

Обзор артефактов стерео видео



Временная синхронизация

Алексей Борисов

Video Group

CS MSU Graphics & Media Lab



Лирическое вступление

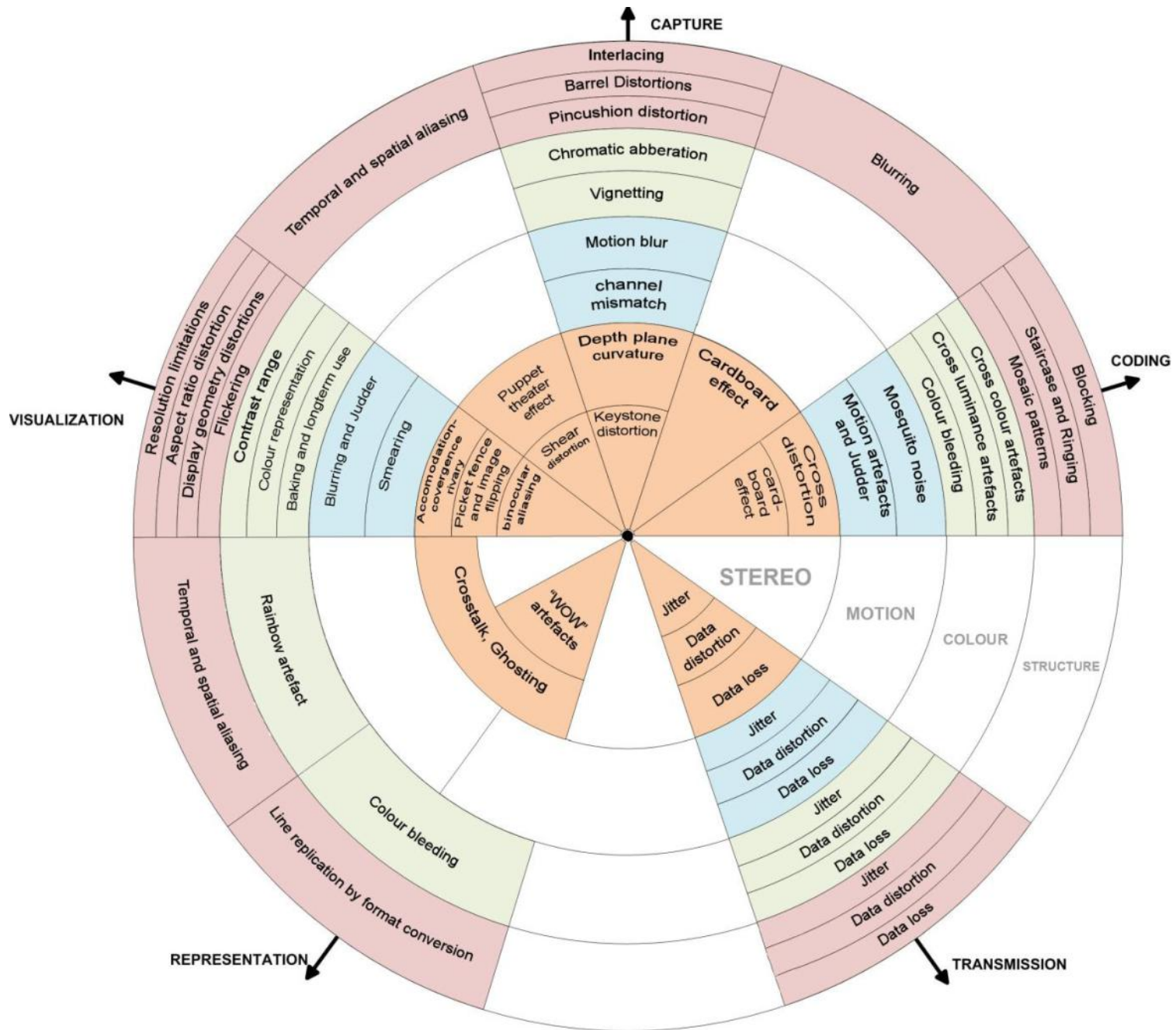
*Stereoscopic images are
easy to do badly,
hard to do well, and
impossible to do correctly*

Фольклор



Содержание

- **Введение**
- Зона комфортного восприятия
- Binocular luster
- Zoom синхронизация
- Временная синхронизация
- Заключение

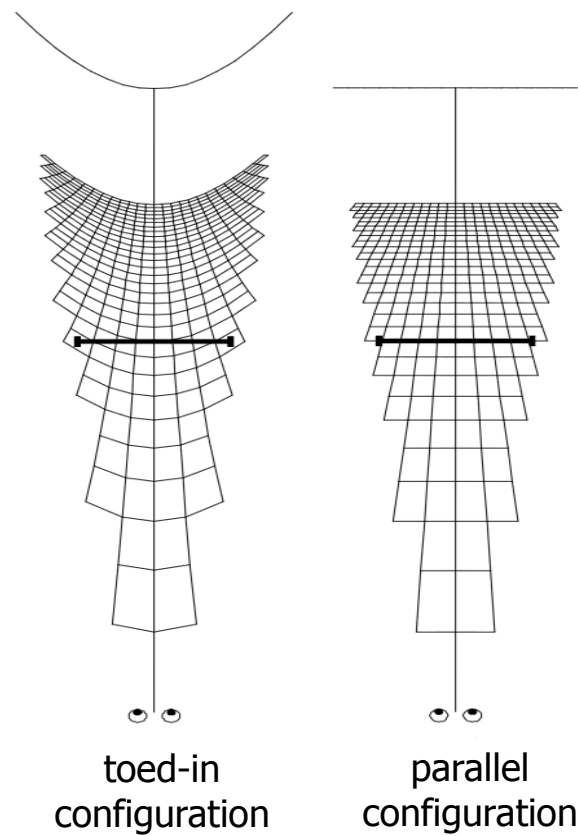


Atanas Boev, Danilo Hollosi, Atanas Gotchev, "Classification of stereoscopic artefacts", MOBILE3DTV technical report, 2008

Введение

Camera configuration (1)

Представление уровней глубины при съёмке параллельной и сходящейся конфигурациями камер



Введение

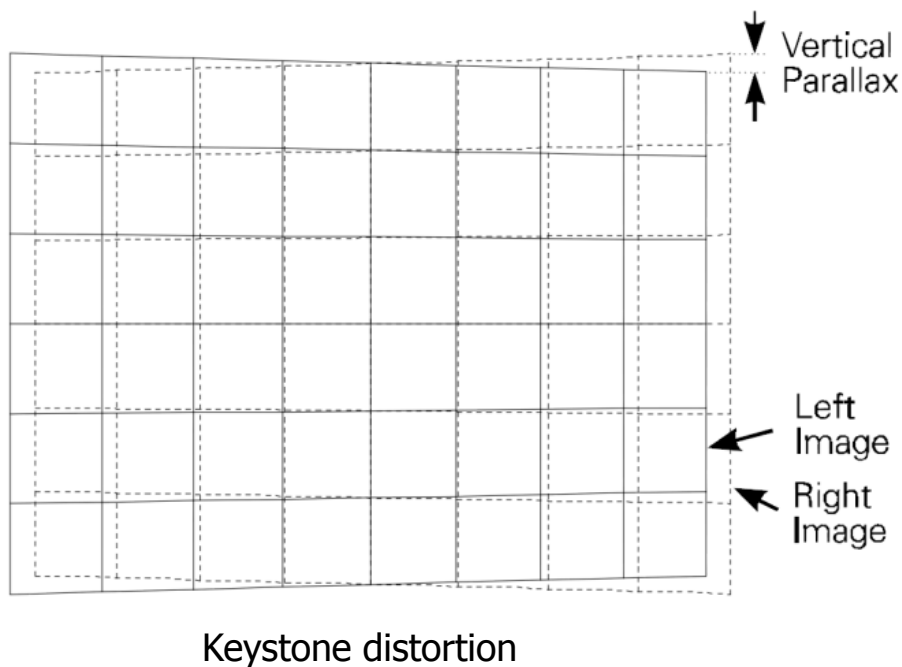
Camera configuration (2)

Camera Configuration	Достоинства	Недостатки
Параллельная	Не требует изменений во время съёмки	HVS (Human Visual System) работает иначе
Сходящаяся	Совместимость с HVS	Необходимость постоянного изменения параметров при съёмке

Введение

Трапецеидальное искажение

При сходящейся конфигурации камер
возникает вертикальный параллакс



Введение

Внешняя симметрия

Temporal Symmetry

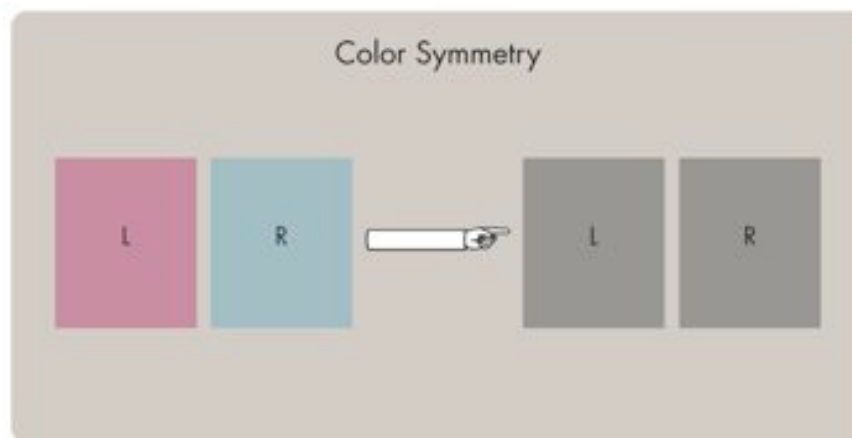


Spatial Symmetry



Введение

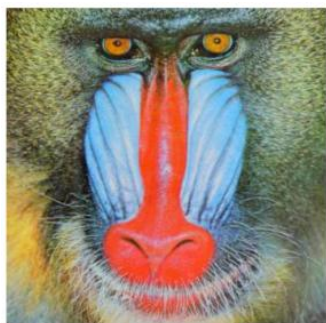
Внутренняя симметрия (1)



Введение

Внутренняя симметрия (2)

