

Компьютерная графика

Лекция 14: рендеринг текста, bitmap-шрифты, векторные шрифты, SDF-шрифты, volume rendering, volume slicing, raymarching

2021

Рендеринг текста

- ▶ Абстрактный текст

Рендеринг текста

- ▶ Абстрактный текст
- ▶ + кодировка \Rightarrow машинное представление текста

Рендеринг текста

- ▶ Абстрактный текст
- ▶ + кодировка \Rightarrow машинное представление текста
- ▶ + шрифт + настройки шейпинга (shaping) \Rightarrow набор глифов (изображений символов) и их координат

Рендеринг текста

- ▶ Абстрактный текст
- ▶ + кодировка \Rightarrow машинное представление текста
- ▶ + шрифт + настройки шейпинга (shaping) \Rightarrow набор глифов (изображений символов) и их координат
- ▶ \Rightarrow нарисованный текст

Кодировки

- ▶ Описывают машинное представление текста, т.е. соответствие последовательностей символов и последовательностей бит

Кодировки

- ▶ Описывают машинное представление текста, т.е. соответствие последовательностей символов и последовательностей бит
- ▶ ASCII: 7 бит (обычно дополняется нулевым старшим битом до 8 бит), первые 32 символа - управляющие (`\r`, `\n`, `tab`, ...), остальные 96 - буквы английского алфавита (большие и маленькие) и прочие символы (различные скобки, арифметические операции, пунктуация, пробел, ...)

Кодировки

- ▶ Описывают машинное представление текста, т.е. соответствие последовательностей символов и последовательностей бит
- ▶ ASCII: 7 бит (обычно дополняется нулевым старшим битом до 8 бит), первые 32 символа - управляющие (`\r`, `\n`, `tab`, ...), остальные 96 - буквы английского алфавита (большие и маленькие) и прочие символы (различные скобки, арифметические операции, пунктуация, пробел, ...)
 - ▶ Многие кодировки совпадают с ASCII в диапазоне 0-127 или 32-127

Кодировки

- ▶ Описывают машинное представление текста, т.е. соответствие последовательностей символов и последовательностей бит
- ▶ ASCII: 7 бит (обычно дополняется нулевым старшим битом до 8 бит), первые 32 символа - управляющие (`\r`, `\n`, `tab`, ...), остальные 96 - буквы английского алфавита (большие и маленькие) и прочие символы (различные скобки, арифметические операции, пунктуация, пробел, ...)
 - ▶ Многие кодировки совпадают с ASCII в диапазоне 0-127 или 32-127
- ▶ Огромное количество в основном 8-битных кодировок для разных алфавитов и систем:
 - ▶ ISO/IEC 8859 - 15 разных вариантов (ISO/IEC 8859-5 для русского языка)
 - ▶ Code page XXX - много разных кодировок для DOS (Code page 866 для русского языка)
 - ▶ Windows code pages (Windows-1251 для русского языка)
 - ▶ KOI-8 и вариации - для русского языка
 - ▶ etc.

Кодировки

- ▶ Описывают машинное представление текста, т.е. соответствие последовательностей символов и последовательностей бит
- ▶ ASCII: 7 бит (обычно дополняется нулевым старшим битом до 8 бит), первые 32 символа - управляющие (\r, \n, tab, ...), остальные 96 - буквы английского алфавита (большие и маленькие) и прочие символы (различные скобки, арифметические операции, пунктуация, пробел, ...)
 - ▶ Многие кодировки совпадают с ASCII в диапазоне 0-127 или 32-127
- ▶ Огромное количество в основном 8-битных кодировок для разных алфавитов и систем:
 - ▶ ISO/IEC 8859 - 15 разных вариантов (ISO/IEC 8859-5 для русского языка)
 - ▶ Code page XXX - много разных кодировок для DOS (Code page 866 для русского языка)
 - ▶ Windows code pages (Windows-1251 для русского языка)
 - ▶ KOI-8 и вариации - для русского языка
 - ▶ etc.
- ▶ Unicode-кодировки

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)
 - ▶ UCS-2: устаревшая, 2 байта на символ, не поддерживает весь unicode

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)
 - ▶ UCS-2: устаревшая, 2 байта на символ, не поддерживает весь unicode
 - ▶ UTF-16: 2 или 4 байта на символ

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)
 - ▶ UCS-2: устаревшая, 2 байта на символ, не поддерживает весь unicode
 - ▶ UTF-16: 2 или 4 байта на символ
 - ▶ UTF-32: 4 байта на символ

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)
 - ▶ UCS-2: устаревшая, 2 байта на символ, не поддерживает весь unicode
 - ▶ UTF-16: 2 или 4 байта на символ
 - ▶ UTF-32: 4 байта на символ
 - ▶ GB 18030: специальная кодировка для китайских иероглифов (но тоже поддерживает весь unicode)

Unicode

- ▶ Unicode - стандарт, описывающий соответствие символов целочисленным кодам в диапазоне 0..10FFFFh исключая D800h..DFFFh (используется для суррогатных пар в UTF-16; всего 1112064 символов), и рекомендации по их интерпретации и визуализации
- ▶ На сегодняшний день описывает 144697 символа
- ▶ Unicode-кодировки:
 - ▶ UTF-8: от 1 до 4 байт на символ, совпадает с ASCII в диапазоне 0..7Fh, самая распространённая сегодня кодировка (95% интернета)
 - ▶ UCS-2: устаревшая, 2 байта на символ, не поддерживает весь unicode
 - ▶ UTF-16: 2 или 4 байта на символ
 - ▶ UTF-32: 4 байта на символ
 - ▶ GB 18030: специальная кодировка для китайских иероглифов (но тоже поддерживает весь unicode)
- ▶ N.B.: один символ unicode **не соответствует** одному видимому символу (*графеме*)

