

Измерение качества видео

Алексей Москвин

*Video Group
CS MSU Graphics & Media Lab*

Содержание доклада

- ◆ Постановка задачи
- ◆ Основы HVS
- ◆ Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system
- ◆ Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs
- ◆ Video classification for video quality prediction
- ◆ No-reference Quality Assessment for JPEG-2000

Постановка задачи

- ◆ Основная цель исследований в области объективной оценки качества видео - это разработка алгоритмов, определяющих его автоматически.
- ◆ Варианты использования
 - Мониторинг качества видео в системах управления передачей его по сети
 - Способ сравнения алгоритмов обработки и сжатия
 - Встроенное в системы обработки средство, позволяющее оптимизировать внутренние параметры

Возможные артефакты (1)



◆ Сжатие

- Блочность (следствие использования DCT)
- Размытие
- Эффект Гиббса (следствие квантования на резких границах)
- Проблемы от МС

Возможные артефакты (2)



- ◆ Ошибки при передаче данных
 - Задержки при передаче по сети
 - Потери пакетов
- ◆ Вид искажений зависит от декодера

Возможные артефакты (3)



◆ Другие ошибки

- Преобразование между цифровой и аналоговой формой
- Уменьшение размера цветowych плоскостей
- Изменение частоты кадров
- Деинтерлейсинг

Популярные метрики

- ◆ Наиболее популярные сейчас алгоритмы PSNR, SSIM, VQM
- ◆ Результаты опроса на doom9.org: Do you beleive to metric result?
 - Only SSIM is a reliable metric test. 9%
 - Only PSNR is a reliable metric test. 0%
 - Both SSIM and PSNR are reliable metric test. 30%
 - I meassure quality only by my own visual perception 61%

Outline



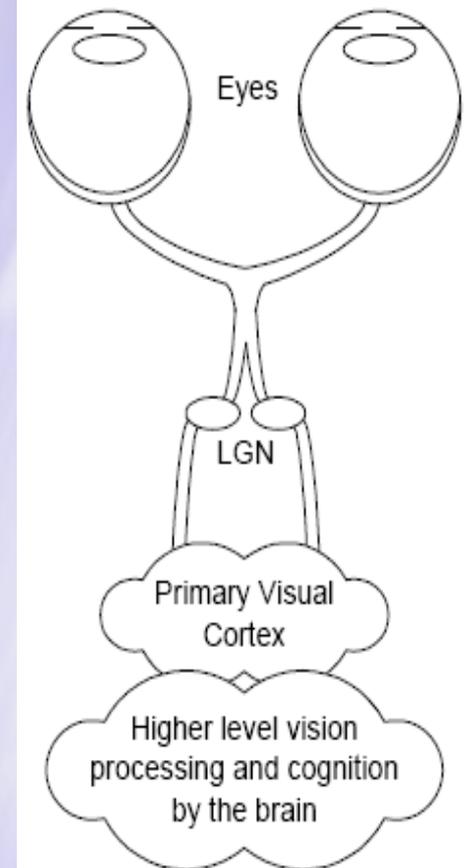
- ◆ Постановка задачи
- ◆ **ОСНОВЫ HVS**
- ◆ Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system
- ◆ Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs
- ◆ Video classification for video quality prediction

HVS (1)



◆ Рецепторы внутри глаза

- Существует 3 типа рецепторов, соответствующих коротким, средним и длинным волнам света
- Распределение рецепторов по сетчатке глаза неравномерно и спадает к краям, кроме того неравномерно и распределение нервных клеток, по которым сигнал поступает в мозг
- На этом этапе проводится свёртка оптического изображения с полем чувствительности рецепторов



HVS (2)



◆ Дальнейшая обработка

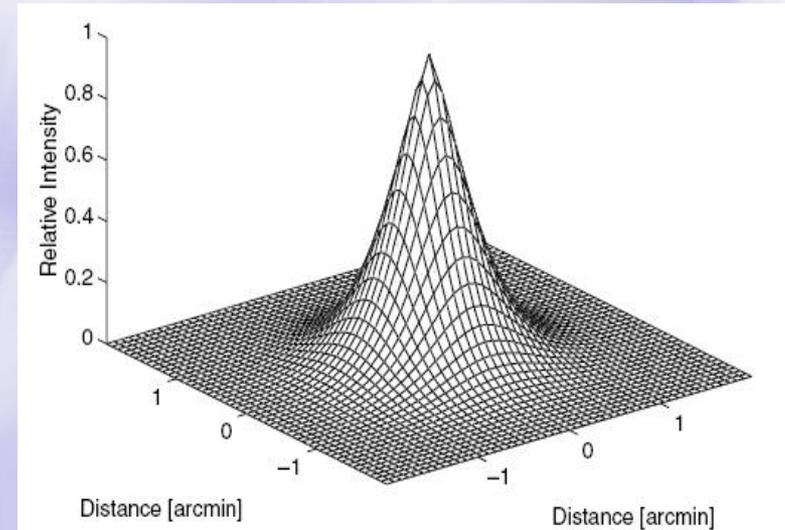
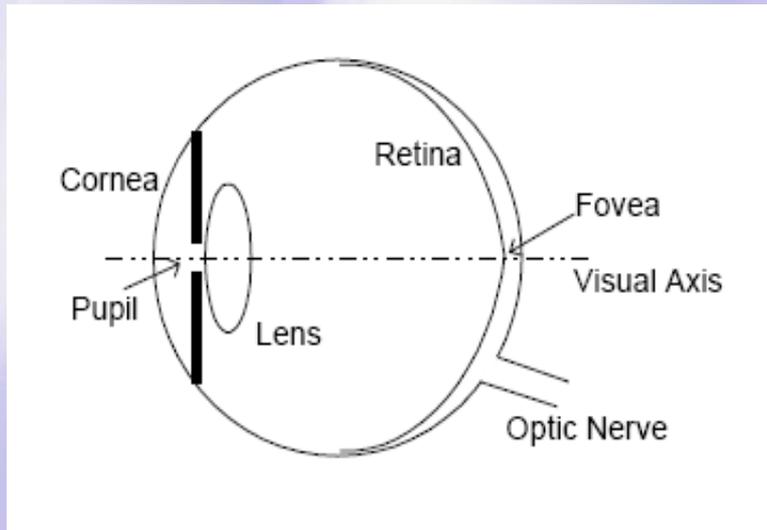
- Следующий этап - латеральное коленчатое ядро (LGN), где производится свёртка полученного с сетчатки сигнала, что приводит к его реорганизации.
- Сигнал, поступающий с LGN-нейронов поступает в кору головного мозга ещё одно преобразование входного потока.
- Сигнал, создаваемый в коре головного мозга передаётся в другие его участки, для анализа движения, распознавания и т.д.