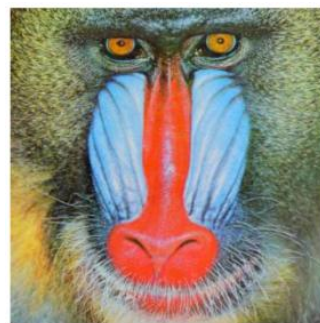
Focus Symmetry



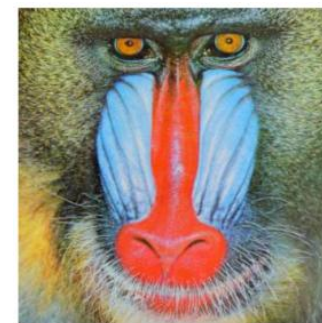
L



R



L



R

Введение

Эффект картонок

Cardboard effect — отсутствие рельефа и объёма



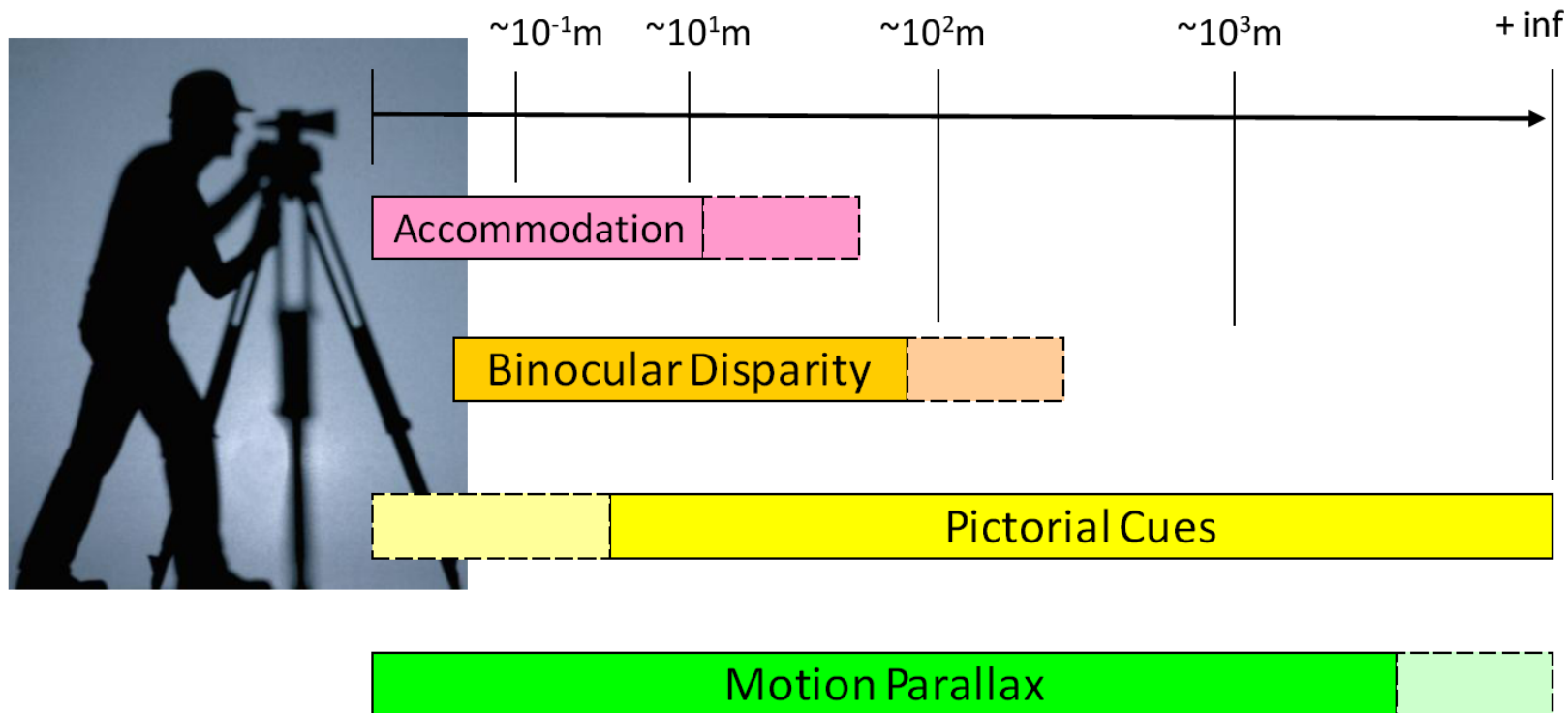
2D изображение



Карта глубины

Введение

Depth cues rivalry (1)

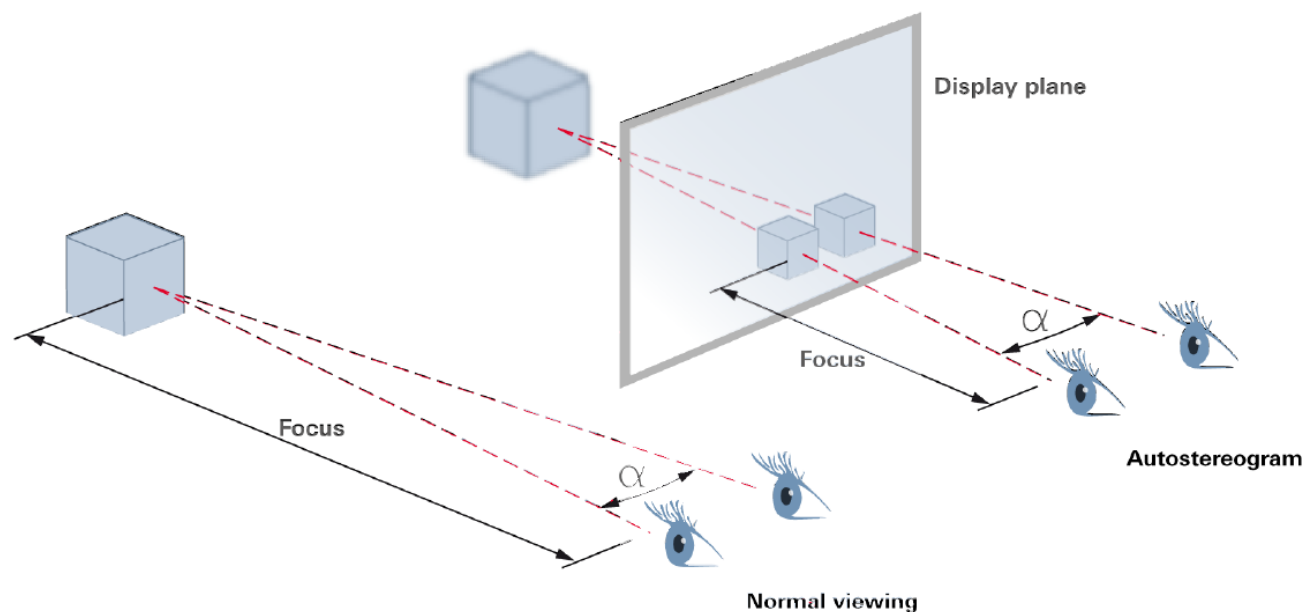


Источники информации о глубине

Введение

Depth cues rivalry (2)

Accommodation-convergence rivalry

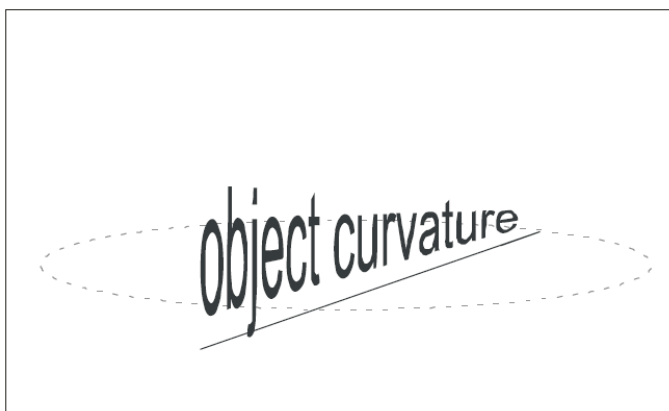


Stephan Reichelt, et al., "Depth cues in human visual perception and their realization in 3D display", Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV

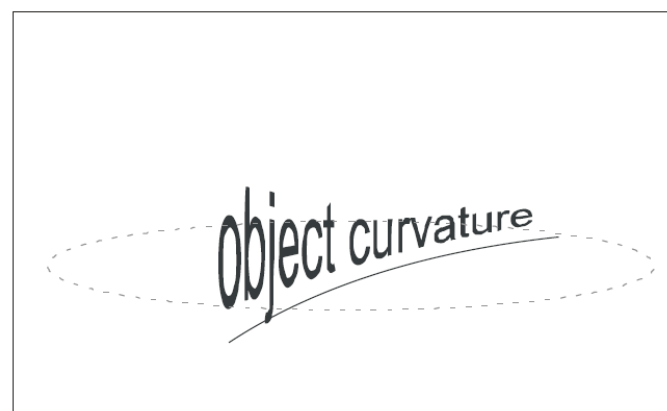
Введение

Depth cues rivalry (3)

Эффект кукольного театра (Perspective-disparity rivalry)



Depth as suggested
by perspective cues



Depth as suggested
by binocular cues



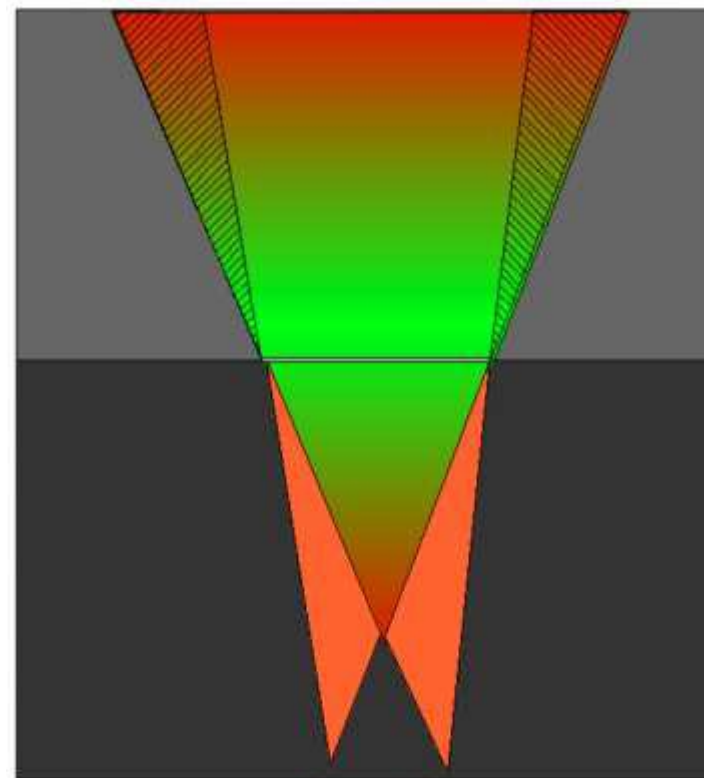
Contents

- Введение
- **Зона комфортного восприятия**
- Binocular luster
- Zoom синхронизация
- Временная синхронизация
- Заключение

Восприятие глубины

Введение

- Серый: невидимая зона
- **Красный: опасная зона**
 - Большая нагрузка на глаза
- **Оранжевый: быстрая зона**
 - Объект виден только одному глазу — нагрузка
- **Зеленый: зона отдыха глаз**
 - Рядом с плоскостью экрана
 - Комфортная для восприятия



