

# Аппаратная реализация на базе FPGA парадигмы Compressive Sensing для захвата изображений

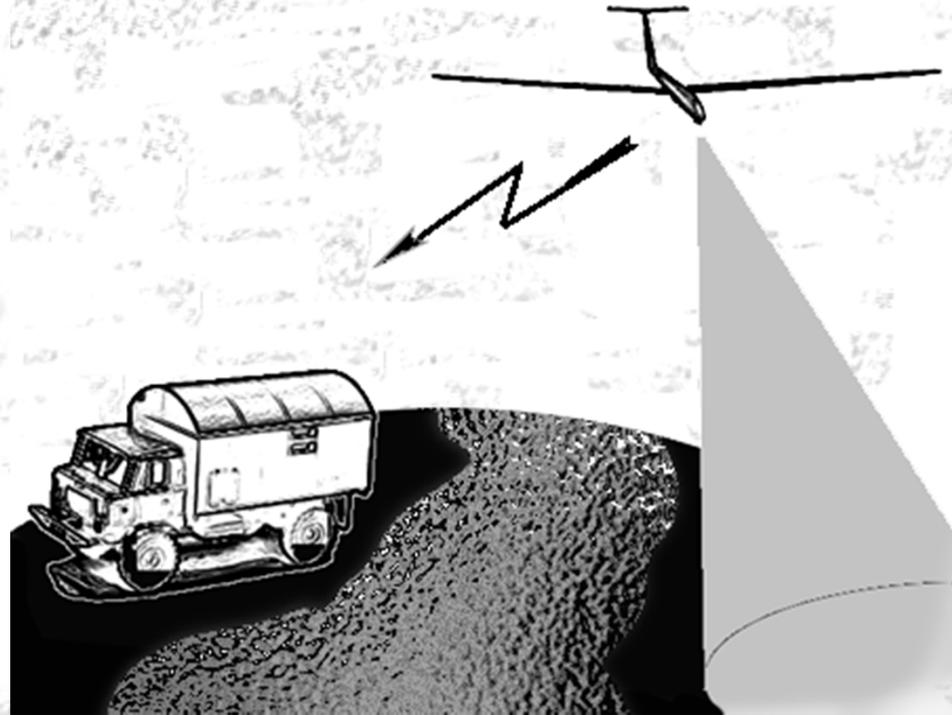
Исаев О.П.

Санкт-Петербургский  
Государственный Университет  
2011 г.



# Задача

Передать информацию от мобильного устройства в стационарный центр обработки информации.



# Проблемы

- Габариты;
- Масса;
- Энергопотребление;
- Стоимость.

# Существующие решения

- Аналоговая камера
  - 720x576 для системы PAL/SECAM;
  - 720x480 для системы NTSC;
  - Низкая стабильность;
  - Невысокая стоимость;
  - Малое энергопотребление.

# Существующие решения

- Цифровая камера
  - 1920x1080 полное HDTV (1080i) ;
  - 1280x720 половинное HDTV (720p) ;
  - Высокая стабильность;
  - Высокая стоимость;
  - Большое энергопотребление.

# Входной сигнал

$$s = \sum_{\gamma \in \Gamma} \alpha_{\gamma} \phi_{\gamma}$$

$$s = \sum_{i=1}^m \alpha_{\gamma_i} \phi_{\gamma_i} + R^{(m)}$$

# Тривиальный словарь

$$s = s_0 \phi_0 + \sum_{\gamma \in \Gamma} (s_\gamma - s_{\gamma-1}) \phi_\gamma,$$

где  $\gamma \in \{0, 1, \dots, n-1\}$ ,

$$\phi_\gamma(t) = \begin{cases} 1, & t \geq \gamma \\ 0, & \text{в другом случае} \end{cases}$$

# Частотный словарь

$$s = s_0 \phi_0 + \sum_{\gamma \in \Gamma} (s_\gamma - s_{\gamma-1}) \phi_\gamma,$$

$\gamma = (\omega, \varphi)$ , где  $\omega \in [0, 2\pi)$ ,  $\varphi \in \{0, 1\}$

$$\phi_{(\omega, 0)} = \cos(\omega t), \quad \omega_k = \frac{2\pi k}{n}, \quad k = 0, \dots, n/2$$

$$\phi_{(\omega, 1)} = \sin(\omega t), \quad \omega_k = \frac{2\pi k}{n}, \quad k = 1, \dots, (n/2) - 1$$

# Временной словарь

$$s = s_0 \phi_0 + \sum_{\gamma \in \Gamma} (s_\gamma - s_{\gamma-1}) \phi_\gamma,$$

$$\gamma = (a, b, \varphi), \quad \psi = 1_{(1/2, 1]} - 1_{[0, 1/2]}, \quad \varphi = 1_{[0, 1]},$$

$$a = 2^j / n, \quad j = j_0, \dots, \log_2(n) - 1$$

$$b_{j,k} = ka_j, \quad k = 0, \dots, 2^j - 1,$$

$$\phi_{(a,b,0)} = \varphi(a(t-b)) \sqrt{a}$$

$$\phi_{(a,b,1)} = \psi(a(t-b)) \sqrt{a}$$

# Прореживание сигнала

- Косинусное преобразование;
- Преобразование Karhunen-Loève;
- Преобразование Уолша-Адамара;
- Преобразование Хаара.