HVS (Центральное и периферийное зрение)

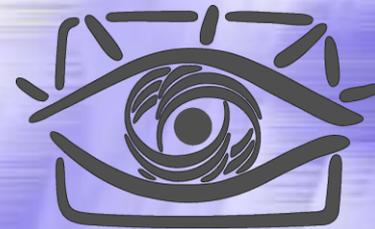


◆ Центральное и периферийное зрение

- Область вокруг точки фиксации глаза поступает с наибольшим разрешением, по мере удаления от этой точки разрешение снижается (используется в очень небольшом числе алгоритмов)



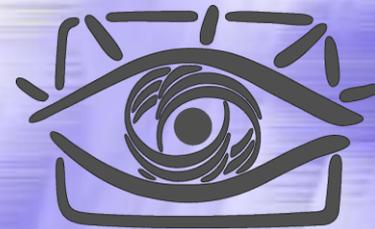
HVS (FWQI)



GRAPHICS & MEDIA LAB
VIDEO GROUP



HVS (FWQI)



GRAPHICS & MEDIA LAB
VIDEO GROUP



HVS (Адаптация к свету)

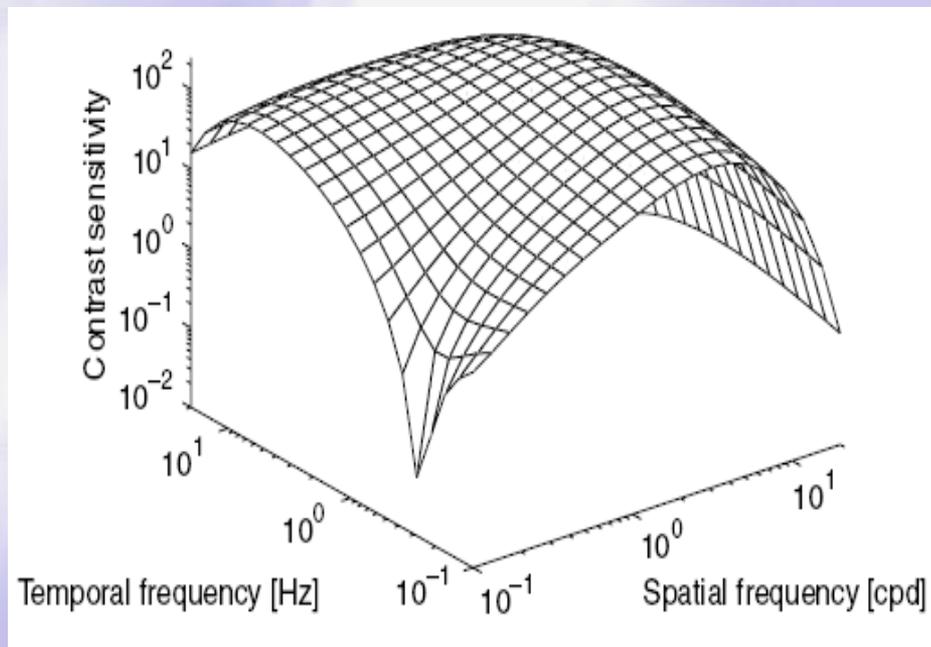
◆ Адаптация к свету

- Интенсивность света, поступающего в глаз регулируется в широких пределах, что приводит к тому, что на сетчатке фиксируется контраст, а не абсолютные значения
- Этапы адаптации
 - ◆ Оптический (ограничение света поступающего на сетчатку)
 - ◆ Химический (в рецепторах)
 - ◆ Адаптация на уровне нейронов
- Weber's Law:
$$C = \frac{\Delta L}{L}$$

HVS (CSF)

◆ CSF

- Чувствительность зависит от частот, присутствующих в входном сигнале (это объясняется полем нейронов в LGN).



- Может моделироваться фильтром или взвешенным суммированием частот.
- CSF зависит от центральности / периферийности зрения, но в большинстве моделей это не учитывается.

HVS (Пространственное маскирование)



◆ Пространственное маскирование

- Наличие одной компоненты в изображении влияет на восприятие другой, расположенной в том же месте
- Наиболее сильно это проявляется когда маска и тестовый сигнал похожи по частоте и ориентации



HVS (Временное маскирование)

◆ Временное маскирование

- Если есть интенсивный сигнал, то перед ним и после него изменяется порог чувствительности.

◆ Адаптация к шаблонам

- Адаптация к шаблонам с, например, сходной частотой ведёт к снижению чувствительности (изменению CSF) около этой частоты.

HVS (Объединение результатов)

◆ Объединение результатов

- Основная проблема: объединение результатов измерений искажений, масок и т.д. по отдельным потокам.
- Чувствительность к искажениям зависит в т.ч. и от распознавания, что трудно моделируется.

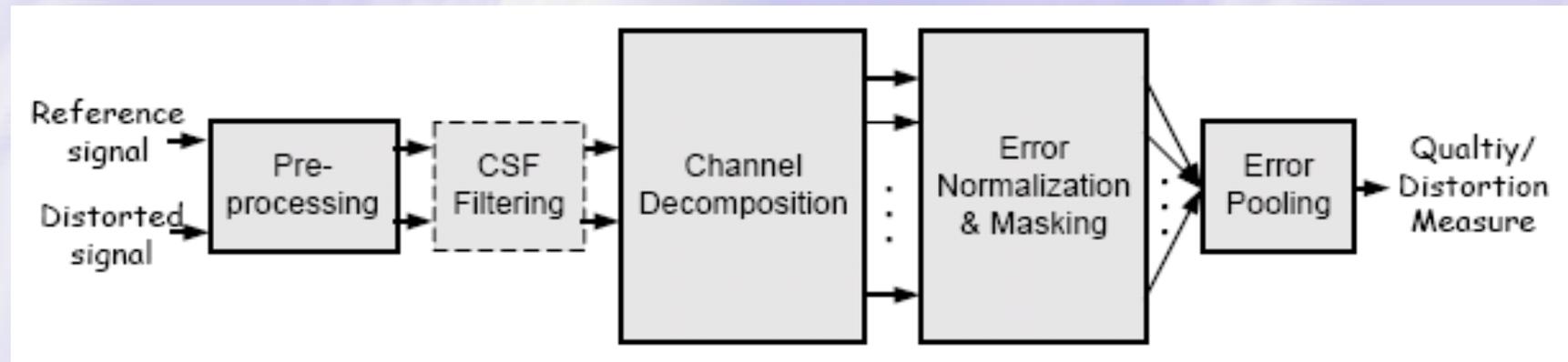
HVS (Описание алгоритма)