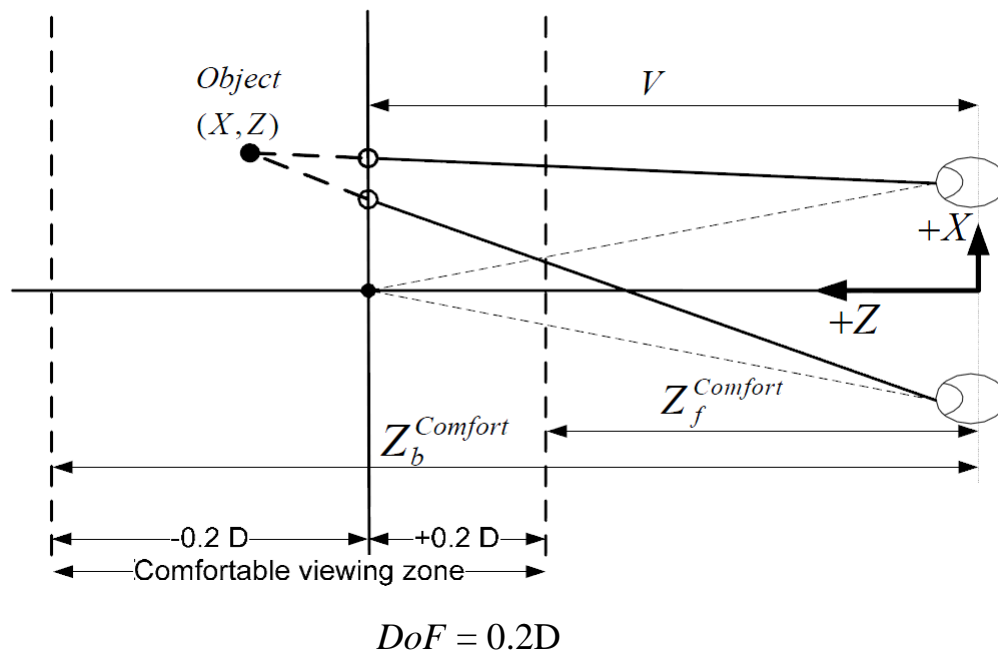
Восприятие глубины

Постановка задачи

Цель: определить зону комфортного восприятия

$$Z_f^{Comfort} = \frac{1}{\frac{1}{V} + DoF}$$

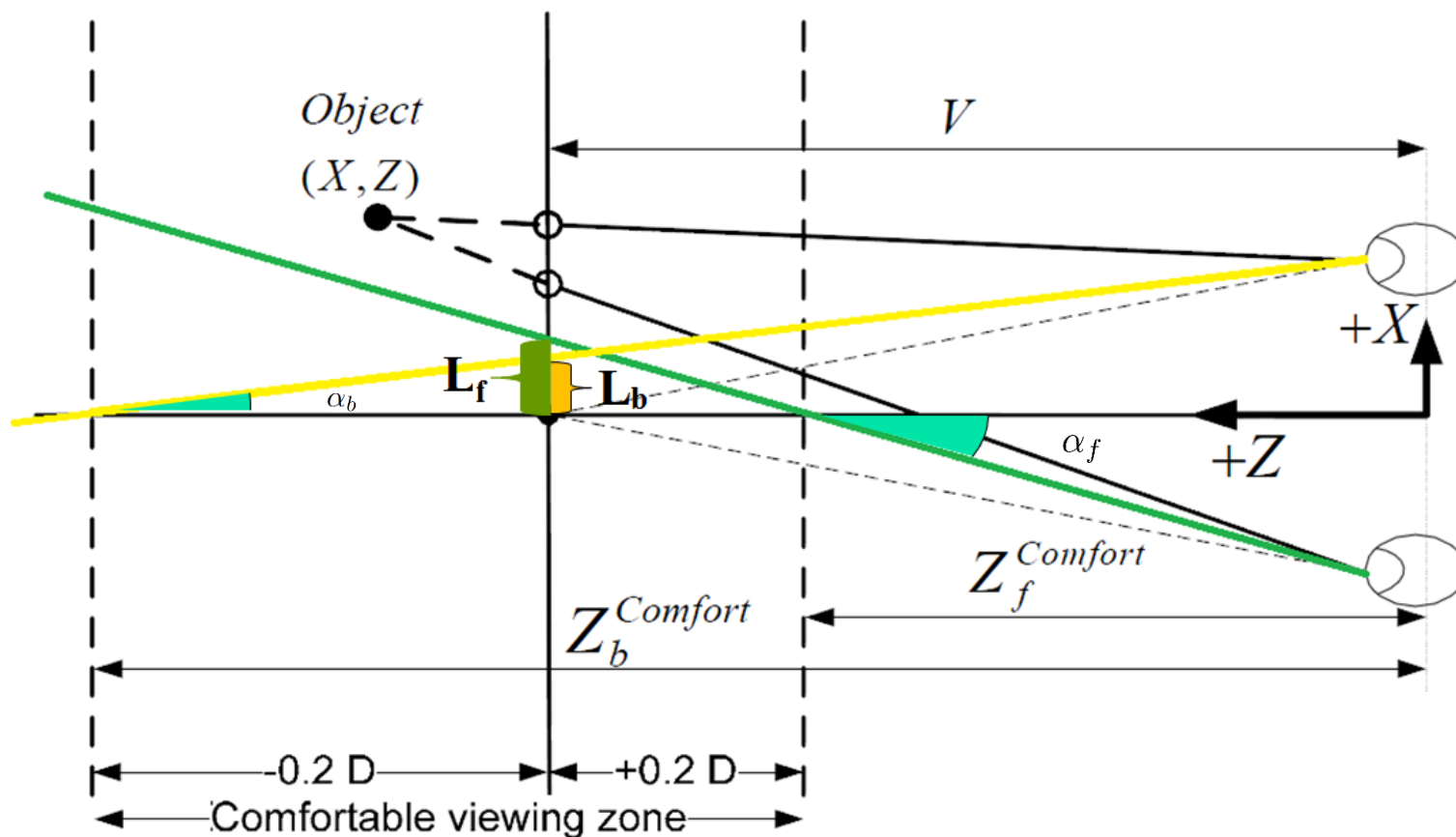
$$Z_b^{Comfort} = \begin{cases} \frac{1}{\frac{1}{V} - DoF} & \text{if } V < DoF^{-1} \\ \infty & \text{otherwise} \end{cases}$$



W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", SPIE, 2011

Восприятие глубины

Максимальный параллакс (1)



W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", SPIE, 2011

Восприятие глубины

Максимальный параллакс (2)

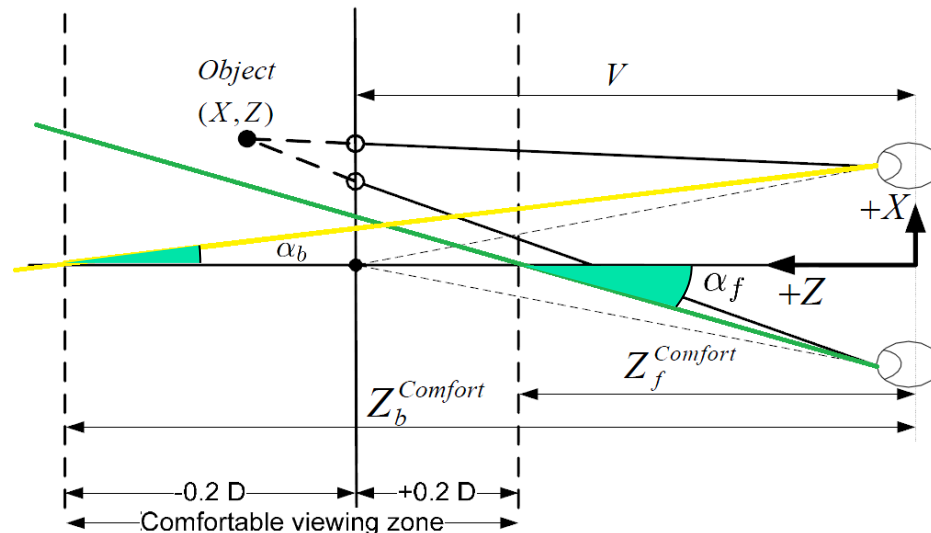
$$\tan \alpha_f = \frac{b}{2Z_f}$$

$$\tan \alpha_b = \frac{b}{2Z_b}$$

$$L_f = 2(V - Z_f) \tan \alpha_f$$

$$L_b = 2(Z_b - V) \tan \alpha_f$$

$$\frac{L_f}{L_b} = \frac{V - Z_f}{Z_b - V} * \frac{Z_b}{Z_f} = \frac{V - \frac{1}{\frac{1}{V} + DoF}}{\frac{1}{\frac{1}{V} - DoF} - V} * \frac{\frac{1}{V} + DoF}{\frac{1}{V} - DoF} = 1$$



W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", SPIE, 2011

Восприятие глубины

Гипотезы

- Для каждой сцены существует свой оптимальный DoF
- Оптимальные параметры съёмки достигаются при наименьшем искажении объекта области интереса (RoI)

Восприятие глубины

Субъективное тестирование (1)

Исследование, цель которого — проверка гипотез + определение доминирующей, если обе верные, но не могут быть выполнены одновременно

- 5 сцен
- 3 значения *DoF*
- 28 испытуемых

Восприятие глубины

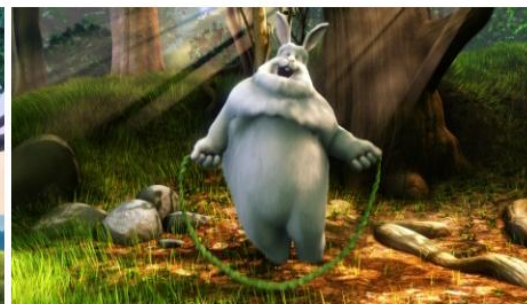
Субъективное тестирование (2)



Scene 1



Scene 2



Scene 3



Scene 4



Scene 5

Сцены, использованные при тестировании

W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", SPIE, 2011

Восприятие глубины

Результаты (1)

Scene Num	2D	$DoF = 0.1$		$DoF = 0.2$		$DoF = 0.3$		$D_s^{RoI} = 1$	
		b (mm)	D_s^{RoI}	b (mm)	D_s^{RoI}	b (mm)	D_s^{RoI}	b (mm)	D_s^{RoI}
1	X	11	0.55	22	1.1	33	1.65	<u>20</u>	<u>1</u>
2	X	43	0.26	<u>85</u>	<u>0.5</u>	145	0.86	168	1
3	X	106	0.59	213	1.20	319	1.79	<u>178</u>	<u>1</u>
4	X	17	0.31	<u>37</u>	<u>0.69</u>	50	0.93	54	1
5	X	638	0.59	1283	1.2	1930	1.8	<u>1069</u>	<u>1</u>

Подчёркнутые значения являются наилучшими согласно проведённому тестированию

b — расстояние между камерами
 D_s^{RoI} — мера искажения объекта RoI

W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", SPIE, 2011

Восприятие глубины

Результаты (2)

- Исследование подтвердило универсальный для всех сцен $DoF = 0,2$
- В большинстве случаев при оптимальных параметрах съёмки искажения объекта RoI незначительны
- Параметры съёмки определяются DoF



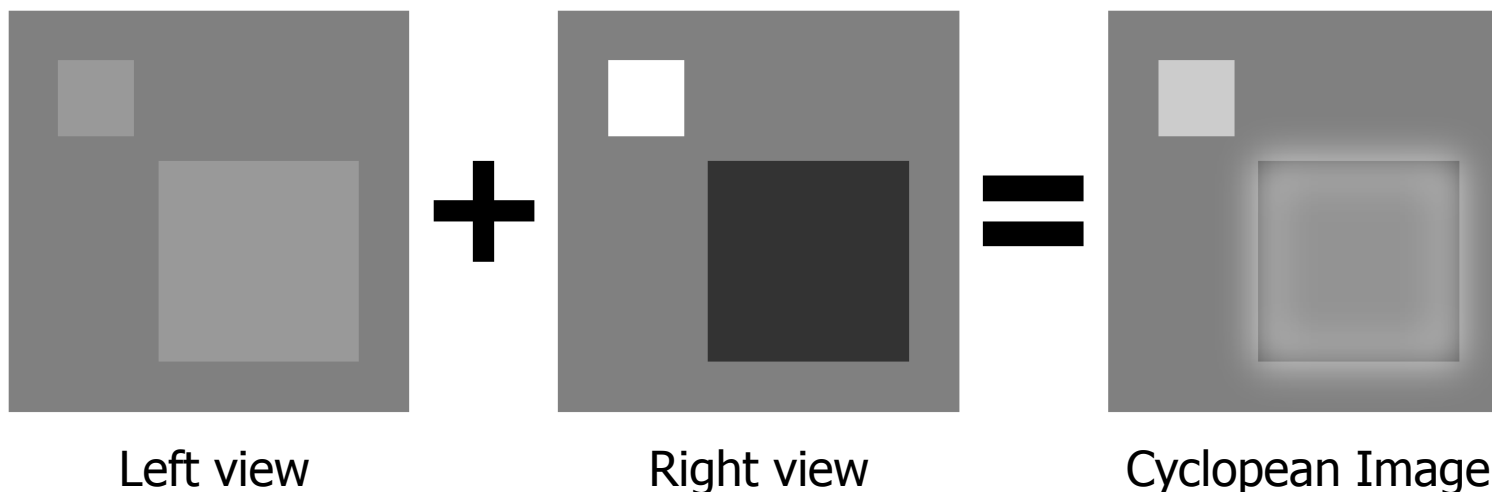
Contents

- Введение
- Зона комфортного восприятия
- **Binocular luster**
- Zoom синхронизация
- Временная синхронизация
- Заключение

Binocular luster

Введение

Binocular luster — визуальный эффект сияния (свечения) объекта, вызванный различными функциями контрастности на левом и правом ракурсах



Binocular luster

Причины возникновения

- Сжатие
- Постобработка
- Конвертация 2D в 3D (?)

