

# ВЕРИФИКАЦИЯ МОДЕЛИ КВАЗИОПТИМАЛЬНОЙ СЕГМЕНТАЦИИ НА ПРИМЕРЕ СОСТАВНОГО ИЗОБРАЖЕНИЯ

М.В. Харинов, [khar@iias.spb.su](mailto:khar@iias.spb.su)

Санкт-Петербургский ин-т информатики и автоматизации РАН

**Область: иерархическая сегментация изображения посредством кусочно-постоянных приближений, минимально отличающихся от изображения по среднеквадратичному отклонению  $\sigma$ .**

## **Развиваемые идеи :**

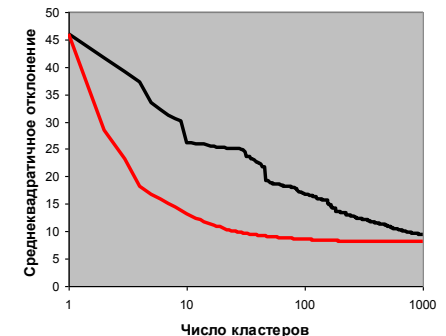
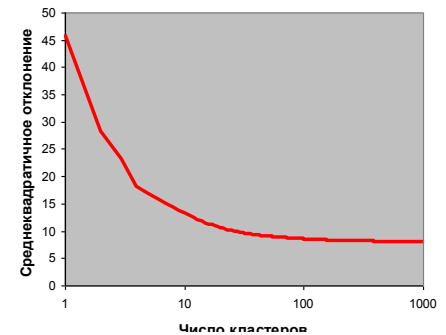
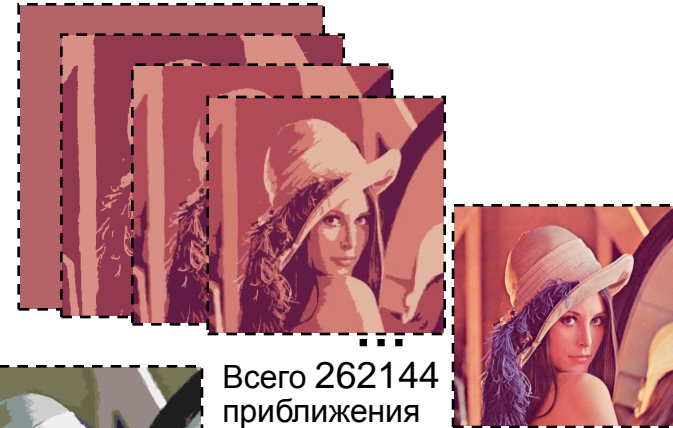
- “Преобразование к удобному для распознавания виду” – Гуревич И.Б., 1985.
- “Применение динамических деревьев в обработке изображений” – СПИИРАН, 1993.
- “Упрощенная модель Мамфорда-Шаха” – Бугаев А.С., Хельвас А.В., 2001.
- “Глобально-локальный анализ ” – Абламейко С.В., Недзвед А.М., 2009.
- “Проблема локализации объектов ” – Визилтер Ю.В., Желтов С.Ю., 2011.
- “Квазиоптимальная разметка изображения” – Мельников П.А., Копылов А.В., 2012.
- “Meanless  $k$ -means” – Двоенко С.Д., 2014.

**Другие:** Ward J.H., Jr.(1965), Tarjan R.E., Sleator D.D.(1975), Mumford D., Shah J.(1985), Koepfler G.(1994), Liao P.-S.(2001), Crisp D.J.(2002), Nock R., Nielsen F.(2004), Arifin A. Z. (2006), etc.