- ◆ Продуманный алгоритм с точки зрения HVS
 - Учёт цветовой и яркостных составляющих
 - Учёт не только центрального, но и периферийного зрения
 - Адаптация к яркости
 - Функция чувствительности контраста
 - Маскирование
 - Объединение итоговых результатов

HVS (Схема алгоритма)

◆ Схема алгоритма



HVS (Субъективные факторы)

- ◆ Субъективные факторы
 - Расстояние до изображения
 - Тип и свойства монитора
 - Точность воспроизведения
 - Соответствие звуковому сопровождению
- ◆ Из-за субъективности все эти факторы нельзя представить в виде точных формул для использования в метриках

Outline

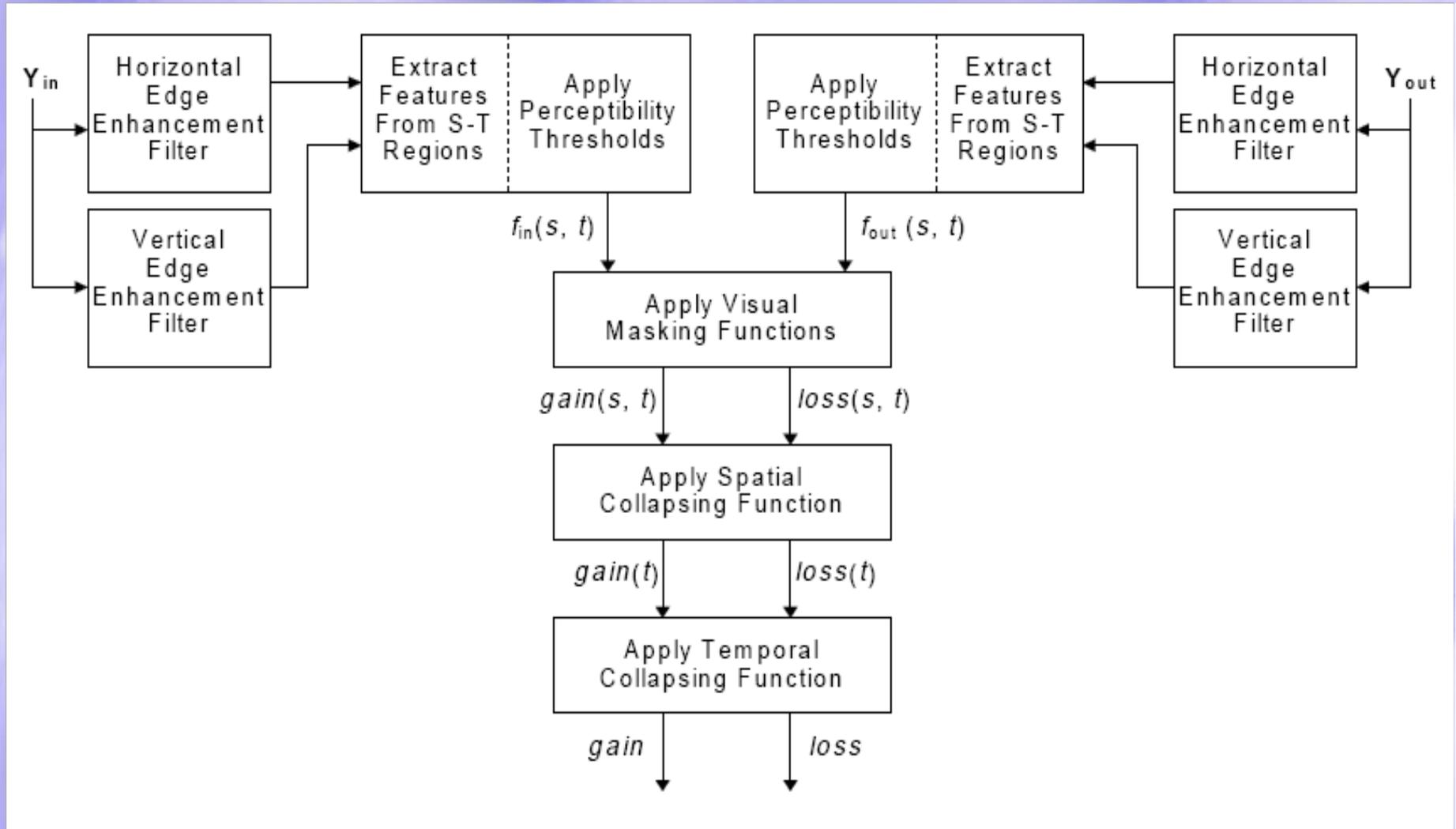
- ◆ Постановка задачи
- ◆ ОСНОВЫ HVS
- ◆ **Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system**
- ◆ Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs
- ◆ Video classification for video quality prediction
- ◆ No-reference Quality Assessment for JPEG-2000

Общая информация

◆ Идея

- Рассчитываются карты изменения границ
- Выделяются пространственно-временные блоки, в которых рассчитываются особенности
- Маскирование
- Усреднение результатов для получения итогового значения

Схема алгоритма



Выделение границ

- ◆ Применение к изображению фильтра Собеля
- ◆ Разделение видео на блоки S-T блоки $8 \text{ px} * 8 \text{ px} * 6$ кадров
- ◆ Для каждой точки в блоке вычислить

