelization strategies // *Fundamenta informaticae*. — 2000. — Vol. 41, no. 1, 2. — P. 187–228.
- [18] Felzenszwalb P.F., Schwartz J.D. Hierarchical matching of deformable shapes // *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on / IEEE*. — 2007. — P. 1–8.
- [19] Borgefors G. Distance transformations in digital images // *Computer vision, graphics, and image processing*. — 1986. — Vol. 34, no. 3. — P. 344–371.
- [20] Egnal G., Mintz M., Wildes R. P. A stereo confidence metric using single view imagery with comparison to five alternative approaches // *Image and vision computing*. — 2004. — Vol. 22, no. 12. — P. 943–957.
- [21] Jähne B., Scharf H., Körkel S. Principles of filter design // *Handbook of computer vision and applications*. — 1999. — Vol. 2. — P. 125–151.
- [22] Rolling guidance filter / Q. Zhang, X. Shen, L. Xu, J. Jia // *Computer Vision-ECCV 2014*. — Springer, 2014. — P. 815–830.
- [23] <http://compression.ru/download/DFMResults.min/films.html>.

Object boundaries inconsistencies detection method for 2D-3D conversion results and depth maps

Stanislav Dolganov, Dmitriy Vatolin

Abstract—The creation of S3D movies by converting 2D captured footage often introduces depth-map inaccuracies. Such artifacts can significantly degrade the viewing experience even if they occur only in unsalient background objects. In this paper we propose a method for detecting foreground objects that are stuck to the background. Our method extracts information about motion in the scene and detects conversion-related discrepancies between motion strength and depth. We demonstrate the performance of the method by applying it to 39 full-length converted 3D movies and by providing the results of our analysis as well as examples of detected problem shots.

Keywords—stereoscopic video, quality assessment, stereo matching, 2D-to-3D conversion, depth estimation