# Двумерное дискретно-косинусное преобразование (DCT-II)

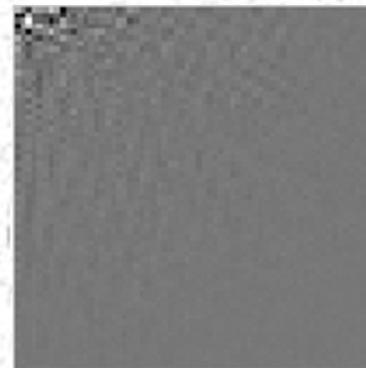
$$G_{i,j} = C_i C_j \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} X_{i,j} \cos \frac{(2i+1)i\pi}{2M} \cos \frac{(2j+1)j\pi}{2N},$$

$$C_i = \begin{cases} \frac{1}{\sqrt{M}} & \text{for } i = 0, \\ \sqrt{\frac{2}{M}} & \text{for } i \neq 0, \end{cases} \quad C_j = \begin{cases} \frac{1}{\sqrt{N}} & \text{for } j = 0, \\ \sqrt{\frac{2}{N}} & \text{for } j \neq 0. \end{cases}$$

# Трансформирование изображения методом DCT-II



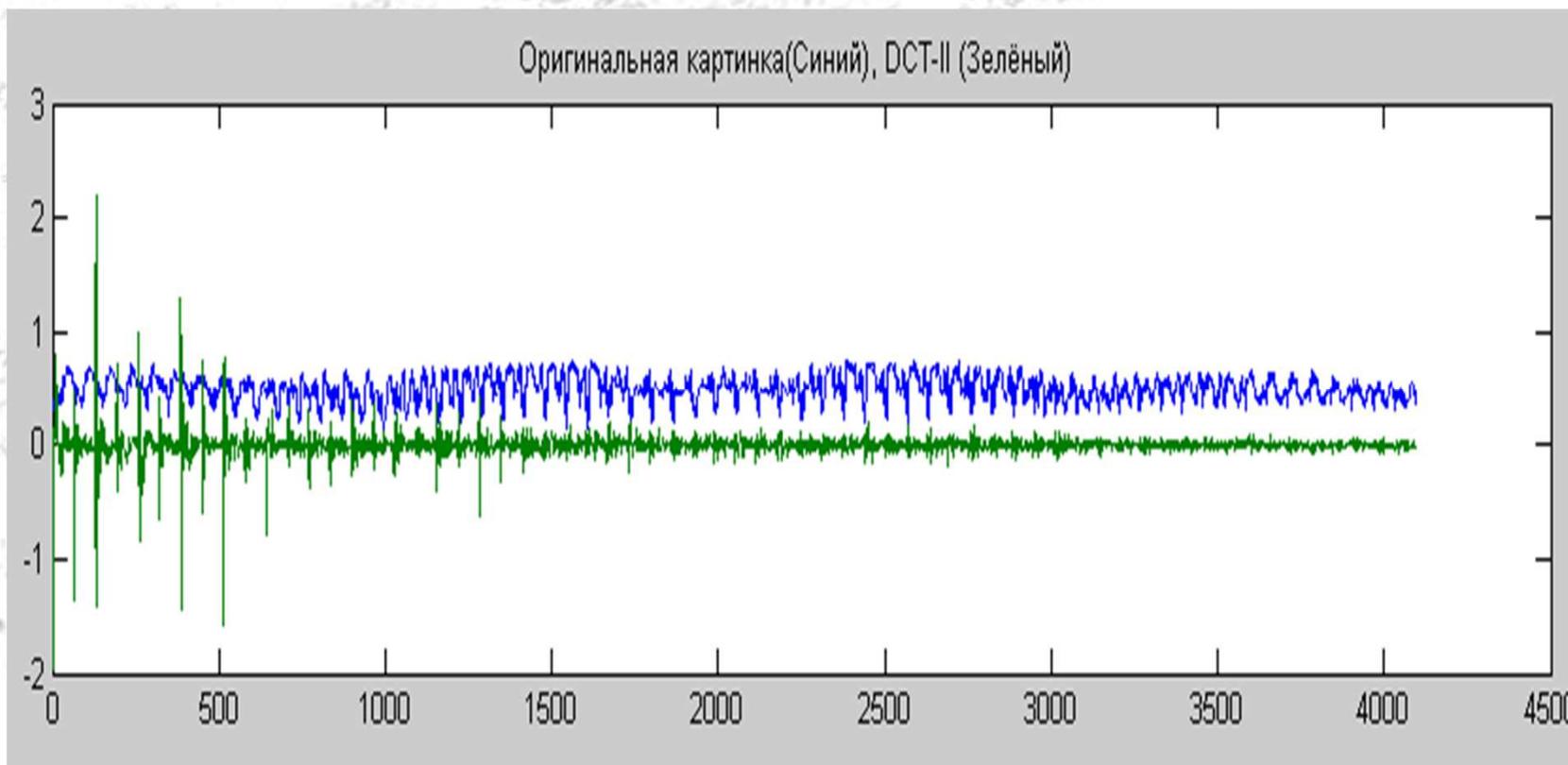
64 x 64



64 x 64

Для оригинальной картинки максимальный элемент равен 0.7572. Для DCT-II:  $G_{11} = 31.9507$ , следующий за ним максимальный элемент равен 2.1962.

# Графическое представление для DCT-II



# Двумерное вейвлет преобразование

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \phi(t - k) + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} d_{j,k} \psi(2^j t - k),$$

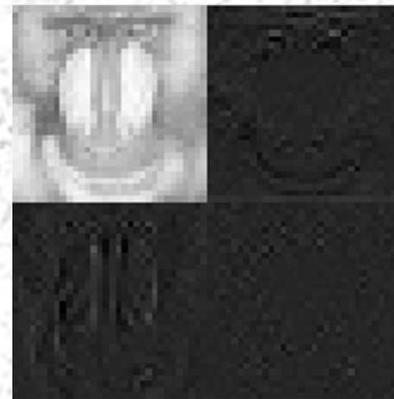
$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 0, & \text{иначе.} \end{cases}$$

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 0.5, \\ -1, & 0.5 \leq t < 1. \end{cases}$$

