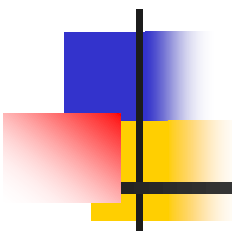


# Обзор подходов к измерению качества 3D видео



---

Денис Сумин

*Video Group*

*CS MSU Graphics & Media Lab*

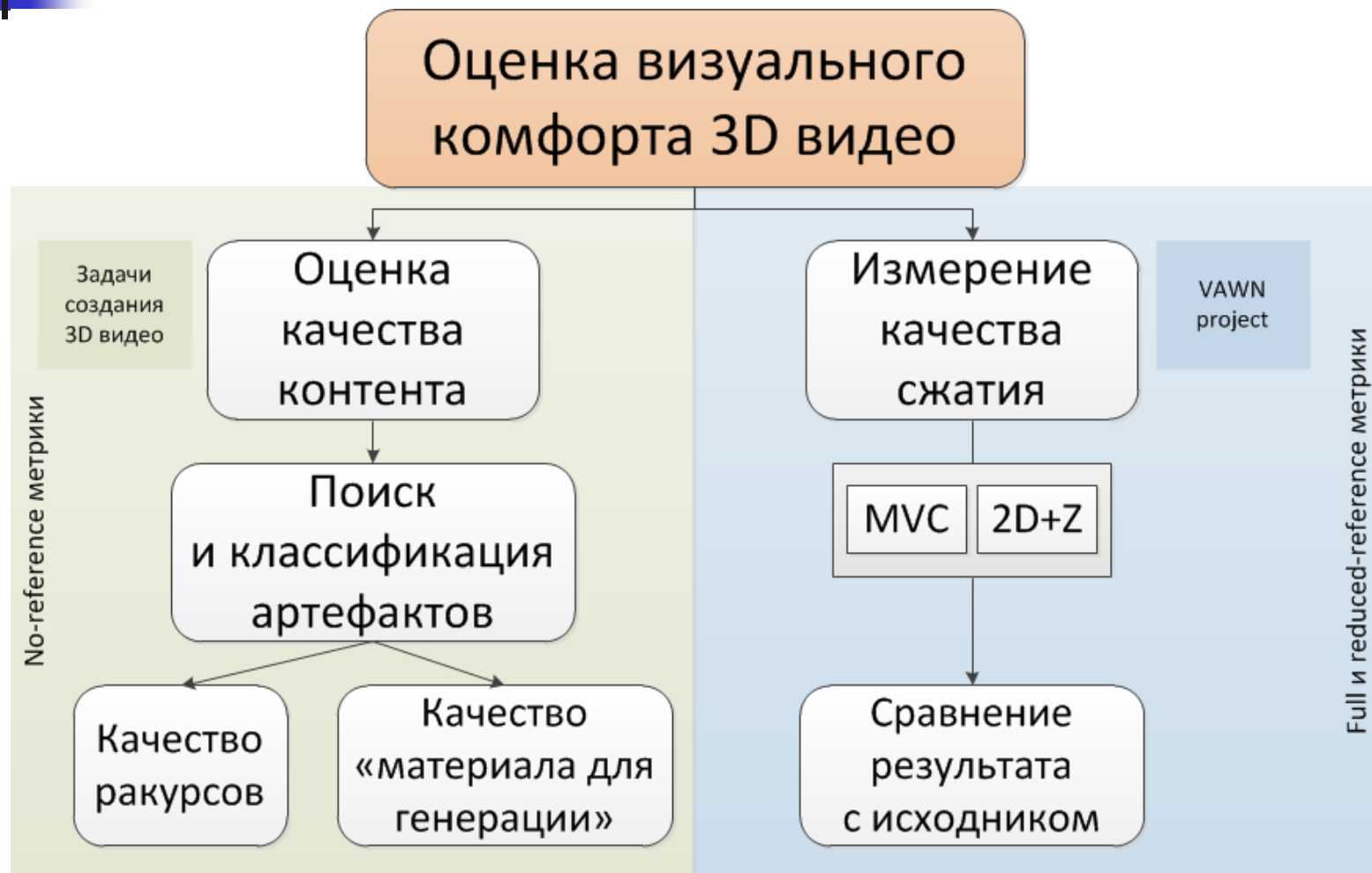


# Содержание

---

- **Введение**
- No-reference метрики
  - Подсчет статистик из карт глубины и движения
  - Оценка согласованности ракурсов стереопары
- Reference метрики
  - Full-reference метрика на основе 3D-DCT
  - Reduced-reference метрика для карт глубины
- Заключение

# Классификация метрик



# Качество контента (1)

Уже умеем:

- находить расхождение ракурсов по цветам, геометрии, времени
- контролировать параллакс

(или смотрите предыдущие серии... ;-)

## Качество контента (2)

Надо находить:

- разный фокус
- проблемы на границах объектов
- проблемы в областях открытий
- простые варианты заполнения
- нестабильное во времени заполнение

Надо измерять качество карт глубины  
для последующей генерации



# Качество сжатия

---

Избыточной информации много —  
хочется хорошей степени сжатия

Возможные проблемы:

- несоответствие артефактов (MVC)
- деградация краев изображения и карты глубины (2D+Z)
- полупрозрачные объекты (2D+Z)

# Содержание

- Введение
- **No-reference метрики**
  - **Подсчет статистик из карт глубины и движения**
  - Оценка согласованности ракурсов стереопары
- Reference метрики
  - Full-reference метрика на основе 3D-DCT
  - Reduced-reference метрика для карт глубины
- Заключение

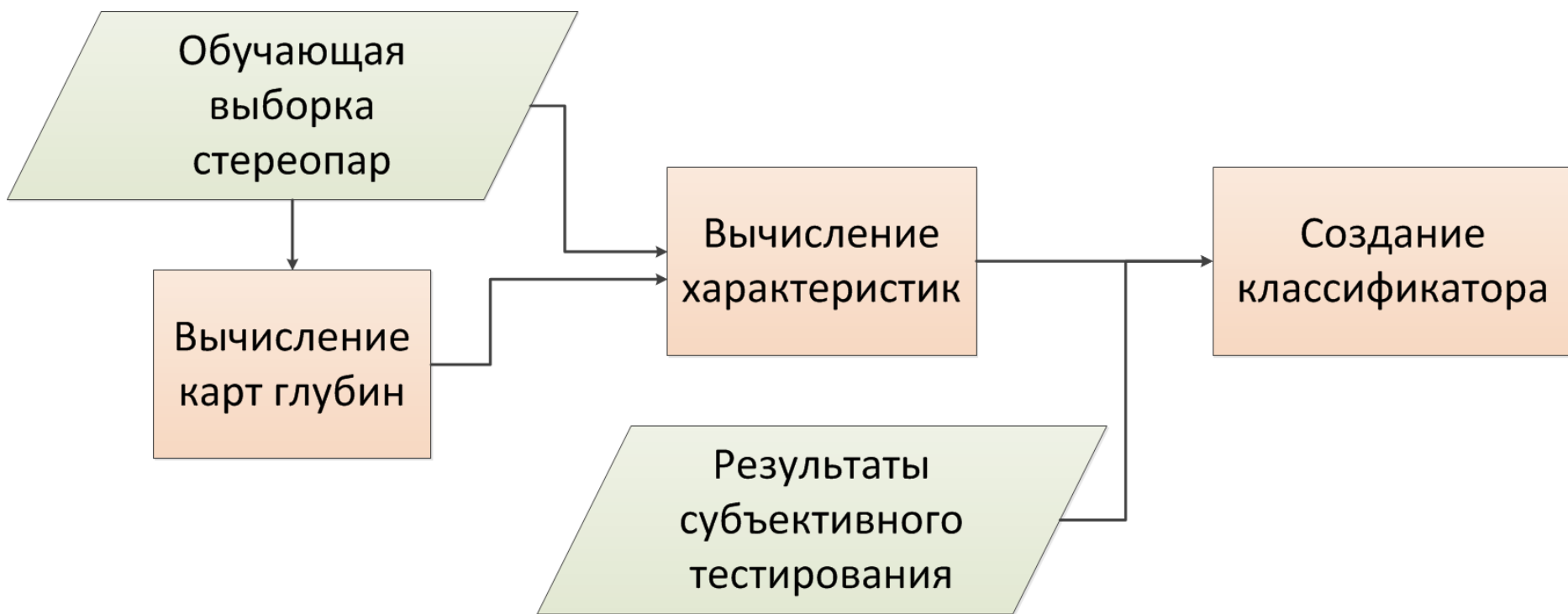
# Подсчет статистик из карт глубины и движения



- No-reference метрика определения качества стереовидео
- Авторы — группа Алан Бовика (одна из самых опытных групп в измерении качества видео)



