

**МИНОБРНАУКИ РОССИИ**  
**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ**  
**ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**  
**«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**  
**Кафедра математического обеспечения и применения ЭВМ**

**ОТЧЕТ**

**по научно-исследовательской работе в осеннем семестре 2022 года**

**Тема: Разработка и анализ подходов по усилению микродвижений на видео**

Студент гр. 7304

Глазунов С.А.

Руководитель

Филатов А.Ю..

Санкт-Петербург

2022

## **СОДЕРЖАНИЕ**

	Введение	3
1.	ПОСТАНОВКА ЗАДАЧИ	4
2.	РЕЗУЛЬТАТЫ РАБОТЫ В ОСЕННЕМ СЕМЕСТРЕ	5
3.	Критерии	10
4.	ОПИСАНИЕ ПРЕДПОЛАГАЕМОГО МЕТОДА РЕШЕНИЯ	11
5.	ПЛАН РАБОТЫ НА ВЕСЕННИЙ СЕМЕСТР	12
	Список использованных источников	13

## **ВВЕДЕНИЕ**

Видеотехника “усиления” движения позволяет нам увидеть небольшие движения, ранее невидимые невооруженным глазом, например, движения вибрирующих крыльев самолета или раскачивание зданий под воздействием ветра, или как кровь течет по человеку. Поскольку движения малы, результаты увеличения подвержены шуму или чрезмерному размытию, или вовсе сложно уловимыми.

Следовательно, есть необходимость в разработке алгоритма, который бы смог:

1. “Усиливать” микро движения различных объектов на видео;
2. С минимальной ошибкой усиливать звук.

## ПОСТАНОВКА ЗАДАЧИ

Целью данного исследования является создание программы для усиления микро движений на видеопотоке.

Предполагаемый функционал:

- Усиления микродвижений;
- Возможность настройки коэффициента усиления;

Для достижения цели были поставлены следующие задачи:

1. Провести обзор аналогов;
2. Если алгоритм основывается на нейронных сетях, то провести отдельный этап сбор датасета для обучения;
3. Разработать алгоритм, который бы позволил усиливать микродвижения на видеопотоке;
4. Реализовать и проанализировать разработанное решение.

Объектом исследования является усиление микродвижений на видеопотоке.

Предметом исследования является пространственных и временных признаков на видеопоследовательности.

Практическая ценность работы: разработанное решение позволит оценивать состояние различных объектов, например, строительного, а также определять пульс человека по видео.

# **РЕЗУЛЬТАТЫ РАБОТЫ В ОСЕННЕМ СЕМЕСТРЕ**

Тема диплома была изменена на 3 семестре магистратуры, так что результаты работы за предыдущий семестр не будут приведены.

## **План на осенний семестр.**

Работа над устранением сильной зависимости качества распознавания от качества звука (микрофона), необходимость дополнительной программная обработка на клиентской и/или серверной стороне;

- Работы над устранением зависимости качества распознавания от быстроты и качества речи;
- Распознавание английских терминов в русской речи;
- Отображение плохо распознаваемых слов в клиентской части приложения.

## **Результаты осеннего семестра.**

Был произведен обзор аналогов, которые имеют отношение к усилению микродвижений на видеопоследовательности.

### **1. Eulerian Video Magnification for Revealing Subtle Changes in the World**

В данной работе использовался метод Эйлеровское усиление видео, которое принимало в качестве входных данных видеопоследовательность, к которой применялось:

1. Пространственное разложение;
2. Временная фильтрация по кадрам;
3. Усиление сигнала на некоторый коэффициент, который является гиперпараметром;
4. Преобразование обработанного сигнала в видеопоследовательность.

Графически данный алгоритм можно рассмотреть на рис. 1.

