Эффективное понижения размерности пространства признаков на основе LASSO-регрессии в задачах   
распознавания эмоций по речи

маг. Краснопрошин Д.В., доц. Вашкевич М.И.

Белорусский государственный университет информатики и радиоэлектроники  
ул. П. Бровки, 6, БГУИР, каф. ЭВС, 220013, Минск, Беларусь,   
e-mail: vashkevich@bsuir.by

В докладе показаны различные техники слияния медицинских изображений в области коэффициентов дискретного вейвлет-преобразования (ДВП), полученных как из известных функций, так и на основе целочисленного полифазного представления банка фильтров. Также предложен вариант слияния, использующий представление НЧ-компонент ДВП в области модуля максимума вейвлет-преобразования. Проведено попарное сравнение методов по метрике взаимной информации с применением статистических тестов. Результаты показали, что использование полифазного банка фильтров, а также представление НЧ-компонент ДВП в области модуля максимума вейвлет-преобразования позволяет получить лучше результаты слияния, чем использование двумерного ДВП с функциями Хаара.

# Введение.

**1. КРАТКОЕ введение в проблематику:**

– Обоснование важности распознавания эмоций в речи для различных областей, включая межличностное взаимодействие и технологии искусственного интеллекта.

– Разъяснение актуальности и использования аудиоданных RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song) в контексте обучения моделей.

– Поставить акцент на важности задачи понижения признакового пространства. Выделить два аспекта: 1) снижение вычислительной сложности алгоритма распознавания и 2) повышение интерпретируемости модели.

**2**. **Обзор литературы:**

– КРАТКИЙ обзор современных исследований по распознаванию эмоций в речи.

– Оценка существующих методов анализа аудиоданных и выявление их преимуществ и недостатков. – Это достаточно сложная задача в рамках доклада на конференцию, лучше поищите информацию по понижению признакового пространства в контексте задачи распознавания эмоций, если найдете что-то то кратко опишите и сошлитесь.

3. Общая архитектура системы распознавания

– краткое описание

4. Процедура отбора признаков (САМОЕ ГЛАВНОЕ, ЧТО МЫ ХОТИМ ОПИСАТЬ – главный акцент!)

– Сказать, что мы решаем многоклассовую задачу;

– Сведенение K-классовой задачи к K задачам бинарной классификации;

– Описание Lasso для ранжирования признаков, отбор сокращенного числа признаков после ранжирования.

– Объединение признаков, полученных для K бинарных классификаторов.

5. Извлечение признаков:

– КРАТКОЕ описание извлечения признаков, включая мел-кепстральные частотные коэффициенты и их производные.

6. Набор данных RAVDESS:

– КРАТКОЕ описание RAVDESS, включая характеристики аудиоданных, эмоциональные состояния и разнообразие речевых сценариев.

– Обоснование выбора данной базы данных для нашего исследования.

- Внедрение статистических моментов (skewness, kurtosis, interquantile range) в матрицу признаков для более полного охвата характеристик аудиосигнала.

6. Результаты экспериментов

– Представление результатов процесса отбора признаков и их влияния на качество классификации эмоций в речи.

– Подробное рассмотрение ключевых изменений в матрице признаков после применения методов отбора.

– Детальное сопоставление размеров исходной и отобранной матриц признаков.

– Обсуждение результатов в свете сохранения высокой точности классификации при снижении размерности.

– Сопоставление использования различных методов классификации (ЛДА, МОВ, деревья?)

:

7. Заключение

– Краткое подведение итогов и ключевых выводов, основанных на анализе результатов исследования.

– Подчеркивание значимости эффективного отбора признаков для оптимизации методов распознавания эмоций в речи.

– Обсуждение потенциальных направлений для будущих исследований, включая расширение набора данных и оптимизацию методов отбора признаков.

На рис. 1 показана общая схема применения многомасштабного преобразования для слияния изображений.