# Общая схема



Дальше — определение качества с помощью полученного классификатора

# Набор характеристик (1)

Из карты глубины  $D$ :

- Среднее значение  $\mu = E[D]$
- Медиана  $med = median(D)$
- Стандартное отклонение  $\sigma = \sqrt{E[(D - \mu)^2]}$

- Коэффициент эксцесса

$$\kappa = E[(D - \mu)^4] / (E[(D - \mu)^2])^2$$

- Коэффициент асимметрии

$$skew = E[(D - \mu)^3] / (E[(D - \mu)^2])^{(3/2)}$$

## Набор характеристик (2)

Теперь все те же, только над производной  $\delta D$ :  
(к глубине применили оператор Лапласа)

- Среднее значение  $\mu_d = E[\delta D]$
- Стандартное отклонение  $\sigma_d = \sqrt{E[(\delta D - \mu_d)^2]}$
- Коэффициент эксцесса  
$$\kappa_d = E[(\delta D - \mu_d)^4] / (E[(\delta D - \mu_d)^2])^2$$
- Коэффициент асимметрии  
$$skew_d = E[(\delta D - \mu_d)^3] / (E[(\delta D - \mu_d)^2])^{(3/2)}$$



# Spatial Indicator (1)

---

## Характеристики:

- Потеря пространственной информации (blurring)
- Сдвиг вертикальных и горизонтальных границ в диагональном направлении и обратно
- Изменения по цвету
- Мера улучшения качества (например, увеличение резкости)
- Движение в блоке (например, дрожащая граница)

## Spatial Indicator (2)

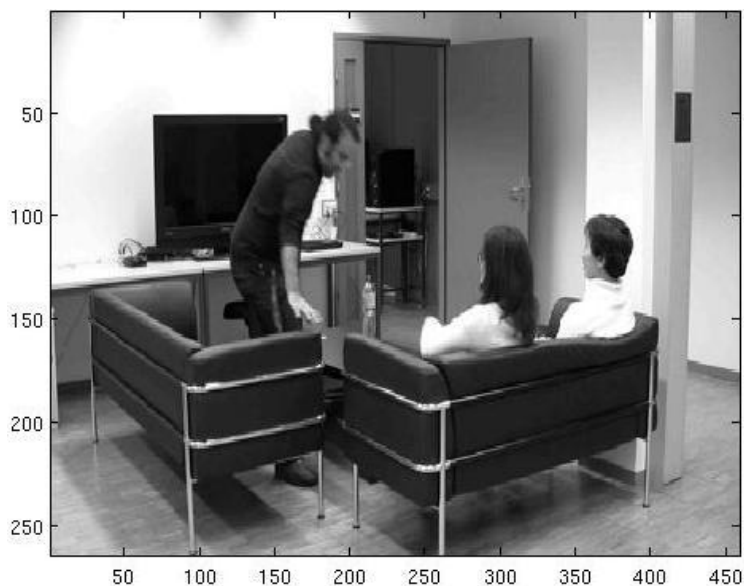
Характеристики суммируются с весами:

$$\begin{aligned} \text{VQM} = & - 0.2097 * \text{si\_loss} + 0.5969 * \text{hv\_loss} + 0.2483 * \text{hv\_gain} \\ & + 0.0192 * \text{chroma\_spread} - 2.3416 * \text{si\_gain} + 0.0431 \\ & * \text{ct\_ati\_gain} + 0.0076 * \text{chroma\_extreme} \end{aligned}$$

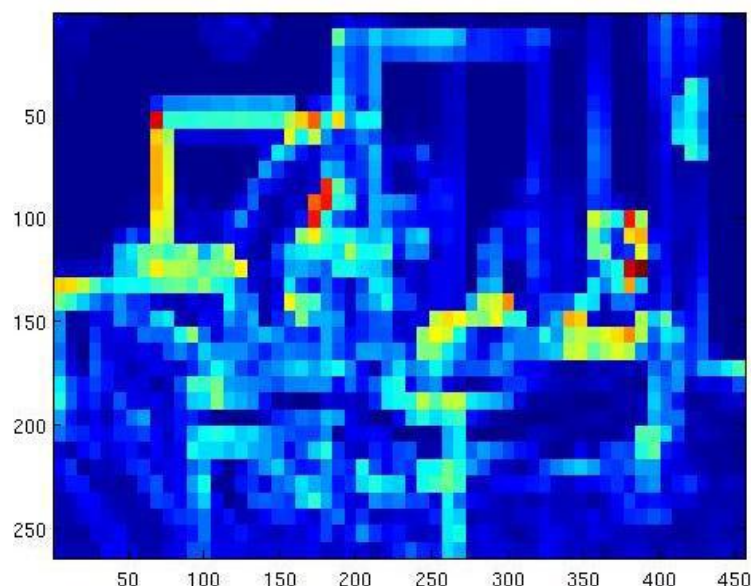
Добавляем вычисление градиента  
по изображению и дисперсии  
непересекающихся блоков 8x8 изображения

# Spatial Activity

Получаем **Spatial Indicator**. Для него считаем всё то же, что и для предыдущих двух наборов



Левый ракурс стереопары



Spatial Activity Map

# Характеристики для видео

- Информация ME для ракурсов и для карт глубины
- Снова медиана, коэффициенты эксцесса и асимметрии для значений межкадровой разницы карт глубины

# Машинное обучение

Пространство: 15 признаков для изображений  
и еще 10 — для видео

Уменьшение размерности:

- PCA — Principal Component Analysis
- FFS — Forward Feature Selection

Соотношение с результатами  
субъективных тестирований:

$$y = \vec{x}^T \vec{\gamma} + \eta$$

$\vec{x}$  — вектор характеристик  
 $\vec{\gamma}$  — вектор весов характеристик  
 $\eta$  — КОНСТАНТА



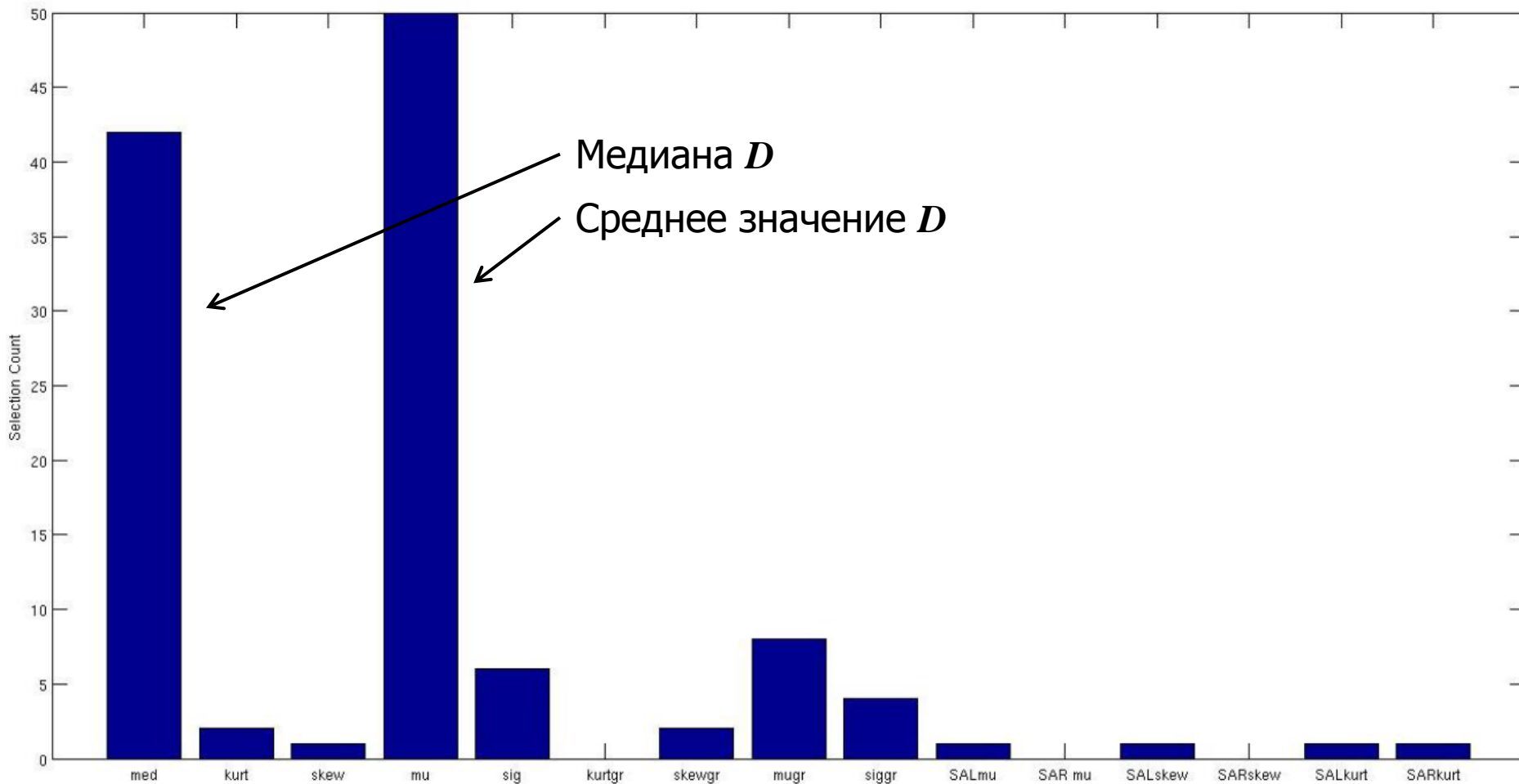
# Результаты для изображений (1)

