

Настоящий блок лекций подготовлен при поддержке «ЦРТ | Группа компаний» и Университета ИТМО





Обработка речевых сигналов Блок 1. Цифровая обработка сигналов

 Цифровой сигнал
 Нейронные сети
 Преобразование Фурье

 Спектр и спектрограмма
 Цифровой фильтр

 Теорема отсчетов
 Частота дискретизации

Владимир Волохов

Научный сотрудник ООО «ЦРТ-инновации», доцент, к.т.н.



### Блок 1. Цифровая обработка сигналов



text-snadow. filter: dropshadow(color= color:#777: header #main-navigation ut li "TEDELET" TOX-shadow Or box-shadows ODX COX -moz-box-shadow hackground-color:#F9F9F9 n ut li accident

Часть 3. Цифровые фильтры

### План лекции

- Что такое цифровой фильтр?
- Что такое линейный инвариантный к сдвигу фильтр?
- Что такое физически реализуемый фильтр?
- Что такое z-преобразование?
- Нерекурсивные и рекурсивные линейные инвариантные к сдвигу фильтры
- Основные характеристики линейных инвариантных к сдвигу фильтров
- Что такое взаимная корреляция и свёртка?
- Теорема о свёртке
- Фильтры как инструмент создания звуковых эффектов

### Цифровые фильтры и их разновидности

- Цифровой фильтр это любая цифровая система, которая согласно заданному алгоритму осуществляет извлечение цифрового сигнала либо его параметров из действующей на входе системы смеси сигнала с помехой
  - Частотно-избирательные фильтры (фильтр нижних частот ФНЧ, фильтр верхних частот ФВЧ, полосовой фильтр ПФ и режекторный фильтр РФ)
  - Фазовые корректоры (всепропускающий фильтр)
  - ...
- Линейный инвариантный к сдвигу фильтр (ЛИС-фильтр)

Линейность: 
$$L\{a_1x_1(n) + a_2x_2(n)\} = a_1L\{x_1(n)\} + a_2L\{x_2(n)\}$$

Инвариантность к сдвигу: 
$$L\{x(n)\} = y(n)$$
,  $L\{x(n-n_0)\} = y(n-n_0)$ 

• Физически реализуемый фильтр (реакция не может возникнуть раньше воздействия)

### Что такое z-преобразование?

Прямое и обратное z-преобразование

#### Прямое z-преобразование

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}, z \in C$$

Z-преобразование является степенным рядом, поэтому сходимость является всегда абсолютной:

$$X(z) < \infty \Leftrightarrow \sum_{n=-\infty}^{\infty} |x(n)z^{-n}| < \infty$$

#### Обратное z-преобразование

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz, n \in \mathbb{Z}$$

### Что такое z-преобразование?

Прямое и обратное z-преобразование

#### Прямое z-преобразование

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}, z \in C$$

Z-преобразование является степенным рядом, поэтому сходимость является всегда абсолютной:

$$X(z) < \infty \Leftrightarrow \sum_{n=-\infty}^{\infty} |x(n)z^{-n}| < \infty$$

#### Обратное z-преобразование

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz, n \in \mathbb{Z}$$

Размышляем о z-преобразовании как о формальном операторе или как о расширении преобразования Фурье дискретного времени:

$$X(z)\Big|_{z=e^{j\widehat{\omega}}}=X(e^{j\widehat{\omega}})$$

### Преобразование Фурье дискретного времени

$$X(e^{j\widehat{\omega}}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\widehat{\omega}n}, \widehat{\omega} \in R$$

### Что такое z-преобразование?

Прямое и обратное z-преобразование

#### Прямое z-преобразование

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}, z \in C$$

Z-преобразование является степенным рядом, поэтому сходимость является всегда абсолютной:

$$X(z) < \infty \Leftrightarrow \sum_{n=-\infty}^{\infty} |x(n)z^{-n}| < \infty$$

#### Обратное z-преобразование

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz, n \in \mathbb{Z}$$

#### Единичный импульс

$$X(z) = \sum_{n=-\infty}^{\infty} \delta(n)z^{-n} = 1 \cdot z^{-0} = 1$$

Единичный скачок
$$x(n) = u(n)$$

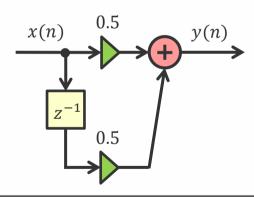
$$X(z) = \sum_{n=-\infty}^{\infty} u(n)z^{-n} = \sum_{n=-\infty}^{\infty} z^{-n}$$

$$= \frac{z}{z-1}, |z| > 1$$

### Нерекурсивные и рекурсивные ЛИС-фильтры

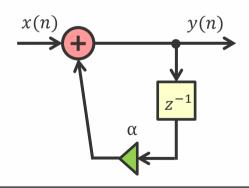
• Арифметическое усреднение

$$y(n) = \frac{x(n) + x(n-1)}{2}$$



Нерекурсивный КИХ-фильтр первого порядка • Система «с памятью» (обратной связью)

$$y(n) = x(n) + \alpha y(n-1)$$



### Основные характеристики ЛИС-фильтров

Арифметическое усреднение

Система «с памятью» (обратной связью)