$$R(i, j, t) = \sqrt{H(i, j, t)^2 + V(i, j, t)^2}$$

$$\theta(i, j, t) = \tan^{-1} \left[\frac{V(i, j, t)}{H(i, j, t)} \right]$$

- где $V(i, j, t)$, $H(i, j, t)$ - результат выделения границ

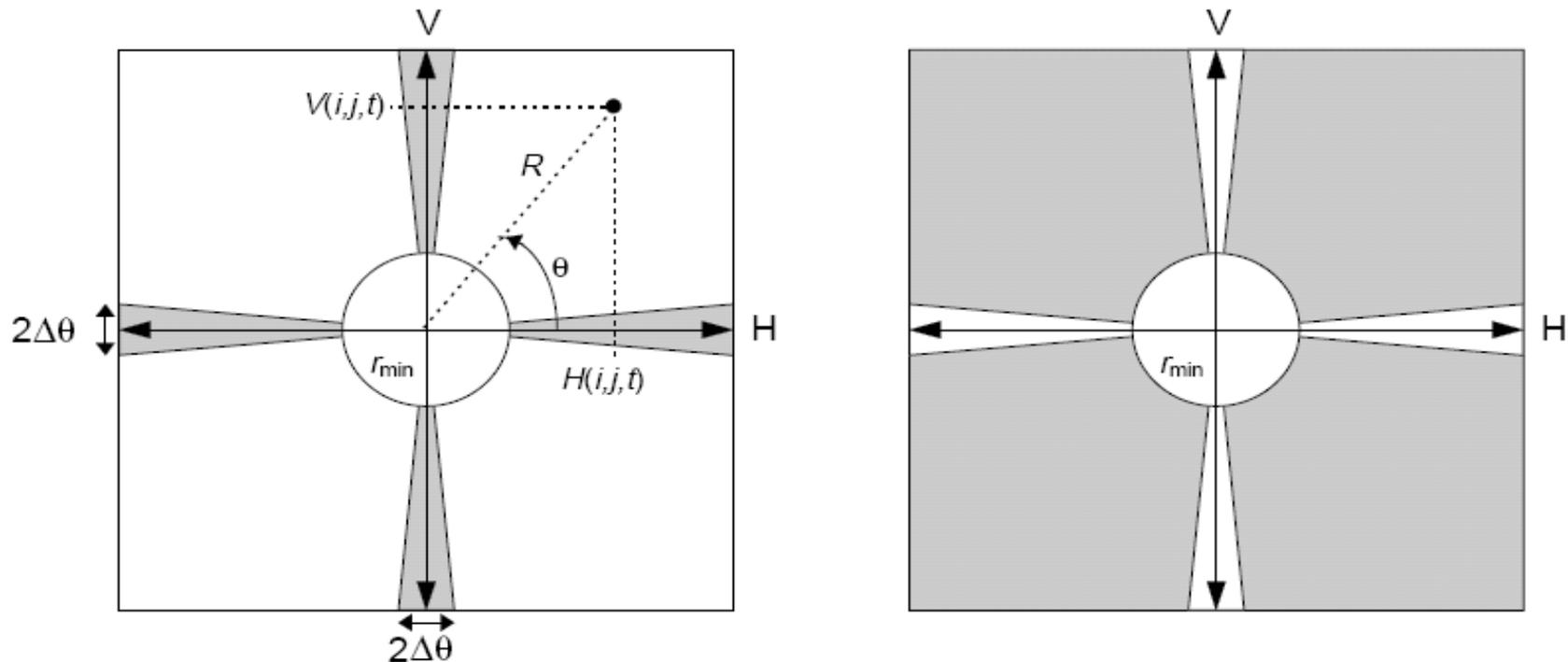
Расчет особенностей (1)

◆ $F1 = \text{stdev}[R(i,j,t)]$

- Если $F1$ оказывается меньше порога (в статье используется $Th1 = 12$), то $F1 = Th1$

Расчет особенностей (2)

- ◆ F2 учитывает изменение ориентации границ



Расчет особенностей (3)

$$HV(i, j, t) = \begin{cases} R(i, j, t) & \text{if } R(i, j, t) \geq r_{\min} \text{ and } m\frac{\pi}{2} - \Delta\theta < \theta(i, j, t) < m\frac{\pi}{2} + \Delta\theta \quad (m = 0, 1, 2, 3) \\ 0 & \text{otherwise} \end{cases}$$

$$\overline{HV}(i, j, t) = \begin{cases} R(i, j, t) & \text{if } R(i, j, t) \geq r_{\min} \text{ and } m\frac{\pi}{2} + \Delta\theta \leq \theta(i, j, t) \leq (m+1)\frac{\pi}{2} - \Delta\theta \quad (m = 0, 1, 2, 3) \\ 0 & \text{otherwise} \end{cases}$$

$$f_2 = \frac{\left\{ \text{mean}[HV(i, j, t)] \right\}_P}{\left\{ \text{mean}[\overline{HV}(i, j, t)] \right\}_P}$$

Расчет особенностей (4)

$$gain(s, t) = pp \left\{ \log_{10} \left[\frac{f_{out}(s, t)}{f_{in}(s, t)} \right] \right\}$$

$$loss(s, t) = np \left\{ \frac{f_{out}(s, t) - f_{in}(s, t)}{f_{in}(s, t)} \right\}$$

- pp - отрицательные значения заменяются на 0,
 np - положительные значения заменяются на 0