Binocular luster

Детектирование при сильном сжатии (1)

Алгоритм:

1. Построение карты полярности для обоих видов
2. Вычитание построенных карт
3. Медианная фильтрация

Binocular luster

Детектирование при сильном сжатии (2)

Contrast Polarity map = *image* \otimes *f*

$$f[k] = \begin{cases} \alpha^{|k-\beta|} & \text{if } 0 \leq k \leq \beta \\ -\alpha^{|k-\beta|} & \text{if } -\beta \leq k < 0 \\ 0 & \text{otherwise} \end{cases}$$

Binocular luster

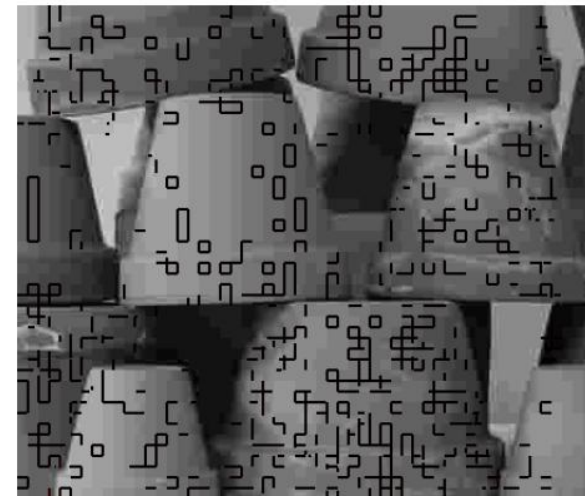
Результаты



Left compressed view



Right compressed view



Right compressed view
with shine detection



Contents

- Введение
- Зона комфортного восприятия
- Binocular luster
- **Zoom синхронизация**
- Временная синхронизация
- Заключение

Zoom synchronization

Введение



Рассинхронизированная стереопара

■ Причины

- Внутрикамерные различия
- Рассинхронизация по времени

■ Эффекты

- Напряжение глаз
- Затруднение синтеза изображения



Синхронизированная стереопара

Zoom synchronization

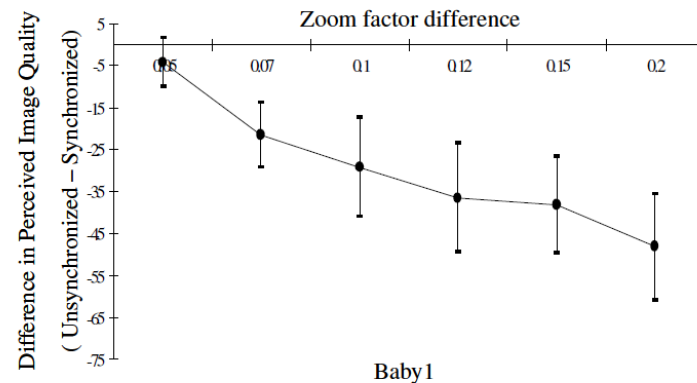
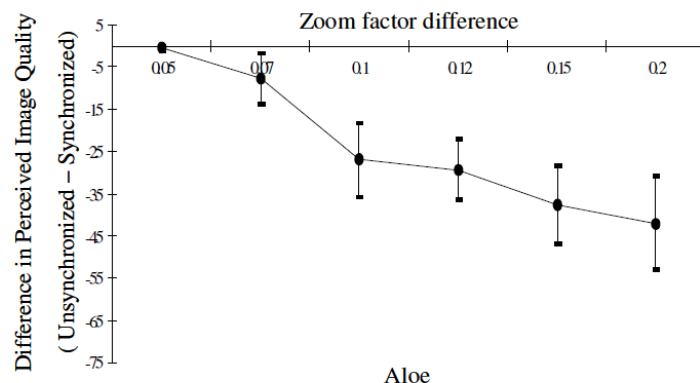
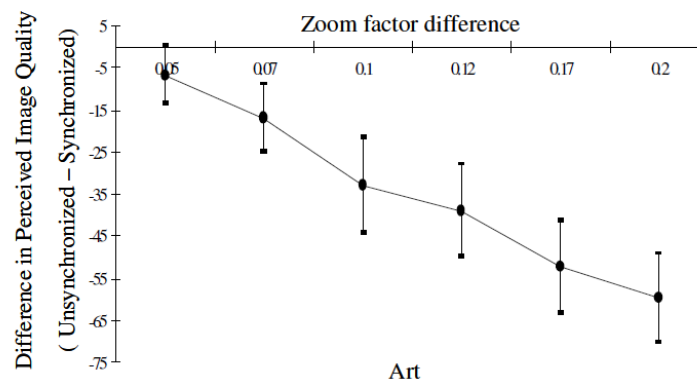
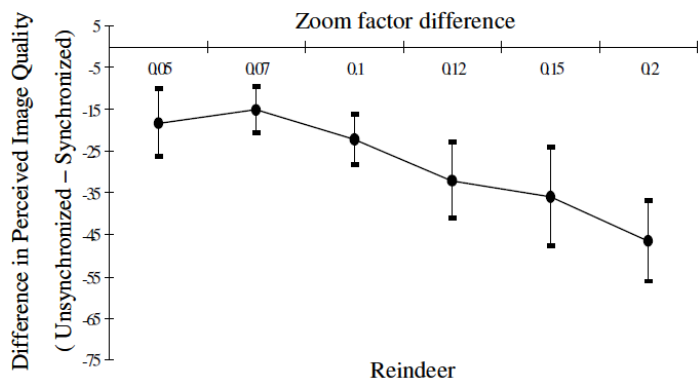
Субъективное оценивание

Исследование по влиянию асинхронного масштабирования на визуальное восприятие зрителя

- Коэффициент масштабирования 5-20%
- 4 последовательности
- 17 испытуемых

Zoom synchronization

Результаты тестирования



Zoom synchronization

Особенности алгоритма

Требования к видео:

- Синхронизация по времени
- Параллельная конфигурация камер
- Отсутствие вертикального параллакса

Zoom synchronization

Модель сцены

$$x_L = x_{L,O} + u_L \quad x_R = x_{R,O} + u_R$$

$$y_L = y_{L,O} + v_L \quad y_R = y_{R,O} + v_R$$

$$x_{L,O} = f_L \frac{X + B/2}{Z} \quad x_{R,O} = f_R \frac{X - B/2}{Z}$$

$$y_{L,O} = f_L \frac{Y}{Z} \quad y_{R,O} = f_R \frac{Y}{Z}$$

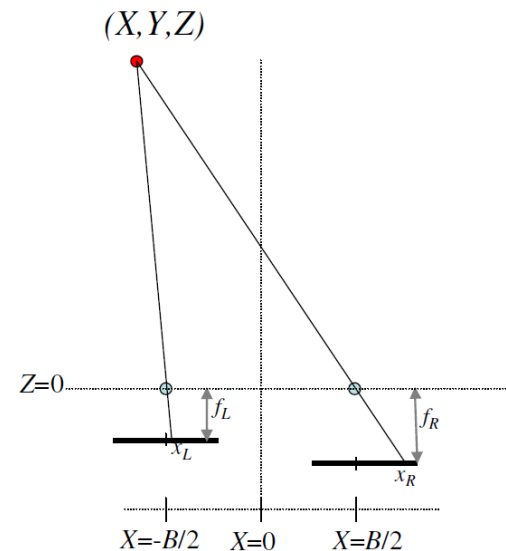


Figure 1: Stereo geometry with parallel cameras

$$y_R = \frac{f_R}{f_L} y_L + v_R - \frac{f_R}{f_L} v_L$$

B

f_L, f_R

x_L, x_R

$x_{L,O}, x_{R,O}$

— расстояние между камерами

— фокальные расстояния

— координаты точки (X, Y, Z)

— относительно центра

Zoom synchronization

Алгоритм

1. Генерация пар (x_L, x_R)
(SIFT detector)
2. Задача линейной регрессии
(решение методом наименьших квадратов)

$$\begin{bmatrix} y_R \\ \vdots \end{bmatrix} = \begin{bmatrix} y_L & 1 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} s \\ t_y \end{bmatrix}$$

3. Итоговое преобразование

$$\begin{bmatrix} x'_L \\ y'_L \end{bmatrix} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} x_L \\ y_L \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad t_x = \frac{W}{2}(1-s)$$

Zoom synchronization

Результаты (1)

До синхронизации



Левый ракурс

Правый ракурс

Zoom synchronization

Результаты (2)

После синхронизации



Левый ракурс

Правый ракурс

Zoom synchronization

Результаты (3)

До синхронизации



Левый ракурс

Правый ракурс

Zoom synchronization

Результаты (4)



После синхронизации



Левый ракурс

Правый ракурс

Zoom synchronization

Выводы

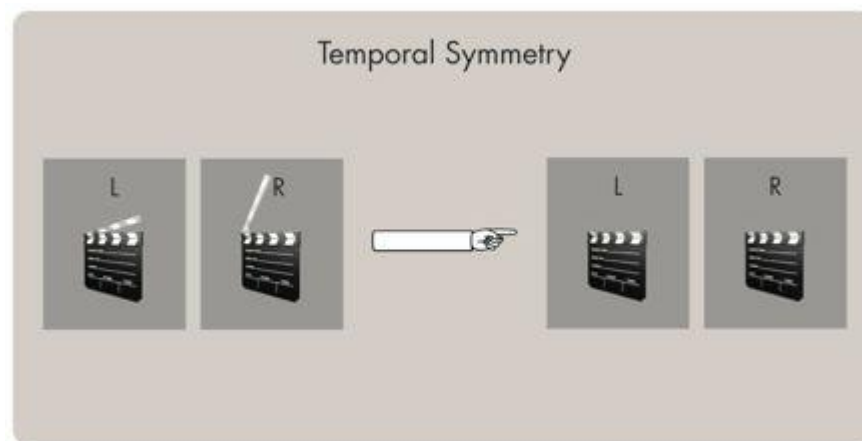
- **Достоинство**
 - Возможность расширения класса применимости путём усложнения модели
- **Недостатки**
 - Point-based
 - Жёсткие ограничения применимости



Contents

- Введение
- Зона комфортного восприятия
- Binocular luster
- Zoom синхронизация
- **Временная синхронизация**
- Заключение

