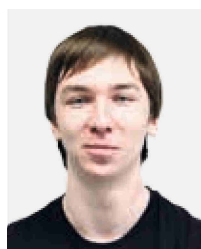


ОБНАРУЖЕНИЕ ВРЕМЕННОГО СДВИГА МЕЖДУ РАКУРСАМИ И ПЕРЕПУТАННОГО ПОРЯДКА РАКУРСОВ в стереофильмах

(7-й отчёт проекта MSU VQMT3D)



В.А. Людвиченко, С.В. Лаврушкин, В.А. Янушковский, Д.С. Ватолин, к.т.н., МГУ им. М.В. Ломоносова, dmitriy@graphics.cs.msu.ru

Аннотация

Сегодня многие зрители испытывают дискомфорт и головную боль при просмотре 3D фильмов. Основной причиной возникающего зрительного дискомфорта является наличие различных артефактов в стереофильме.

В работе предложены методы автоматического обнаружения двух типов артефактов в стереовидео: временного сдвига между ракурсами и перепутанного порядка ракурсов. Так как эти артефакты сложно обнаружить вручную, но легко исправить, то предложенные методы позволяют практически исключить возможность их появления в фильмах. Предложенный метод обнаружения перепутанного порядка ракурсов был апробирован на 105 современных фильмах, а метод обнаружения временного сдвига – на 26. Наиболее значительные артефакты, обнаруженные в этих фильмах, приведены в статье. Метод обнаружения перепутанных ракурсов имеет на 17% большую точность (по метрике AUROC) и в 3,5 раза меньшую вычислительную сложность, чем ранее предложенный подход.

Ключевые слова: стереовидео, анализ изображений, временной сдвиг, перепутанный порядок ракурсов.

DETECTION OF TEMPORAL SHIFT AND SWAPPED VIEWS IN S3D MOVIES

V. Lyudvichenko, S. Lavrushkin, V. Yanushkovskiy, D. Vatin

Abstract

Many people experience the visual discomfort and a headache caused by watching S3D movies. The main cause of the visual discomfort is various artifacts in S3D films.

In this work we propose automatic methods for detection of two artifacts' types: temporal shift between views and swapped views. As these artifacts are hard to detect, but easy to fix, then proposed methods practically allow to eliminate such artifacts from S3D films.

The methods of swapped views and temporal shift detection were tested on 105 and 26 modern movies, respectively. Most significant detected artefacts were presented in the article. The method of swapped views detection on 17% more accurate (by AUROC metric) and on 3.5 time less time-consuming than previously proposed approach.

Keywords: stereovideo, image analysis, temporal shift, swapped views.

Введение

За последние пять лет доступность и распространённость технологий демонстрации стереоскопического видео существенно возросла. В 2013 году вышли в прокат 50 стереофильмов со средним бюджетом 130 млн. дол-

ларов. В 2014 году уже вышли в прокат более 40 стереофильмов. Количество залов, оборудованных для показа стереофильмов, находится на самом большом уровне в истории кинематографа. Только в Москве работают не менее 75 кинотеатров 3D. Доля продаваемых телеви-

зоров с поддержкой 3D в России превысила 25%, в Азии приближается к 75%. Несмотря на это, интерес общественности к 3D переживает спад, основной причиной которого является низкое качество производимого стерео [13–15]. Это делает актуальным, как никогда раньше, создание инструментов обеспечения контроля качества создаваемых стереоскопических видеоматериалов.

Существует множество артефактов, вызывающих ухудшение визуального восприятия и зрительный дискомфорт при просмотре стереофильмов у части зрителей [1, 2]. В данной статье представлены автоматические методы выявления следующих двух типов артефактов:

- временной сдвиг между видеопотоками для левого и правого ракурса;
- перепутанный порядок ракурсов.

При временном сдвиге действия, отображаемые в одном видеопотоке (ракурсе), опережают те же действия в другом видеопотоке. Под сценами с перепутанным порядком ракурсов мы будем понимать сцены, в которых на месте правого канала находится левый, и наоборот.

Чаще всего временной сдвиг между ракурсами возникает на этапе съёмки кинофильма из-за несовершенства съёмочного оборудования. Например, временной сдвиг может возникнуть при использовании малобюджетных систем стереозаписи с независимыми камерами из-за рассинхронизации момента начала записи. В таком случае временной сдвиг будет неизменным на протяжении всей сцены. Даже одновременное начало записи при использовании таких систем может привести к неравномерному сдвигу во времени вследствие различной частоты записи кадров у двух камер. Неаккуратность или невнимательность при монтаже стереофильма являются основными причинами появления перепутанного порядка ракурсов.