Усреднение значений

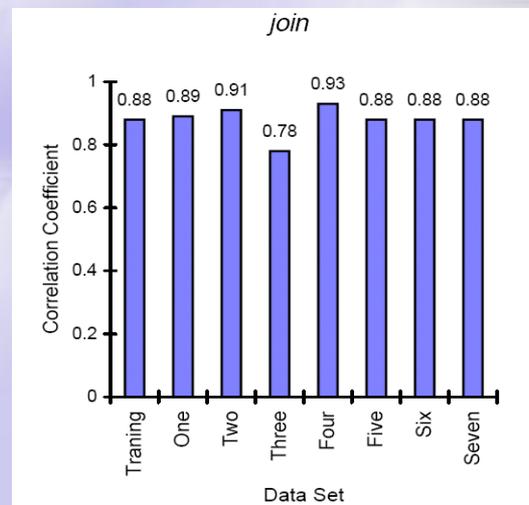
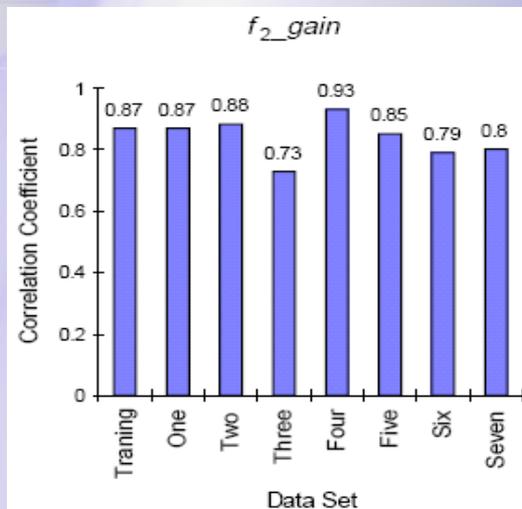
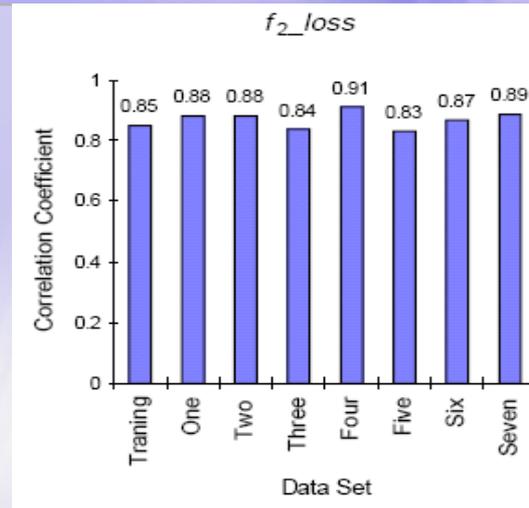
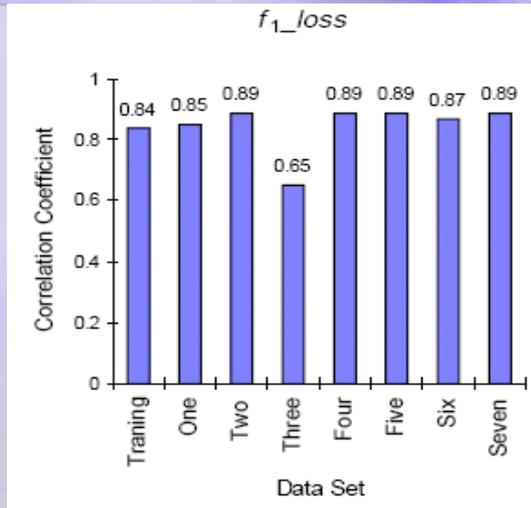
- ◆ $Res = 0.38 * f1_loss + 0.39 * f2_loss - 0.23 * f2_gain$
- ◆ Пространственное усреднение
 - В качестве значения для кадра в каждый момент t выбирается среднее значение из 5% блоков с максимальным изменением особенностей
- ◆ Временное усреднение
 - Среднее значение по обработанным кадрам каждые 10 секунд

Тестирование

◆ Данные

- 7 наборов данных, более 40 экспертов для оценки результатов на каждом наборе
- Использовалось несколько вариантов оценок (5, 7, 11-бальные)
- В процессе тестирования были подобраны оптимальные размеры ST-блоков и коэффициенты в вычислении итоговой суммы

Тестирование (2)



Outline

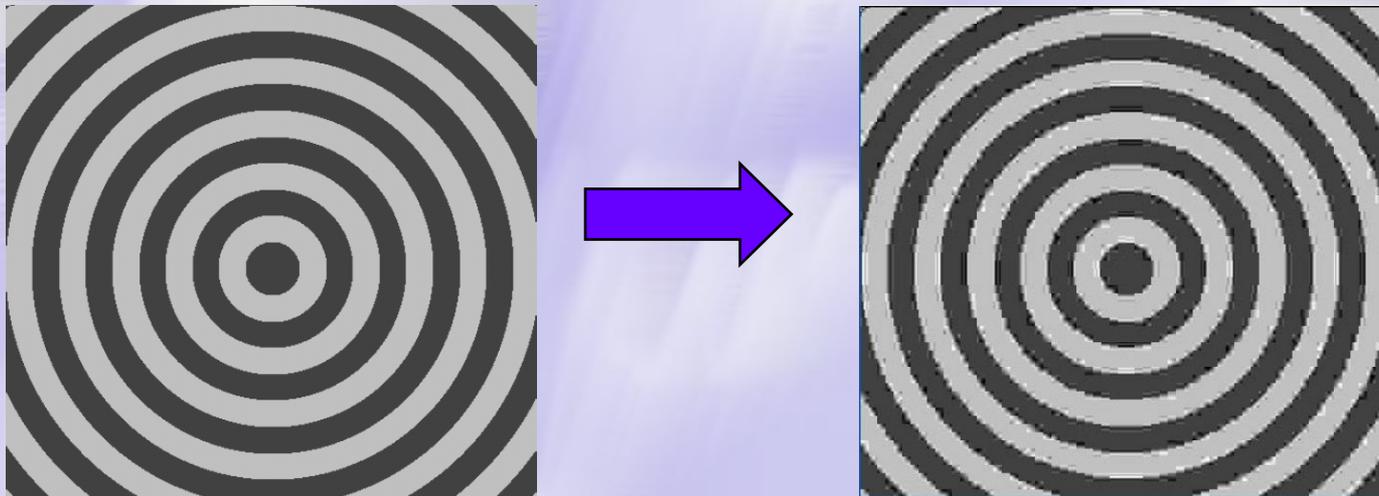


- ◆ Постановка задачи
- ◆ ОСНОВЫ HVS
- ◆ Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system
- ◆ **Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs**
- ◆ Video classification for video quality prediction
- ◆ No-reference Quality Assessment for JPEG-2000

