

Метод создания связанных сцен из художественного видео

Студент: Домнин Егор Олегович, ИУ7-43М

Руководитель: Рудаков Игорь Владимирович

Цель и задачи работы

Цель работы – разработка метода создания связанных сцен из художественного видео.

Задачи работы:

- анализ предметной области, выделение основных подходов к созданию связанных сцен;
- создание алгоритма для решения поставленной задачи;
- описание специфических особенностей предложенных методов;
- исследование наиболее важных аспектов разработанного программного продукта.

Предметная область



Постановка задачи



Разбиение видео на планы

- Монтажные склейки бывают двух видов
- Склейка встык
- Плавный переход



Границы планов можно





Первый план первая сцена



Первый план первая сцена

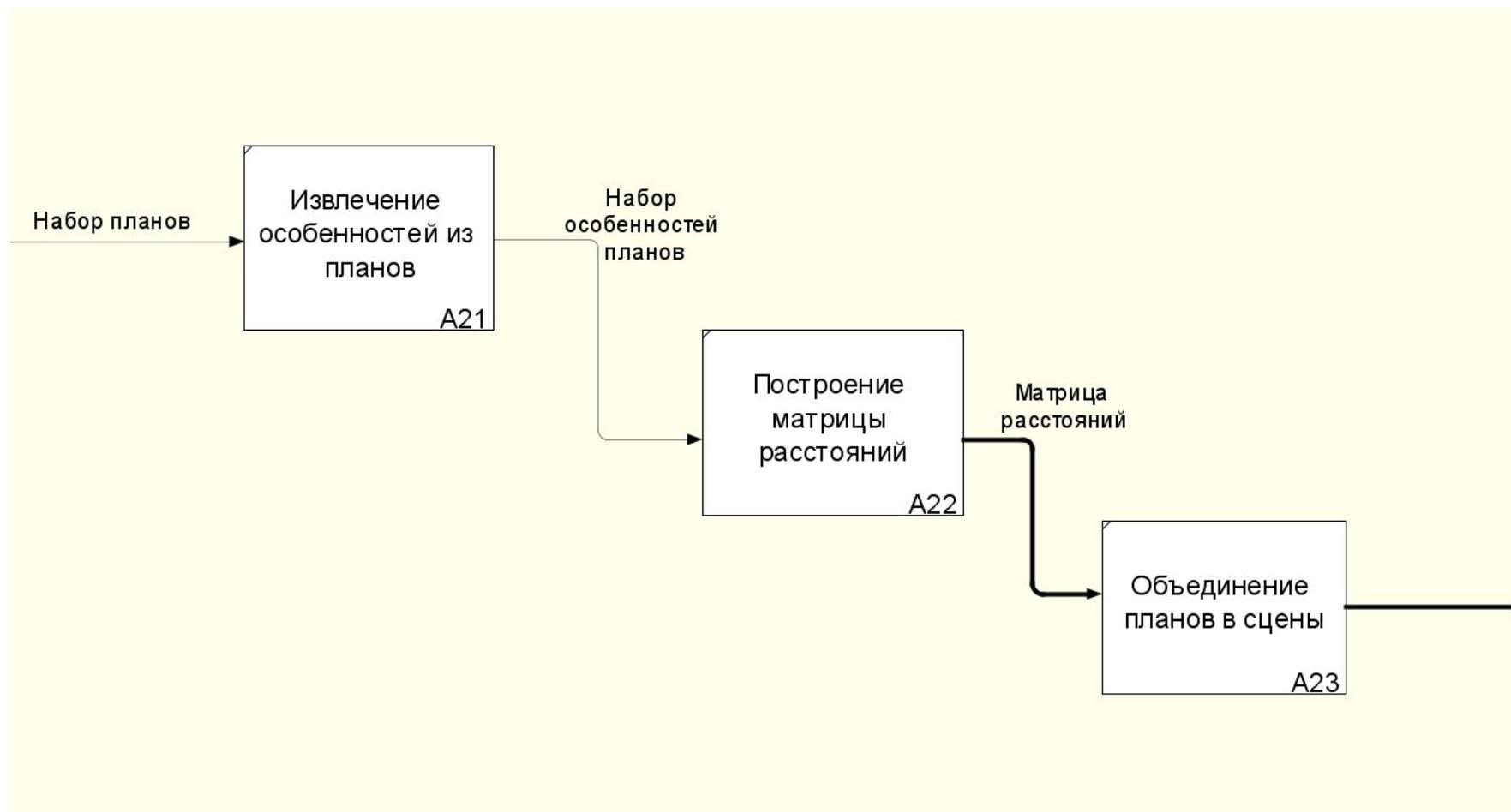


Второй план первая сцена

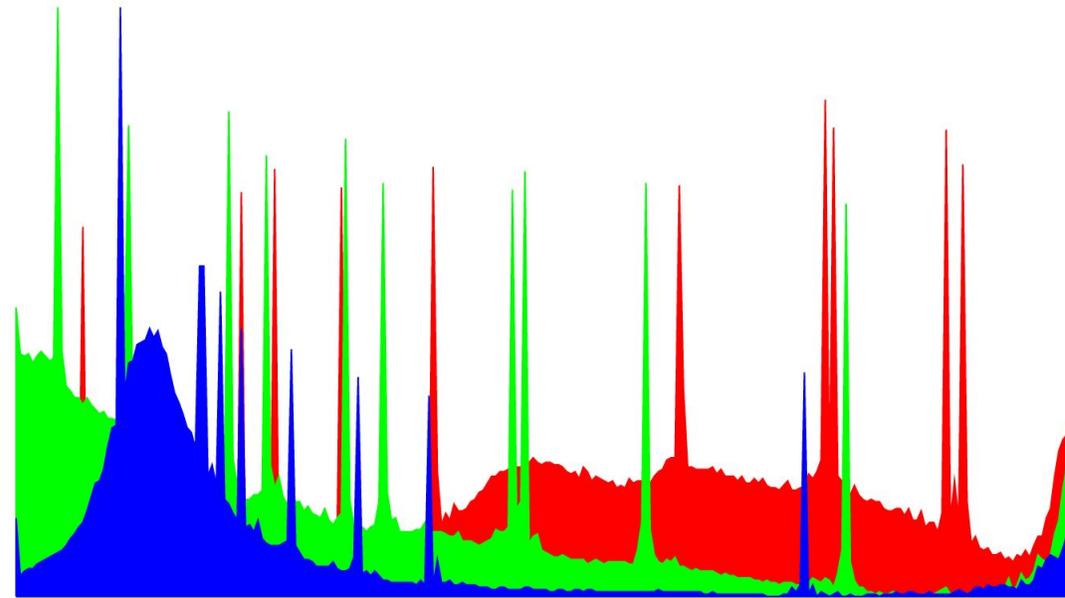
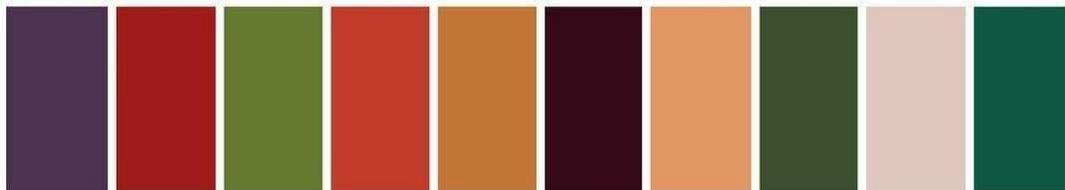