Очевидно, что зрительная система человека не имеет опыта восприятия информации об окружающем мире, содержащей подобные искажения. Поэтому появление данных артефактов в фильмах крайне нежелательно.

Исследования качества стереовидео

Основная масса исследований в области оценки качества стереовидео направлена на выявление искажений, возникших на этапе его сжатия и передачи [11, 12], ввиду того, что аналогичная задача для одноканального видео хорошо изучена. Вопросам выявления искажений, возникающих на этапе создания стереовидео, уделено внимание в относительно малом числе работ. В частности, нашим коллективом был опубликован ряд работ [1, 2], в которых затронуты различные аспекты качества стереовидео и предложены методы оценки следующих характеристик:

- оценка максимального горизонтального параллакса,
- оценка вертикального параллакса,
- оценка соответствия по цвету,
- оценка геометрического соответствия,

- оценка соответствия по фокусу.

На сегодняшний день авторам известна одна работа, описывающая алгоритм выявления перепутанных ракурсов [3], но данный подход имеет относительно низкую точность и высокую вычислительную сложность, что не позволяет использовать его для анализа полнометражных фильмов.

Для обнаружения временного сдвига могут быть использованы методы пространственно-временного выравнивания последовательностей 2D видео [8, 9]. Но они не могут обеспечить оценку временного сдвига между ракурсами стереовидео с точностью менее одного кадра [8] или имеют высокую вычислительную сложность и склонны к ошибкам второго рода из-за того, что не используют особенности стереовидео [9].

Алгоритм обнаружения перепутанных ракурсов

Для проверки корректности порядка ракурсов необходимо оценить параллакс объектов сцены и убедиться, что объекты с меньшим параллаксом расположены ближе к камере, чем объекты с большим параллаксом.

Алгоритмы оценки расстояния до объектов, основанные на бинокулярных признаках, неприменимы для данной задачи, так как перепутанные ракурсы приведут к инвертированной оценке глубины. Поэтому для построения карты расстояния до объектов необходимо использовать монокулярные признаки. Стоит отметить, что для решения данной задачи достаточно знать относительное расположение объектов в сцене, а не точную карту глубины. Для этого могут использоваться так называемые алгоритмы упорядочения глубины (depthordering), которые извлекают из монокулярной видеопоследовательности информацию о взаимном перекрытии объектов [6] и могут оценить их расположение относительно камеры [7]. Однако качество и скорость работы современных алгоритмов упорядочения глубины не позволяют использовать их для обнаружения перепутанных ракурсов на практике.

В целом, довольно сложно выделить единый критерий классификации сцен по наличию перепутанных ракурсов. Поэтому в предложенном методе мы используем композицию четырёх базовых алгоритмов. Каждый из них классифицирует сцену, основываясь на том, что для анализируемой сцены верно определённое предположение. Точность классификации отдельно взятого алгоритма максимальна, когда для сцены выполнено соответствующее предположение данного алгоритма. Конечное решение о наличии перепутанных ракурсов в сцене принимается путём взвешенного голосования базовых алгоритмов-классификаторов. Ниже приведены описания используемых классификаторов. Далее во всех приведённых алгоритмах для сопоставления стереоракурсов и оценки карт диспаратета мы используем блочный алгоритм с поблочной метрикой SND [4], устойчивой к цветовым искажениям между ракурсами.



Рис. 1. Пример кадра с «выпадающим» объектом по центру и оценка карты глубины. Исходное изображение взято из фильма «Avatar»

Анализ областей перекрытия

Если в сцене присутствуют области перекрытия, то верно следующее: в левом ракурсе все области открития всегда находятся слева от объекта переднего плана, а в правом ракурсе – справа. Чтобы проверить соответствие сцены данному утверждению, на первом шаге алгоритм производит сопоставление стереоракурсов и находит области перекрытия с помощью метода LRC (left-right consistency check) [5]. Помимо этого, для каждой области перекрытия определяется её *смежная область*.