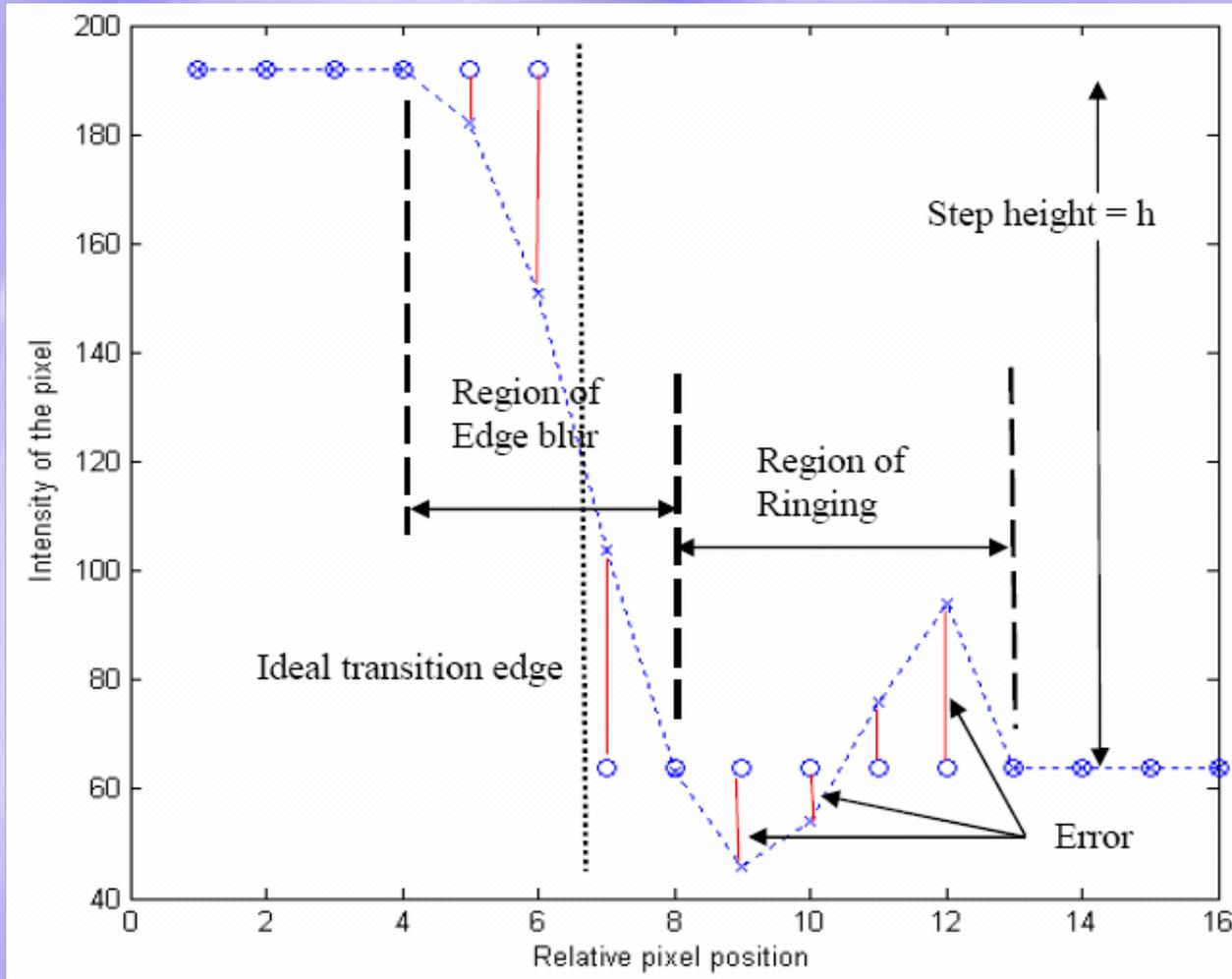
Общая информация

◆ Идея

- Алгоритм измеряет специфические артефакты JPEG-2000: размытие и ringing
- В качестве тестовых данных используются специально сгенерированные изображения



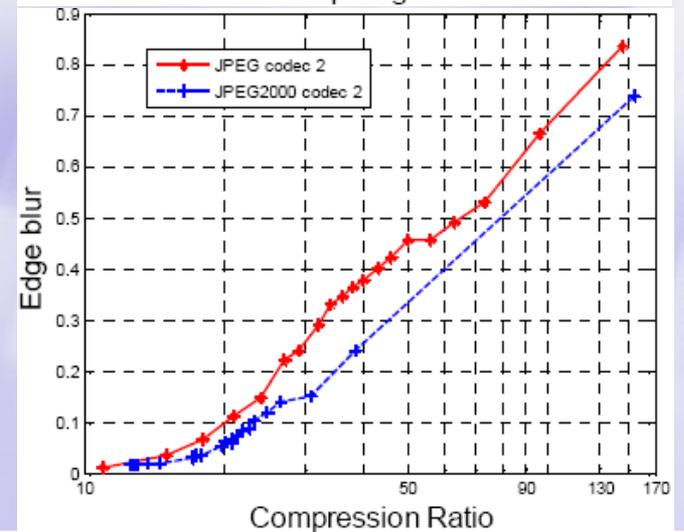
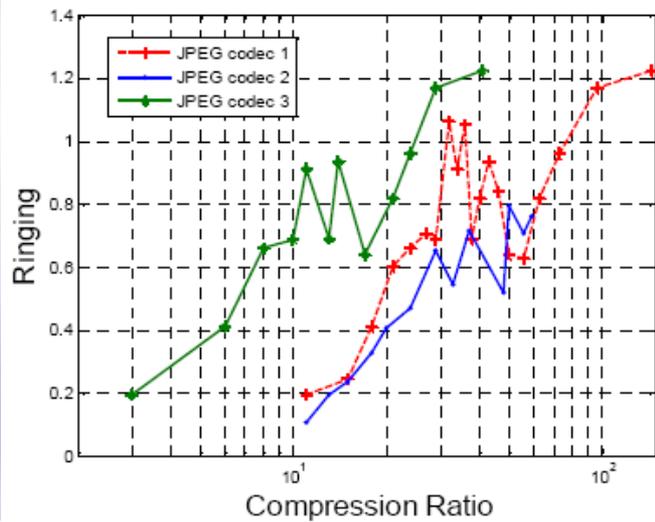
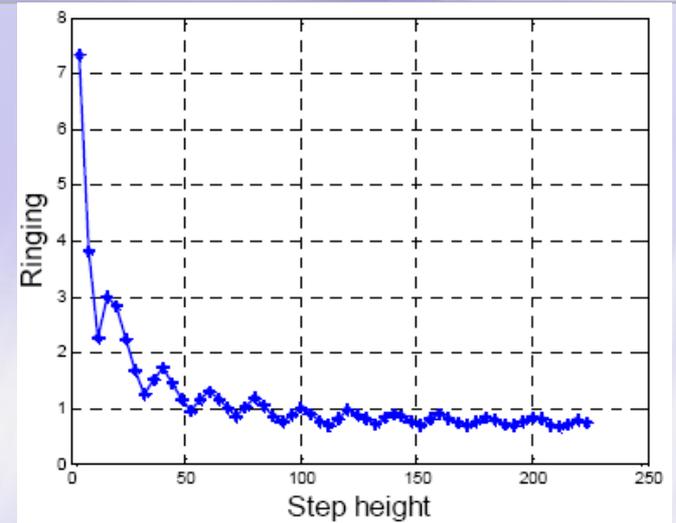
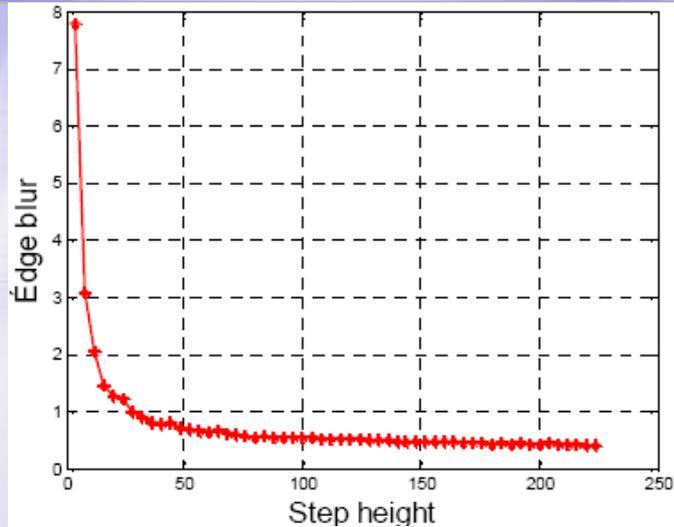
Описание алгоритма