# Трансформирование изображения методом DWT-II



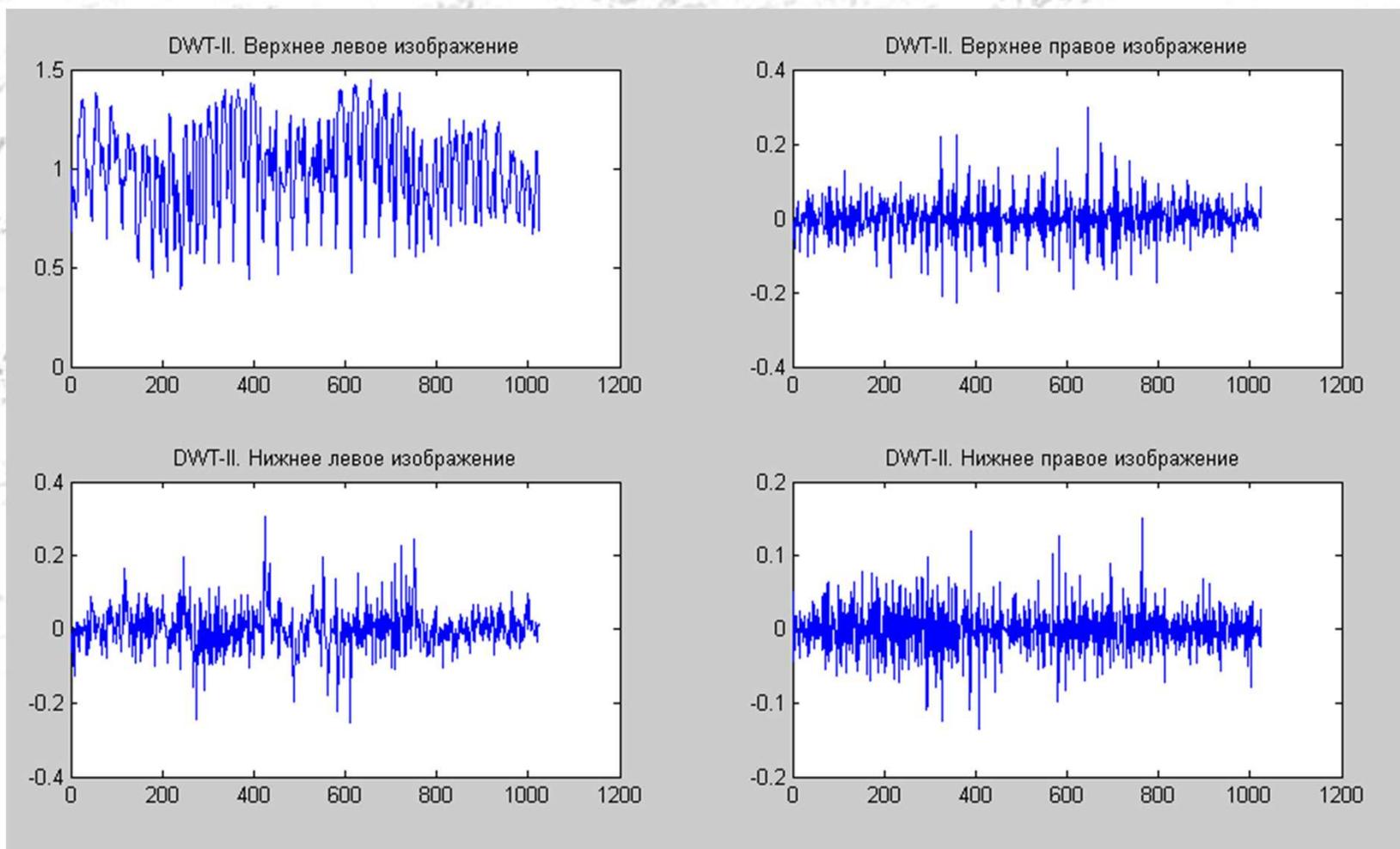
64 x 64



4 x 32 x 32

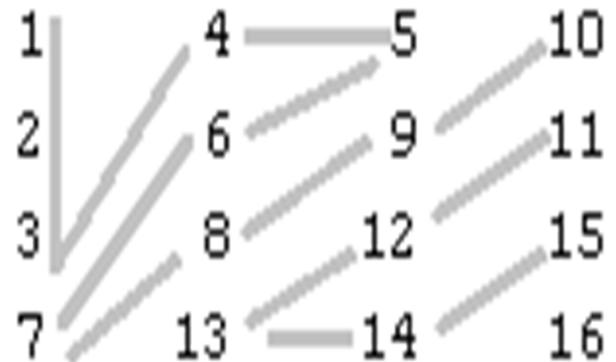
Максимальный элемент = 1.4462.

# Графическое представление для DWT-II



# Сжатие изображения (JPEG)

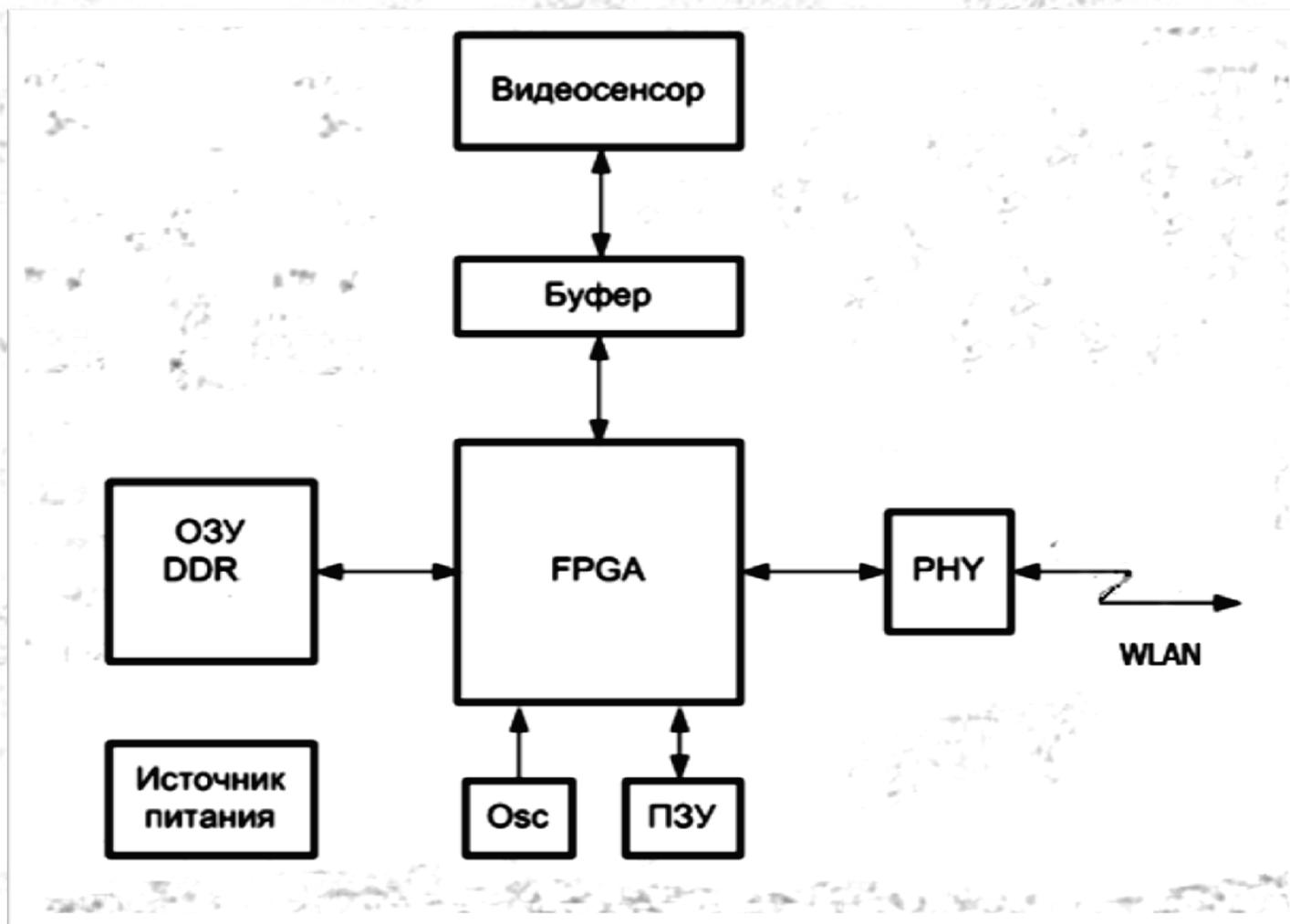
- Разбиение на 8×8 блоки;
- 2D-DCT;
- Квантование;
- Zig-Zag сканирование;
- Энтропийное сжатие;
- Упаковка данных в стандартный контейнер.