Для левого ракурса смежная область – это область, лежащая слева от области перекрытия, для правого – справа. Если ракурсы не были перепутаны, то области перекрытия и их смежные области соответствуют одним и тем же объектам заднего плана, в противном случае области перекрытия соответствуют объектам заднего плана, а их смежные области – переднего. При выполнении предположения о том, что в анализируемой сцене объекты переднего и заднего плана имеют различные цвета, наличие перепутанных ракурсов можно проверить, сравнив распределение цветов в областях перекрытия и их смежных областей. Для этого на втором шаге алгоритм вычисляет цветовые гистограммы областей перекрытия и их смежных областей, разреженные до 16^3 градаций. Если для какой-то области перекрытия расстояние Кульбака-Лейблера между такими гистограммами больше 0,15, то считается, что ракурсы перепутаны.

Анализ перспективы

В фильмах часто можно встретить различные изображения ландшафта с небольшим количеством объектов переднего плана. Для таких сцен карту глубины можно аппроксимировать вертикальным градиентом: внизу кадра находятся объекты с наименьшей глубиной, вверх – с наибольшей.

Алгоритм вычисляет линейную функцию $d(y)$, аппроксимирующую средние значения карты диспаратета в каждой её строке (строки нумеруются сверху вниз, т. е. $y = 0$ соответствует верхней строке изображения). Если $d(y)$ является убывающей, то считается, что правильный порядок ракурсов соблюден.

Анализ «выпадающих» объектов

Обычно объекты, к которым приковано наибольшее внимание зрителей, находятся ближе к центру кадра и «выпадают» из экрана, т. е. имеют отрицательный параллакс. Сначала алгоритм по карте диспаратета пытается найти объект в центре кадра с наибольшим абсолютным значением параллакса. Если такой объект обнаружен, и он имеет положительный параллакс, то считается, что ракурсы были перепутаны.

Анализ распределения диспаратета

Создатели стереофильмов, как правило, стараются компоновать кадр таким образом, чтобы области отрицательного параллакса составляли треть зоны комфортного восприятия стерео, а регионы положительного параллакса – две трети. Таким образом, предполагается, что в левом ракурсе будет больше положительных значений диспаратета, чем отрицательных, в правом ракурсе – наоборот. Алгоритм, использующий данное предположение, был ранее описан в [3].

Алгоритм обнаружения временного сдвига между ракурсами

В качестве признака наличия временной рассинхронизации между ракурсами было выбрано наличие несоответствий скоростей движения объектов в разных ракурсах. Данный признак позволяет эффективно выявлять как неконстантные временные сдвиги, так и константные, при наличии в сцене объекта с изменяющимся вектором скорости.

Для полного анализа стереокартины часто оказывается полезным оценить величину обнаруженного временного сдвига. Например, зная, что величина временного сдвига равна 0,9 кадра, можно нивелировать заметность данного артефакта, сдвинув один из ракурсов на 1 кадр, тем самым сократив сдвиг до 0,1 кадра. Чтобы полностью устранить заметность нецелых временных сдвигов, необходимо задействовать алгоритмы изменения частоты кадров [16]. В ходе проведённых исследований было выявлено, что в реальных фильмах временные сдвиги почти всегда не превосходят одного кадра.

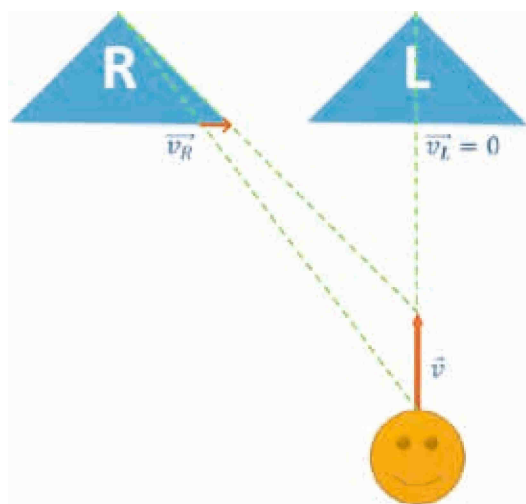


Рис. 2. Пример различия величины горизонтальной проекции вектора скорости в разных ракурсах