$$Blur = \frac{\sum |Error|_{blur_region}}{Nh}$$

$$Ringing = \frac{\sum |Error|_{ringing_region}}{Nh}$$

Результаты



Outline

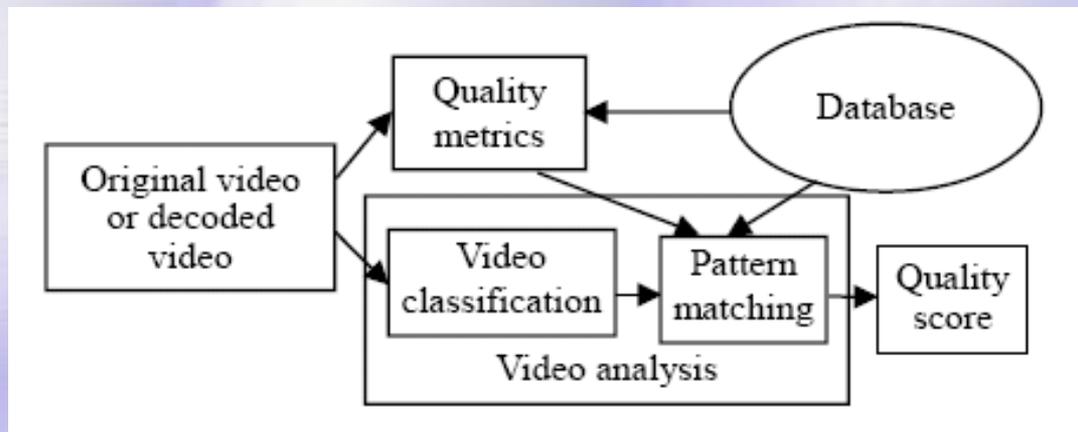


GRAPHICS & MEDIA LAB
VIDEO GROUP

- ◆ Постановка задачи
- ◆ ОСНОВЫ HVS
- ◆ Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system
- ◆ Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs
- ◆ **Video classification for video quality prediction**
- ◆ No-reference Quality Assessment for JPEG-2000

Идея алгоритма

- ◆ Классифицировать блоки в кадре
- ◆ Поиск в базе наиболее похожих по характеристикам последовательностей
- ◆ Вычисление оценки качества



Пространственная классификация блоков (1)

- ◆ Каждый блок получает одну из следующих характеристик
 - Гладкий
 - Тонкая текстура
 - Текстура
 - Граница



Пространственная классификация блоков (2)

$$F_{S, \text{Var}} = \frac{1}{KM} \sum_{k=1}^K \sum_{m=1}^M \left\{ \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N [x_m^k(i, j) - \bar{x}_m^k]^2 \right\}$$
$$T_{E, m}^k = \left\{ \sum_{i=1}^N \sum_{j=1}^N (i, j) \neq (1, 1) [\hat{H}(i, j)]^2 [x_m^k(i, j)]^2 \right\}^{1/2}$$

x_m^k - DCT-коэффициент в позиции (i,j)

$H(i,j)$ - функция чувствительности

Пространственная классификация блоков (3)

◆ Вычисление особенностей

$$F_{s,flat} = \frac{1}{K} \sum_{k=1}^K \frac{S^k (flat_area)}{M}$$

$$F_{s,edge} = \frac{1}{K} \sum_{k=1}^K \frac{S^k (edge_area)}{M}$$

$$F_{s,texture} = \frac{1}{K} \sum_{k=1}^K \frac{S^k (texture_area)}{M}$$

$$F_{s,fine-texture} = \frac{1}{K} \sum_{k=1}^K \frac{S^k (finetexture_area)}{M}$$

- $S(A)$ - количество блоков, классифицированных как A
- K -количество кадров
- M -количество блоков в кадре

Временная классификация блоков (1)



- ◆ Используется блочный ME и вычисляется ошибка компенсации, после чего блоки классифицируются по ошибке (SAD или MSE)
 - zero
 - low
 - medium
 - high
- ◆ Используется небольшой размер окна для оптимизации

Временная классификация блоков (2)

◆ Вычисление особенностей

$$F_{t,zero} = \frac{1}{K} \sum_{k=1}^K \frac{S^k(zero)}{M}$$

$$F_{t,low} = \frac{1}{K} \sum_{k=1}^K \frac{S^k(low)}{M}$$

$$F_{t,medium} = \frac{1}{K} \sum_{k=1}^K \frac{S^k(medium)}{M}$$

$$F_{t,high} = \frac{1}{K} \sum_{k=1}^K \frac{S^k(high)}{M}$$

- $S(A)$ - количество блоков, классифицированных как А
- K -количество кадров
- M -количество блоков в кадре

Вычисление особенностей



◆ Вычисление особенностей

- По полученному набору особенностей (F_s , F_t) могут быть вычислены новые, например $F_s(\text{texture}) + F_s(\text{edges})$ (т.к. изменения в текстуре и на границах наиболее заметны)
- В предлагаемом методе выбирается 6 особенностей
 - ◆ Количество гладких блоков
 - ◆ Количество блоков с текстурой + количество блоков с границами
 - ◆ Количество блоков с каждым видом движения

Классификация видео

$$i_best = \arg \min_i \{D_i\} = \arg \min_i \left(\sum_{j=1}^n w_j \|f_j^{(i)} - f_j^{(test)}\| \right), \quad \sum_{j=1}^n w_j = 1$$

- ◆ $F(i)$ - вектор особенностей последовательности из базы данных, $F(test)$ - вектор исследуемой последовательности
- ◆ Веса $w(j)$ подбираются экспериментально и зависят от битрейта

$$Q = \frac{\sum_{i=1}^k t_i Q_i}{\sum_{i=1}^k t_i}$$

Q_i - оценка качества для i -го видео
 t_i - расстояние между $F(test)$ и $F(i)$

Предсказываемые метрики

- ◆ В качестве основной эталонной метрики используется VQM (NTIA/ITS)
- ◆ Тестирование
 - Составлена база последовательностей
 - Для каждой последовательности строится вектор особенностей
 - Каждая последовательность выбирается как тестовая, для неё из базы выбираются наиболее близкие, оценивается значение VQM
 - Предсказанное значение VQM сравнивается с рассчитанным