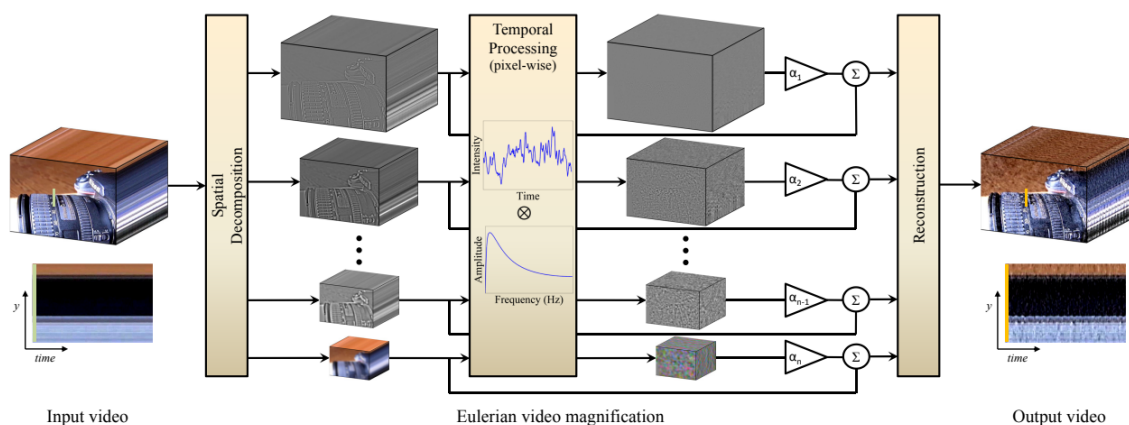


Рис. 1 – алгоритм Эйлеровского усиления видео

Система сначала разбивает входную видеопоследовательность на различные пространственные полосы частот и применяет один и тот же временной фильтр ко всем полосам. Затем отфильтрованные пространственные полосы усиливаются на заданный коэффициент  $\alpha$ , добавляются обратно к исходному сигналу и сворачиваются для получения выходного видео.

Данный подход используется как для “усиления цвета”, так и для усиления микродвижений на видео.

Результаты, полученные автором, по производительности следующие – 45 fps на CPU на изображении 640 на 480. Какое именно оборудование использовалось не было уточнено.

Данный подход стал опорным для многих последующих работ, которые вошли в список аналогов

## 2. Learning-based Video Motion Magnification

В данной работе автор стремится обучить фильтр на примерах, используя глубокие конволюционные нейронные сети. В качестве входных данных используются синтетические данные, и используется двухкадровый вход для обучения. Данная работа нацелена на нелинейное решение – обучение нейронной сети. В нейронной сети автор заметил, что временные фильтры могут быть использованы с полученными данными до умеренного увеличения, позволяя осуществлять частотный выбор движения на основе частоты.

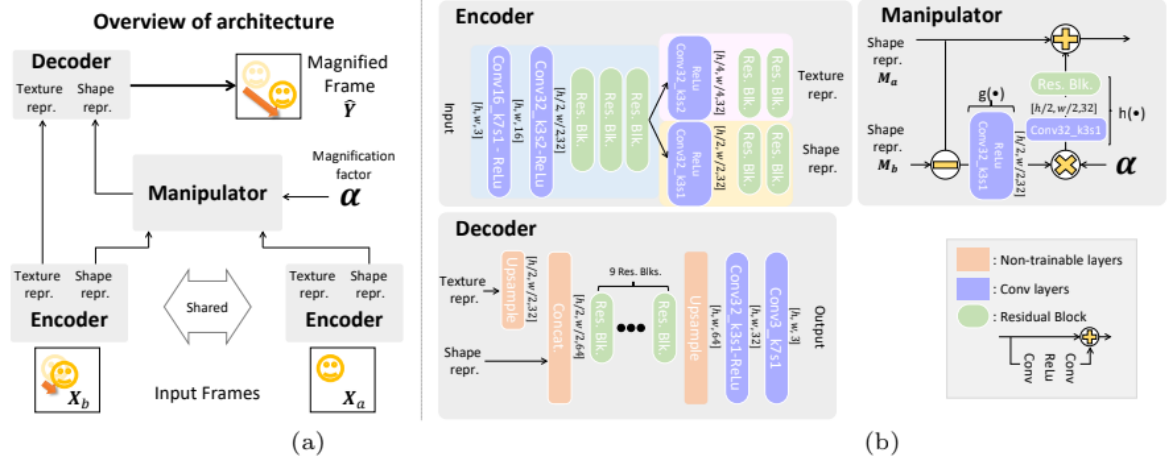


Рис.2 – Архитектура нейронной сети.

Нейронная сеть состоит из 3 основных частей:

1. Декодера;
2. Манипулятора;
3. Энкодера

Как и в предыдущем аналоге в данном решении использования коэффициент усиления (Манипулятор). Энкодер и декодер полностью сверточные. Целью данной работы является уменьшение шумов и артефактов, которые возникают при других методах усиления микродвижений на видео.

В качестве входных данных используются синтетические данные. Предоставлен код для обучения нейронной сети и датасет, на котором тестировалось решение.

### 3. Video Magnification in the Wild Using Fractional Anisotropy in Temporal Distribution

Данная работа нацелена на решение проблемы появления артефактов, шумов при усилении движений на видео. В данной работе для борьбы с лишними “тонкими” движениями, которые не несут смысла, применяется метод, использующий фракционную анизотропию (ФА) для обнаружения только значимых изменений на видео.

Фракционная анизотропия используется для определения областей (см. рис 3 правое нижнее изображение), в которых стоит производить операцию усиления видео. В качестве алгоритма усиления видео используется алгоритм, описанный в 1-ом аналоге.