# Сжатие изображения (JPEG2000)

- Разбиение на 8×8 блоки;
- 2D-DWT;
- Квантование;
- Разделение каналов;
- Энтропийное сжатие;
- Упаковка данных в стандартный контейнер.

# Типовое аппаратное решение



# Ресурсы для реализации JPEG

Метод	Умножитель	Сдвиг	Сумматор
Прямое одномерное DCT	64	0	56
Chen	16	0	26
Zoeffler	11	0	29
Zoe Huang	12	5	29
AAN*	13	6	29
AAN unscaled	5	0	29

Систематизация от W. van Dyck, T. Berndorfer and all.

\**AAN algorithm* предложен в 1988 Arai, Agui and Nakajima

# Аппаратные ресурсы для реализации JPEG

Операция	Умножитель	BRAM	Регистр	LUTs
Выравнивание на входе	0	60 кБит	180	160
2D-DCT	10	18 кБит	580	1000
Квантование	1	0	35	24
Кодер переменной длины	0	18 кБит	660	700
ROM с табличными данными	0	18 кБит	0	0
<b>ИТОГО:</b>	<b>11</b>	<b>144 кБит</b>	<b>1455</b>	<b>1884</b>

На кристалле Spartan3 XC3S1000 фирмы XILINX.  $F_{\max} = 80$  МГц. Реализация W. van Dyck and all. (640x480)

# Аппаратные ресурсы для реализации JPEG2000

Операция	Умножитель	RAM	LEs
Выравнивание на входе	0	8 кБит	0
2D-DWT	12+1?	104 кБит	4400
RAM буфер	0	1 MBit	0
Квантование	2	0	1000
Разделение каналов	0	0	2000
RAM буфер	0	24 кБит	0
Энтропийный кодер	0	16 кБит	2000
<b>ИТОГО:</b>	<b>15</b>	<b>152 кБит</b>	<b>9400</b>

Систематизация от O. Cantineau и B. Jentz, BARCO SILEX. IP-ядро VA112JPEG2000E для кристалла Altera EP1S20FC484C5.

$F_{\max} = 113$  МГц (640x480, 1280x720)

# Сравнение ортогональных методов

## DCT

- Сложность алгоритма  $O(N \log N)$ ;
- Простой в реализации;
- Сжимает до 10-20 раз.

## DWT

- Сложность алгоритма  $O(N)$
- Ресурсоёмкий из-за необходимости подбора коэффициентов фильтров
- Потенциально сжимает на 20-200% лучше, чем JPEG.