Третий план вторая сцена

Методы динамической оптимизации

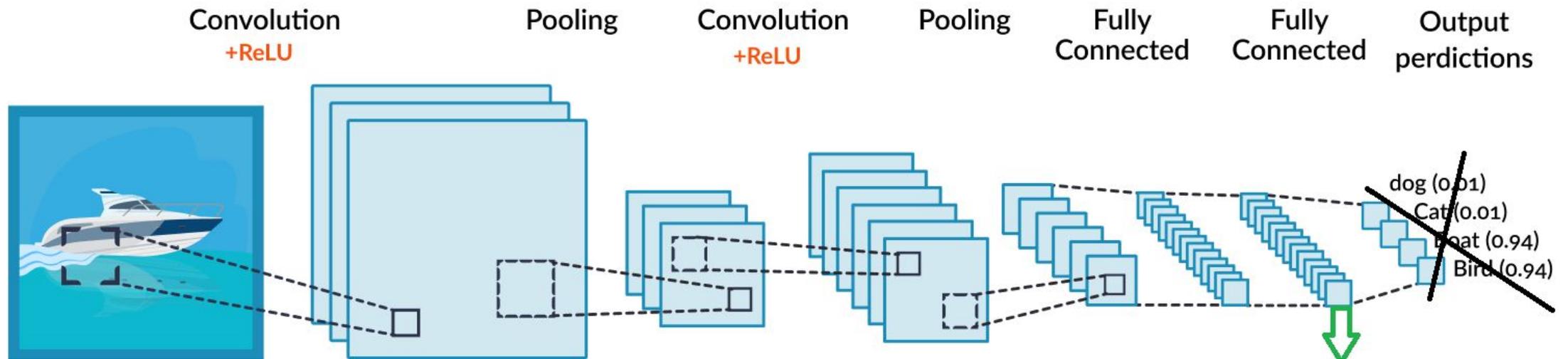


Особенности планов



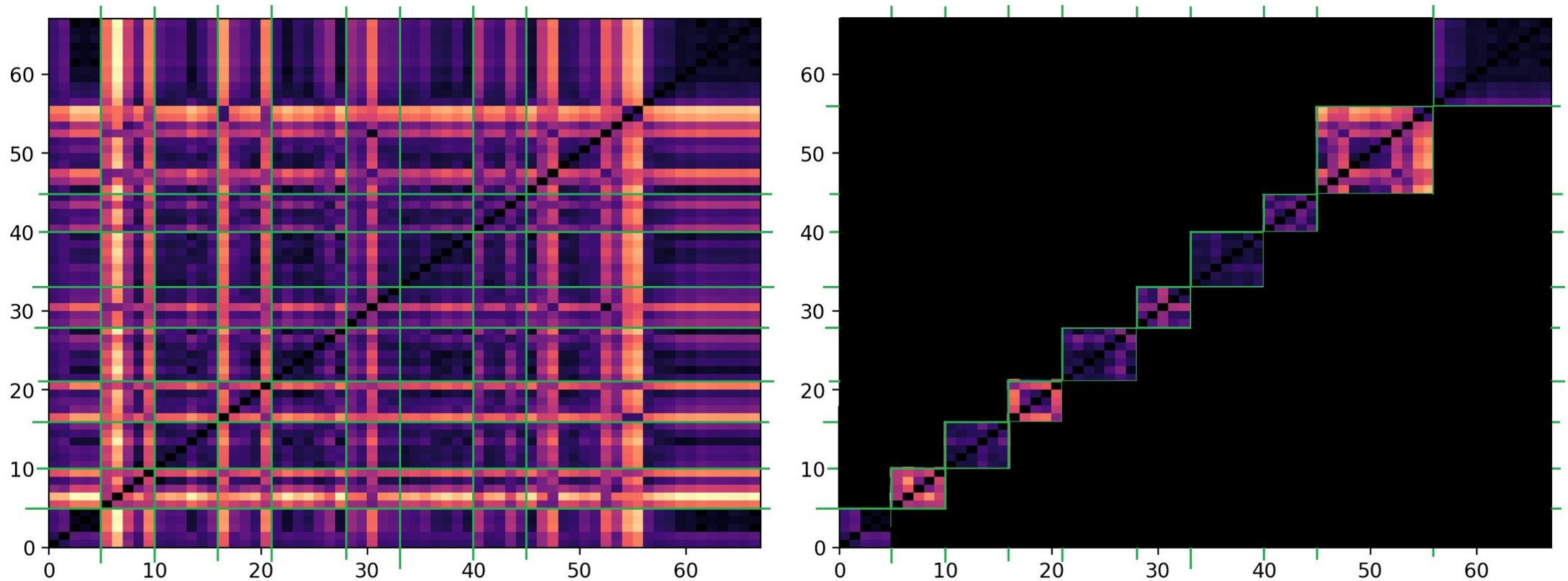
RGB гистограмма

Неявные особенности



Извлечение вектора особенностей из классифицирующей свёрточной нейронной сети

Матрица расстояний планов



Пример матрицы расстояний планов с выделенными на ней связными сценами

Метод динамической оптимизации аддитивной функции

- Пусть $D(j, i) = D(x_j, x_i)$ функция расстояния планов с индексами i и j , тогда задачу объединения планов в сцены можно определить как задачу минимизации функции $H_D^{K,N}$, где K – количество планов, N – количество сцен

$$H_D^K(t) = \sum_{i=1}^K \left\{ \frac{\left(\sum_{j=t_{i-1}+1, j'=t_{i-1}+1}^{t_i} D(x_j, x_{j'}) \right)}{S_i} \right\}$$

