

# **Тема: Разработка программного средства генерации видеоклипа на основе аудиозаписи (Шифр ПС ГВОА)**

**Руководитель:** к.т.н., Капитанов Андрей Иванович

**Исполнитель:** ст. гр. ПИН-42 Пилюгина Олеся Игоревна

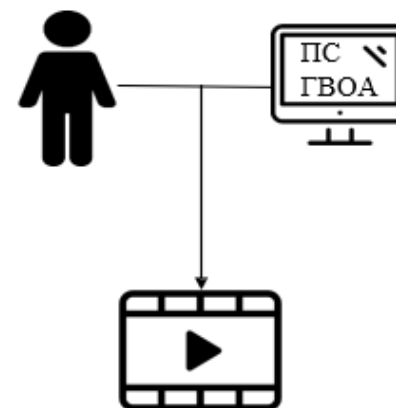
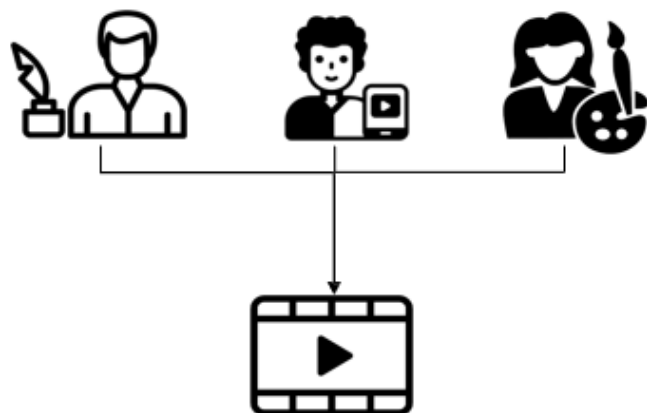
**Цель:** сокращение временных затрат при создании видеопродукции за счет алгоритма анализа аудиозаписи.

## **Задачи:**

- исследование предметной области;
- сравнительный анализ существующих программных решений;
- выбор языка и среды разработки;
- разработка алгоритма ПС ГВОА;
- разработка пользовательского интерфейса;
- программная реализация ПС ГВОА;
- тестирование ПС ГВОА;
- разработка руководства оператора.

## Исследование предметной области

До разработки ПС ГВОА	После разработки ПС ГВОА
Участие нескольких высококвалифицированных специалистов	Участие одного человека
Использование большого количества различных ПО	Единое программное средство
Большие временные и финансовые затраты. Создание видеоролика занимает минимум неделю	Сокращение временных и финансовых затрат. Создание видеоролика занимает в среднем 1,5 часа



→ Условные обозначения:  
Стрелки означают процесс создания

## Обзор аналогичных программных решений

Параметры	Программное средство					ГБОУ
	Deform Stable Diffusion <sup>1</sup>	Gen-2 <sup>2</sup>	txt2video <sup>3</sup>	Animate Diff <sup>4</sup>	WZRD <sup>5</sup>	
Открытый исходный код	+	—	+	+	—	+
Контроль графики	+	—	—	—	—	+
Обработка аудио	—	—	—	—	+	+
Обработка текста	+	+	+	+	—	+
Качество HD и выше	+	+	—	+	+	+
Русская локализация	—	—	—	—	—	+

Источники информации:

<sup>1</sup> <https://github.com/deform-art/deform-stable-diffusion>

<sup>2</sup> <https://runwayml.com/ai-magic-tools/gen-2/>

<sup>3</sup> <https://github.com/deform-art/sd-webui-text2video>

<sup>4</sup> <https://github.com/continue-revolution/sd-webui-animatediff>

<sup>5</sup> <https://app.wzrd.ai/intro>

Условные обозначения:

+ – наличие функционала

— – отсутствие функционала

## Выбор языка программирования

Критерии выбора	Язык программирования			
	Java	Python	C++	R
Взаимодействие со Stable Diffusion	—	+	—	—
Опыт использования, лет	0,5	2	3	0
Применяется на предприятии	—	+	+	—
Интерпретируемый язык	—	+	—	+
Наличие необходимых библиотек	—	+	+	—