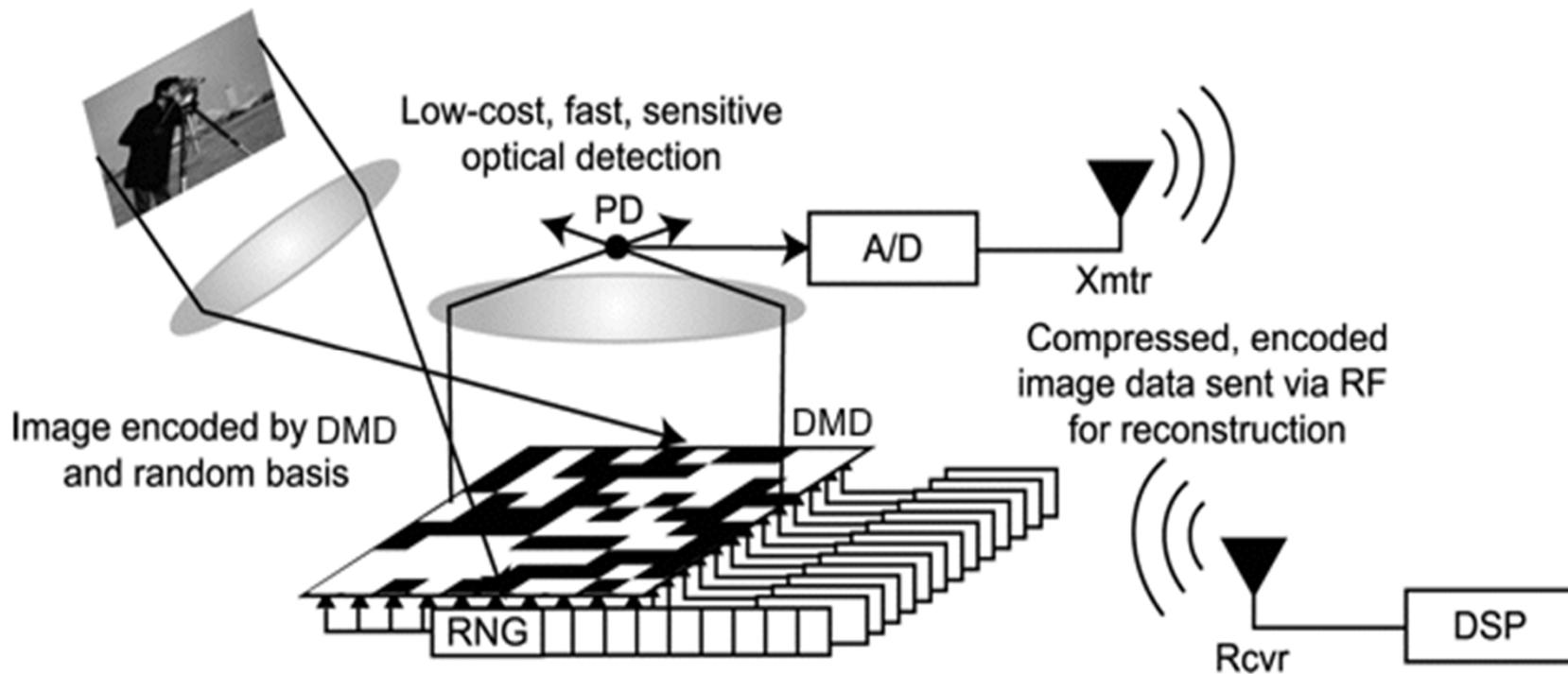
# Прямое получение сжатой информации

- Выравнивание потоков;
- Ортогональные преобразования;
- Энтропийное кодирование;
- Передача данных;
- Обработка информации;
- Принятие решения.

# Прямое получение сжатой информации

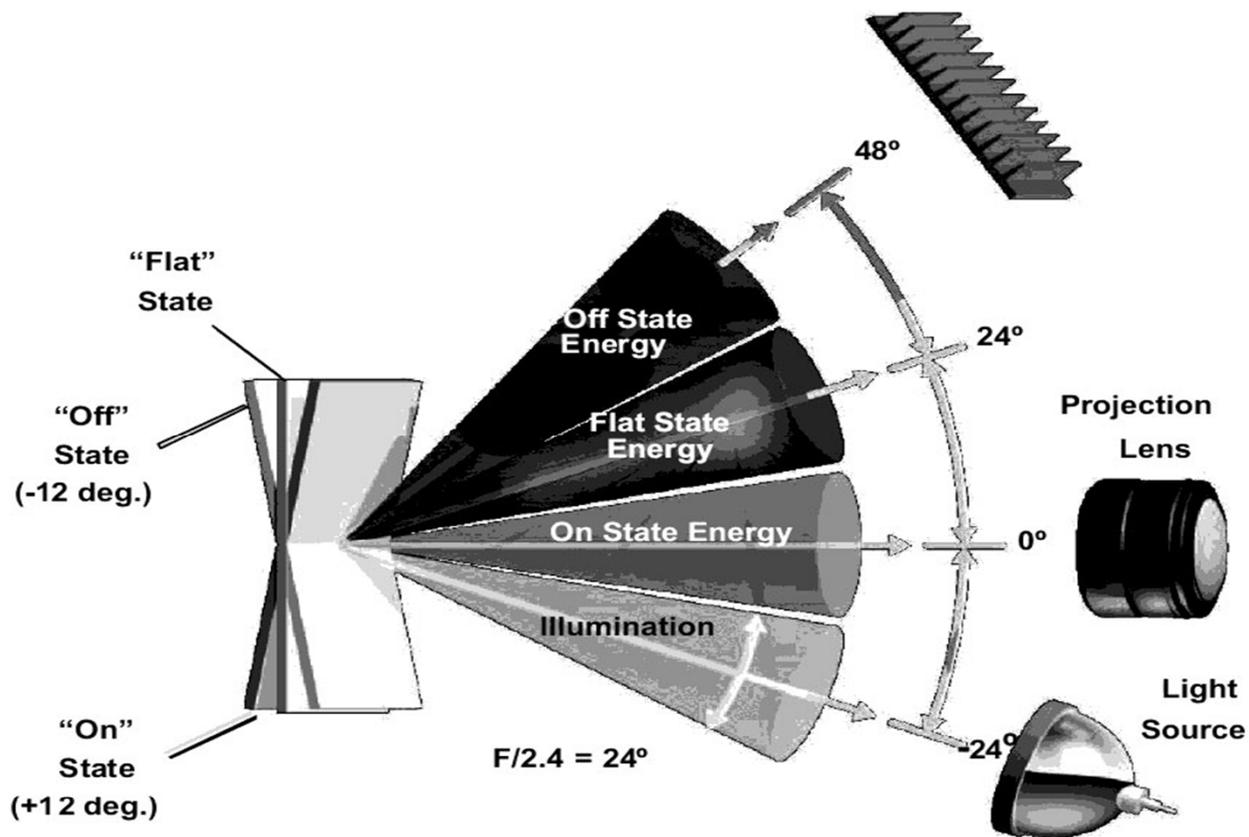
- ~~Выравнивание потоков;~~
- ~~Ортогональные преобразования;~~
- ~~Энтропийное кодирование;~~
- Передача данных;
- Обработка информации;
- Принятие решения.

# Однопиксельная камера



R. Baraniuk, K. Kelly and all.

# Принцип работы DMD



Производитель Texas Instruments

(<http://focus.ti.com/docs/prod/folders/print/dlp5500.html>)

# Compressive Sensing

Можно ли восстановить информацию полностью из существенно неполной входной информации?