Временная синхронизация



Причина — дороговизна использования дополнительного аппаратного обеспечения

Временная синхронизация

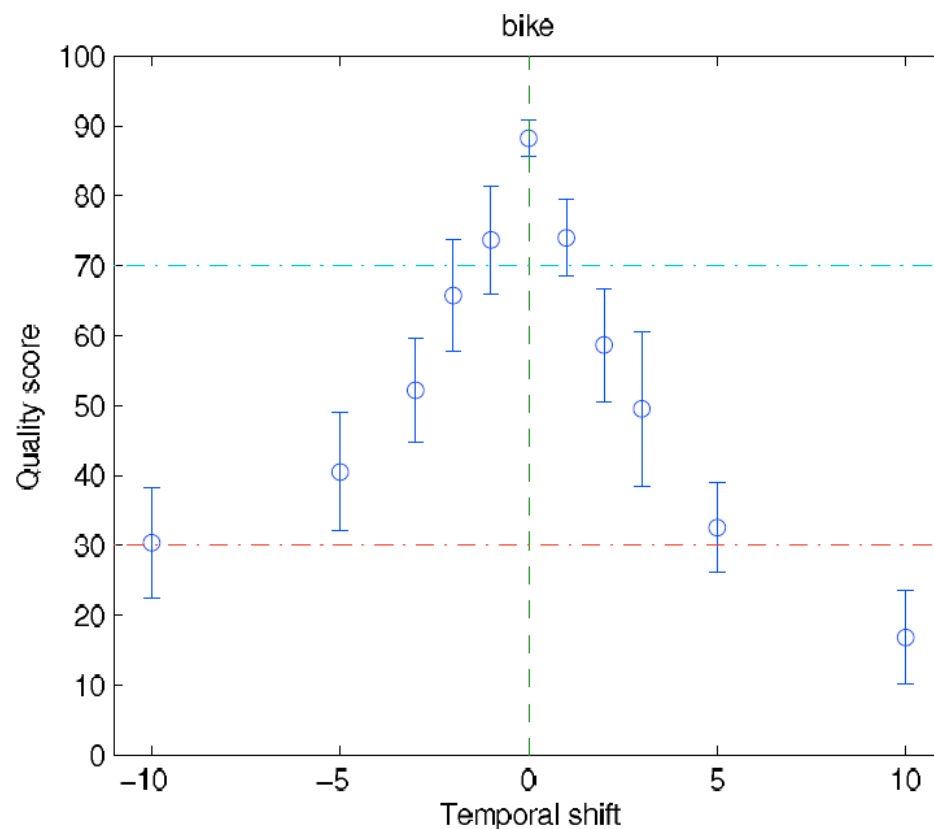
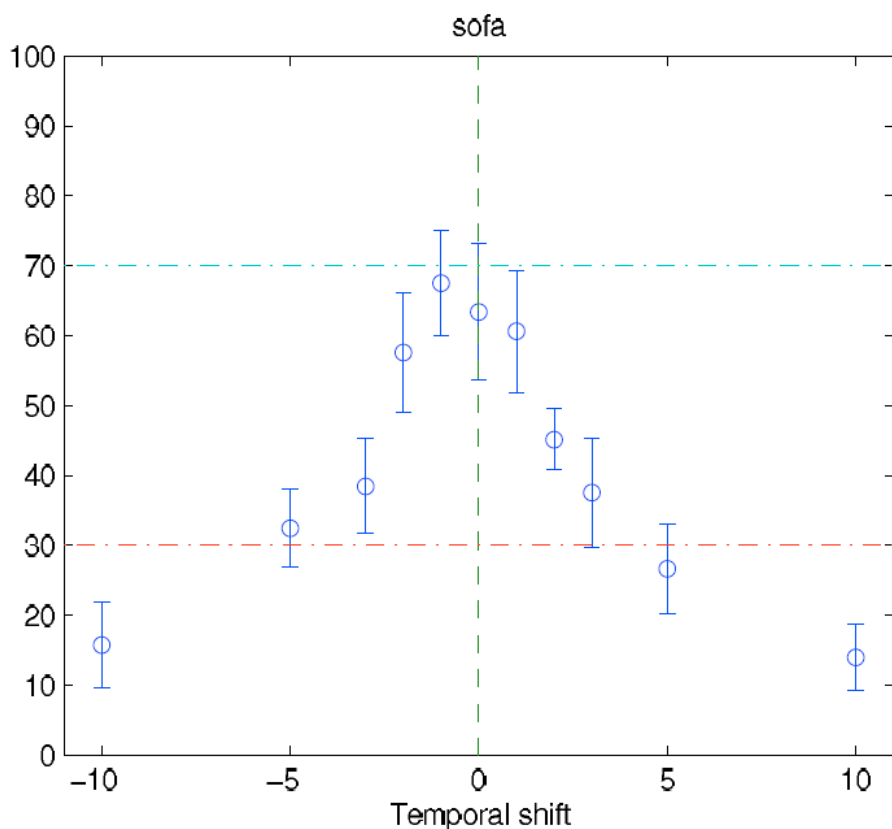
Субъективное оценивание

Исследование по влиянию рассинхронизации по времени на визуальное восприятие зрителя

- Смещение ± 10 кадров
- Видео 25 fps
- 66 тестов
- 20 испытуемых

Временная синхронизация

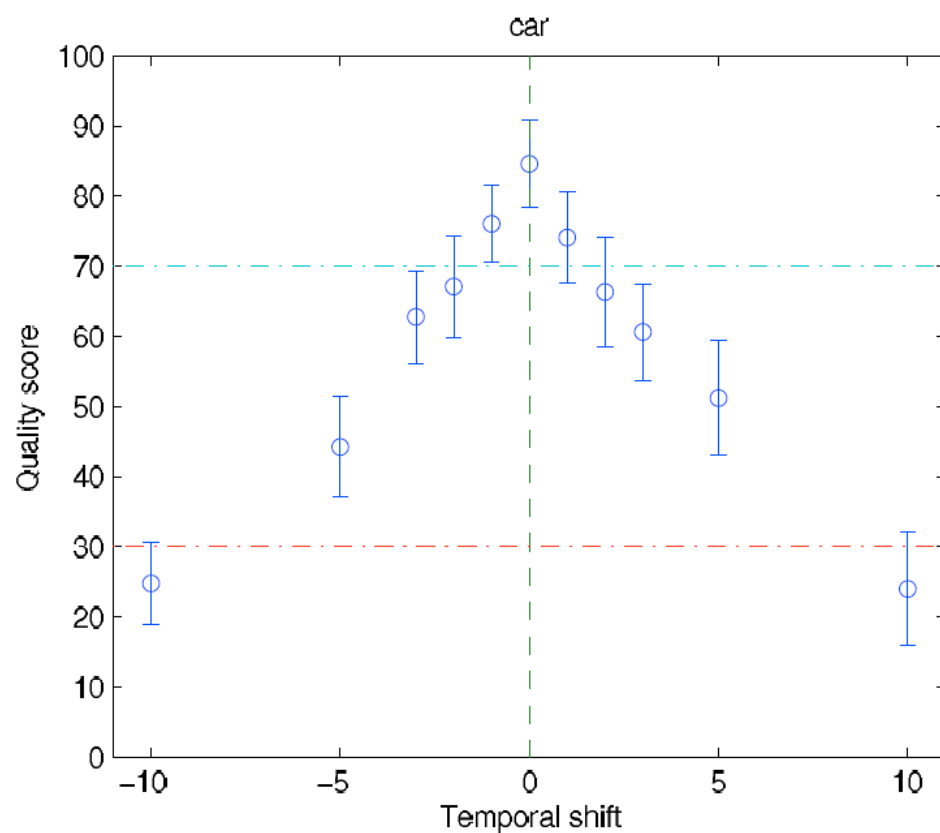
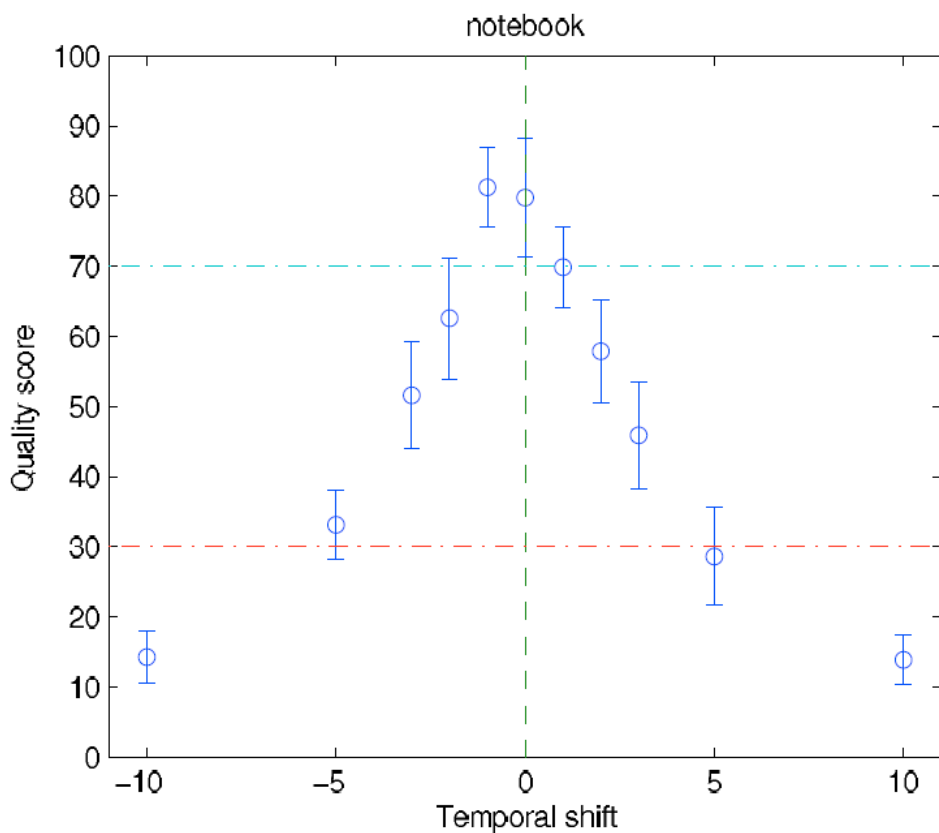
Результаты тестирования (1)



Lutz Goldmann, Jong-Seok Lee, Touradj Ebrahimi, "Temporal synchronization in stereoscopic video: influence on quality of experience and automatic asynchrony detection", ICIP, 2010

Временная синхронизация

Результаты тестирования(2)



Lutz Goldmann, Jong-Seok Lee, Touradj Ebrahimi, "Temporal synchronization in stereoscopic video: influence on quality of experience and automatic asynchrony detection", ICIP, 2010

Временная синхронизация

Выводы

Результаты:

$\tau < 80 \text{ ms}$ — Good

$\tau > 200 \text{ ms}$ — Bad

В пересчёте на fps:

Frame rate	Frame shift (good)	Frame shift (bad)
25 fps	2	5
50 fps	4	10



Contents

- Введение
- Зона комфортного восприятия
- Binocular luster
- Zoom synchronization
- **Временная синхронизация**
 - Subjective quality test
 - **Distribution based sync**
 - Inflection points based sync
- Заключение

Distribution based sync

Алгоритм

1. Поиск точек с наибольшими пространственно-временными искажениями на обоих видах
2. Построение гистограмм распределений найденных точек
3. Поиск сдвига, минимизирующего разницу гистограмм

Distribution based sync

Space-time interest points (1)