Method	Mean	Standard deviation
PCA	0.79	0.08
FFS	0.86	0.11

3D Image QoE: значения Spearman's Rank Ordered Correlation Coefficient (SROCC)

QoE — Quality of Experience

# Результаты для изображений (2)

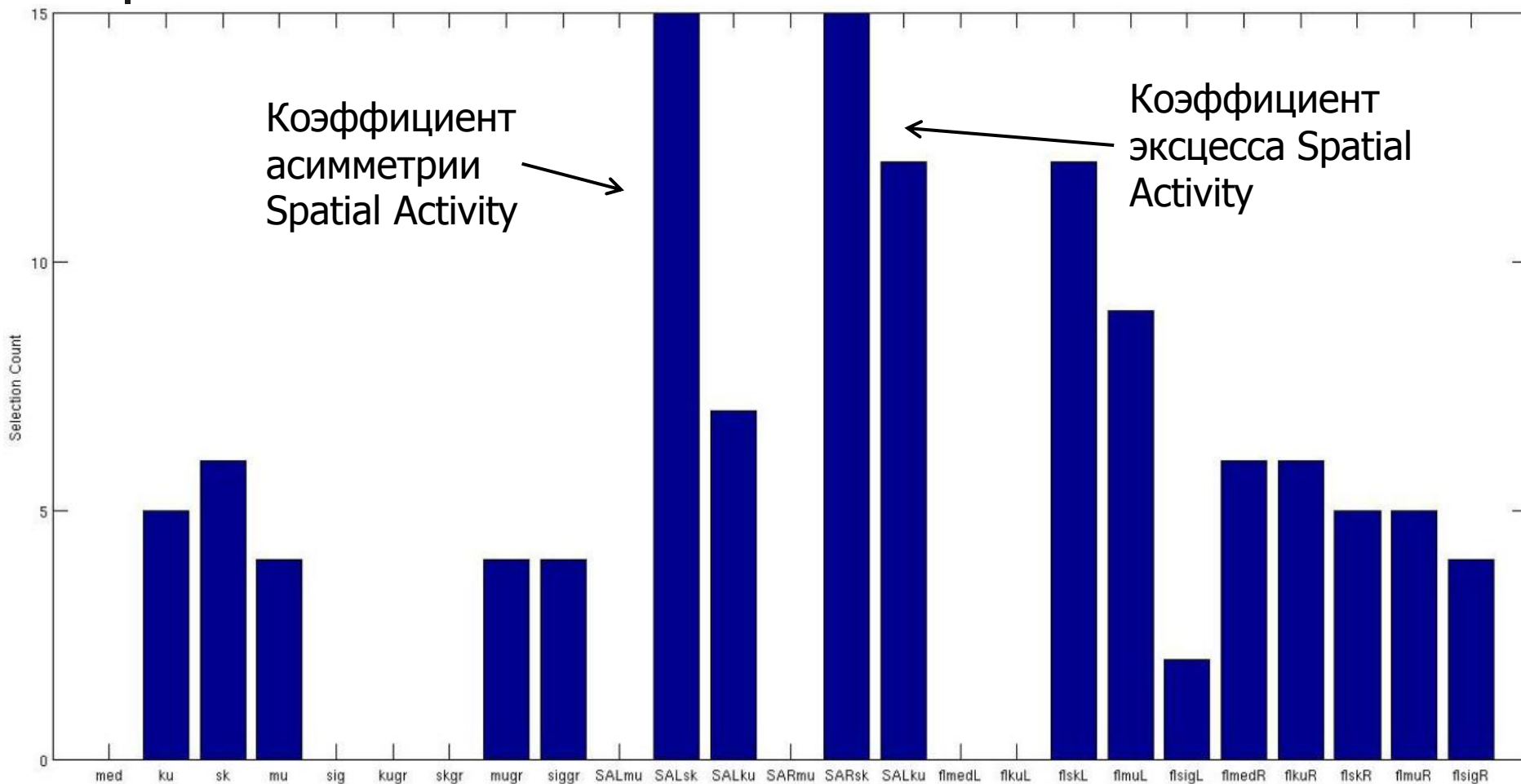


# Результаты для видео (1)

Method	Mean	Standard deviation
PCA	0.76	0.25
FFS	0.68	0.28

3D video QoE: значения SROCC

# Результаты для видео (2)





# Выводы

---

- Достоинства
  - Интересные features для машинного обучения
  - Хорошие результаты
- Недостатки
  - Нет классификации артефактов

# Содержание

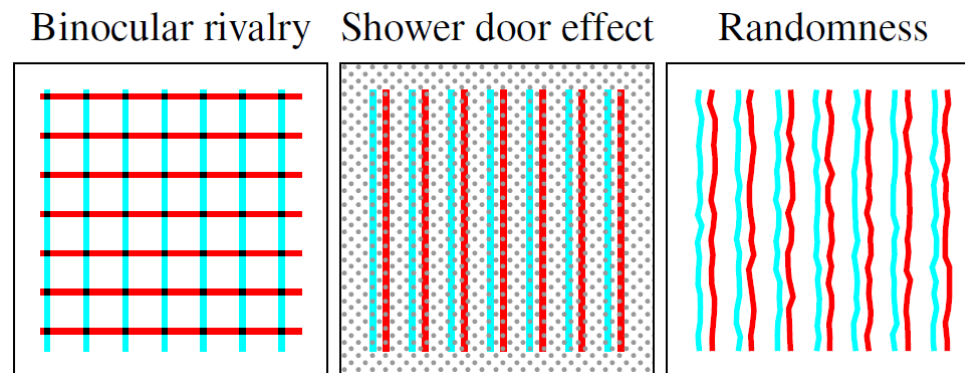
- Введение
- **No-reference метрики**
  - Подсчет статистик из карт глубины и движения
  - **Оценка согласованности ракурсов стереопары**
- Reference метрики
  - Full-reference метрика на основе 3D-DCT
  - Reduced-reference метрика для карт глубины
- Заключение

# Оценка согласованности ракурсов стереопары

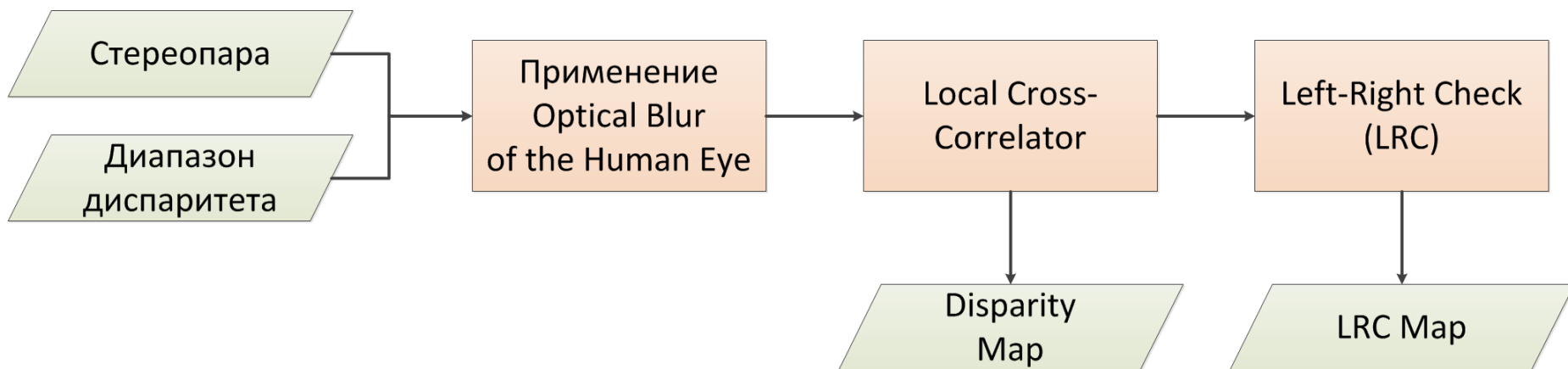


Система сравнивает ракурсы стереопары и пытается выявить следующие артефакты:

- binocular rivalry
- shower door effect
- randomness



# Схема работы



Численная оценка качества получается из анализа карты LRC (Left-Right Consistency Check)



# Пример входных данных

6 стереопар  
(3 из Middlebury)

19 фильтров  
PhotoShop

Chalk & Charcoal	Palette Knife Stamp
Halftone Pattern	Find Edges
Rough Pastels	Photocopy
Craquelure	Texturizer
Ocean Ripple	Glass
Spatter	Poster Edges
Cutout	Grain
Paint Daubs	Reticulation
Stained Glass	
Diffuse Glow	



# Optical Blur of the Human Eye



Препроцессинг:

Функция рассеяния точки (ФРТ)  
для среднестатистического глаза ( $d_{\text{зрачок}} = 3$  мм)

$$h(x, y) = a \cdot g_{s_1}(x, y) + (1 - a) \cdot g_{s_2}(x, y),$$

где  $g_s$  — двумерное Гауссово размытие  
со стандартным отклонением  $s$ :

$$g_s(x, y) = (2\pi s^2)^{-1} \cdot e^{-(x^2 + y^2) / 2s^2},$$

$$a = 0.583, s_1 = 0.433 \text{ arcmin}, s_2 = 2.04$$

# Local Cross-Correlator (1)

Корреляция между окнами с диспаратетом  $d$ :

$$c(d) = \frac{\sum_{(x,y) \in W} (L(x,y) - \mu_L) (R(x-d,y) - \mu_R)}{\sqrt{\sum_{(x,y) \in W} (L(x,y) - \mu_L)^2} \sqrt{\sum_{(x,y) \in W} (R(x-d,y) - \mu_R)^2}}$$

$L$  и  $R$  — интенсивности пикселей левого и правого ракурсов

$\mu_L$  и  $\mu_R$  — средние интенсивности пикселей внутри окна

$W$  — размер окна. Для нашей задачи  $W=5$  (пикселей)

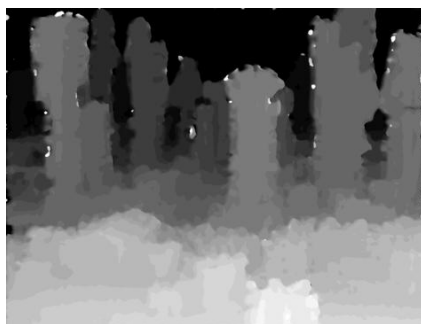
# Local Cross-Correlator (2)