# МОДЕЛЬ КВАЗИОПТИМАЛЬНОЙ СЕГМЕНТАЦИИ

## Основные положения:

1. Изображение состоит из изображений, которые задаются кусочно-постоянными приближениями и именуются «объектами»
2. Объекты составляют бинарную иерархию
3. Последовательность приближений представима в виде изображения, называемого «инвариантным представлением»
4. Любая тройка вложенных друг в друга изображений описывается выпуклой последовательностью значений суммарной квадратичной ошибки  $E = 3N\sigma^2$  приближения от изображения
5. Любая иерархия изображений по определенному алгоритму преобразуется в иерархию, описываемую выпуклой последовательностью значений  $E$



## БАЗОВЫЕ ОПЕРАЦИИ МИНИМИЗАЦИИ $E$

$$E(1 \cup 2 \cup 3 \cup \dots \cup j) = \sum_{p=1}^j E_p + \frac{1}{2} \cdot \frac{\sum_{p=1}^j \sum_{q=1}^j n_p n_q \|I_p - I_q\|^2}{\sum_{p=1}^j n_p}$$

$n_p, n_q$  - число пикселей в кластерах,  $I_p, I_q$  - «вектора» средних яркостей

**1) Слияние (*merge*):**  $\Delta E_{merge} \equiv E(1 \cup 2) - E(1) - E(2) = \frac{n_1 n_2}{n_1 + n_2} \|I_1 - I_2\|^2 \geq 0$

**2) Разделение (*divide*):**  $\Delta E_{divide}(1) = -\Delta E_{merge}(2, 3) \leq 0$  , где  $1 = 2 \cup 3$

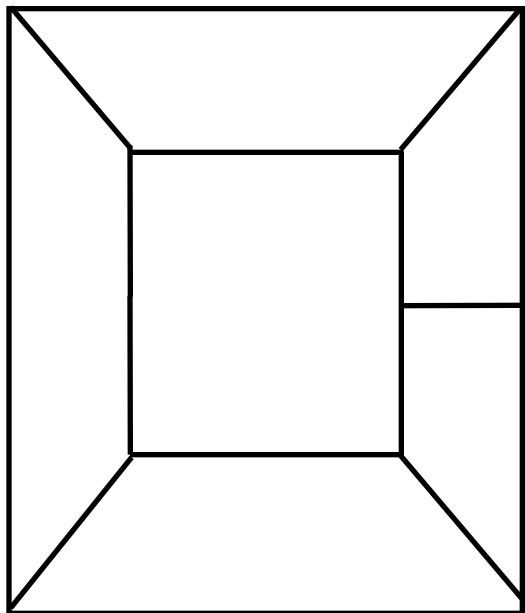
**3) Дробление (*split*):**  $\Delta E_{split} = -\frac{kn_1}{n_1 - k} \|I - I_1\|^2 \leq 0$

**4) Коррекция (*correct*):**  $\Delta E_{correct} = \frac{\|I - I_2\|^2}{\frac{1}{k} + \frac{1}{n_2}} - \frac{\|I - I_1\|^2}{\frac{1}{k} - \frac{1}{n_1}}$   $\Delta E_{correct} \geq 0 \Leftrightarrow \|I - I_1\| \leq \alpha \cdot \|I - I_2\|$ ,  
 $\alpha \equiv \sqrt{\frac{n_2(n_1 - k)}{n_1(n_2 + k)}} < 1$

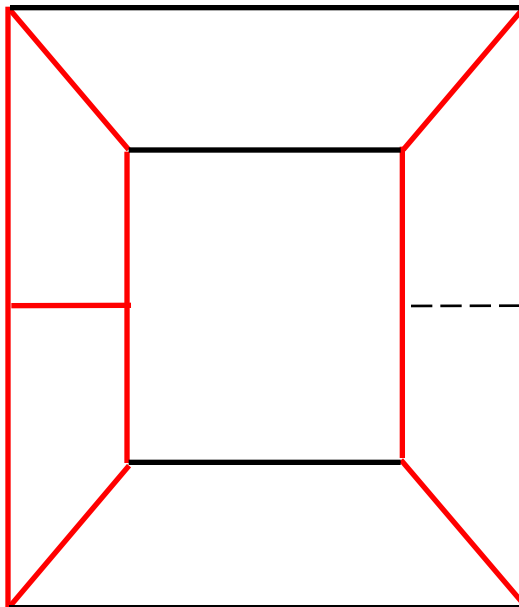
$k, n_1, n_2$  - число пикселей в кластерах,  $I, I_1, I_2$  - «вектора» средних яркостей