ЛИС-фильтры описываются линейным разностным уравнением с постоянными коэффициентами

$$y(n) = \frac{x(n) + x(n-1)}{2}$$

$$y(n) = x(n) + \alpha y(n-1), \ |\alpha| < 1$$

Импульсная характеристика (ИХ) является откликом ЛИС-фильтра на единичный импульс

Конечная ИХ 
$$h(n) = \frac{1}{2}\delta(n) + \frac{1}{2}\delta(n-1)$$

$$h(n) = \delta(n) + \alpha h(n-1) = \alpha^n u(n)$$
 Бесконечная ИХ

Переходная характеристика (ПХ) является откликом ЛИС-фильтра на единичный скачок

$$g(n) = \frac{1}{2}u(n) + \frac{1}{2}u(n-1)$$

$$g(n) = u(n) + \alpha g(n-1) = \left(\frac{1}{1-\alpha} - \frac{\alpha^{n+1}}{1-\alpha}\right) u(n)$$

### Основные характеристики ЛИС-фильтров

Арифметическое усреднение

• Система «с памятью» (обратной связью)

Передаточная функция (ПФ) фильтра является отношением z-образа сигнала на его выходе к z-образу сигнала на его входе при нулевых начальных условиях

$$Y(z) = \frac{1}{2}X(z) + \frac{1}{2}X(z) \cdot z^{-1}$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{2} + \frac{1}{2}z^{-1} = \frac{z+1}{2z}$$

#### Свойство линейности

$$Z(\alpha x(n) + \beta y(n))$$
  
=  $\alpha Z(x(n)) + Z(\beta y(n))$ 

#### Свойство задержки

$$x(n) \leftrightarrow X(z)$$
  
 $x(n-1) \leftrightarrow X(z) \cdot z^{-1}$ 

$$Y(z) = X(z) + \alpha Y(z) \cdot z^{-1}$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - \alpha z^{-1}} = \frac{z}{z - \alpha}$$

Нули – корни числителя ПФ, полюсы – корни знаменателя ПФ

Ноль:  $z_0 = -1$ ; полюс:  $z_{\times} = 0$ 

Ноль:  $z_0 = 0$ ; полюс:  $z_{\times} = \alpha$ 

### Основные характеристики ЛИС-фильтров

Арифметическое усреднение

Система «с памятью» (обратной связью)

Частотная характеристика (ЧХ) фильтра представляет собой передаточную функцию в точках на единичной окружности:  $H(z)|_{z=e^{j\widehat{\omega}}}=H(e^{j\widehat{\omega}})$ 

$$H(e^{j\widehat{\omega}}) = \frac{Y(e^{j\widehat{\omega}})}{X(e^{j\widehat{\omega}})} = \frac{1}{2} + \frac{1}{2}e^{-j\widehat{\omega}} = \frac{e^{j\widehat{\omega}} + 1}{2e^{j\widehat{\omega}}}$$

$$H(e^{j\widehat{\omega}}) = \frac{Y(e^{j\widehat{\omega}})}{X(e^{j\widehat{\omega}})} = \frac{1}{1 - \alpha e^{-j\widehat{\omega}}} = \frac{e^{j\widehat{\omega}}}{e^{j\widehat{\omega}} - \alpha}$$

Амплитудно-частотная характеристика (АЧХ) является модулем частотной характеристики

ФНЧ

$$|H(e^{j\widehat{\omega}})| = \sqrt{(1 + \cos(\widehat{\omega}))/2}$$

$$|H(e^{j\widehat{\omega}})| = \frac{1}{\sqrt{1 - 2\alpha\cos(\widehat{\omega}) + \alpha^2}}$$
  $\Phi$ HY:  $\alpha > 0$   $\Phi$ BY:  $\alpha < 0$ 

Фазочастотная характеристика (ФЧХ) является аргументом частотной характеристики

$$\arg\left(H\left(e^{j\widehat{\omega}}\right)\right) = -\tan^{-1}(\sin(\widehat{\omega})/(1+\cos(\widehat{\omega}))) \qquad \arg\left(H\left(e^{j\widehat{\omega}}\right)\right) = -\tan^{-1}(\alpha\cdot\sin(\widehat{\omega})/(1-\alpha\cdot\cos(\widehat{\omega})))$$



### Что происходит с гармоническим сигналом в ЛИС-фильтре?

Рассмотрим ЛИС-фильтр с АЧХ и ФЧХ, описываемыми функциями  $|H(e^{j\widehat{\omega}})|$  и  $\arg \left(H(e^{j\widehat{\omega}})\right)$ , соответственно

#### Сигнал на входе ЛИС-фильтра

$$x(n) = A_1 \cos(\widehat{\omega}_1 n + \varphi_1) + A_2 \cos(\widehat{\omega}_2 n + \varphi_2)$$