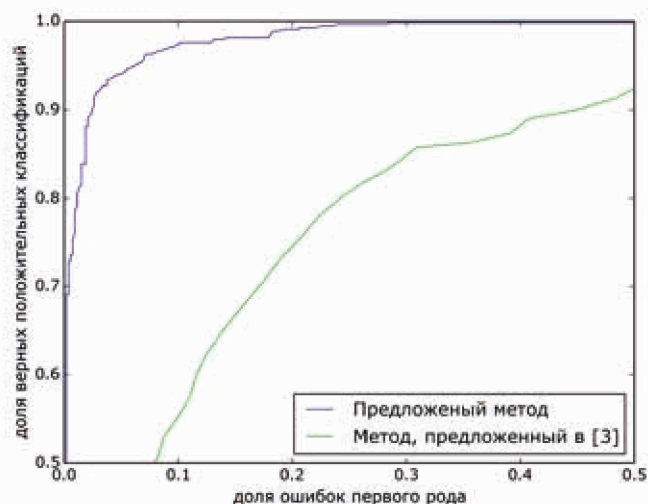


Рис. 3. ROC-кривая предложенного метода и метода, описанного в [3]

Табл. 1. Статистика фильмов, в которых были обнаружены сцены с перепутанным порядком (ППР)

Название фильма	Премьера (год)	Бюджет (млн. долл.)	Обнаружено сцен с ППР	Общая длительность сцен с ППР (с)	Длительность фильма (с)	Длительность сцен с ППР (% от общей длительности фильма)
The Child's Eye	2010	4,50	15	57,456	5822,85	0,99%
The Nutcracker in 3D	2010	90	9	28,946	6480,14	0,45%
3D Sex and Zen: Extreme Ecstasy	2011	2,5	9	23,108	6775,43	0,34%
Spy Kids 3D: Game Over	2003	39	5	10,302	5063,05	0,20%
Hugo	2011	170	2	10,261	7581,03	0,14%
Sharks 3D	2004	5	1	8,926	3073,07	0,29%
SAW 3D	2010	20	3	6,674	5405,31	0,12%
The Last Airbender	2010	150	2	6,590	6193,18	0,11%
Dark Country	2009	4	4	5,756	5287,28	0,11%
Creature from the Black Lagoon	1954	—	2	5,422	4745,99	0,11%
Ghost Rider: Spirit of Vengeance	2012	57	1	4,630	5731,76	0,08%
Stalingrad	2013	30	1	4,296	7848,79	0,05%
Avatar	2009	237	1	3,336	9701,69	0,03%
Spy Kids: All the Time in the World in 4D	2011	27	1	2,962	5265,09	0,06%
Harry Potter and the Deathly Hallows: Part 2	2011	125	1	2,878	7826,94	0,04%
The Chronicles of Narnia: The Voyage of the Dawn Treader	2010	155	1	2,586	6760,75	0,04%
Conan the Barbarian	2011	70	1	1,168	6742,23	0,02%
Step Up 3D	2010	30	1	0,876	6431,38	0,01%
Clash of the Titans	2010	125	1	0,709	6306,3	0,01%
Drive Angry	2010	50	1	0,668	6271,26	0,01%
Bait	2012	31	1	0,584	5587,83	0,01%
A Very Harold & Kumar 3D Christmas	2011	19	1	0,501	5379,37	0,01%
The Three Musketeers	2011	75	1	0,500	6631,54	0,01%

Так как временные сдвиги являются достаточно маленькими, то мы можем использовать информацию о несоответствии скоростей движения объектов для оценки величины сдвига. Пусть вертикальная проекция вектора скорости объекта в кадре t левого ракурса равна $v_L(t) = v(t)$ пикселей за кадр, и $v_R(t) = v(t + \Delta t)$ в правом ракурсе, где Δt – величина временного сдвига.

Тогда,
 $v(t + \Delta t) \approx v(t) + v'(t) \Delta t$, $v'(t) \approx v_L(t + 1) - v_L(t)$,
 следовательно

$$\Delta t \approx \frac{v_R(t) - v_L(t)}{v_L(t+1) - v_L(t)}.$$

В приведённых формулах мы используем вертикальную составляющую вектора движения (скорости), так как горизонтальная компонента проекции вектора скорости объекта на плоскость экрана в разных ракурсах может быть различна (рис. 2). Предложенный алгоритм вычисляет поля векторов движения для левого и правого видеопотока, производит сопоставление стереоракурсов и обнаружение областей перекрытия, после чего вычисляется средняя величина временного сдвига по всем областям, кроме областей перекрытия и областей, имеющих нулевой вектор движения.