Для построения disparity map применяется принцип winner-take-all

$$d_L(x, y) = \arg \max_{d \in D} c(d)$$



Source



$d_L$



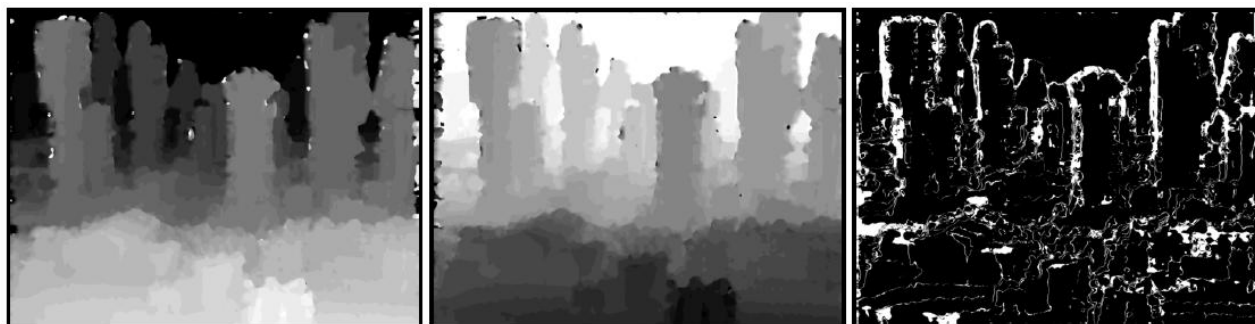
$d_R$

# Left-Right Check (LRC)

$d_L$  — disparity map левого ракурса

$d_R$  — disparity map правого ракурса

$$\left| d_L(x, y) + d_R(x - d_L(x, y), y) \right| < T_{LRC}, T_{LRC} = 1$$

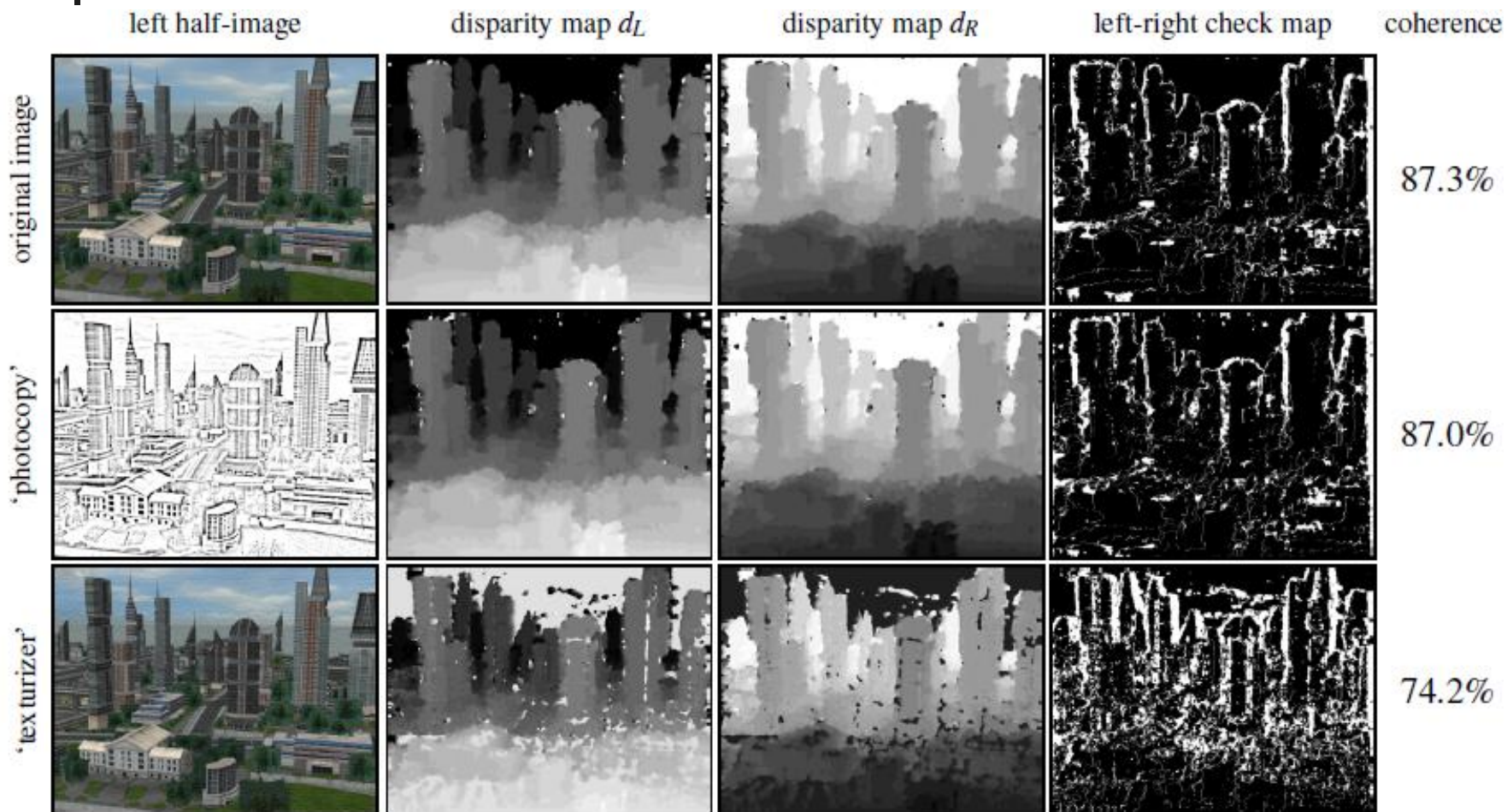


$d_L$

$d_R$

LRC map

# Что получается?



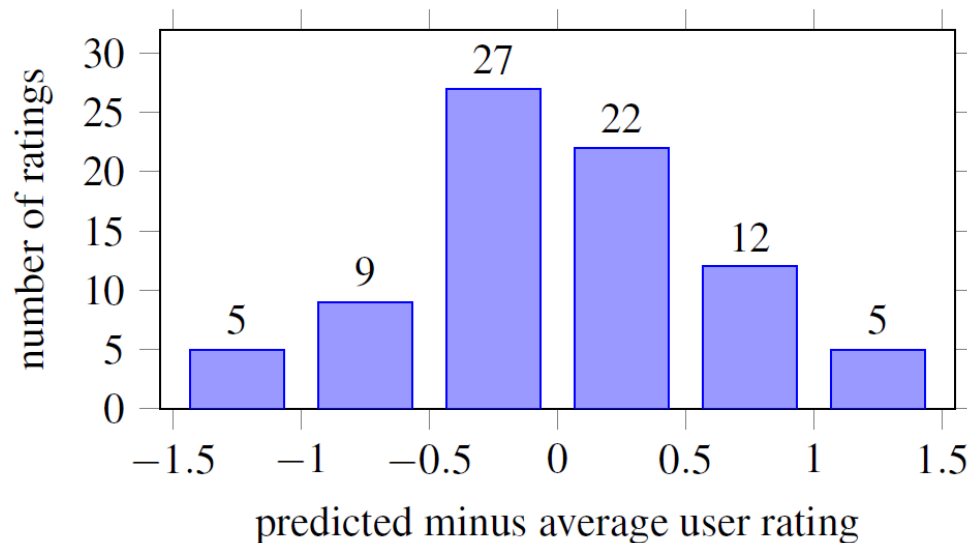
# Сравнение с субъективным тестированием



Вводится соотношение между изначальными данными и результатами субъективного тестирования

$$x' = 4.36 \cdot x - 0.07$$

Утверждается  
точность  
оценки в 88 %



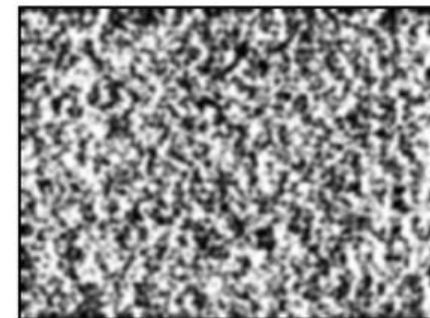
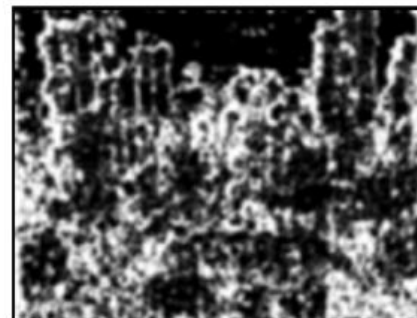
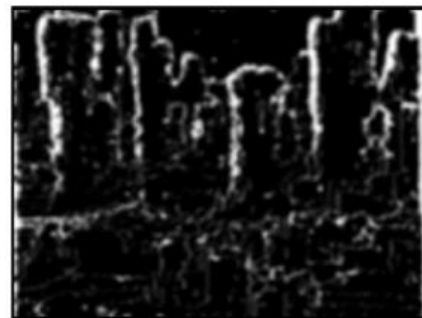
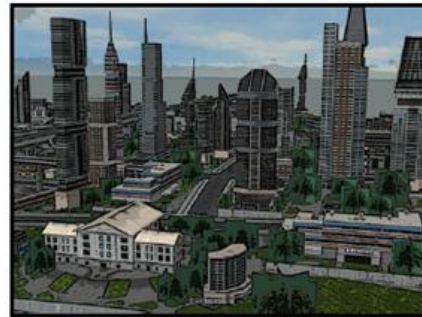
# Применение (1) Binocular Rivalry

Original image

“Poster Edges”

“Palette knife”

“Stained glass”



Left half-image

Smoothed LRC map

К LRC-карте применяется размытие, чтобы уменьшить влияние артефактов в областях occlusion



# Original image



# “Pallette knife” effect



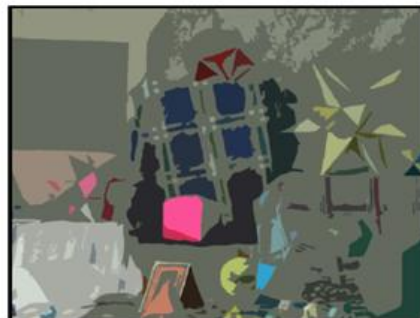
# Применение (2)

## Shower Door Effect

Original image



“Cutout”



“Glass”

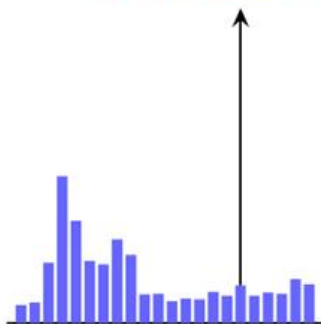


“Texturizer”

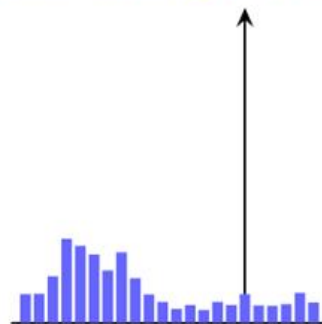


Left half-image

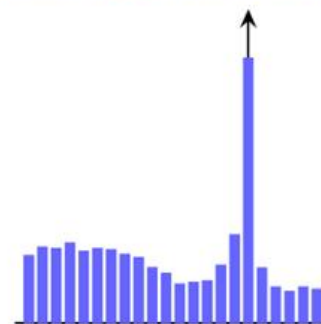
Number of pixels



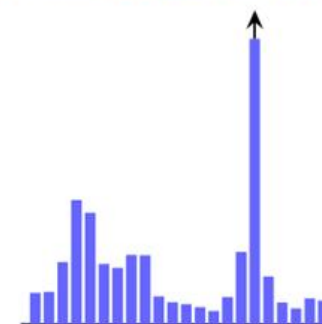
Disparity



Disparity



Disparity



Disparity

Для карт диспаратетов строятся гистограммы. Пики вдоль оси  $O_y$  означают проявление Shower Door Effect



# Original image



# "Glass" filter



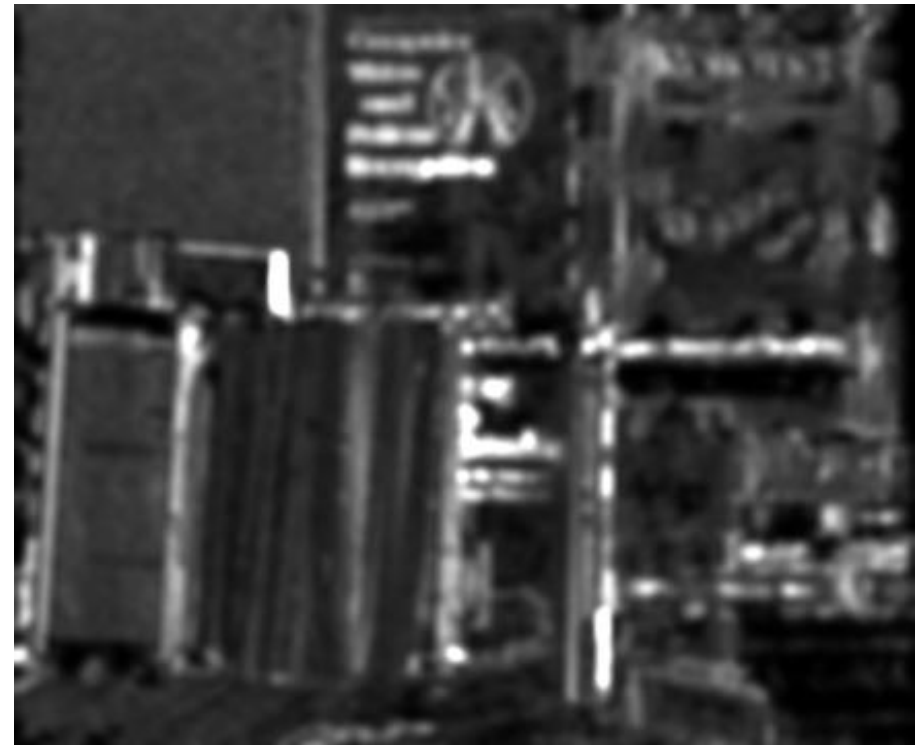
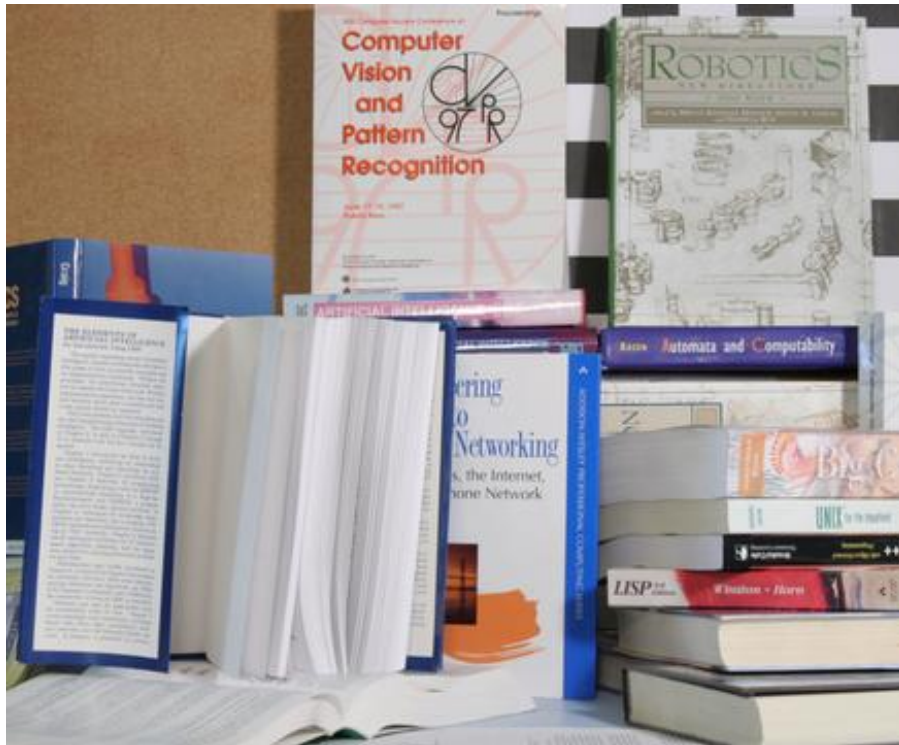
# Применение (3)