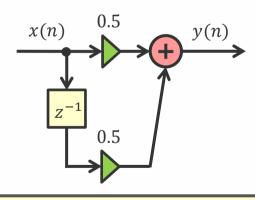
#### Сигнал на выходе ЛИС-фильтра

$$y(n) = A_1 |H(e^{j\widehat{\omega}_1})| \cos(\widehat{\omega}_1 n + \varphi_1 + \arg(H(e^{j\widehat{\omega}_1})))$$
$$+ A_2 |H(e^{j\widehat{\omega}_2})| \cos(\widehat{\omega}_2 n + \varphi_2 + \arg(H(e^{j\widehat{\omega}_2})))$$

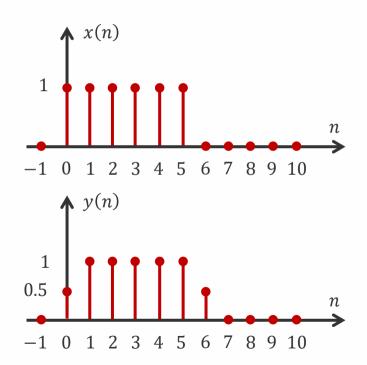


• Арифметическое усреднение

$$y(n) = \frac{x(n) + x(n-1)}{2}$$



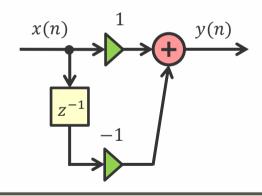
Нерекурсивный КИХ-фильтр первого порядка, нижних частот!



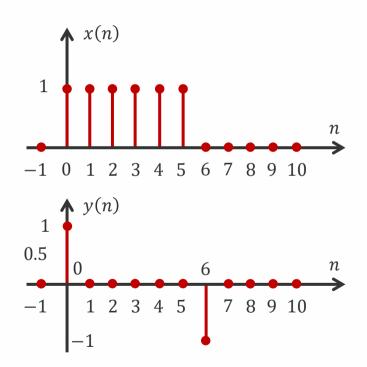


Вычисление «производной» сигнала

$$y(n) = x(n) - x(n-1)$$



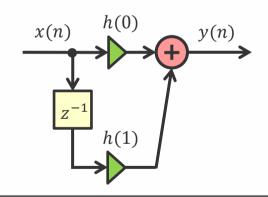
Нерекурсивный КИХ-фильтр первого порядка, верхних частот!





Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



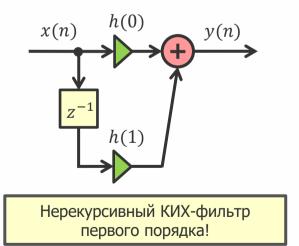
Нерекурсивный КИХ-фильтр первого порядка!

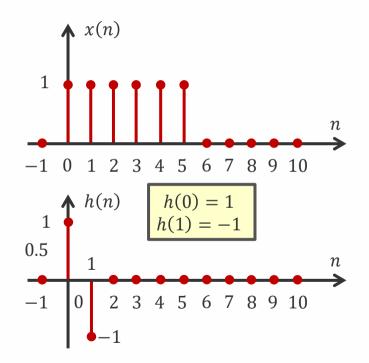
Описать работу ЛИС-фильтра можно через скалярное произведение двух последовательностей бесконечной длины!