Рис. 4. Визуализация кадра из фильма «Harry Potter and the Deathly Hallows: Part 2» с перепутанным порядком ракурсов



Рис. 5. Визуализация кадра из фильма «3D Sex and Zen: Extreme Ecstasy» с перепутанным порядком ракурсов



Рис. 6. Визуализация кадра из фильма «A Very Harold&Kumar 3D Christmas» с перепутанным порядком ракурсов

Табл. 2. Список фильмов, проанализированных на наличие временного сдвига между ракурсами (ВСМД), и их характеристики. В последнем столбце приведена средняя величина временного сдвига среди всех сцен фильма с ВСМД

Название фильма	Премьера (год)	Бюджет фильма (млн. долл.)	Метод производства (съёмка или конвертация)	Обнаружено сцен с ВСМД	Длительность сцен с ВСМД (с)	Средняя величина ВСМД (с)
Drive Angry	2010	50	съёмка	30	99,60	0,36
Resident Evil: Afterlife	2010	60	съёмка	5	12,89	0,17
Pina	2011	4,16	съёмка	1	4,25	0,16
Step Up 3D	2010	30	съёмка	2	2,59	0,07
Step Up Revolution	2012	33	съёмка	2	5,09	0,05
Pirates of the Caribbean: On Stranger Tides	2011	140	съёмка	2	16,43	0,05
Hugo	2011	170	съёмка	2	8,17	0,03
A Very Harold and Kumar 3D Christmas	2011	19	съёмка	1	1,58	0,01
The Three Musketeers	2011	75	съёмка	0	0,00	0,00
Galapagos: The Enchanted Voyage	1999	—	съёмка	0	0,00	0,00
Dolphin Tale	2011	37	съёмка	0	0,00	0,00
Sanctum	2011	30	съёмка	0	0,00	0,00
Resident Evil Retribution	2012	65	съёмка	0	0,00	0,00
The Amazing Spiderman	2012	230	съёмка	0	0,00	0,00
Alice in Wonderland	2010	200	конвертация	0	0,00	0,00
Clash of the Titans	2010	125	конвертация	0	0,00	0,00
Conan the Barbarian	2011	70	конвертация	0	0,00	0,00
Green Lantern	2011	200	конвертация	0	0,00	0,00
Gulliver's Travels	2010	112	конвертация	0	0,00	0,00
The Chronicles of Narnia: The Voyage of the Dawn Treader	2010	155	конвертация	0	0,00	0,00
Immortals	2011	75	конвертация	0	0,00	0,00
The Avengers	2012	220	конвертация	0	0,00	0,00
Wrath of Titans	2012	150	конвертация	0	0,00	0,00
Abraham Lincoln: Vampire Hunter	2012	69	конвертация	0	0,00	0,00
Avatar	2009	237	съёмка	0	0,00	0,00
Into the Deep	1994	—	съёмка	0	0,00	0,00

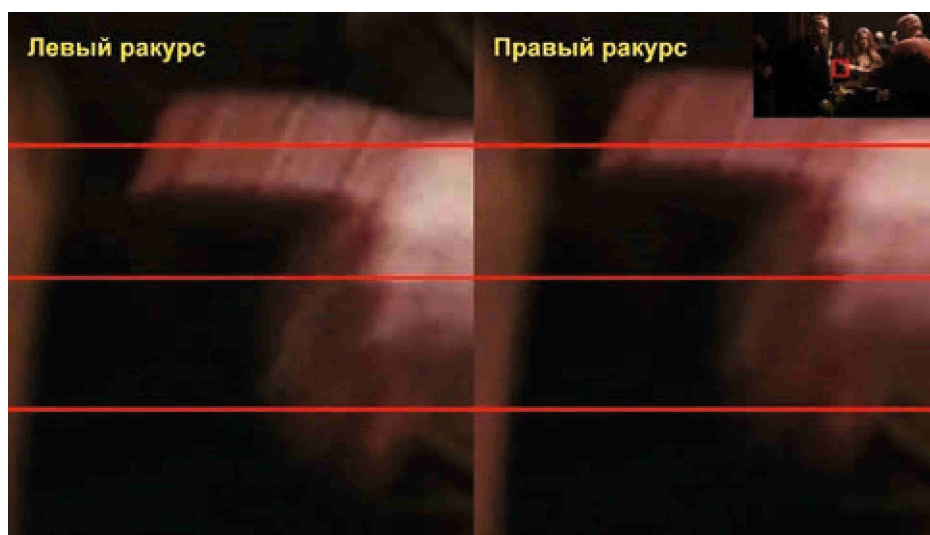


Рис. 7. Пример временного сдвига между ракурсами, обнаруженного в фильме «Drive Angry»