## Randomness

Левый ракурс

Original image

Image cross-check



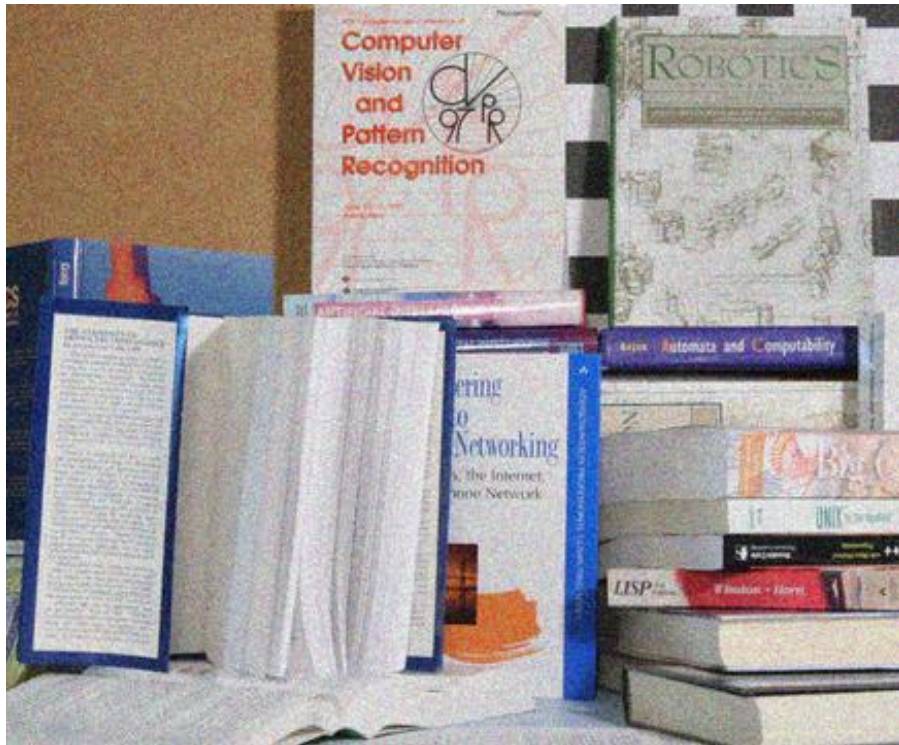
# Применение (3)

## Randomness

“Grain” filter

Левый ракурс

Image cross-check





# Выводы

---

- Достоинства
  - Алгоритм учитывает особенности зрительной системы человека
  - Предполагается возможность нахождения конкретных артефактов
- Недостатки
  - Результаты работы представлены на синтетических изображениях





# Содержание

---

- Введение
- No-reference метрики
  - Подсчет статистик из карт глубины и движения
  - Оценка согласованности ракурсов стереопары
- **Reference метрики**
  - **Full-reference метрика на основе 3D-DCT**
  - Reduced-reference метрика для карт глубины
- Заключение

# Full-reference метрика на основе 3D-DCT

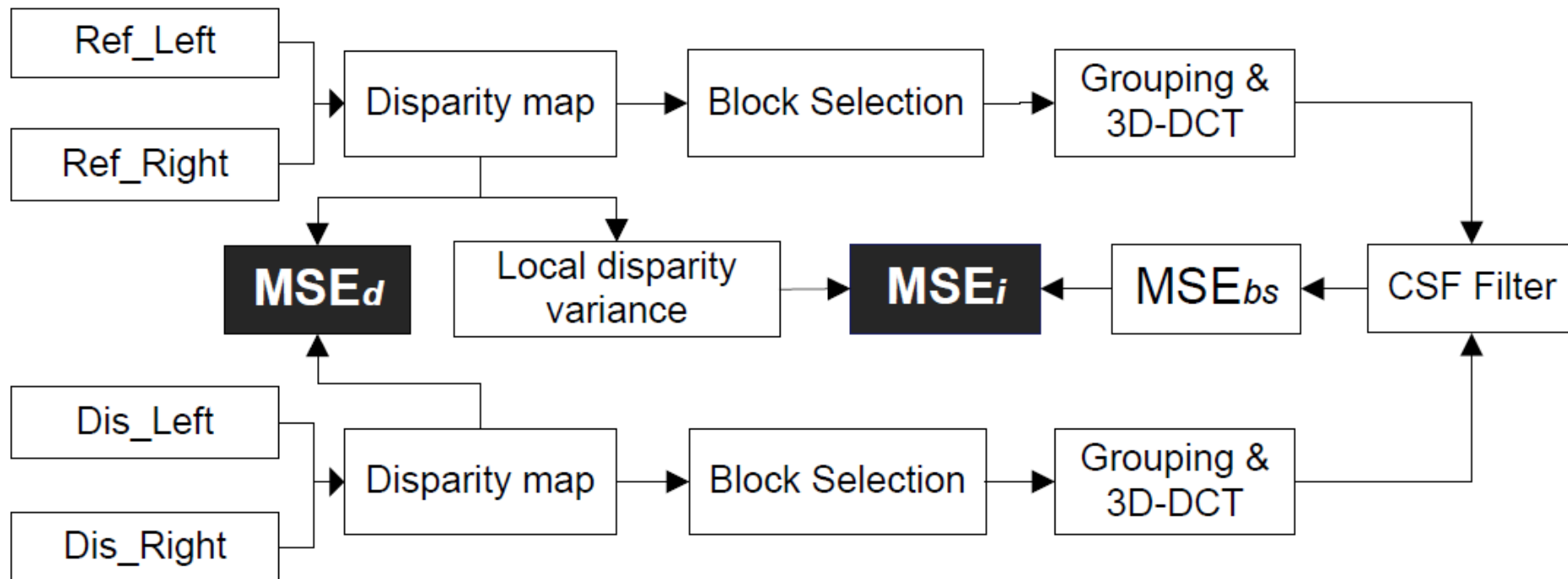


3D-DCT — трехмерное дискретное косинусное преобразование

Авторы — Tampere University of Technology.  
Они же создали одну из лучших классификаций артефактов в стереовидео

Обозначается упор на Mobile 3DTV content,  
но по сути подход достаточно общий

# Общая схема



MSE — Mean Squared Error

CSF — Contrast Sensitive Function

# Карты глубины и их оценка

Нормализация карты глубины относительно устройства для просмотра стерео: учёт комфортного параллакса

$R_{ij}$  и  $D_{ij}$  — карты глубины reference и distorted изображений после нормализации

$$MSE_d = \frac{1}{\#\Omega} \sum_{i,j \in \Omega} (R_{ij} - D_{ij})^2$$

$\#\Omega$  — мощность области (количество блоков)