# СХЕМА ГЛОБАЛЬНОГО И ЛОКАЛЬНОГО УЛУЧШЕНИЯ КАЧЕСТВА СЕГМЕНТАЦИИ

Приближение

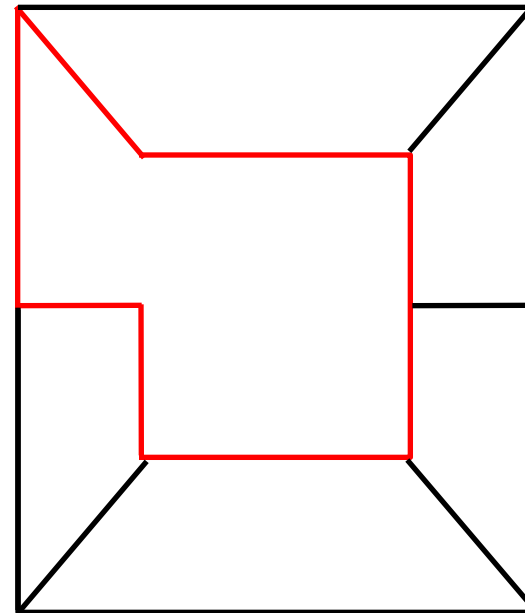


Глобальное улучшение



Операции *merge/divide*  
**SI-метод**

Локальное улучшение



Операция *correct*  
**K-meanless –метод**

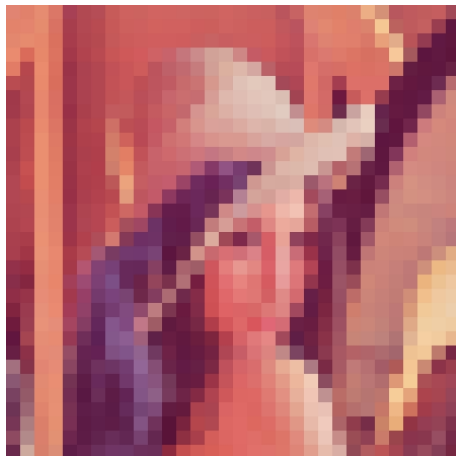
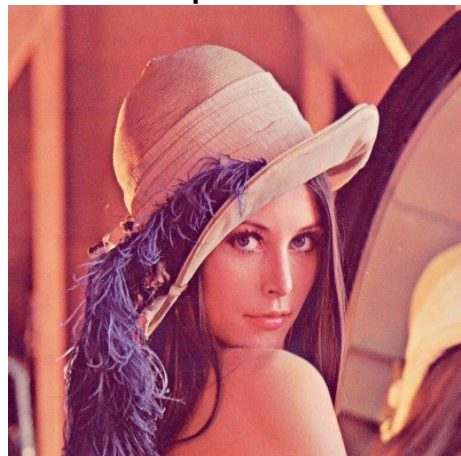
Критерии однородности:  $\Delta E_{divide}(1) + \Delta E_{merge}(2, 3) \geq 0$

$\Delta E_{correct} \geq 0$

# УЛУЧШЕНИЕ И ИСПРАВЛЕНИЕ СЕГМЕНТАЦИИ

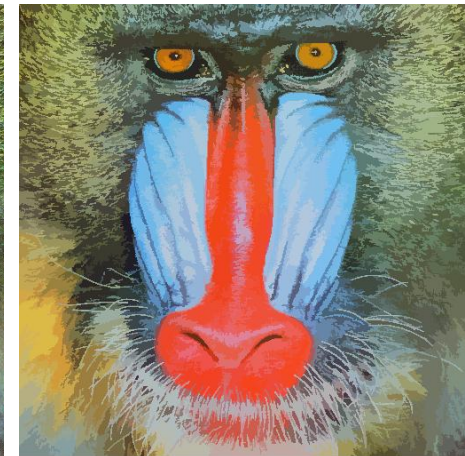
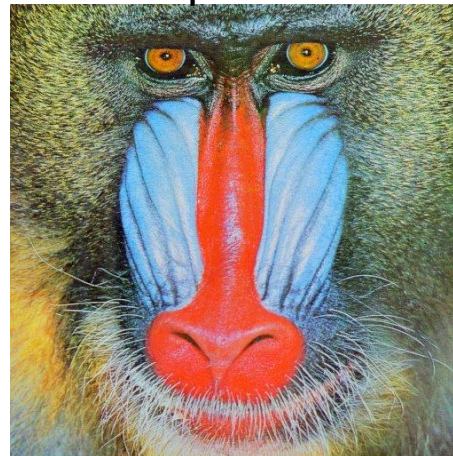
Пример приближений из 1024 сегментов