#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$

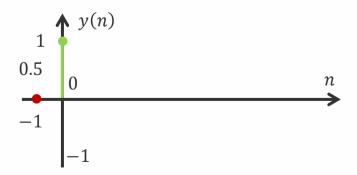


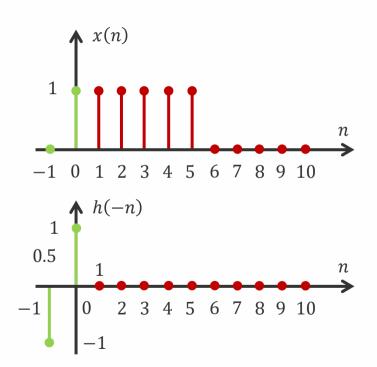




#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$

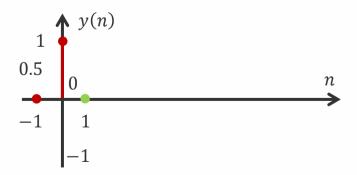


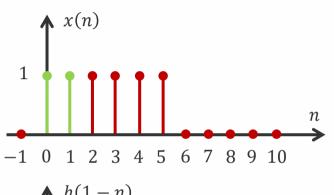


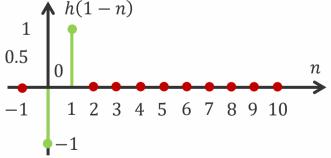


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



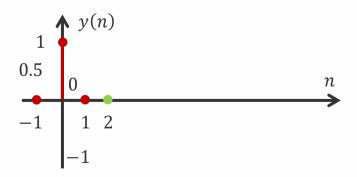


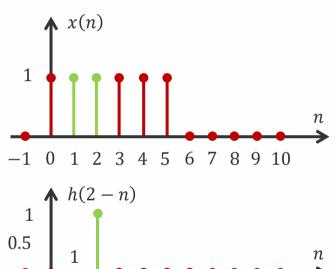


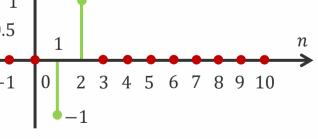


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



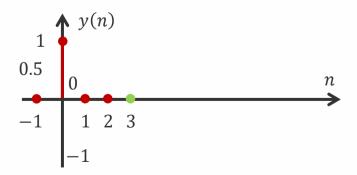


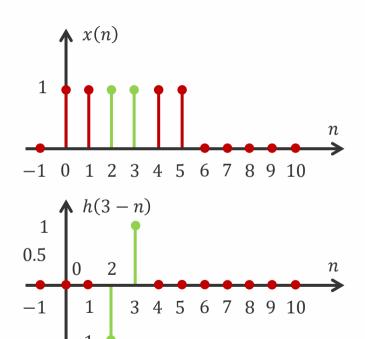




#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$

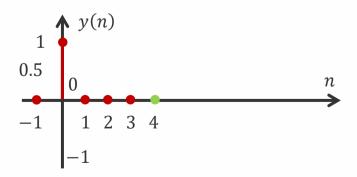


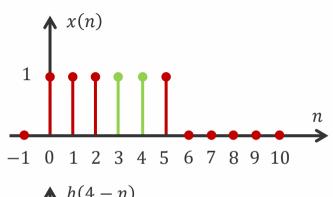


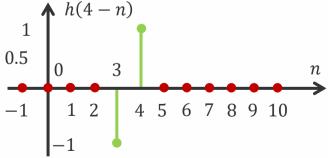


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



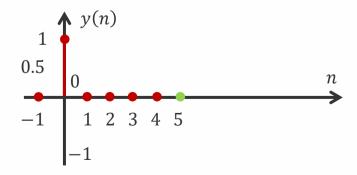


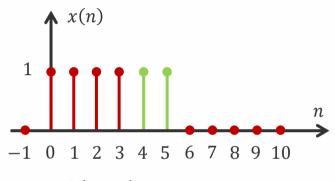


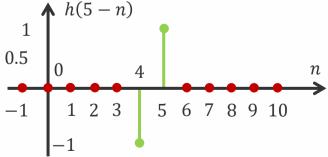


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



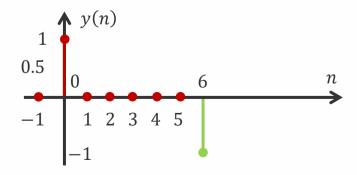


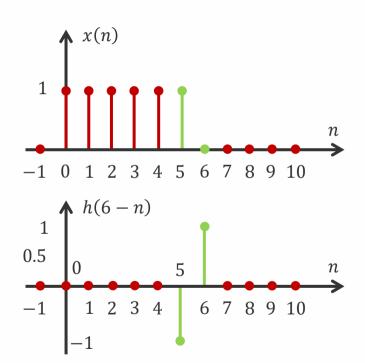




#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$

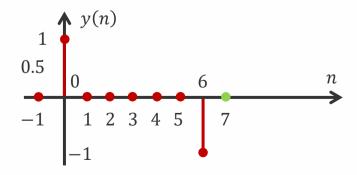


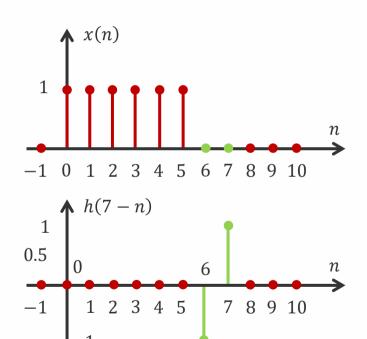




#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$

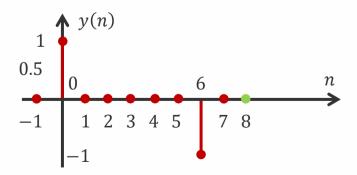


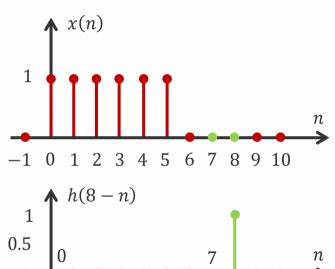


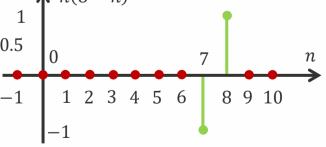


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



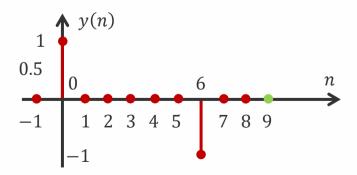


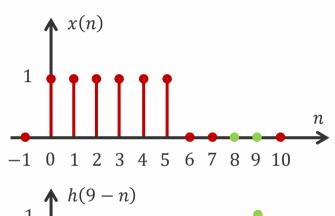


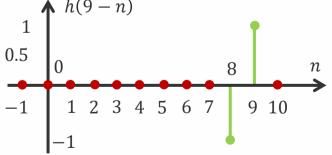


#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$