Количество измерений  $m = 4S \log \frac{N^2}{S}$

# Получение измерений

$$X = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1N} \\ X_{21} & X_{22} & \dots & X_{2N} \\ X_{31} & X_{32} & \dots & X_{3N} \\ \vdots & \vdots & \ddots & \vdots \\ X_{M1} & X_{M2} & \dots & X_{MN} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2N} \\ \phi_{31} & \phi_{32} & \dots & \phi_{3N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1} & \phi_{m2} & \dots & \phi_{mN} \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

где  $\phi_{mN}$  – матрица  $1 \times MN$

$$Y = \Phi x$$

где  $Y$ - матрица измерений  $m \times N$ ,  $\Phi$ - матрица измерений  $1 \times MN$ ,  $X$ - входной двумерный массив данных размера  $M \times N$ ,  $x$  - вектор столбец  $MN \times 1$ , полученный из  $X$  простой перестановкой строк в столбцы.

# Выбор матрицы измерения

- Удовлетворять Restricted Isometric Property;
- Не должна быть когерентной.

# Восстановление информации

$$\min \|\hat{x}\|_{\ell_1}, \Phi \hat{x} = Y$$

# Basis Pursuite

$$\min_{\alpha_{\gamma_i}} \left\| s - \sum_{i=1}^m \alpha_{\gamma_i} \phi_{\gamma_i} \right\|_{\ell_1} \leq R^{(m)} \rightarrow 0$$

2004 г. S.S. Chen, D. L. Donoho и M. A. Saunders

# Алгоритм для программируемой логической интегральной схемы (FPGA)

- Матрица измерений  $\Phi$  из  $\{0,1\}$  со случайным распределением Бернулли: вероятность следования '1' в матрице равна  $P_a$ .
- Моделирование на MATLAB

# Измерения при разных $P_a$



$P_a = 0.001$



$P_a = 0.5$



$P_a = 0.01$



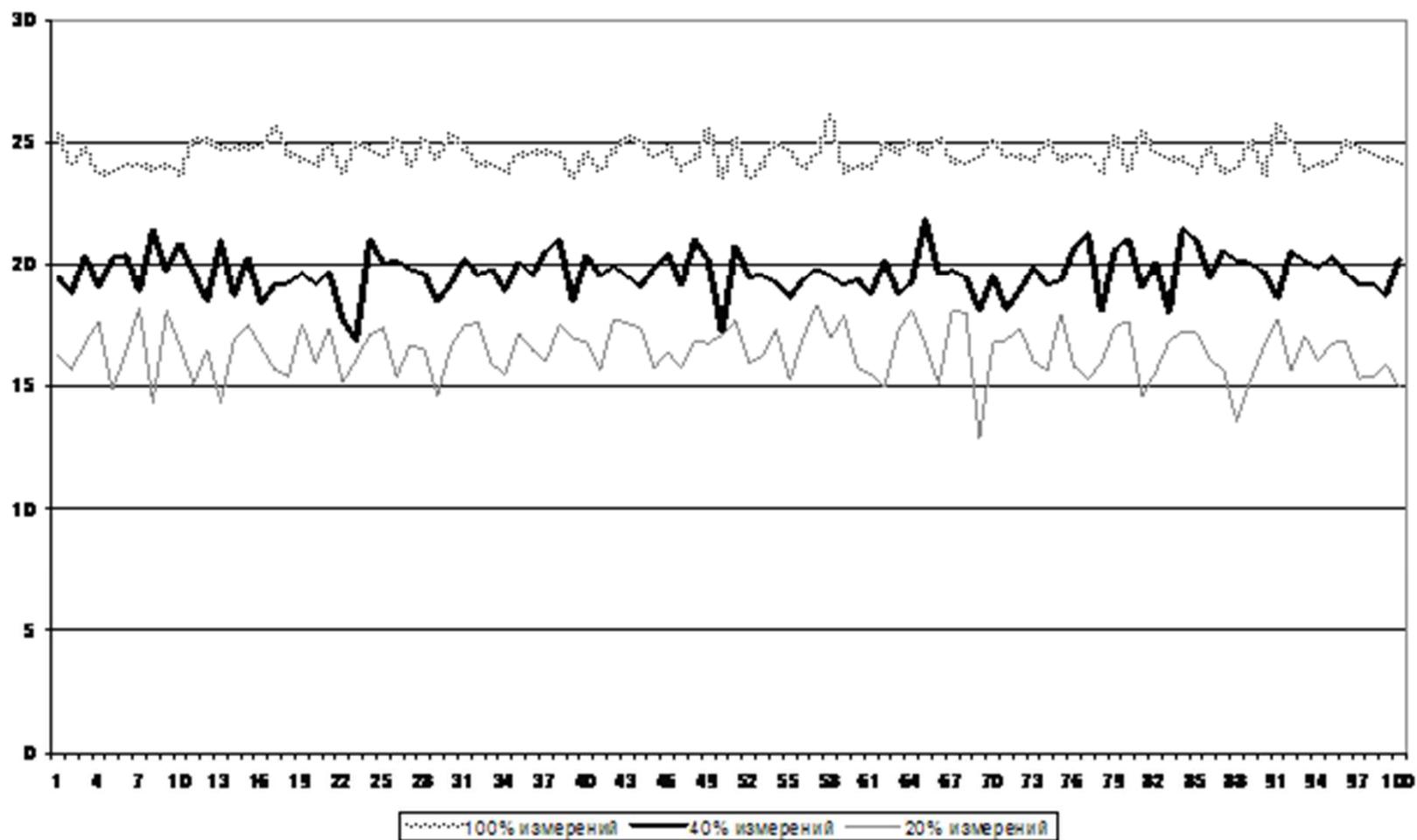
$P_a = 0.8$

На рисунках представлены: Оригинал 100%, 40% и 20% измерений

# Численные данные при разных $m$

Размер изображения, $pix$	$32 \times 32$	$32 \times 32$	$32 \times 32$
Погрешность восстановления, $\epsilon$	0.01	0.01	0.01
Среднеквадратичная ошибка	0.0014	0.0067	0.0112
Пиковое отношение сигнал/шум, $dB$	28.5235	21.7098	19.5096
Отношение сигнал/шум, $dB$	25.7745	18.9609	16.7607
Количество измерений, $m$	1024	409	204
Вероятность распределения, $P_a$	50%	50%	50%