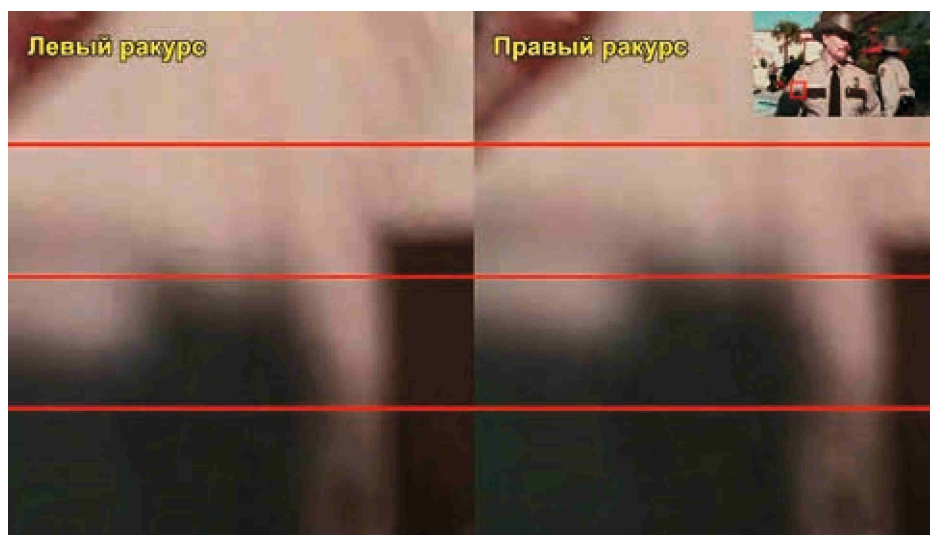


Рис. 8. Пример временного сдвига между ракурсами, обнаруженного в фильме «Drive Angry»



Рис. 9. Пример временного сдвига между ракурсами, обнаруженного в фильме «Resident Evil: After life»

Результаты

Для тестирования алгоритма обнаружения перепутанных ракурсов была составлена тестовая выборка из 1000 сцен, взятых из современных стереофильмов (в 500 сценах порядок ракурсов соблюден, в оставшихся 500 сценах порядок был искусственно изменён). Предложенный метод имеет на 17% большую точность классификации по метрике AUROC (рис. 3) и в 3,5 раза меньшую вычислительную сложность, чем ранее предложенный подход [3]. Предложенным алгоритмом было проанализировано 105 различных стереофильмов, в 23 из которых было обнаружено 65 сцен с неправильным порядком ракурсов общей длительностью 189 секунд. Более подробные данные, полученные в ходе анализа, представлены в табл. 1. Также на рис. 4–6 приведены кадры в формате красно-бирюзового анаглифа из сцен, в которых был обнаружен неправильный порядок ракурсов.

В ходе тестирования алгоритма обнаружения временного сдвига между ракурсами было выявлено 45 сцен в 8 фильмах с временным сдвигом более 0,05 кадра, общая продолжительность которых составила 150 секунд. Большинство обнаруженных сцен имели константный временной сдвиг, величина максимального временного сдвига составила 0,92 кадра. Всего было проанализировано 26 фильмов. Таким образом, 18 фильмов не содержали заметных временных сдвигов. Такие хорошие показатели данных фильмов можно объяснить тем, что половина из них была произведена с помощью технологии конвертации из 2D-фильмов. Результаты проведённого тестирования представлены в табл. 2. На рис. 7–9 приведе-

ны визуализации кадров из сцен, в которых был обнаружен временной сдвиг. На этих визуализациях показаны два увеличенных фрагмента изображений левого и правого ракурса одного кадра. Оба фрагмента имеют одинаковое вертикальное выравнивание, таким образом, временной сдвиг можно обнаружить по вертикальному смещению отдельного объекта между левым и правым фрагментом.