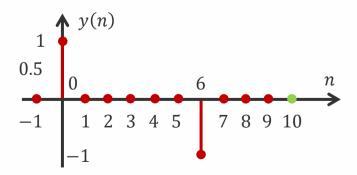


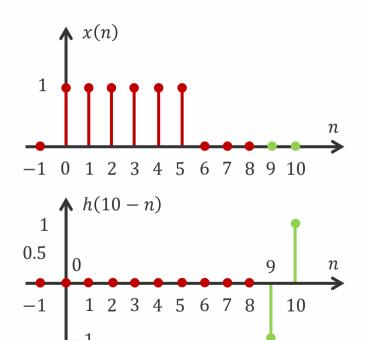




#### Вычисление отклика фильтра

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n-1)$$





### Что такое взаимная корреляция и свёртка?

■ Взаимная корреляция сигналов x(n) и h(n):

$$y(n) = \sum_{u=-\infty}^{\infty} h(u)x(n+u)$$

• Свёртка представляет собой взаимную корреляцию, в которой один из двух сигналов инвертируется по временной оси перед началом выполнения операции:

$$y(n) = \sum_{u=-\infty}^{\infty} h(u)x(n-u)$$



### Взаимная корреляция и свёртка для двумерного случая

■ Взаимная корреляция изображений x(n,m) и h(n,m):

$$y(n,m) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} h(u,v)x (n+u,m+v)$$

• Свёртка представляет собой взаимную корреляцию, в которой одно из двух изображений зеркально отражается по горизонтали и вертикали перед началом выполнения операции

$$y(n,m) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} h(u,v)x (n-u,m-v)$$



# Пример выполнения взаимной корреляции для изображений

0	0	0	0	0	0
0	90	90	90	90	0
0	90	90	90	90	0
0	90	0	90	90	0
0	90	90	90	90	0
0	0	0	0	0	0

Фрагмент исх. изобр.

ядро (маска) = ИХ					
1	1	1			
1	1	1			
1	1	1			

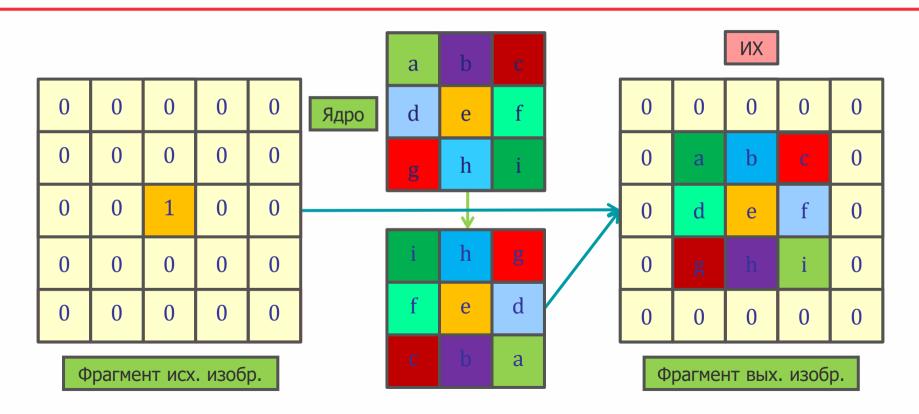
Для ЛИС-систем

X	X	X	X	X	X
X	40	60	60	40	X
X	50	80	80	60	X
X	50	80	80	60	X
X	30	50	50	40	X
X	X	X	X	X	X

Ядро фильтра

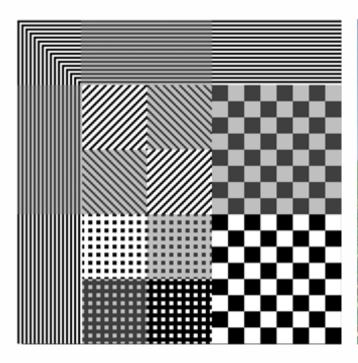
Фрагмент вых. изобр.

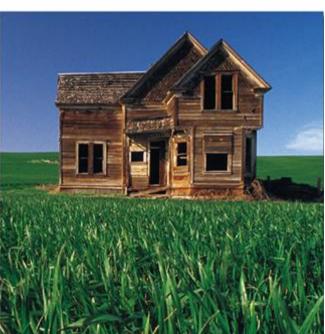
### Пример выполнения свёртки для изображений





# Пример выполнения свёртки: исходные изображения

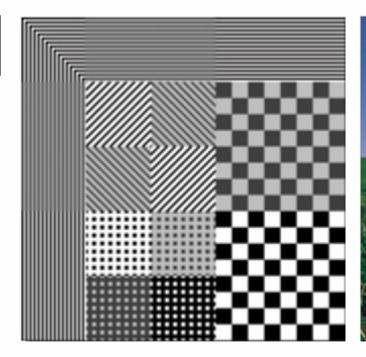


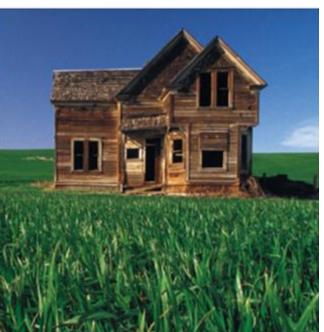




## Пример выполнения свёртки: размытие с маской 3x3

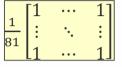
 $\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ 