Для разработки ПС ГВОА выбран язык программирования Python

Условные обозначения:

+ – наличие функционала

– – отсутствие функционала

## Выбор среды разработки

Критерии выбора	IDE			
	Visual Studio Code <sup>1</sup>	Sublime Text <sup>2</sup>	PyCharm <sup>3</sup>	Spyder <sup>4</sup>
Интеграция с системами контроля версий	+	—	+	—
Встроенная отладка	+	—	+	+
Бесплатный доступ	+	—	—	+
Опыт использования, лет	3	0	2	0
Минимальные требования к ОЗУ ≤ 1 Гб	+	+	—	+
	ОЗУ = 1 Гб	ОЗУ = 512 Мб	ОЗУ = 8 Гб	ОЗУ = 1 Гб

Для разработки ПС ГВОА выбрана среда разработки Visual Studio Code.

Источники информации:

<sup>1</sup> <https://code.visualstudio.com/docs>

<sup>2</sup> <https://www.sublimetext.com/>

<sup>3</sup> <https://www.jetbrains.com/help/pycharm/getting-started.html>

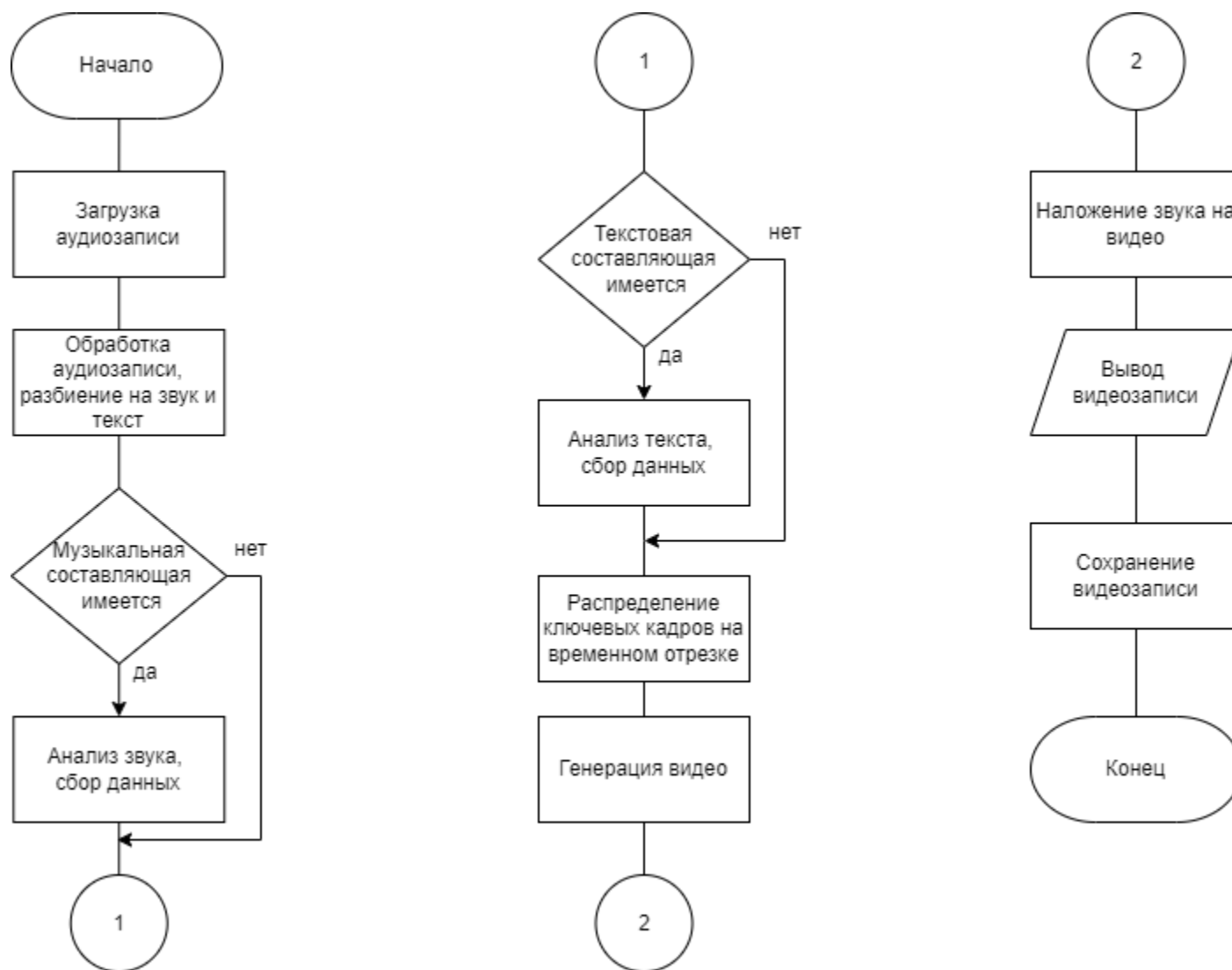
<sup>4</sup> <https://docs.spyder-ide.org>