- Пространство

- детектор Харриса

$$H = \det(\eta) - k \text{trace}^2(\eta) \quad \eta = g(\cdot, \cdot, \sigma) * \begin{pmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{pmatrix}$$

- Пространство + время

$$H = \det(\eta) - k \text{trace}^3(\eta) \quad \eta = g(\cdot, \cdot, \cdot, \sigma) * \begin{pmatrix} I_x^2 & I_x I_y & I_x I_t \\ I_y I_x & I_y^2 & I_y I_t \\ I_x I_t & I_y I_t & I_t^2 \end{pmatrix}$$

I_x, I_y, I_t — изменение интенсивности по x, y, t

$$g(\cdot, \cdot, \sigma) = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Distribution based sync

Space-time interest points (2)

- Значимость точки

$$k = \frac{\det(\eta)}{\text{trace}^3(\eta)} = \frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)^3}$$

$\lambda_1, \lambda_2, \lambda_3$ - собственные значения η

- Выбираются 200 точек максимальной значимости:

$$\min(\lambda_1, \lambda_2, \lambda_3) > T$$

Distribution based sync

Space-time interest points (3)



Красным квадратом выделена область,
содержащая space-time interest points

Jingyu Yan, Marc Pollefeys, "Video synchronization via space-time interest point distribution", Proceedings of Advanced Concepts for Intelligent Vision Systems, 2004

Distribution based sync

Space-time interest points (4)



Красным квадратом выделена область,
содержащая space-time interest points

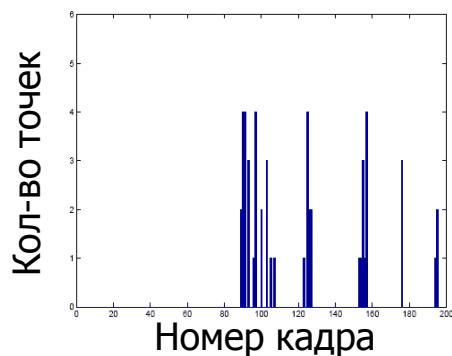
Jingyu Yan, Marc Pollefeys, "Video synchronization via space-time interest point distribution", Proceedings of Advanced Concepts for Intelligent Vision Systems, 2004

Distribution based sync

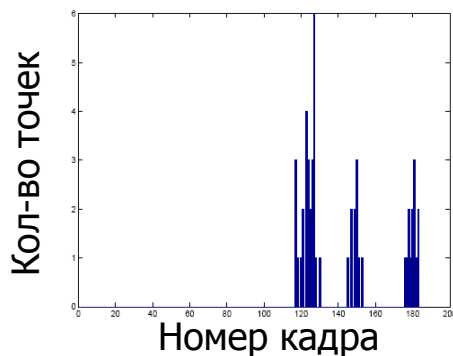
Корреляционная функция

$$C(t) = \sum_{j=t_0 - \Delta t}^{t_0 + \Delta t} \eta_j * \zeta_{j+t}$$

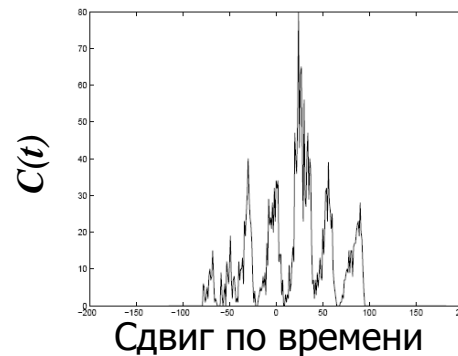
Распределение точек интереса для левого ракурса



Распределение точек интереса для правого ракурса



Корреляционная функция (максимум соответствуют найденному сдвигу по времени)



$[\eta_1, \eta_2, \dots, \eta_n], [\zeta_1, \zeta_2, \dots, \zeta_n]$

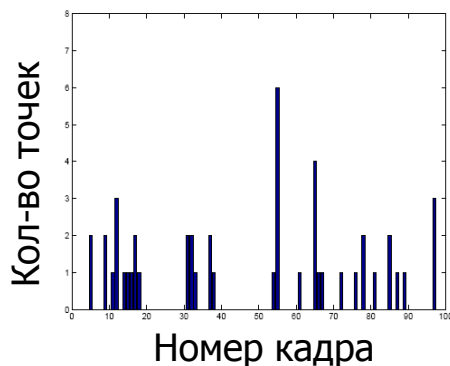
— распределение пространственно-временных точек левого/правого вида

Jingyu Yan, Marc Pollefeys, "Video synchronization via space-time interest point distribution", Proceedings of Advanced Concepts for Intelligent Vision Systems, 2004

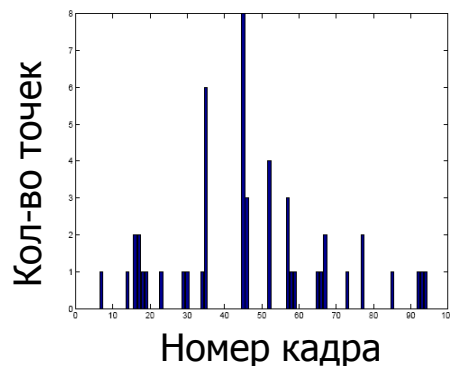
Distribution based sync

Результаты

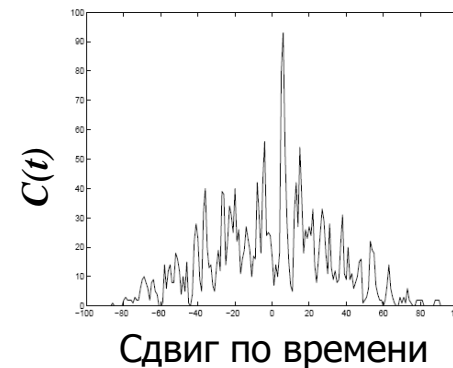
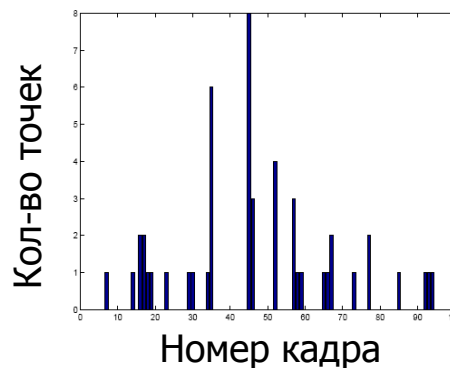
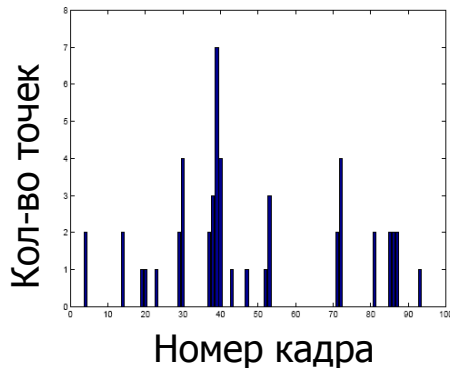
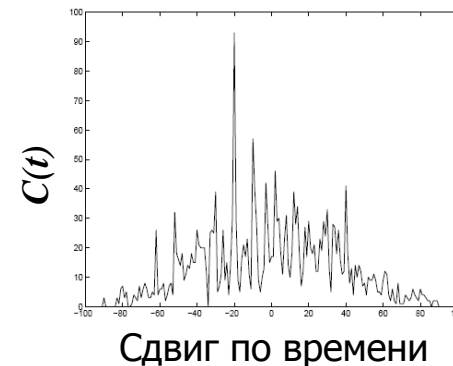
Распределение точек интереса
для левого ракурса



Распределение точек интереса
для правого ракурса



Корреляционная функция
(максимум соответствуют
найденному сдвигу по времени)



Jingyu Yan, Marc Pollefeys, "Video synchronization via space-time interest point distribution", Proceedings of Advanced Concepts for Intelligent Vision Systems, 2004

Distribution based sync

Выводы

- Достоинство
 - Скорость
- Недостаток
 - Невозможно синхронизировать видео с зацикленным движением



Contents

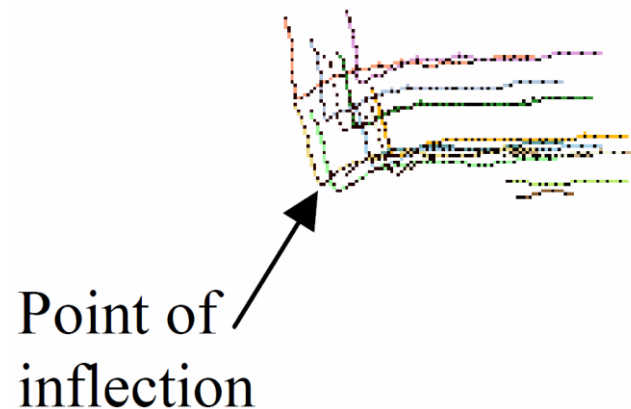
- Введение
- Зона комфортного восприятия
- Binocular luster
- Zoom синхронизация
- **Временная синхронизация**
 - Subjective quality test
 - Distribution based sync
 - **Inflection points based sync**
- Заключение

Inflection Points based sync

Введение



Point of inflection (PoI) — точка резкого изменения траектории



Идея: сопоставление не траекторий,
а ключевых точек траекторий

Inflection Points based sync

Алгоритм (1)



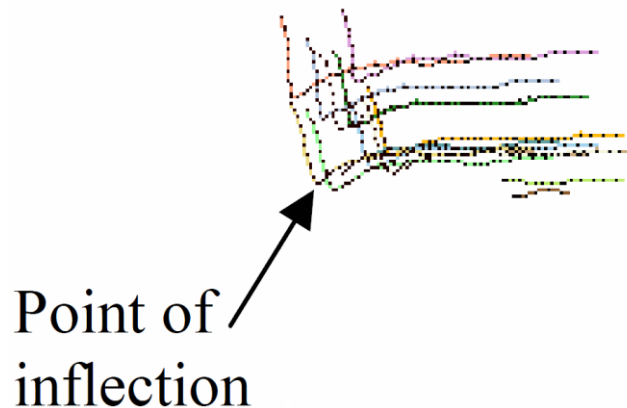
1. Вычисление PoI
2. Сопоставление на основе PoI
3. Уточнение результатов с помощью эпиполярной геометрии

Inflection Points based sync

Алгоритм (2)