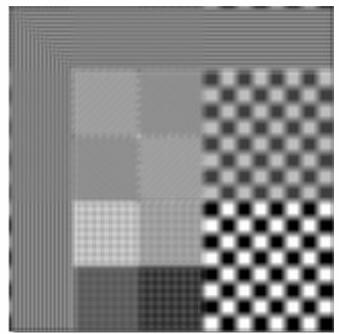


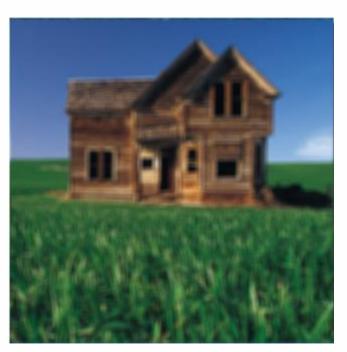




## Пример выполнения свёртки: размытие с маской 9х9



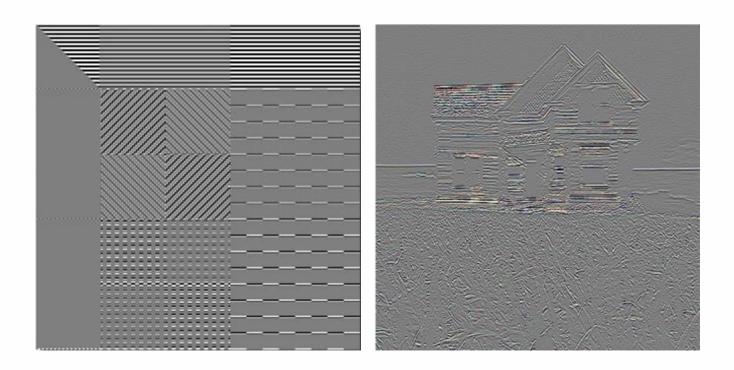






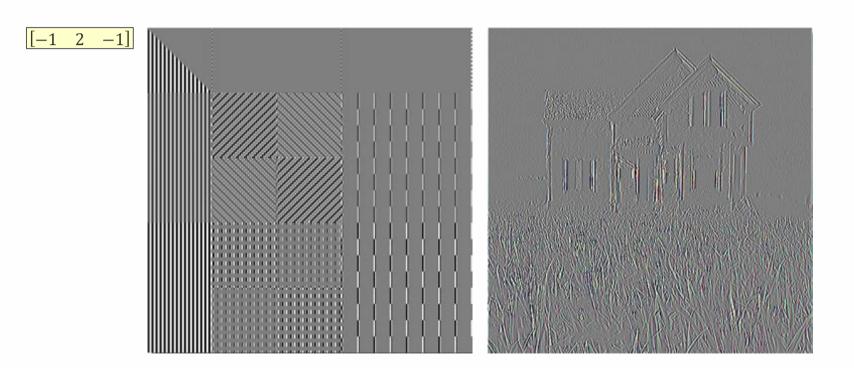
# Пример выполнения свёртки: вертикальные разности



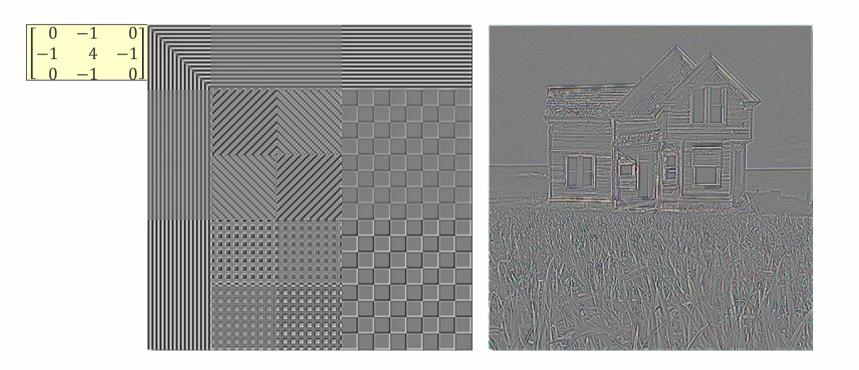




# Пример выполнения свёртки: горизонтальные разности

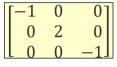


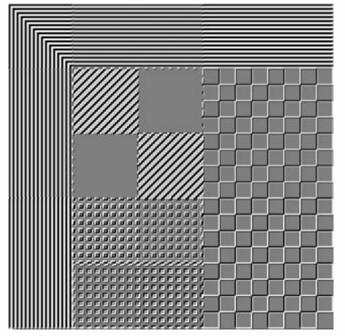
# Пример выполнения свёртки: вертикальные + горизонтальные разности