Условные обозначения:

+ – наличие функционала

— – отсутствие функционала

## Схема алгоритма



# Экранная форма пользовательского интерфейса

## ВИЗУАЛИЗИРУЙТЕ ВИДЕОКЛИП В ВАШЕМ СТИЛЕ

🎵 Загрузите аудиофайл

Поместите Аудио Здесь  
- или -  
Нажмите, чтобы загрузить

🖼️ Результат



Выберите стиль видео

☒ Аниме ☐ Живопись ☐ Черно-белый рисунок ☐ Реализм

Выберите графику

☒ 3D ☐ 2D

☐ НЕ учитывать смысл текста при генерации

Выберите разрешение

☐ 1280 x 720 ☐ 854 x 480 ☒ 512 x 512 ☐ 640 x 360 ☐ 426 x 240

Прервать

Запуск

# Тестирование ПС ГВОА

Номер теста	Стиль	Графика	Учитывать текст	Разрешение	Песня	Тест пройден
1	Аниме	2D	+	512 x 512	веселая	+
2	Живопись	2D	+	1280 x 720	веселая	+
3	Черно-белый рисунок	2D	+	426 x 240	грустная	+
4	Реализм	3D	+	640 x 360	агрессивная	+
5	Аниме	3D	+	854 x 480	грустная	+
6	—	—	—	512 x 512	агрессивная	+
7	—	—	—	854 x 480	веселая	+
8	—	—	—	1280 x 720	грустная	+



## Тестирование ПС ГВОА

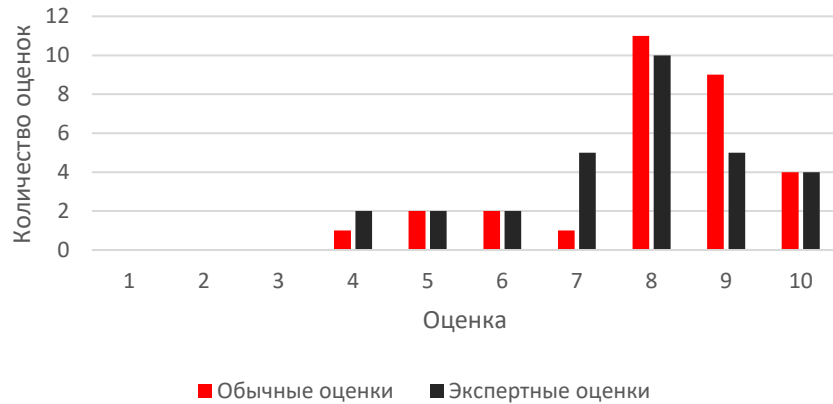
Общее количество исполняемых операторов (строк кода)	Количество операторов, которые не были выполнены во время тестирования (пропущены)	Процент покрытия кода
491	33	93%

В процессе разработки ПС ГВОА было написано 16 тестов. Тестирование и отладка ПС ГВОА производились в среде разработки VS Code при помощи ее встроенных средств и библиотеки pytest.