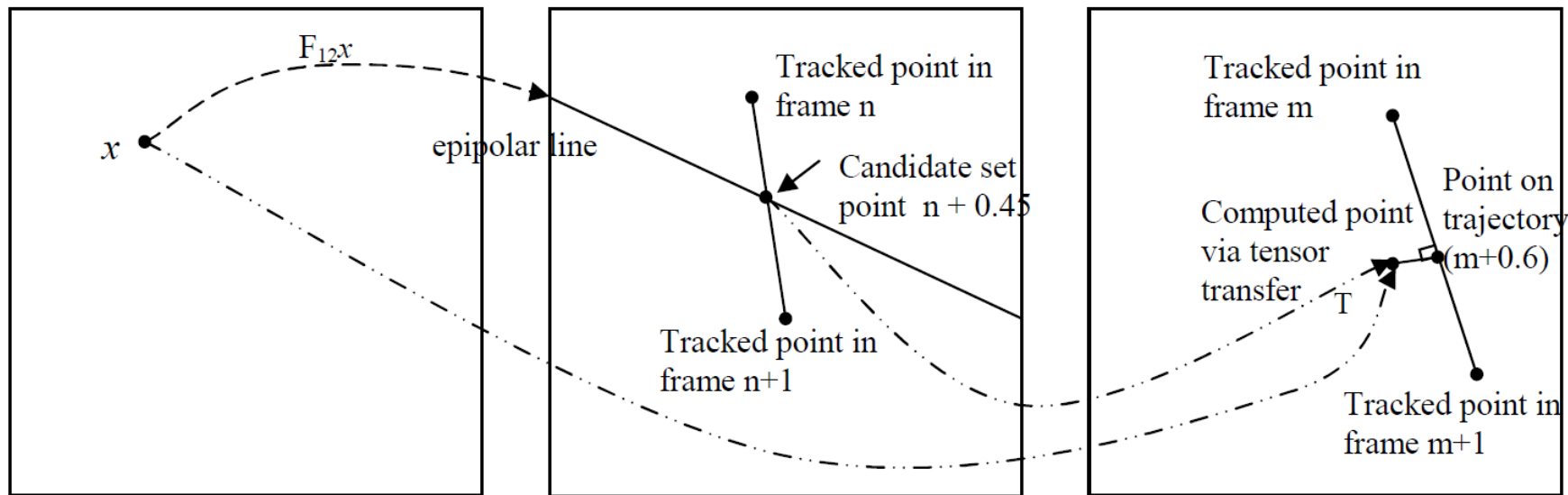


Авторы используют feature tracker для построения траекторий, по которым вычисляются PoI



Inflection Points based sync

Алгоритм (3)



Уточнение результатов

Inflection Points based sync

Результаты (1)



Camera (#)	Gross sync (frame #)	Gross shift	Precise sync (frame #)	Precise full frame shift	Ground Truth (shift)
1	141	0	141	0	0
2	146	5	146.05	5	5
3	151	10	150.97	10	10

Результаты алгоритма на синтезированных данных

Inflection Points based sync

Результаты (2)



Camera (#)	Gross sync (frame #)	Precise sync (frame #)	Manual sync (frame #)
1	127	126.75	127
2	228	229.33	229
3	160	159	159

Результаты алгоритма на реальных данных

Inflection Points based sync

Выводы



- Достоинства
 - Отсутствие детерминированной ошибки сопоставления траекторий
 - Возможен переход от точек к областям (использование ME или OF)
- Недостаток
 - Point-based алгоритм



Содержание

- Введение
- Зона комфортного восприятия
- Binocular luster
- Zoom синхронизация
- Временная синхронизация
- **Заключение**

Заключение

Мои результаты

Stereo Quality Assessment Tool (SQAT) =
= VirtualDub filter + MATLAB script

Осуществляется контроль:

- максимального параллакса
- пространственной симметрии
- соответствия по цвету

Заключение

Мои результаты (2)

Контроль максимального параллакса



Заключение

Мои результаты (2*)

Контроль максимального параллакса



Channel mismatch

Заключение

Мои результаты (3)

Контроль максимального параллакса

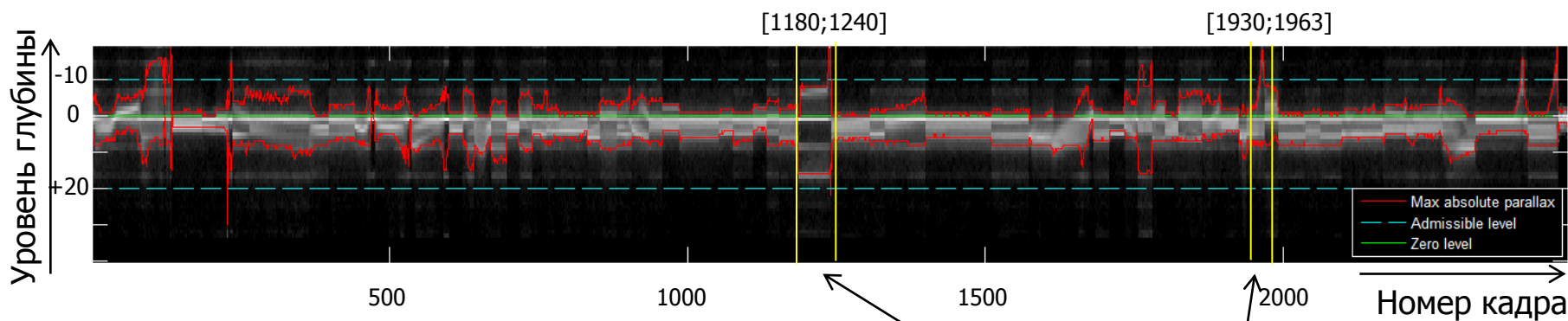


График распределения количества пикселей по уровням глубины

Максимальный параллакс

Заключение

Мои результаты (4)

Контроль максимального параллакса



Заключение

Мои результаты (5)

Контроль параллакса

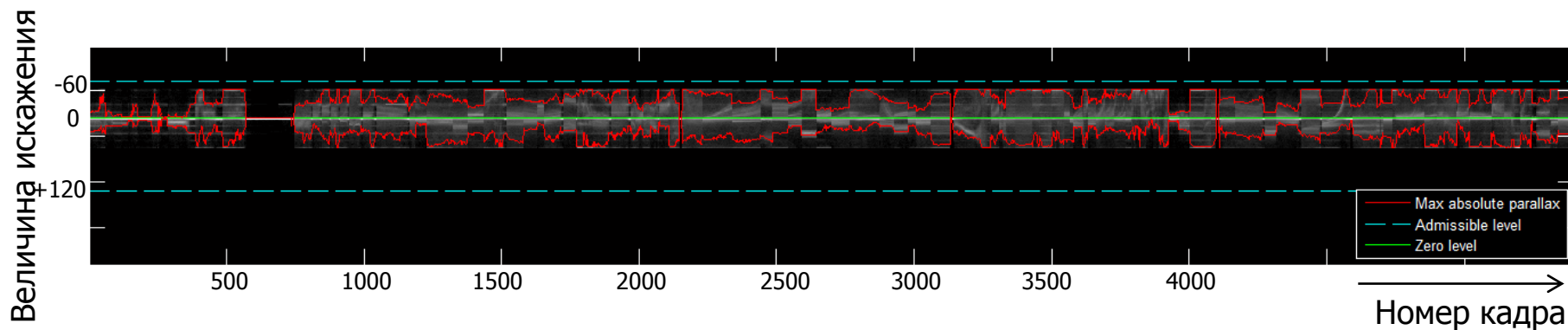


График распределения количества пикселей по уровням искажений

Заключение

Мои результаты (6)

Контроль пространственной симметрии

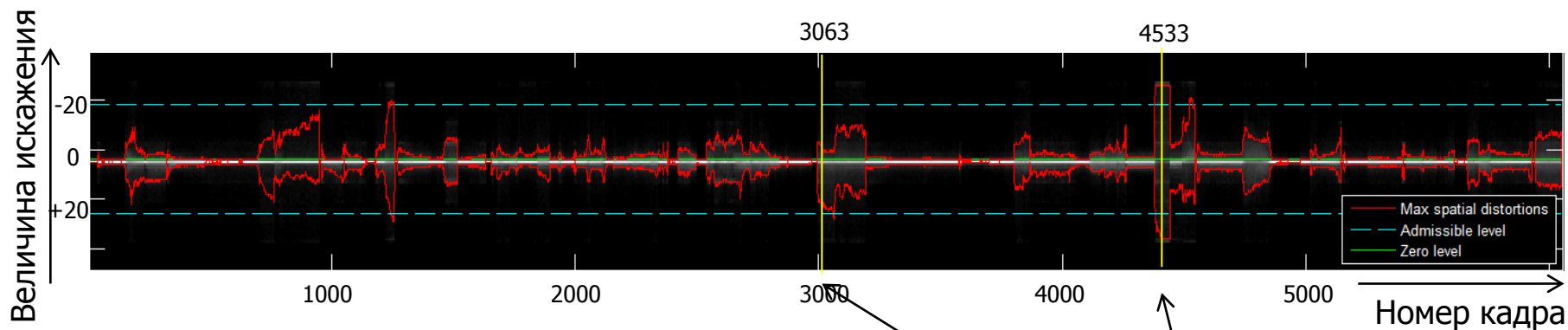


График распределения количества пикселей по уровням искажений

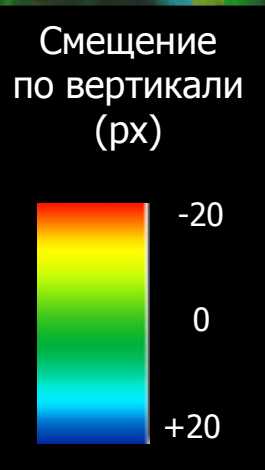
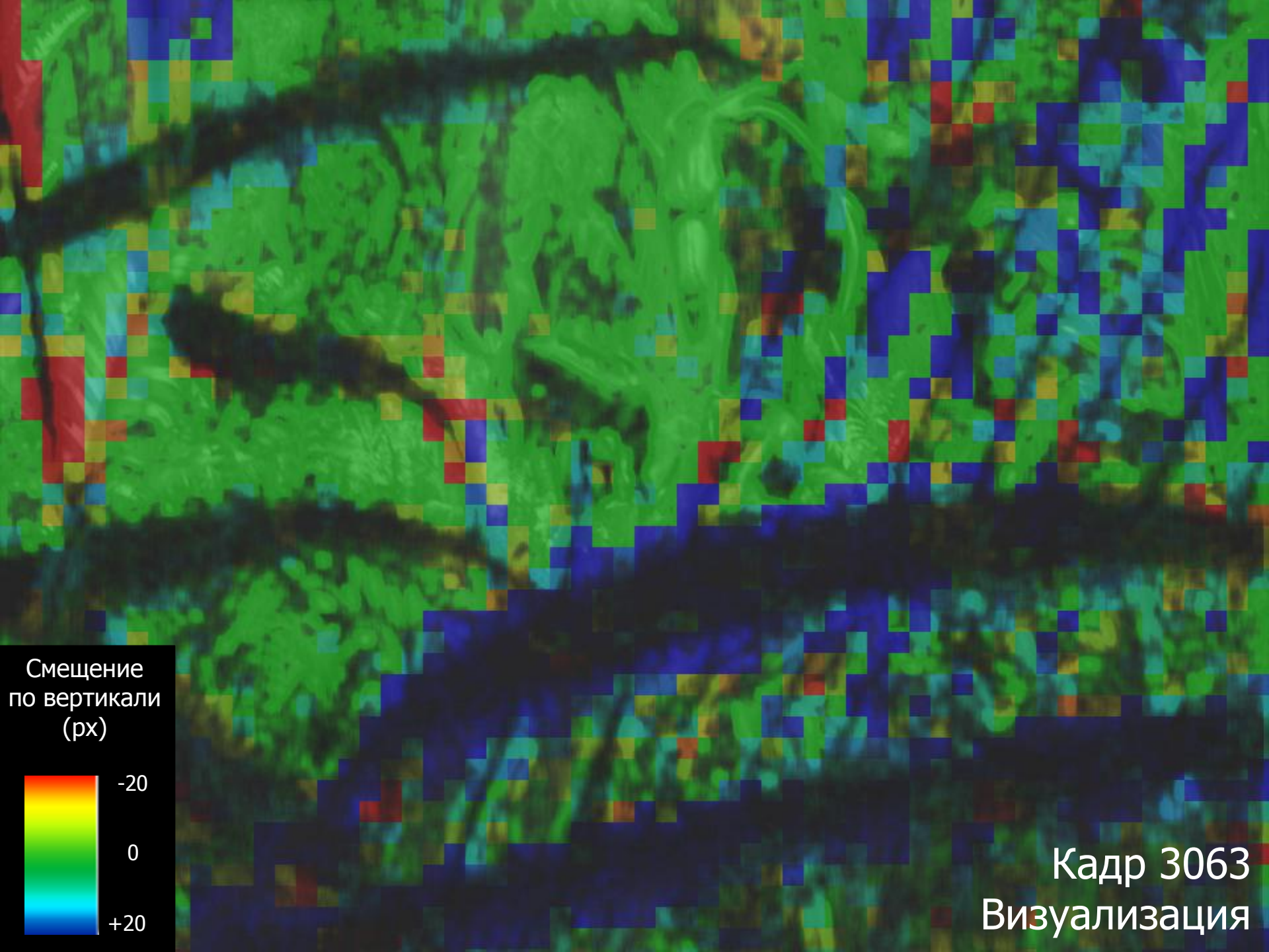
**Максимальные
искажения**