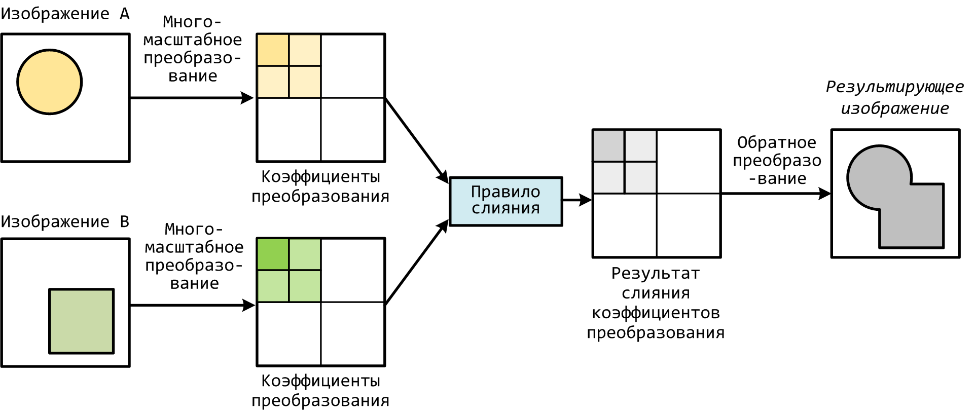


Рисунок 1 – Схема слияния изображения

# Заголовок 1.

Общий принцип построения правил слияния заключается в обеспечении максимальной сохранности всех важных признаков, таких как границы и края. В [5] предложен следующие правила слияния:

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |

где – это соответствующие коэффициенты ДВП изображения .

Согласно (1) НЧ-компоненты, отвечающие за крупные детали объединяются усреднением. Выражение (2) показывает, что слияние ВЧ-коэффициентов ДВП выполняется путем расчета максимума.

# Разреженное представление изображения в пространстве максимумов модуля вейвлет-преобразования.

Результаты экспериментов.

# Заключение.

**Литература**

1. Хлесткин А. Ю. Методы формирования мультимодального изображения //Научные труды SWorld. – 2014. – Т. 4. – №. 4. – С. 14-18.
2. Li S. et al. Pixel-level image fusion: A survey of the state of the art // Information Fusion. – 2017. – vol. 33. – P. 100-112.
3. De Silva D. et al. Wavelet based edge feature enhancement for convolutional neural networks //Eleventh International Conference on Machine Vision (ICMV 2018). – SPIE, 2019. – P. 751-760.
4. Structurally orthogonal finite precision FPGA implementation of block-lifting-based quaternionic paraunitary filter banks for L2L image coding / N.A. Petrovsky, E.V. Rybenkov, A.A. Petrovsky // Digital Signal Processing (DSP’2017): Proc. 22nd Int. Conf., London, UK, Aug. 23-25. – 2017 – 5 P.
5. Qu G., Zhang D., Yan P. Medical image fusion by wavelet transform modulus maxima //Optics Express. – 2001. – Т. 9. – №. 4. – С. 184-190.
6. Дворкович, В.П. Новый подход к использованию вейвлет-фильтров при обработке изображений / В.П. Дворкович, А.В. Гильманшин // Цифровая обработка сигналов – 2008 – № 1 – С. 37–42.
7. Парфенюк М., Петровский А.А. Параунитарные банки фильтров на основе алгебры кватернионов: теория и применение // Цифровая обработка сигналов – 2008 – № 1 – C. 22-36.
8. Qu G., Zhang D., Yan P. Information measure for performance of image fusion //Electronics letters. – 2002. – vol. 38. – №. 7. – P. 313–315.
9. The whole brain atlas [Электронный ресурс]. – 2023. – Режим доступа: <http://www.med.harvard.edu/aanlib/home.html>

FUSION OF MEDICAL IMAGES BASED ON THE   
DISCRETE WAVELET TRANSFORM

Petrovsky N.A., Kuis I.S. Vashkevich M.I.

Belarusian State University of Informatics and Radioelectronics

6, P. Brovki str., Computer Engineering Department, 220113, Minsk, Belarus,  
e-mail: {nick.petrovsky, vashkevich[}@bsuir.by](mailto:%7d@bsuir.by)

The report shows various techniques for fusion of medical images in the domain of discrete wavelet transform (DWT) coefficients obtained both from Haar functions and based on an integer polyphase representation of the filter bank. A fusion method is also proposed that uses the representation of the low-frequency components of the DWT in the domain of the wavelet modulus maximum. A pairwise comparison of methods based on the metric of mutual information with the use of statistical tests was carried out. The results showed that the use of a polyphase filter bank decomposition, as well as the representation of the low-frequency DWT components in the wavelet modulus maximum domain, allows to obtain better fusion results than using a two-dimensional DWT with Haar functions.