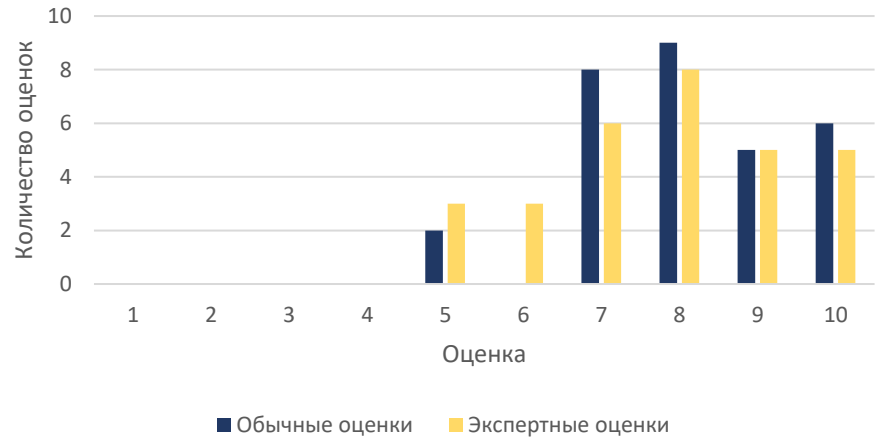
Покрытие кода измерялось при помощи инструмента Coverage.

# Экспертные оценки

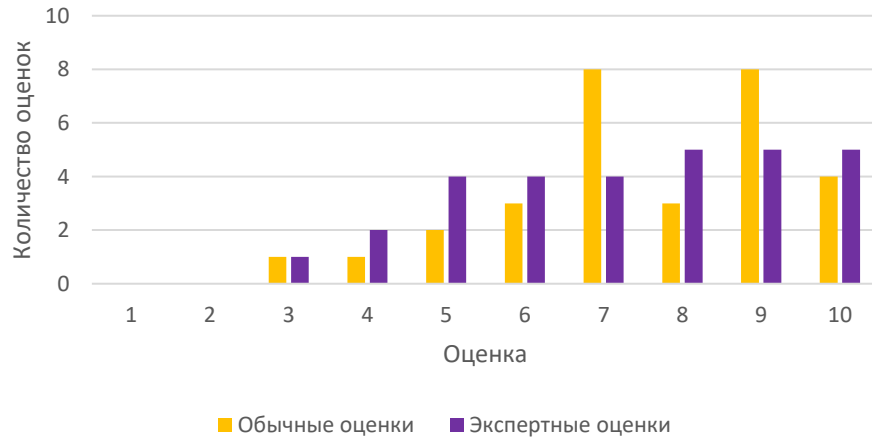
## Диаграмма оценок соответствия фрагментов видео тексту песни