# Статистика сжатий и восстановлений



# Оценка аппаратных затрат

Для реализации аппаратного кодера использовался язык VHDL. Оценка ресурсов, требуемых для реализации в FPGA для кристалла EP2C5 Cyclon II фирмы ALTERA, при работе с одним блоком размером 64 x64 пикселя:

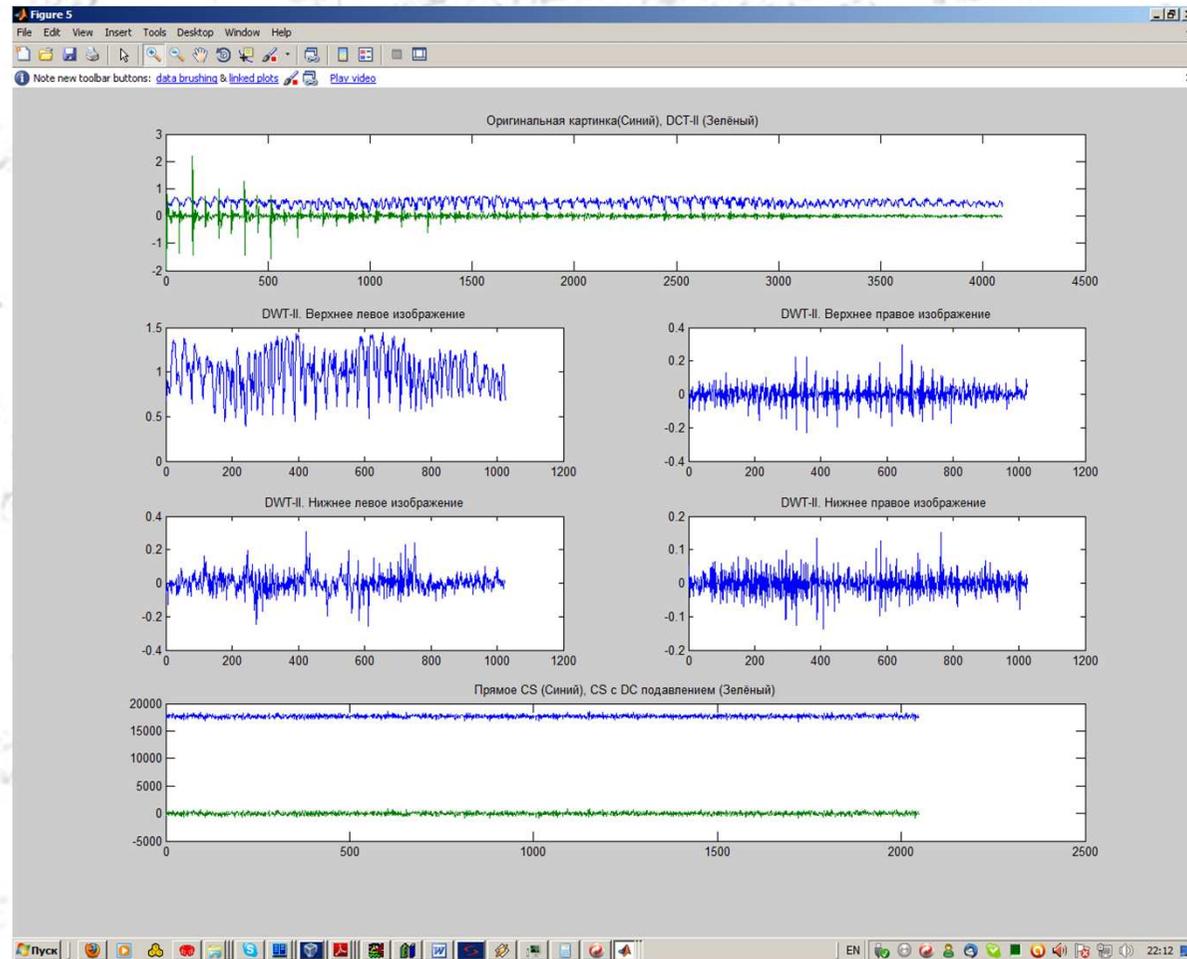
- количество логических элементов --- 165;
- количество регистров --- 95;
- объем памяти --- 64 кБит.

# Реализация второй версии упаковщика



- Восстановленная картинка (справа) PSNR = 33dB с размером 65кБ (энтропийный энкодер 7ZIP), при размере оригинала 113кБ.