# Пример выполнения свёртки: диагональные разности

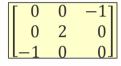


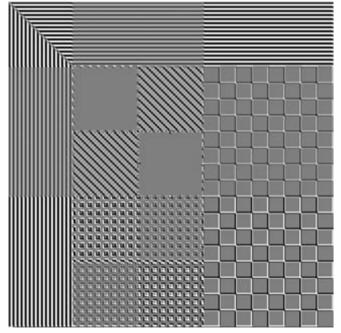






# Пример выполнения свёртки: диагональные разности

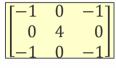


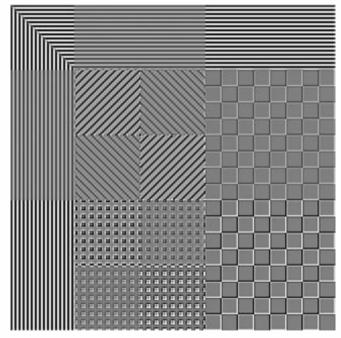






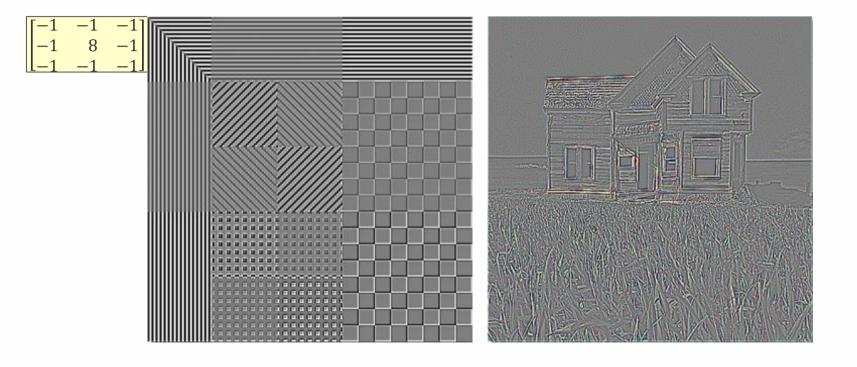
# Пример выполнения свёртки: диагональные + диагональные разности







## Пример выполнения свёртки: вертикальные + горизонтальные + диагональные разности



### Теорема о свёртке

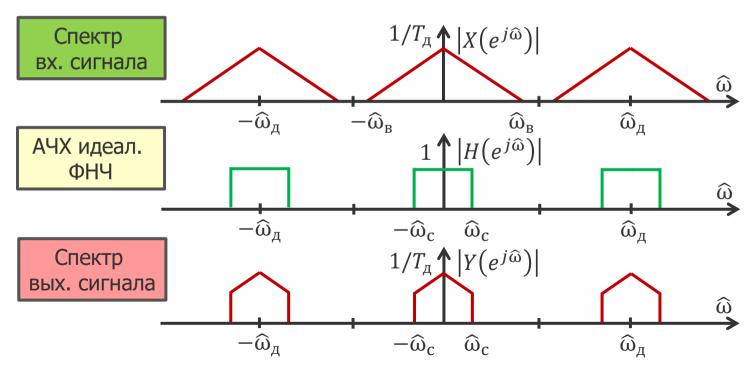
• Свёртка двух сигналов (изображений) во временной области (в пространственной области) соответствует умножению их образов Фурье в частотной области

$$y(n) = \sum_{u = -\infty}^{\infty} h(u)x(n - u)$$
$$Y(e^{j\widehat{\omega}}) = H(e^{j\widehat{\omega}}) \cdot X(e^{j\widehat{\omega}})$$

$$y(n,m) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} h(u,v)x (n-u,m-v)$$
$$Y(e^{j\widehat{\omega}_{1}},e^{j\widehat{\omega}_{2}}) = H(e^{j\widehat{\omega}_{1}},e^{j\widehat{\omega}_{2}}) \cdot X(e^{j\widehat{\omega}_{1}},e^{j\widehat{\omega}_{2}})$$

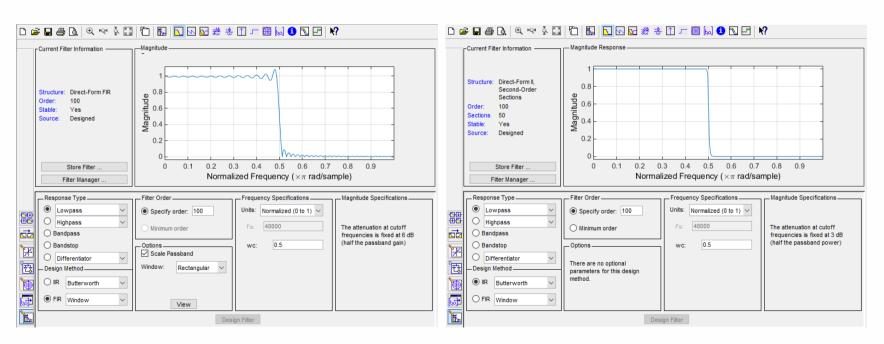
## Пример фильтрации в частотной области