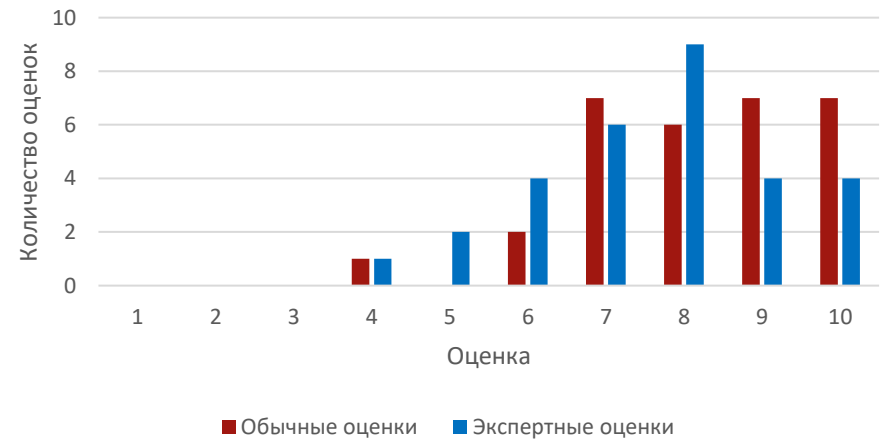
## Диаграмма оценки красоты видео



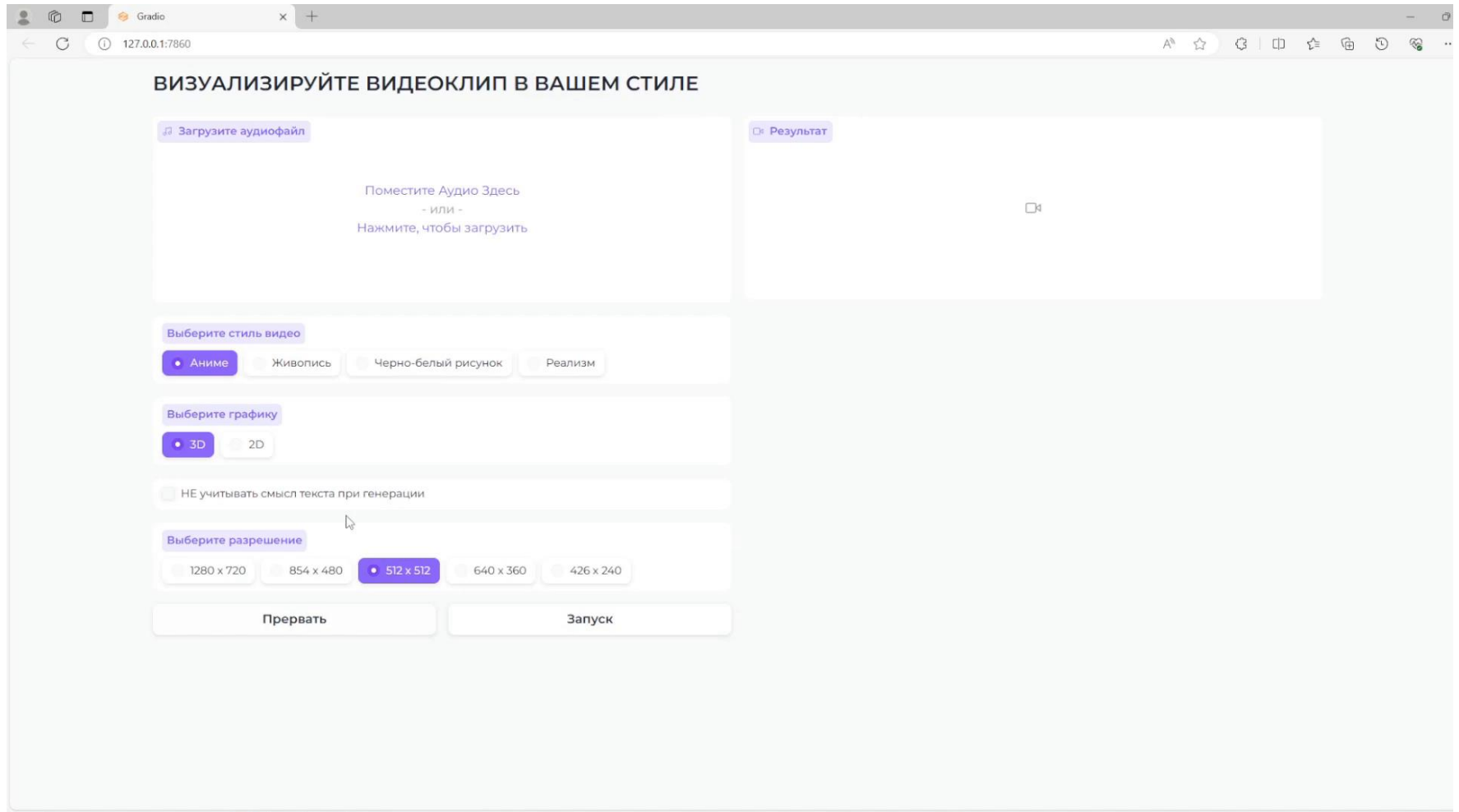
## Диаграмма оценки оригинальности видео