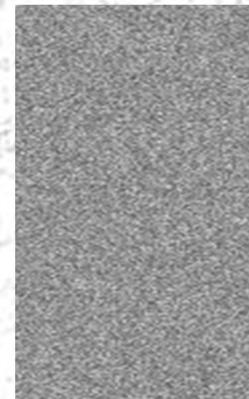
# Общий вид трёх методов сжатия



# Трансформирование изображения методом Compressive Sensing



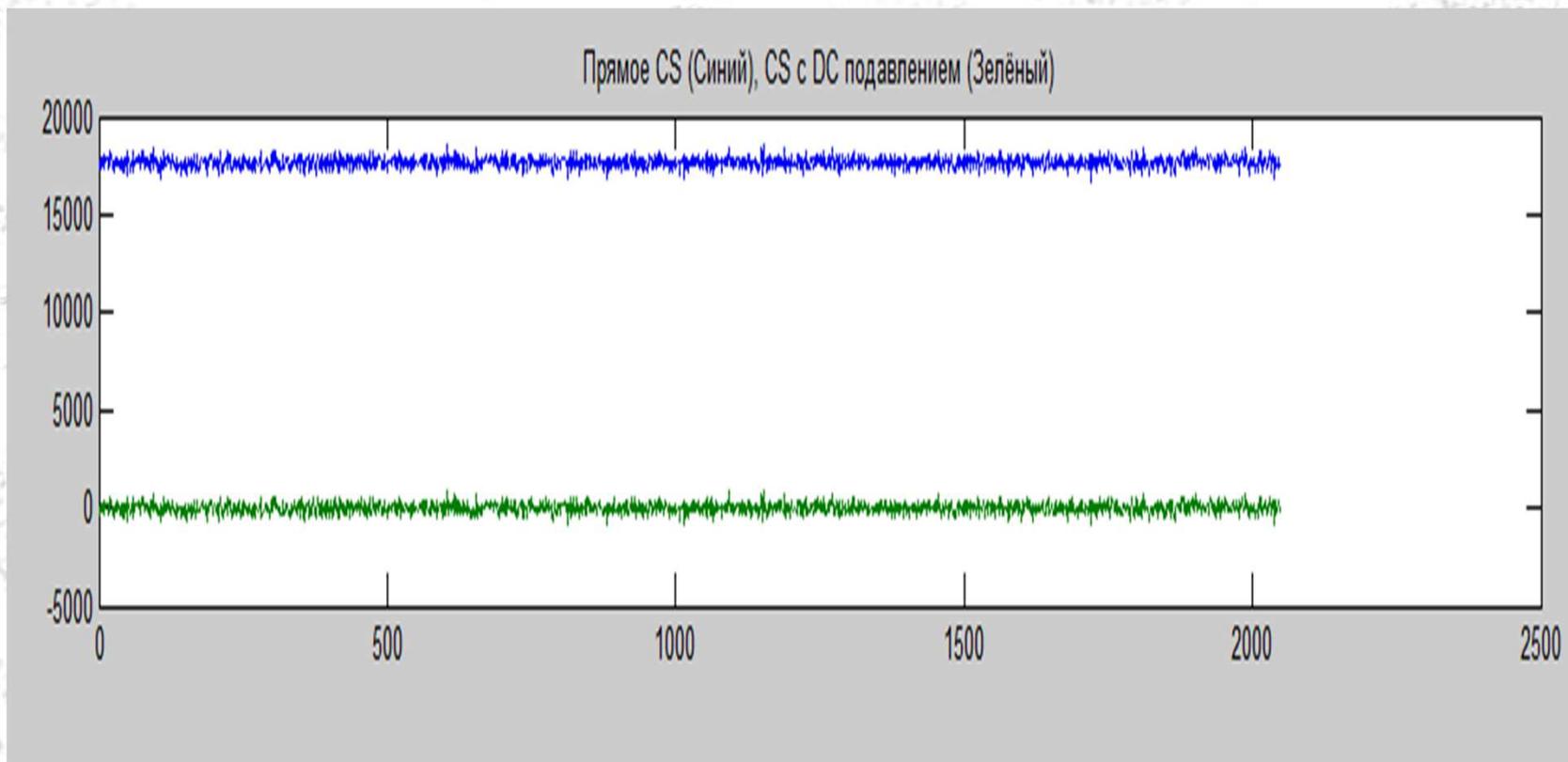
64 x 64



64x 32

Максимальный элемент не более 12 бит.

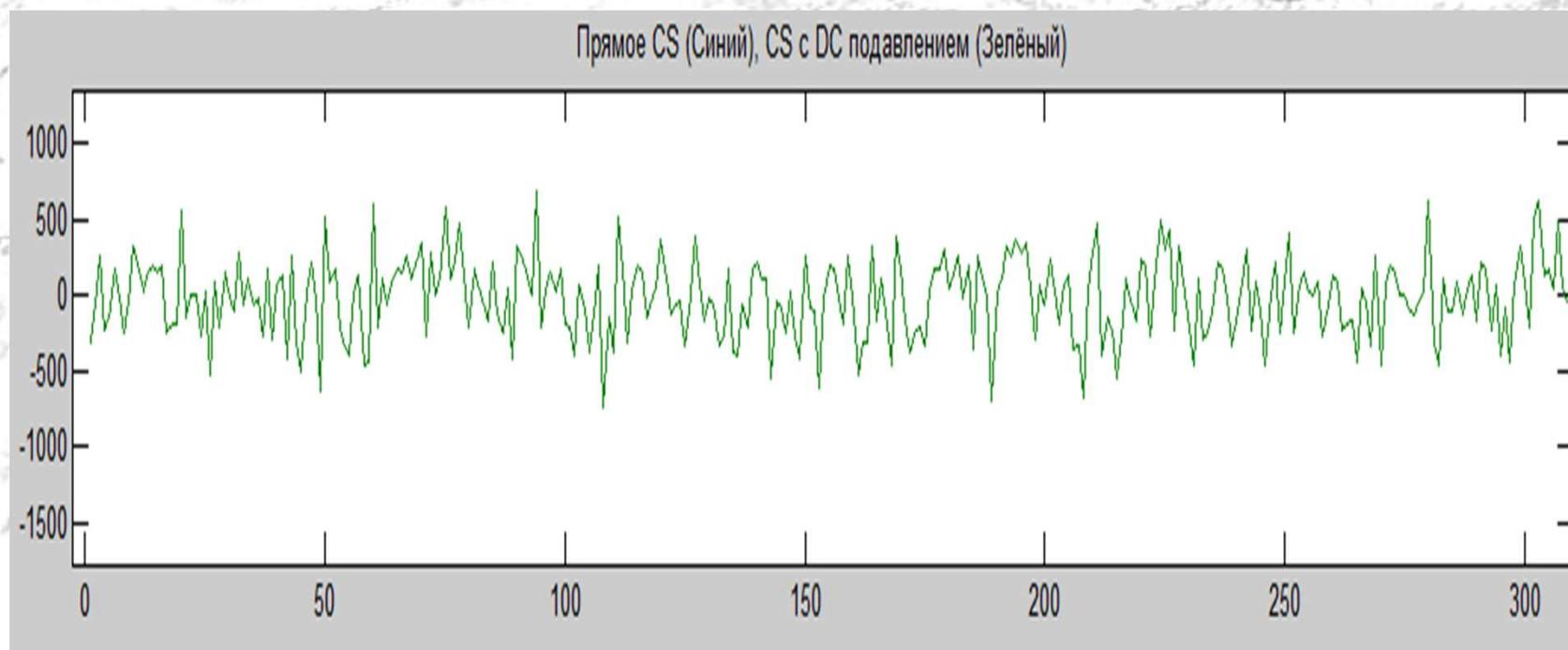
# Метод на базе Compressive Sensing



Распределение измеренных значений (синий), и  
устранение постоянной составляющей

# Метод на базе Compressive Sensing

- Оценка разрядности данных



Спасибо за внимание!



[oleg.p.isaev@gmail.com](mailto:oleg.p.isaev@gmail.com)

Таруса'2011