• Рассмотрим идеальный фильтр нижних частот



### Реальные фильтры выглядят иначе!

• Примеры фильтров, синтезированных в FDATool пакета MATLAB





## Фильтры как инструмент создания звуковых эффектов

### • Простое эхо

$$y(n) = \frac{ax(n) + bx(n-N) + cx(n-2N)}{a+b+c}$$

Равноотстоящие повторения, масштабированные на различный вес для моделирования затухания во времени

«Натуральное» эхо

$$y(n) = x(n) + \alpha y(n - M)$$

$$y(n) = x(n) + \alpha \cdot \sum_{u = -\infty}^{\infty} h(u)y(n - M - u)$$

Настоящее эхо имеет бесконечное число повторений

Настоящее эхо содержит низкочастотные особенности



## Фильтры как инструмент создания звуковых эффектов

 Реверберация – суперпозиция большого числа отражённых сигналов с различными задержками и амплитудами

$$y(n) = -\alpha x(n) + x(n - M) + \alpha y(n - M)$$

Простейший вариант моделирования реверберации возможен с использованием всепропускающего фильтра

- Реверберацию можно промоделировать путем свёртки сигнала с импульсным откликом конкретного помещения, рассматривая помещение как ЛИС-систему
  - Импульсный отклик можно измерить естественным образом, хлопнув воздушный шарик под потолком какой-либо комнаты и записав импульсный отклик в некоторой ее точке
  - Импульсный отклик можно промоделировать с использованием некоторого программного обеспечения, например pyroomacoustics (Pyroomacoustics: A Python package for audio room simulations and array processing algorithms [Scheibler R., Bezzam E., Dokmanić I., 2018])



## Фильтры как инструмент создания звуковых эффектов

- Некоторые гитарные эффекты, которые не принадлежат к ЛИС-системам
  - Дисторшн ограничение сигнала по амплитуде

$$y(n) = \operatorname{trunc}(ax(n))/a$$

• Тремоло — применение к сигналу амплитудной модуляции с использованием медленно изменяющейся во времени гармонической компоненты, приводящей к периодическому изменению громкости звука

$$y(n) = (1 + \cos(\widehat{\omega}_0 n)/G) \cdot x(n)$$

• Фланжер – применение к сигналу процедуры синусоидальной задержки во времени

$$y(n) = x(n) + x(n - \lfloor d \cdot (1 + \cos(\widehat{\omega}_0 n)) \rfloor)$$

### Благодарности

- В лекции использовались материалы следующих книг:
  - Солонина А.И. и др. Основы цифровой обработки сигналов, 2005
  - Prandoni P., Vetterli M. Signal processing for communications, 2008
  - Hayes M.H. Schaum's outlines of digital signal processing, 2011
  - Николенко С., Кадурин А., Архангельская Е. Глубокое обучение, 2018
- В лекции использовались материалы следующих онлайн-курсов:
  - Oppenheim A.V. Digital signal processing (RES.6-008), 1975. Massachusetts Institute of Technology
  - Peters R.A. Image processing (EECE/CS 253), 2007. Vanderbilt University School of Engineering
  - Prandoni P., Vetterli M. Digital signal processing, 2014. Coursera's online course of École polytechnique fédérale de Lausanne
  - Ng A. Convolutional neural networks, 2019. Coursera's online course of Stanford University
  - Ng A. Sequence models, 2019. Coursera's online course of Stanford University
- Отдельное спасибо коллегам по работе, без которых бы эта лекция не состоялась!



### Группа компаний ЦРТ

#### O HAC

В группу компаний ЦРТ входят компании ЦРТ, ЦРТ-инновации и SpeechPro.

**ЦРТ** – российская компания, разработчик инновационных систем в сфере технологий синтеза и распознавания речи, анализа аудио- и видеоинформации, распознавания лиц, голосовой и мультимодальной биометрии.

**ЦРТ-инновации** — научно-исследовательская компания, передовой разработчик голосовых и бимодальных биометрических систем. Резидент Фонда «Сколково». Области научно-исследовательской деятельности компании: биометрия по голосу и лицу, распознавание речи, анализ больших данных.

SpeechPro – представительство Группы ЦРТ в Северной Америке с главным офисом в Нью-Йорке. SpeechPro взаимодействует с клиентами и партнерами ЦРТ из США и Канады.

#### КОНТАКТНАЯ ИНФОРМАЦИЯ

### Санкт-Петербург

Адрес: 194044, г. Санкт-Петербург, ул. Гельсингфорсская, 3-11, лит. Д

Телефон: (+7 812) 325-88-48 Факс: (+7 812) 327-92-97

Эл. почта: stc-spb@speechpro.com

#### Москва

Адрес: Москва, ул. Марксистская, д.3, стр.2,

Бизнес-центр «Таганский» Телефон: +7 (495) 669-74-40 Факс: +7 (495) 669-74-44

Эл. почта: stc-msk@speechpro.com