## Диаграмма оценки видео в целом



# Демонстрация работы программного средства



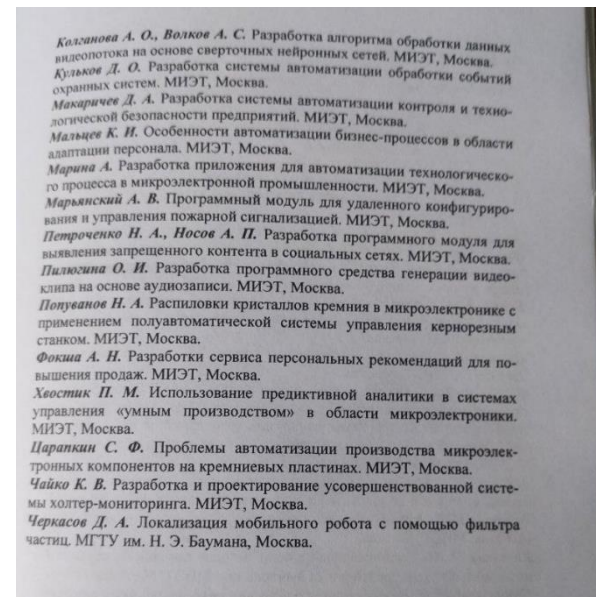
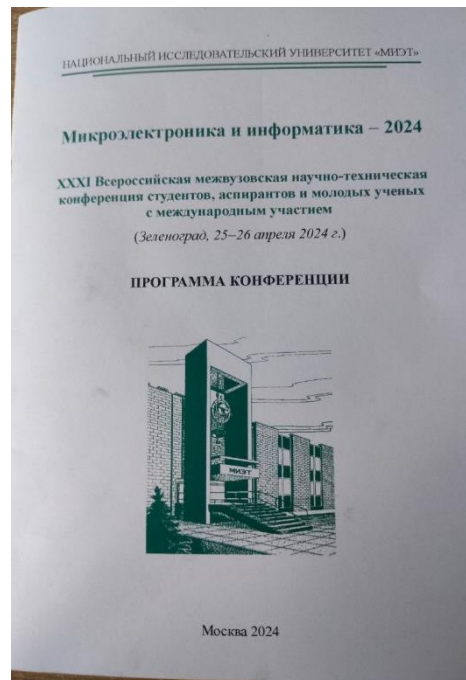
## Демонстрация работы программного средства



Beyond the emerald sea, above the serever skies  
За изумрудным морем, над серебристым небом

# Апробация

Результаты разработки апробированы в рамках научно-практической конференции: Пилюгина О. И. Разработка программного средства генерации видеоклипа на основе аудиозаписи. МИЭТ, Москва. XXXI Всероссийская межвузовская научно-техническая конференция студентов, аспирантов и молодых учёных «Микроэлектроника и информатика-2024». Стр. 27.



## Результаты работы

Видео	Ручной подход				ПС ГВОА
	Эксперт №1	Эксперт №2	Эксперт №3	В среднем	
<b>Живопись</b> <b>1280 x 720 3D</b> Длина – 1,50 ч	17 ч	10 ч	13 ч	~ 13 ч	1 ч 53 мин
<b>Реализм</b> <b>640 x 360 3D</b> Длина – 2,33 ч	19 ч	14 ч	16 ч	~ 16 ч	1 ч 48 мин
<b>Аниме</b> <b>512 x 512 2D</b> Длина – 2,18 ч	15 ч	11 ч	10 ч	~ 12 ч	1 ч 41 мин
<b>Черно-белый рисунок</b> <b>426 x 240 2D</b> Длина – 2,18 ч	12 ч	8 ч	6 ч	~ 9 ч	0 ч 35 мин

## Результаты работы

- исследована предметная область;
- проведен сравнительный анализ существующих программных решений;
- проведен выбор языка и среды разработки;
- разработан алгоритм ПС ГВОА;
- разработан пользовательский интерфейс;
- проведена программная реализация ПС ГВОА;
- проведено тестирование ПС ГВОА;
- разработано руководство оператора.

**Спасибо за внимание!**