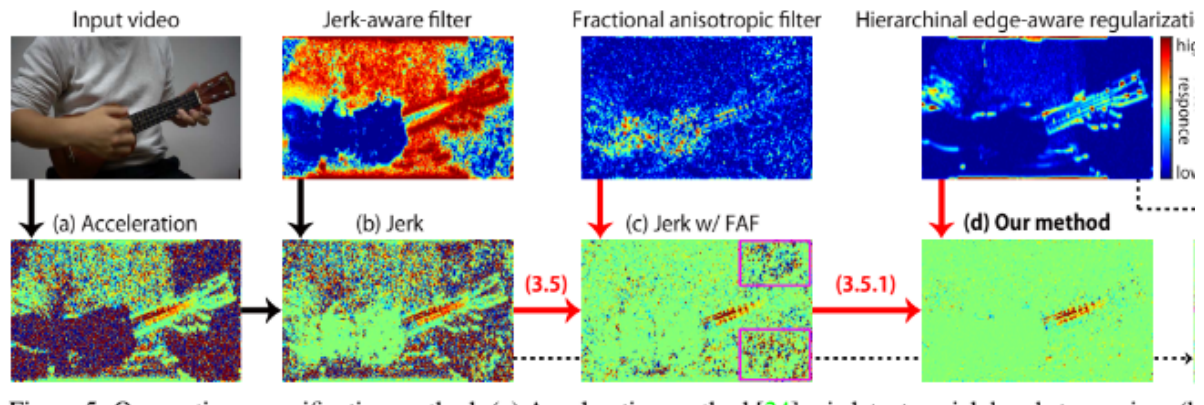


Рис. 3 – примеры усиления с шумом.

Главный недостаток данного решения – нет гарантий, что ФА не будет брать только неважные микродвижения. Пример микродвижений, которые ФА может “заблокировать” – пульс и другие.

#### 4. Motion magnification multi-feature relation network for facial microexpression recognition

Данная работа использовалась для определения мимики лица, используя усиления микродвижений. В основе решения лежит нейронная сеть. В качестве основы была взята нейронная сеть Resnet50. MMFRN(см. рис.4) состоит из 3 частей:

1. Encoder – усиление видео;
2. Manipulator – выявление высокоуровневых признаков, используя Resnet50;
3. Decoder – модуль, который объединяет результат.