Заключение

Предложены алгоритм выявления перепутанных ракурсов и алгоритм обнаружения сдвига во времени, таким образом, был расширен список искажений стереовидео, выявляемых в рамках проекта VQMT3D (Video Quality Measurement Tool 3D) [10], развиваемого лабораторией компьютерной графики и мультимедиа МГУ им. М.В. Ломоносова. Основное преимущество предложенных методов заключается в достаточно

высокой скорости работы, что позволило целиком обрабатывать с их помощью полнометражные фильмы, в т.ч. провести анализ перепутанных ракурсов для 105 фильмов.

Оба метода были протестированы на видеоданных, взятых из современных стереофильмов, доказав свою практическую применимость. Новый метод обнаружения перепутанных ракурсов показывает значительные улучшения относительно метода, разработанного нашим коллективом ранее [3], как в плане вычислительной сложности, так и точности классификации. Сегодня крайне мало публикаций по определению временного сдвига, по крайней мере, найти аналоги, способные работать на полнометражных стереофильмах, авторам не удалось.

Данные, полученные в ходе проведённого анализа, планируется опубликовать в 8-м и 9-м отчётах проекта VQMT3D [10]. ■

ЛИТЕРАТУРА

1. Voronov A., Vatolin D., Sumin D., Napadovsky V., and Borisov A. «Methodology for stereoscopic motion-picture quality assessment», Proc. SPIE 8648, Stereoscopic Displays and Applications XXIV, vol. 8648, pp. 864810-1–864810-14, March 2013. doi:10.1117/12.2008485
2. Ватолин Д.С., Воронов А.А., Нападовский В.В., Борисов А.В. «Исследование артефактов в стереокино и примеры анализа кинокартин», Международная ежегодная научно-техническая конференция «Запись и воспроизведение объёмных изображений в кинематографе и других областях», с. 190–203, Москва, 2012.
3. Alexey Shestov, Alexander Voronov, and Dmitriy Vatolin. «Detection of swapped views in stereo image», 22st GraphiCon International Conference on Computer Graphics and Vision, pp. 23–27, 2012.
4. Jung-Jae Yu, Hae-Dong Kim, Ho-Wook Jang, Seung-Woo Nam. «A hybrid color matching between stereo image sequences», 3DTV Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTV-CON), pp.1, 4, 16–18 May 2011.
5. Jian Sun, Yin Li, Sing Bing Kang, Heung-Yeung Shum. «Symmetric stereo matching for occlusion handling», Computer Vision and Pattern Recognition, 2005.
6. Bergen L., Meyer F. «A novel approach to depth ordering in monocular image sequences», Proceedings. IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 536, 541, 2000.
7. Smith P., Drummond Tom, Cipolla R. «Layered motion segmentation and depth ordering by tracking edges», IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 4, pp. 479, 494, April 2004.
8. Y. Caspi and M. Irani. «Spatio-Temporal Alignment of Sequences», IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, pp. 1409–1424.
9. B. Meyer, T. Stich, M.A. Magnor, and M. Pollefeys. «Subframe Temporal Alignment of Non-Stationary Cameras», British Machine Vision Conference, 2008, pp. 1–10.
10. Проект Video Quality Measurement Tool 3D, <http://compression.ru/video/vqmt3d/>
11. Huynh-Thu Q., Callet P.L., and Barkowsky M. «Video quality assessment: From 2D to 3D – challenges and future trends», IEEE International Conference on Image Processing (ICIP), pp. 4025–4028, 2010.
12. Boev A., Hollosi D., Gotchev A., and Egiazarian K. «Classification and simulation of stereoscopic artifacts in mobile 3DTV content», Proc. SPIE Stereoscopic Displays and Applications XX, 2009.
13. «3-D TV is Officially Dead (For Now) and This is Why it Failed», <http://spectrum.ieee.org/techtalk/consumer-electronics/audiovideo/3d-tv-is-officially-dead-for-now-and-this-is-why-it-failed/>.
14. «James Cameron says: 3D TV will come, but it requires better tech and content», <http://venturebeat.com/2010/10/27/james-cameron-says-3d-tv-will-come-but-it-requires-better-techand-content/>.
15. Lambooj M. «Stereoscopic displays and visual comfort: a review», SPIE Newsroom, 2007.
16. Choi B., Lee S., Ko S. «New frame rate up-conversion using bi-directional motion estimation», IEEE Transactions on Consumer Electronics, 2000.