# Дисперсия локального диспаратета (1)

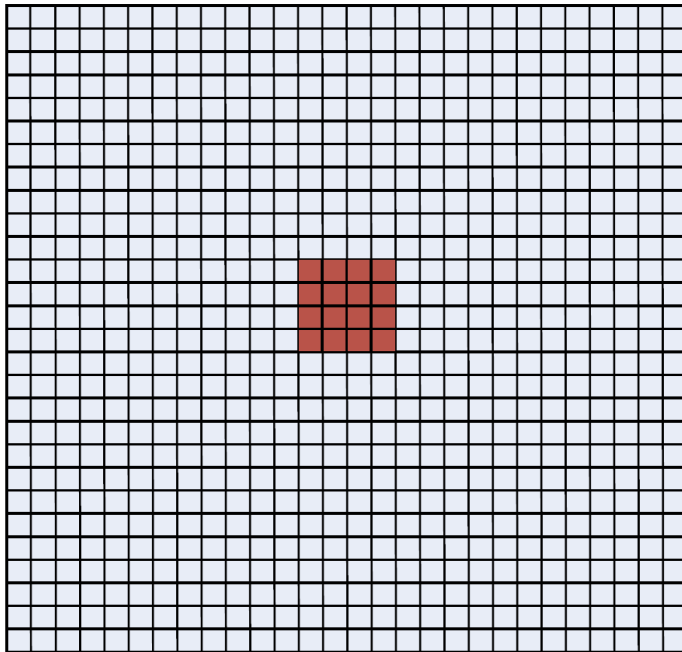
Считаем для блоков размера  $N \times M$  таких, что блок полностью проецируется на центральную ямку сетчатки глаза

$$\sigma_d^2(k) = \frac{1}{\#\Omega_k - 1} \sum_{i,j \in \Omega_k} (\mu(k) - R_{ij}(k))^2$$

$\mu(k)$  — дисперсия  $k$ -го блока  $N \times M$

# снова означает мощность

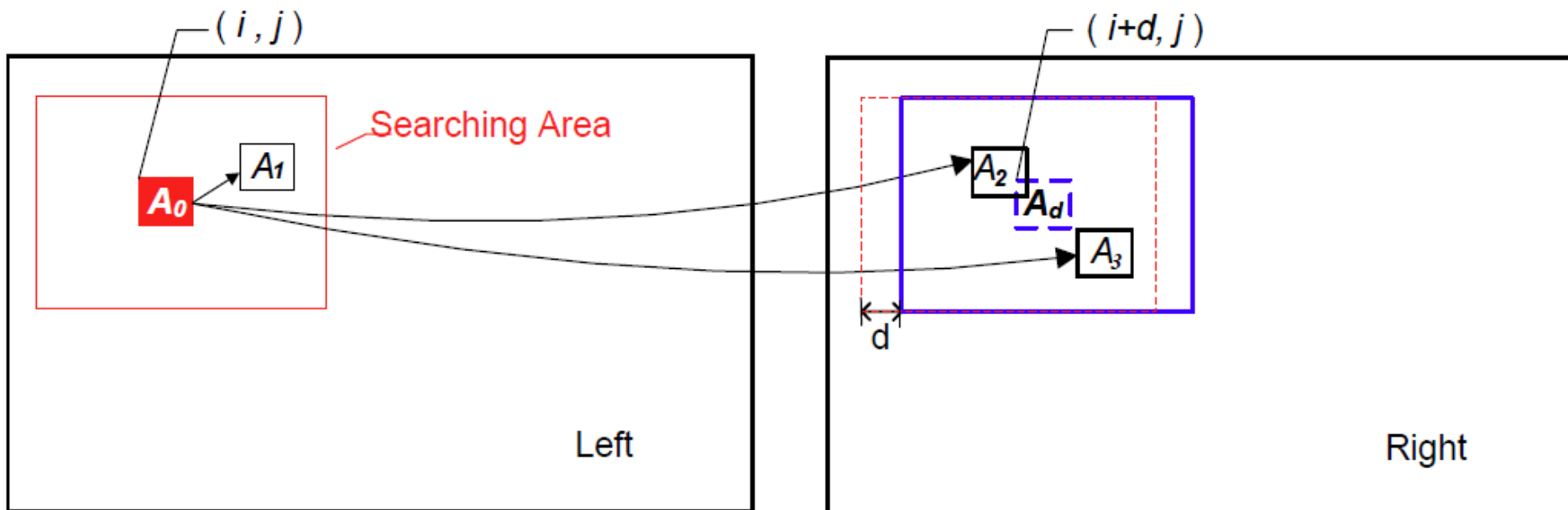
# Дисперсия локального диспаратета (2)



Локальная дисперсия характеризует изменение глубины вокруг центрального блока 4x4, что позволяет оценивать его визуальное качество

Центральный блок 4x4 и блок, для которого подсчитывается дисперсия локального диспаратета (28x28 для мобильных устройств)

# Block Selection and Matching



Получаем набор блоков  $A = \{A_0, A_1, A_2, A_3\}$  для reference изображения и соответствующий набор  $B$  для distorted изображения. Они группируются в два 3D массива

# 3D-DCT и применение CSF

4 блока размера 4x4 складываются в кубик, к ним применяется 3D-DCT преобразование

Применение Contrast Sensitivity Function:

Table I. Set of CSF masking coefficients for 4x4 DCT

1.6084	2.5735	1.0723	0.5046
1.8382	1.6084	0.6434	0.3730
1.4297	0.6955	0.3785	0.2499
0.5252	0.3299	0.2499	0.2145



# Вычисление метрики

$$MSE_{bs} = \frac{1}{64} \sum_{i,j=1}^4 \sum_{n=0}^3 \omega_n T_{i,j}^2 (U_{i,j,n} - V_{i,j,n})^2$$

$$MSE_{3dbs}(k) = MSE_{bs}(k) \frac{MSE_{bs}(k)}{MSE_{bs}(k) + \alpha \sigma_d^2(k)}$$

$U = \{U_{i,j,n}\}$ ,  $V = \{V_{i,j,n}\}$  — 3D-DCT коэффициенты  
для кубов  $A$  и  $B$ .  $i, j = 1..4$ ,  $n = 0..3$

$T_{i,j}$  — CSF-коэффициенты

$\alpha$  — весовой коэффициент

# Тестовые последовательности



Шесть последовательностей различных типов

# Результаты

Table IV Spearman correlations on 3D Video Database II

	MRSC	MVC	Simlu- cast	V+D	ALL
PSNR	0.0757	0.3278	0.3678	0.0504	0.2542
MSSIM	<b>0.7313</b>	0.7270	0.7130	0.6557	0.6327
SSIM	0.3983	0.5870	0.5426	0.2052	0.4427
UQI	0.4226	0.4513	0.3635	0.1739	0.2998
NRMSE	0.4522	0.6078	0.5609	0.3157	0.4128
PSNR-HVS	0.3991	0.4226	0.4183	0.0939	0.2266
PSNR-HVS-M	0.5209	0.5052	0.4765	0.1061	0.2449
Global Comb.	0.5617	0.8026	0.6991	0.5217	0.6455
PHSD	0.6487	<b>0.8104</b>	<b>0.7783</b>	<b>0.7374</b>	<b>0.7999</b>



# Выводы

---

- Достоинства
  - Высокие результаты на реальных последовательностях
  - Учет особенностей зрения человека
- Недостаток
  - Метод не новый, достаточно много взято из метрики для 2D видео тех же авторов

# Содержание

- Введение
- No-reference метрики
  - Подсчет статистик из карт глубины и движения
  - Оценка согласованности ракурсов стереопары
- **Reference метрики**
  - Full-reference метрика на основе 3D-DCT
  - **Reduced-reference метрика для карт глубины**
- Заключение