Изображение

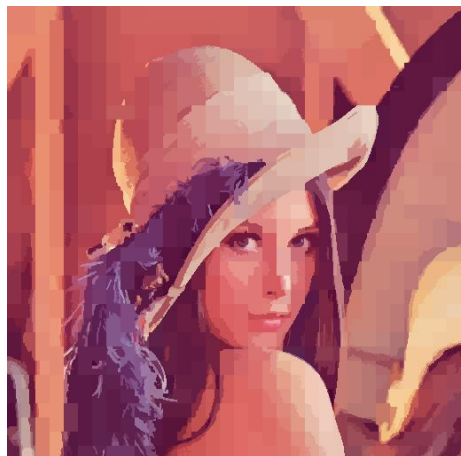


$\sigma = 21,94260$

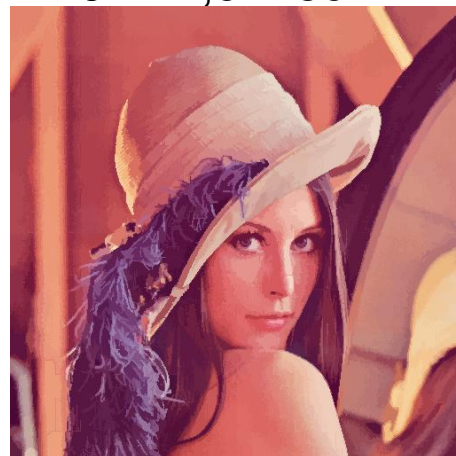
Изображение



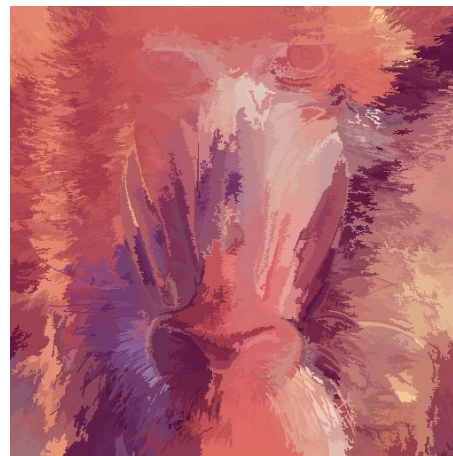
$\sigma = 18,92676$



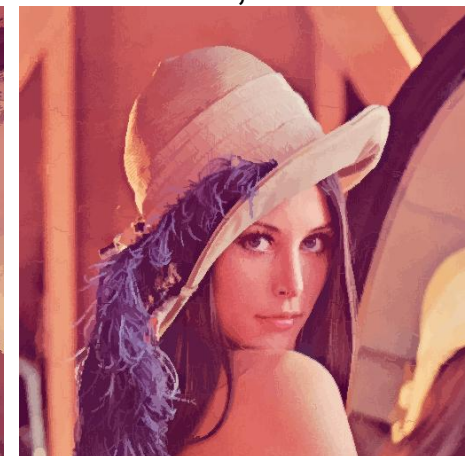
$\sigma = 11,95288$



$\sigma = 8,19082$



$\sigma = 31,53760$

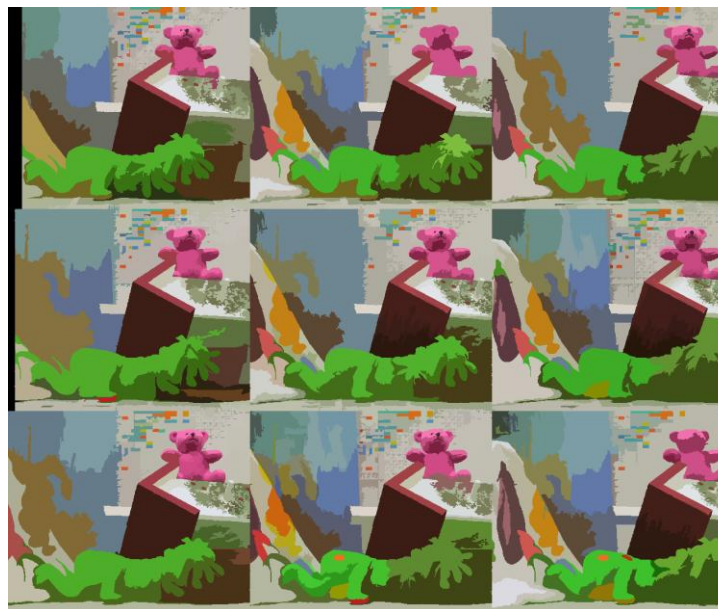


$\sigma = 7,79065$



# ВЫДЕЛЕНИЕ ОБЪЕКТОВ НА СОСТАВНОМ ИЗОБРАЖЕНИИ

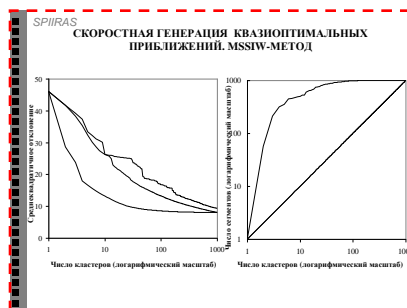
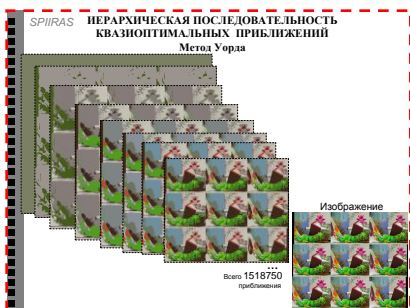
Изображение



$$\sigma = 23,64707$$



$$\sigma = 15,81120$$

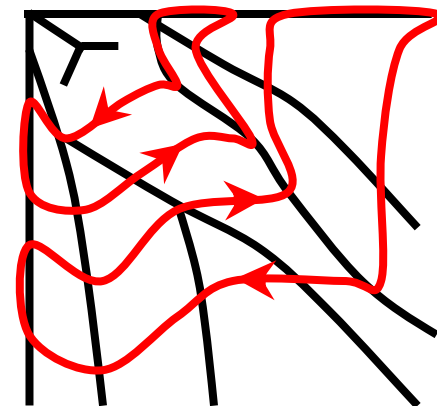


# АППАРАТ ВЫЧИСЛЕНИЙ В ТЕРМИНАХ «ДИНАМИЧЕСКОЙ» СЕТИ

«Динамическую» сеть составляют:

- **Динамические деревья Слейтора-Тарьяна**
- **Ассоциативные списки пикселей, заданные в виде циклов**

которые обеспечивают скоростные вычисления с иерархией кластеров пикселей (сегментов) и последовательностями приближений изображения в объеме RAM.



## Особенности:

1. Метаданные, описывающие иерархию кластеров пикселей, и сами кластеры задаются на одном и том же множестве координат;
2. Для минимизации  $E$  или  $\sigma$  реализуется обобщенный механизм обратимых вычислений, при котором состояние вычислительной системы в любой момент времени восстанавливается не однозначно.

## ВЫВОДЫ

- В докладе представлена модель сегментации изображения посредством аппроксимации оптимальных приближений иерархическими последовательностями квазиоптимальных приближений
- Модель сегментации имеет очевидную интерпретацию операций с кластерами пикселей и сегментами изображения, скоростное выполнение которых поддерживается благодаря сетевой структуре на основе динамических деревьев и ассоциативных циклов
- Судя по результатам сегментации, модель проще всего применить для детектирования расстояний до объектов сцены и др. задачах обработки стереопар
- В более развитом применении модель может служить основой унифицированного инструментария для выполнения выделения конкретных объектов программистами без специального опыта
- Актуальность инструментариев для создания приложений в области ИИ определяется тем, что подобные решения интенсивно разрабатываются за рубежом, например, в четырехлетнем проекте PPAML (Probabilistic Programming for Advancing Machine Learning) американского агентства DARPA, который завершится в 2017 г.