Кадр 3063
Левый ракурс



Кадр 3063
Правый ракурс



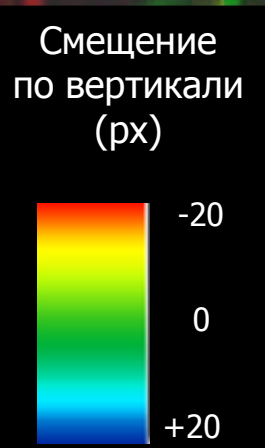
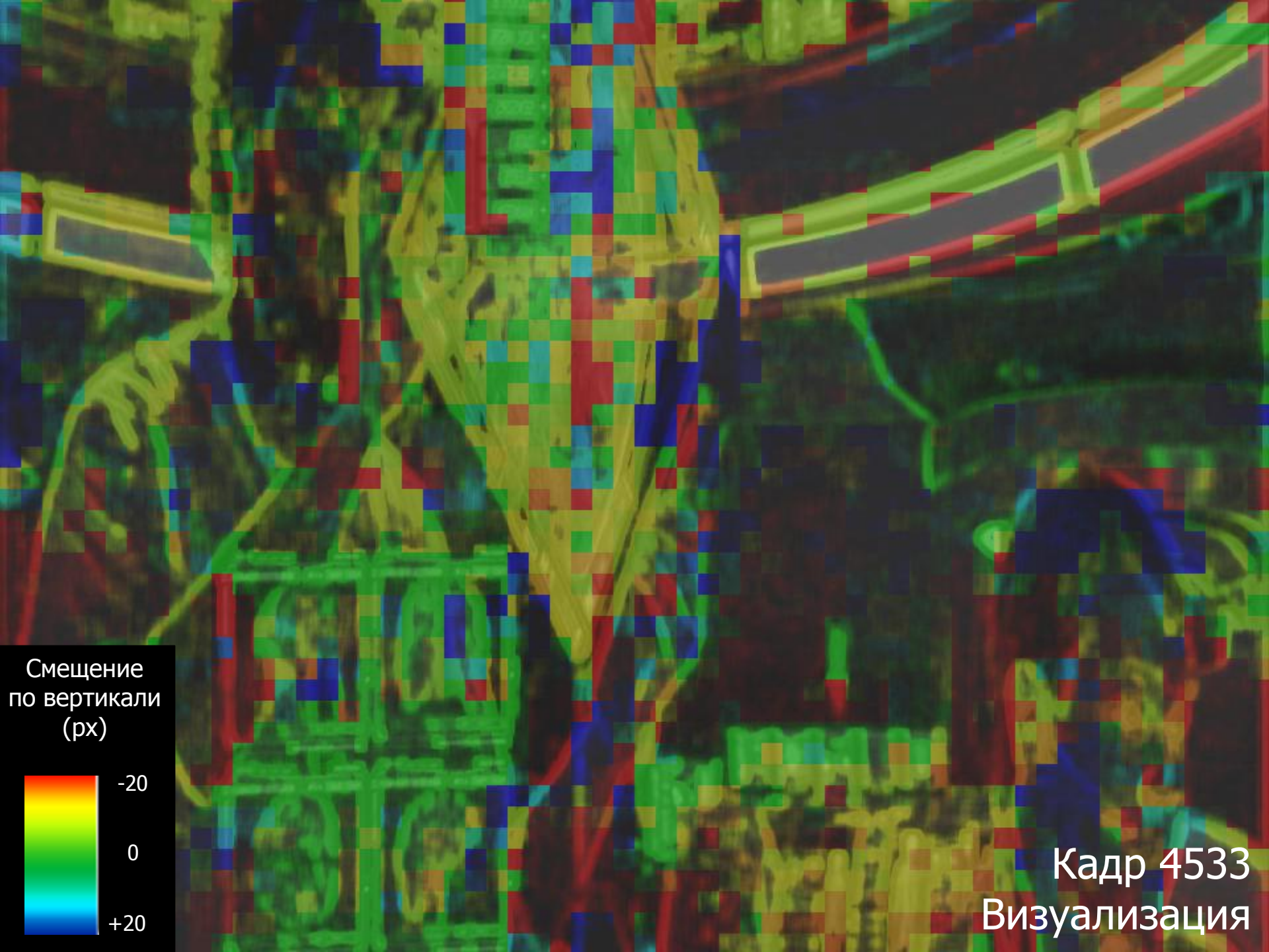
Кадр 3063
Визуализация



Кадр 4533
Левый ракурс



Кадр 4533
Правый ракурс



Кадр 4533
Визуализация

Заключение

Мои результаты (13)

Контроль соответствия по цвету

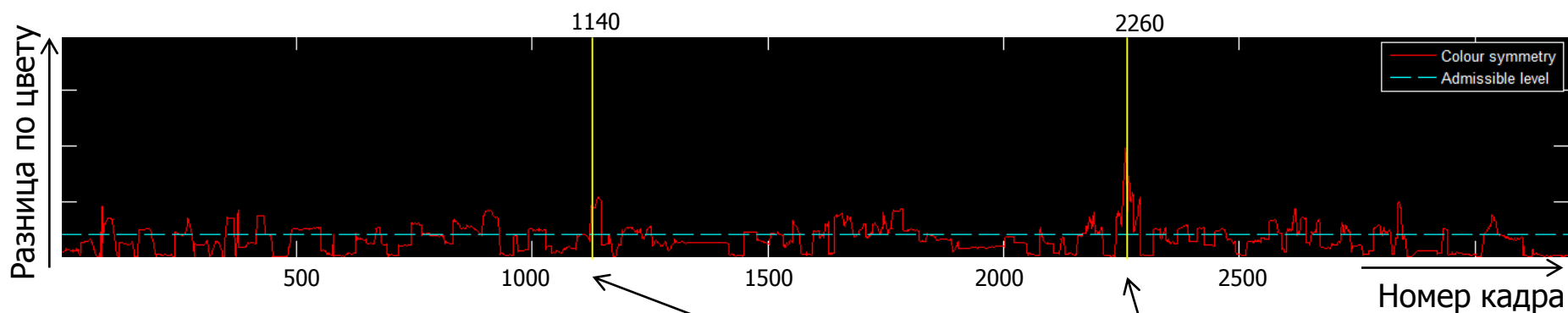


График соответствия/несоответствия
по цвету между кадрами

**Максимальные
искажения**



Кадр 1140
Левый ракурс



Кадр 1140
Правый ракурс



Кадр 2260
Левый ракурс



Кадр 2260
Правый ракурс



Заключение

Были освещены:

- Базовые артефакты стерео видео
- 3 субъективных тестирования по влиянию различных артефактов
- Алгоритм детектирования binocular luster
- 3 алгоритма синхронизации стереопары

Заключение

Дальнейшие планы (1)

- **Временная синхронизация**
 - Реализовать алгоритм на основе распределения space-time interest points
 - Разработать алгоритм, основанный на идее points of inflection (PoI)
- **Zoom синхронизация**
 - Доработать предложенный алгоритм
- **Binocular luster**
 - Разработать алгоритм, способный детектировать артефакт, возникший не обязательно при сжатии

Заключение

Дальнейшие планы (2)

- Нечестное стерео (сдвиг по времени)
 - Разработать алгоритм, определяющий некорректный способ построения стереопары
- Восприятие глубины
 - Разработать алгоритм, сопоставляющий информацию о глубине, полученную из разных depth cues

Литература (1)

1. W. Chen, J. Fournier, M. Barkowsky, P. Le Callet, "New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone", Society of Photographic Instrumentation Engineers, 2011
2. Can Bal, Ankit K. Jain, Truong Q. Nguyen "Detection and removal of binocular luster in compressed 3D images", International Conference on Acoustics, Speech, and Signal Processing, 2011
3. Colin Doutre, Mahsa T. Pourazad, Alexis M. Tourapis, et al. "Correcting Unsynchronized Zoom in 3D Video", International Symposium on Circuits and Systems, 2011
4. Lutz Goldmann, Jong-Seok Lee, Touradj Ebrahimi, "Temporal synchronization in stereoscopic video: influence on quality of experience and automatic asynchrony detection", the International Conference on Image Processing, 2010

Литература (2)

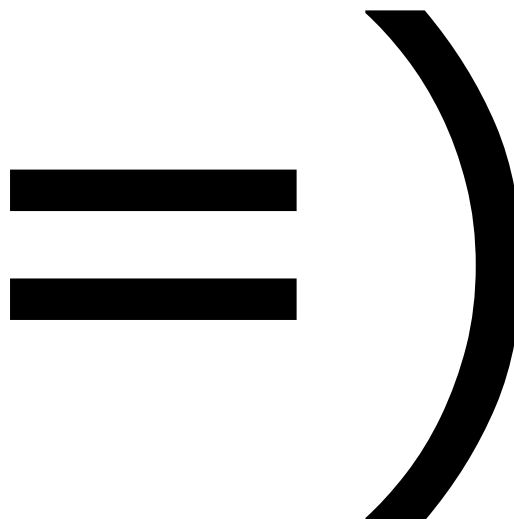
5. Jingyu Yan, Marc Pollefeys, "Video synchronization via space-time interest point distribution", Proceedings of Advanced Concepts for Intelligent Vision Systems, 2004
6. Anthony Whitehead, Robert Laganierie et al. "Temporal Synchronization of Video Sequences in Theory and in Practice", Motion and Video Computing, Workshop on Applications of Computer Vision, 2005
7. Andrew Woods, Tom Docherty, Rolf Koch, "Image Distortions in Stereoscopic Video Systems", Processing of Society of Photographic Instrumentation Engineers, 1993
8. Atanas Boev, Danilo Hollosi, Atanas Gotchev, "Classification of stereoscopic artefacts", MOBILE3DTV technical report, 2008
9. Burazerovic et al., "Automatic depth profiling of 2D cinema and photographic images", the International Conference on Image Processing , 2009

Литература (3)

10. Stephan Reichelt, et al., "Depth cues in human visual perception and their realization in 3D display", Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV
11. Bernard Mendiburu, "Tutorial on Stereoscopic Imaging", Digital Cinema Summit, National Association of Broadcasters, 2009



А где слайд с большим знаком вопроса?



Лаборатория компьютерной графики и мультимедиа



Видеогруппа — это:

- Выпускники в аспирантурах Англии, Франции, Швейцарии (в России в МГУ и ИПМ им. Келдыша)
- Выпускниками защищено 5 диссертаций
- Наиболее популярные в мире сравнения видеокодеков
- Более 3 миллионов скачанных фильтров обработки видео