# Reduced-reference метрика для карт глубины

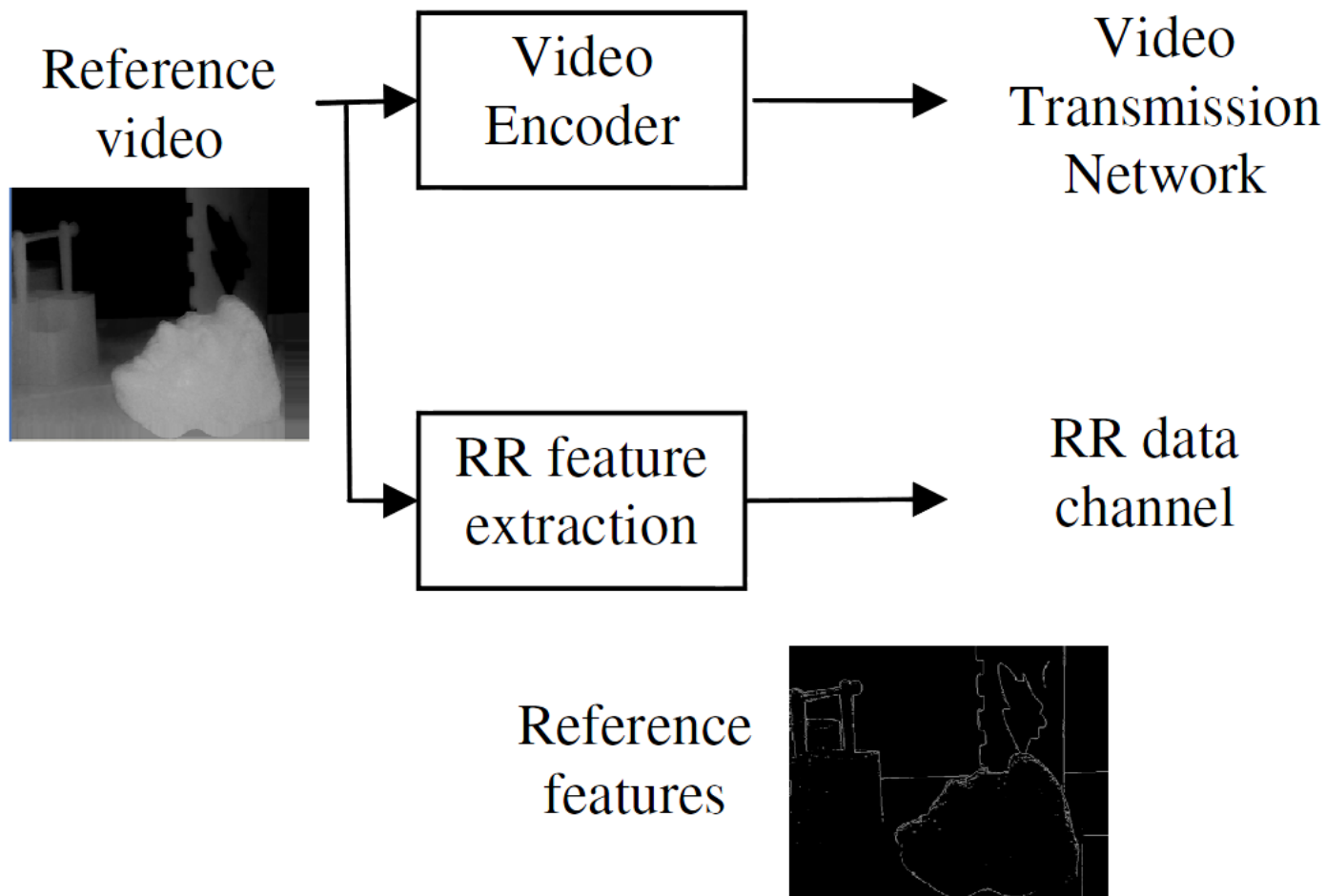


Задача: контролировать качество карты глубины после передачи через беспроводные сети

Авторы: Wireless, Multimedia and Networking Research Group из University of Kingston

Reduced-reference: для передачи reference требуется немного ресурсов

# Общая схема





# Идея

---

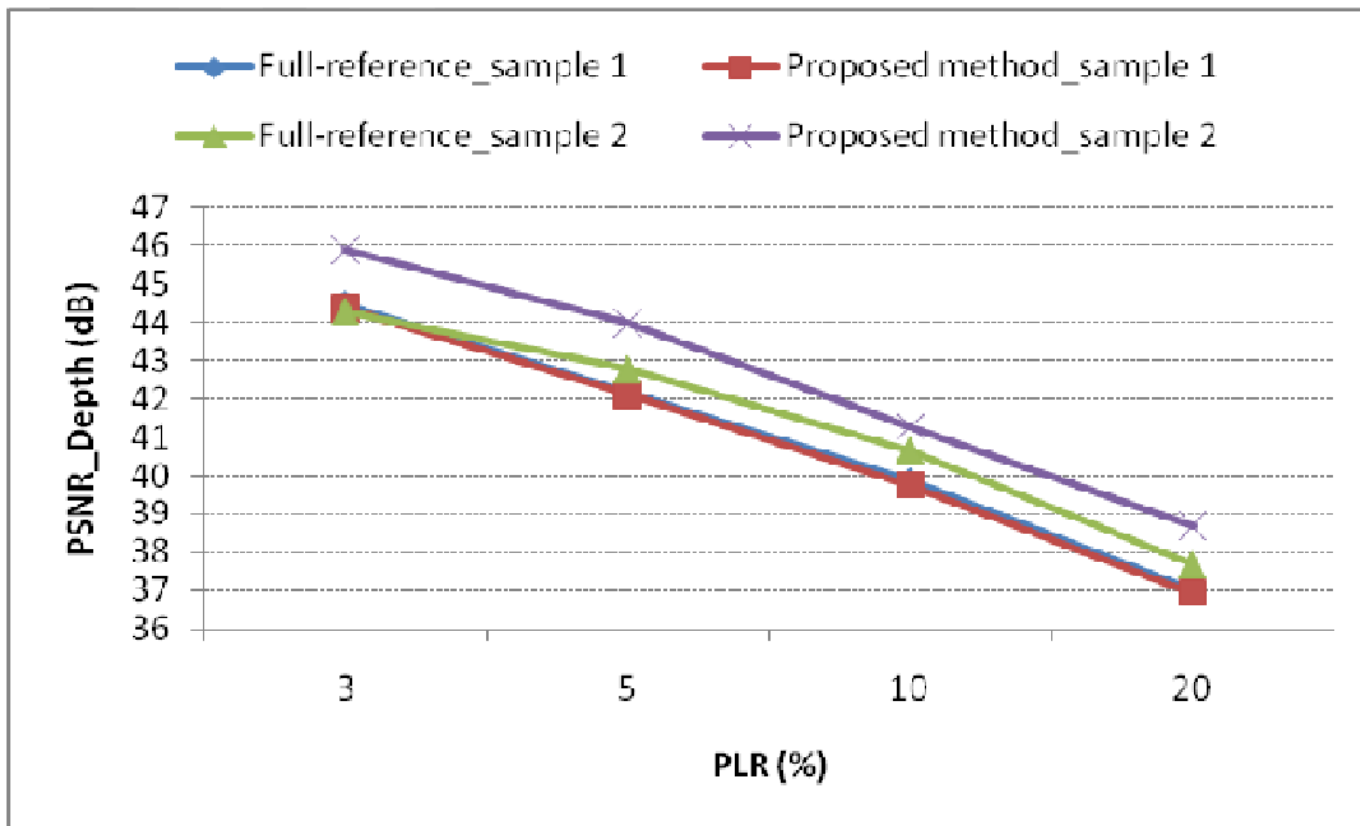
(очень простая...)

- Вычисляем карту краев для карты глубины
- Экспериментально получается

$$PSNR_{Depth\_Map} = f(PSNR_{Binary\_Edge\_Mask})$$



# Результаты



PLR — Packet Loss Rates



# Выводы

---

- Достоинства
  - Авторы обещают малые затраты на передачу reference при адекватной функциональности
- Недостатки
  - Очень простой метод, не учитывается исходное 2D изображение



# Вариант улучшения

---

При передаче 2D+Z видео можно использовать, например, информацию о краях и 2D канала, и Z канала.



# Содержание

---

- Введение
- No-reference метрики
  - Подсчет статистик из карт глубины и движения
  - Оценка согласованности ракурсов стереопары
- Reference метрики
  - Full-reference метрика на основе 3D-DCT
  - Reduced-reference метрика для карт глубины
- **Заключение**



# Заключение

---

- Метрик и подходов очень мало
- Артефакты не классифицируются
- Почти все подходы расширяют ранее известные техники работы с 2D видео

Все плохо. Надо сделать, чтобы было хорошо

# Литература

1. "Algorithmic assessment of 3D quality of experience for images and videos". A. Mittal, A.K. Moorthy, J. Ghosh and A.C. Bovik. IEEE Digital Signal Processing Workshop, Sedona, Arizona, January 04-07, 2011
2. "Predicting Stereoscopic Viewing Comfort Using a Coherence-Based Computational Model". Christian Richardt, Lech Świrski, Ian P. Davies, Neil A. Dodgson. ACM, 2011
3. "Validation of a New Full Reference Metric for Quality Assessment of Mobile 3DTV Content". Lina Jin, Atanas Gotchev, Atanas Boev, Karen Egiazarian. EUSIPCO 2011: 19th European Signal Processing Conference
4. "Reduced-reference quality metric for 3D depth map transmission". C. Hewage, M.G. Martini. 3DTV Conference 2010, Tampere, Finland, 7-9 June 2010
5. "A new standardized method for objectively measuring video quality". M.H. Pinson, S. Wolf. IEEE Transactions on Broadcasting. 2004

# 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