Результаты



Sequence	<i>VQM</i> prediction
Akiyo	0.9976
Bridge-close	0.8437
Bridge-far	0.9371
Canoa	0.9656
Coastguard	0.9456
Container	0.9982
Crew	0.8128
F1	0.9359
Flower	0.9761
Football	0.9610
Foreman	0.9589

Outline



GRAPHICS & MEDIA LAB
VIDEO GROUP

- ◆ Постановка задачи
- ◆ ОСНОВЫ HVS
- ◆ Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system
- ◆ Objective Evaluation of Edge Blur and Ringing Artefacts: JPEG 2000 Image Codecs
- ◆ Video classification for video quality prediction
- ◆ **No-reference Quality Assessment for JPEG-2000**

Идея алгоритма

- ◆ Алгоритм не требует оригинального изображения
- ◆ Накапливаем выборку из оригинальных и сжатых картинок
- ◆ Выделяем набор особенностей на тестируемом изображении и оцениваем их принадлежность к классу сжатых или оригинальных примеров
- ◆ По этой вероятности получаем оценку качества

Накопление статистики

- ◆ Набор оригинальных и сжатых изображений
- ◆ Проходим по всем точкам на границе объектов
- ◆ Растягиваем окно вокруг $X(i,j)$ в вектор
- ◆ Если это оригинальное изображение, то заносим его в $S(+)$, иначе - в $S(-)$
- ◆ Метод главных компонент для уменьшения размерности векторов в $S(+)$ и $S(-)$

Искажение в данной точке

- ◆ Для каждой точки в тестируемом изображении, которая лежит на границе

$$Ld(i, j) = \frac{(P(x(i, j) | S(-)) * (P(S(-)) / P(S(+))))}{P(x(i, j) | S(+)) + P(x(i, j) | S(-)) * (P(S(-)) / P(S(+)))}$$

- ◆ Эта мера ошибки никак не учитывает её характер, важно только, чтобы она была такой же, как и на тренировочных изображениях

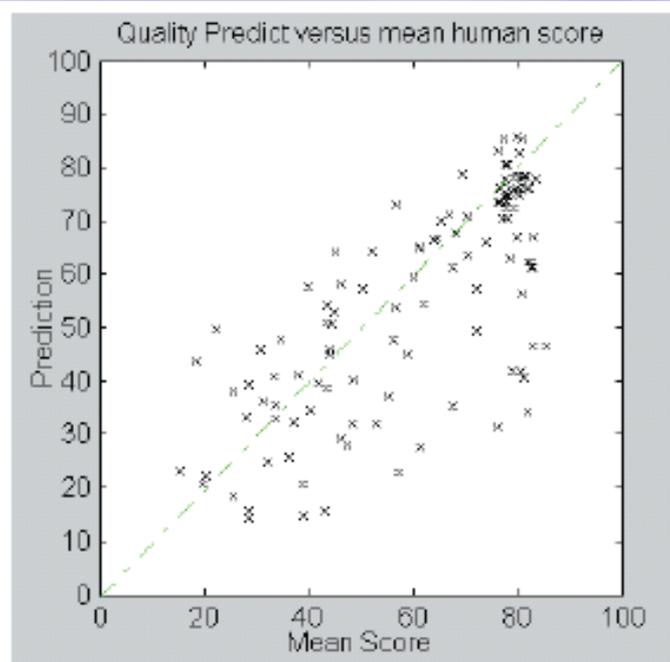
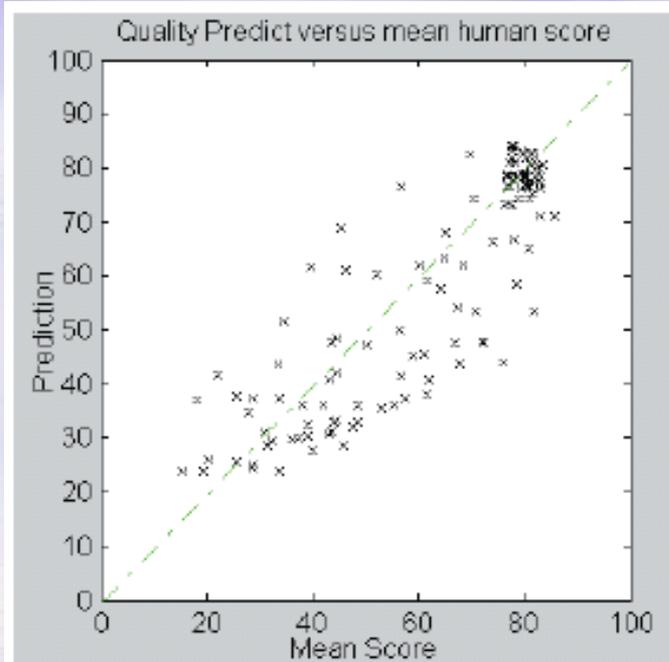
Метрика искажения

$$Dm = \frac{1}{N_{edge}} \sum_{edge_p} Ld(i, j)$$

$$Q = a + bDm^c$$

- ◆ Параметры подбирались экспериментально (29 оригинальных изображений, 227 всего, 25 наблюдателей, шкала оценок с 5 вариантами)

Результаты



◆ Корреляция с оценками экспертов 85%

Статьи



GRAPHICS & MEDIA LAB
VIDEO GROUP

- ◆ **Objective Evaluation of Edge Blur and Ringing Artefacts: Application to JPEG and JPEG 2000 Image Codecs** G. A. D. Punchihewa, D. G. Bailey, and R. M. Hodgson
- ◆ **Video classification for video quality prediction** LIU Yu-xin, KURCEREN Ragip2 BUDHIA Udit
- ◆ **NO-REFERENCE QUALITY ASSESSMENT FOR JPEG2000 COMPRESSED IMAGES** *Hanghang Tong¹, Mingjing Li, Hong-Jiang Zhang, Changshui Zhang*
- ◆ **Digital Video Quality Vision Models and Metrics** *Stefan Winkler*
- ◆ **OBJECTIVE VIDEO QUALITY ASSESSMENT** *Zhou Wang, Hamid R. Sheikh and Alan C. Bovik*
- ◆ **Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system** *Stephen Wolf, Margaret H. Pinson*