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Bitmap-шрифты: глиф - готовое изображение (bitmap)

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Битмар-шрифты: глиф - готовое изображение (bitmap)
 - ▶ Векторные шрифты: глиф описывается как геометрическая фигура

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Битмар-шрифты: глиф - готовое изображение (bitmap)
 - ▶ Векторные шрифты: глиф описывается как геометрическая фигура
 - ▶ SDF-шрифты: глиф описывается с помощью *signed distance field* (SDF)

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Bitmap-шрифты: глиф - готовое изображение (bitmap)
 - ▶ Векторные шрифты: глиф описывается как геометрическая фигура
 - ▶ SDF-шрифты: глиф описывается с помощью *signed distance field* (SDF)
- ▶ Современные форматы шрифтов (.ttf - TrueType, .otf - OpenType) - векторные, описывают границу глифа как набор отрезков и квадратичных кривых Безье (т.е. 2-ого порядка)

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Bitmar-шрифты: глиф - готовое изображение (bitmap)
 - ▶ Векторные шрифты: глиф описывается как геометрическая фигура
 - ▶ SDF-шрифты: глиф описывается с помощью *signed distance field* (SDF)
- ▶ Современные форматы шрифтов (.ttf - TrueType, .otf - OpenType) - векторные, описывают границу глифа как набор отрезков и квадратичных кривых Безье (т.е. 2-ого порядка)
- ▶ Bitmar и SDF шрифты часто строятся по векторным шрифтам

Шрифты

- ▶ Содержит набор *глифов* (изображений символов в каком-либо виде) и правил их использования
- ▶ Виды шрифтов:
 - ▶ Битмар-шрифты: глиф - готовое изображение (bitmap)
 - ▶ Векторные шрифты: глиф описывается как геометрическая фигура
 - ▶ SDF-шрифты: глиф описывается с помощью *signed distance field* (SDF)
- ▶ Современные форматы шрифтов (.ttf - TrueType, .otf - OpenType) - векторные, описывают границу глифа как набор отрезков и квадратичных кривых Безье (т.е. 2-ого порядка)
- ▶ Битмар и SDF шрифты часто строятся по векторным шрифтам
- ▶ Freetype - самая распространённая библиотека для чтения векторных шрифтов; умеет растеризовать в битмар и (с версии 2.11.0, июль 2021) в SDF

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:
 - ▶ Настройки шейпинга: направление (слева-направо, справа-налево, сверху-вниз, снизу-вверх), размер шрифта, межбуквенное расстояние, стиль (жирный, курсив, и т.п.)

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:
 - ▶ Настройки шейпинга: направление (слева-направо, справа-налево, сверху-вниз, снизу-вверх), размер шрифта, межбуквенное расстояние, стиль (жирный, курсив, и т.п.)
 - ▶ Hinting: применение дополнительных преобразований к векторному глифу в зависимости от разрешения

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:
 - ▶ Настройки шейпинга: направление (слева-направо, справа-налево, сверху-вниз, снизу-вверх), размер шрифта, межбуквенное расстояние, стиль (жирный, курсив, и т.п.)
 - ▶ Hinting: применение дополнительных преобразований к векторному глифу в зависимости от разрешения
 - ▶ Kerning: изменение расстояния между соседними глифами

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:
 - ▶ Настройки шейпинга: направление (слева-направо, справа-налево, сверху-вниз, снизу-вверх), размер шрифта, межбуквенное расстояние, стиль (жирный, курсив, и т.п.)
 - ▶ Hinting: применение дополнительных преобразований к векторному глифу в зависимости от разрешения
 - ▶ Kerning: изменение расстояния между соседними глифами
 - ▶ Лигатуры: последовательность несвязанных символов, представленная одним глифом (ff, fi, <=>)

Шейпинг

- ▶ Процесс преобразования последовательности символов в набор отпозиционированных глифов
- ▶ Может включать в себя:
 - ▶ Настройки шейпинга: направление (слева-направо, справа-налево, сверху-вниз, снизу-вверх), размер шрифта, межбуквенное расстояние, стиль (жирный, курсив, и т.п.)
 - ▶ Hinting: применение дополнительных преобразований к векторному глифу в зависимости от разрешения
 - ▶ Kerning: изменение расстояния между соседними глифами
 - ▶ Лигатуры: последовательность несвязанных символов, представленная одним глифом (ff, fi, <=>)
- ▶ harfbuzz - одна из самых распространённых библиотек для шейпинга текста
- ▶ Freetype позволяет сделать шейпинг, но хуже, чем harfbuzz

abcfgop AO *abcfgop*
abcfgop AO *abcfgop*

維基百科
維基百科國際
維基百科
維基百科國際

abcfgop

abcfgop

Kerning

AV Wa
No kerning

AV Wa
Kerning applied

Лигатуры

$AE \rightarrow \text{Æ}$	$ij \rightarrow \text{ij}$
$ae \rightarrow \text{æ}$	$st \rightarrow \text{ſt}$
$OE \rightarrow \text{Œ}$	$ft \rightarrow \text{ft}$
$oe \rightarrow \text{œ}$	$et \rightarrow \text{\&}$
$ff \rightarrow \text{ff}$	$fs \rightarrow \text{\beta}$
$fi \rightarrow \text{fi}$	$ffi \rightarrow \text{ffi}$