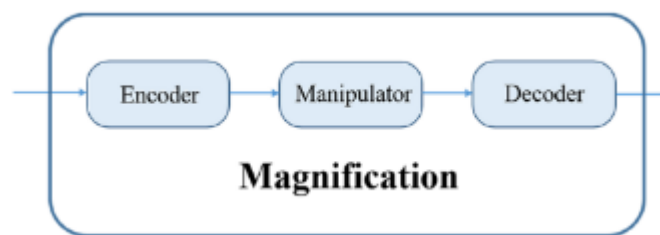


Рис. 4 – общая схема нейронной сети

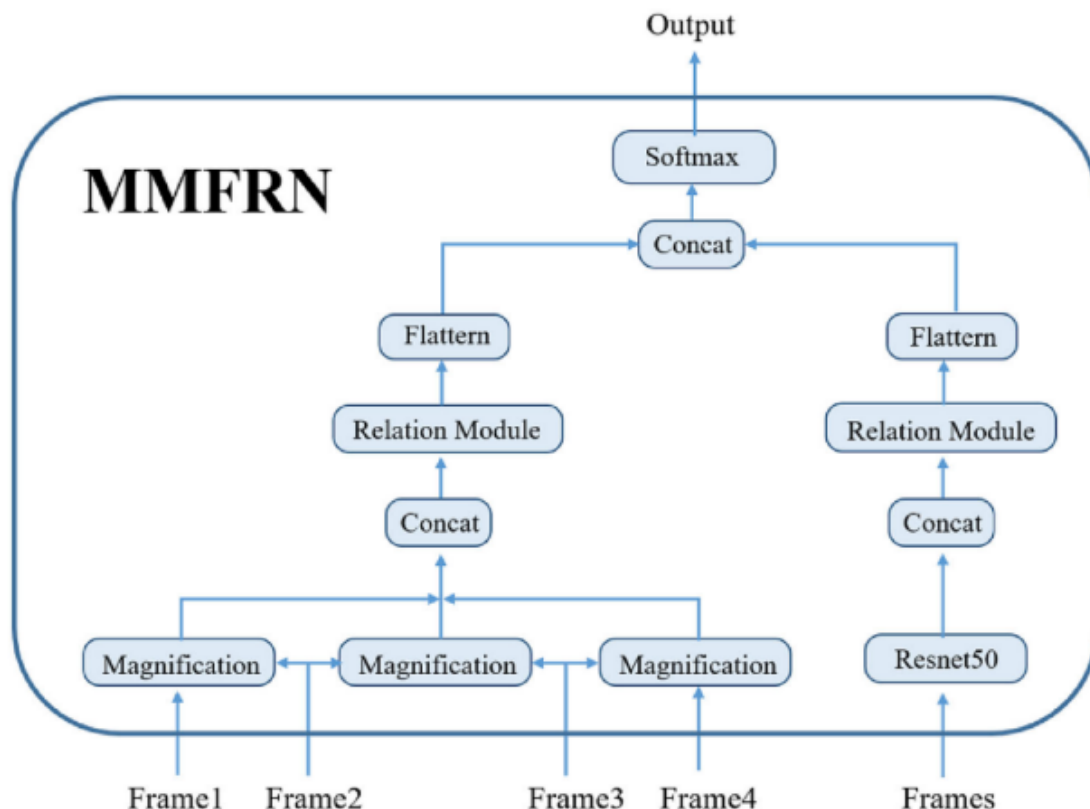


Рис.5 – описание MMFRN

Данное решение предоставляет датасет CASME 2.

## 5. Eulerian video magnification: a review

В данной работе используются подходы, которые уже были описаны в предыдущих аналогах. Полезность данной работы заключается в более детальном повествовании о временной и пространственной фильтрации. Процесс временной фильтрации применяется к серии временных пикселей в каждой пространственной полосе в пространственной пирамиде для выделения интересующих частотных полос. Временной полосовой фильтр используется

для извлечения движений или сигналов, которые должны быть усилены. В зависимости от приложения, используемого в алгоритме, пользователи должны иметь возможность контролировать частоту полосу частот. Однако в некоторых случаях частотная полоса может быть выбрана автоматически.

## Критерии

Для сравнения аналогов создадим критерии:

1. Использование нейронных сетей в решении для усиления видео;
2. Использование эйлеровского подхода для решения;
3. Наличие датасета;
4. Количество кадров в секунду.

Таблица №1 – сравнение аналогов по критериям

#	Аналог №1	Аналог №2	Аналог №3	Аналог №4	Аналог №5	Аналог №6
K1	-	+	-	-	+	-
K2	+	-	+	+	-	+
K3	-	+	-	-	-	-
K4	45	-	-	-	-	-

## ОПИСАНИЕ ПРЕДПОЛАГАЕМОГО МЕТОДА РЕШЕНИЯ

Предполагаемое решение представляет из себя программный модуль, который состоит из:

1. Предобработка видеопотока, используя ФА для уменьшения ошибок, которые появляются из за естественного шума на видео и так далее.
2. Нейронная сеть, которая будет принимать на вход последовательность изображений.

Данный модуль будет обернут в python библиотеку. Для увеличения производительности модель будет конвертирована в формат onnx, после чего в engine файл для запуска через TensorRT.

Использование программного решения для “усиления” микродвижений на видео в качестве отдельного компонента архитектуры позволяет легко изменить/заменить его в случае необходимости, без изменения кода приложения, которое будет использовать.

# ПЛАН РАБОТЫ НА ВЕСЕННИЙ СЕМЕСТР

Направления для дальнейшей разработки:

- Реализовать фильтр, который позволит выделять только значимые движения;
- Обучить нейронную сеть для усиления микродвижений на Pytorch
- Запуск на TensorRT
- Привести в формат библиотеки для использования сторонними приложения;

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Wu H. Y. et al. Eulerian video magnification for revealing subtle changes in the world //ACM transactions on graphics (TOG). – 2012. – Т. 31. – №. 4. – С. 1-8.
2. Oh T. H. et al. Learning-based video motion magnification //Proceedings of the European Conference on Computer Vision (ECCV). – 2018. – С. 633-648.
3. Takeda S. et al. Video magnification in the wild using fractional anisotropy in temporal distribution //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – С. 1614-1622.
4. Wei Y., Gracheva N., Tudor J. A non-invasive subtle pulse rate extraction method based on Eulerian video magnification //UK Workshop on Computational Intelligence. – Springer, Cham, 2019. – С. 461-471.
5. Zhang J. et al. Motion magnification multi-feature relation network for facial microexpression recognition //Complex & Intelligent Systems. – 2022. – С. 1-14.
6. Shahadi H. I. et al. Eulerian video magnification: A review //Indones. J. Electr. Eng. Comput. Sci. – 2020. – Т. 18. – С. 799-811.
7. Zhang Y., Pinte S. L., Van Gemert J. C. Video acceleration magnification //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2017. – С. 529-537.
8. Das R., Negi G., Smeaton A. F. Detecting deepfake videos using Euler video magnification //Electronic Imaging. – 2021. – Т. 2021. – №. 4. – С. 272-1-272-7.
9. Laha S. et al. Analysis of Video Retinal Angiography With Deep Learning and Eulerian Magnification //Frontiers in Computer Science. – 2020. – Т. 2. – С. 24.