- Но значение целевой функции не содержит информации о разбиении
- Инициализируем таблицы C, J, E :
- $C(n, 1) = \sum_{j=n}^N \sum_{j'=n}^N D(x_j, x_{j'})$
- $J(n, 1) = N$
- $E(i, i') = \sum_{j=i}^{i'} \sum_{j'=i}^{i'} D(x_j, x_{j'})$

Метод динамической оптимизации аддитивной функции

Е может быть рассчитана рекурсивно при запросе как

$$E(i, i') \\ = E(i - 1, i') + E(i, i' - 1) - E(i - 1, i' - 1) + D(x_i, x_{i'}) + D(x_{i'}, x_i)$$

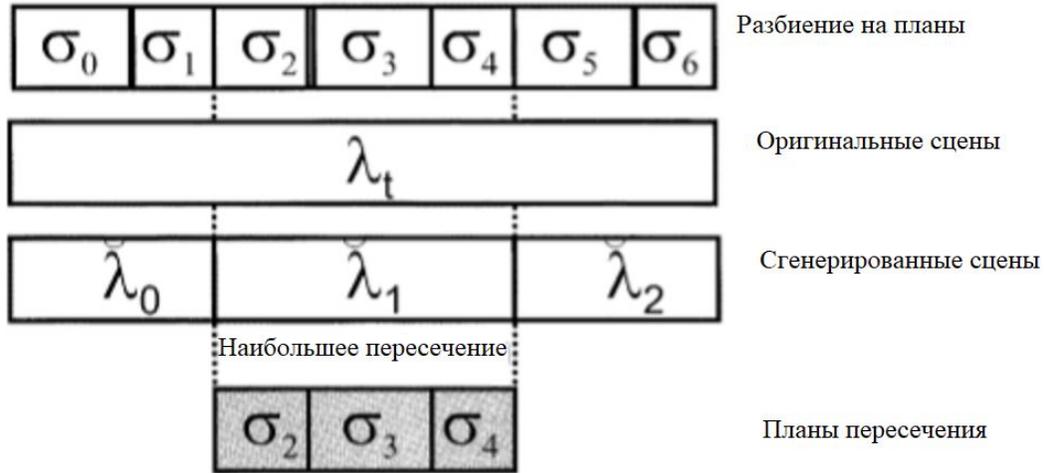
Тогда

$$C(n, k) = \min_i \{E(i, i') + C(i + 1, k - 1)\}$$

$$J(n, k) = \operatorname{argmin}_i \{E(i, i') + C(i + 1, k - 1)\}$$

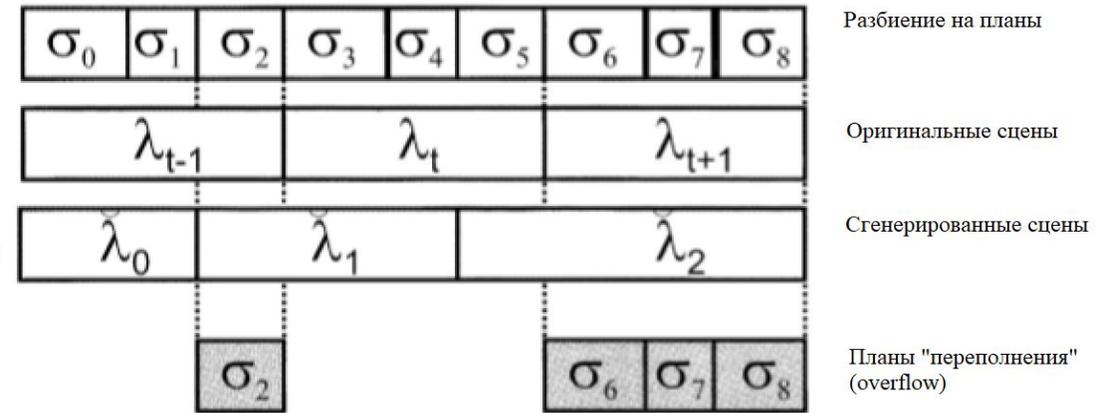
А граница i -той сцены находится как $t_i^* = J(t_{i-1} + 1, K - i + 1)$

Оценка результатов работы



$$C(L_{new}, L_{orig}) = \frac{\sum^n \maxintersec(L_{new}, L_{orig_i})}{\sum^n len(L_{orig_i})}$$

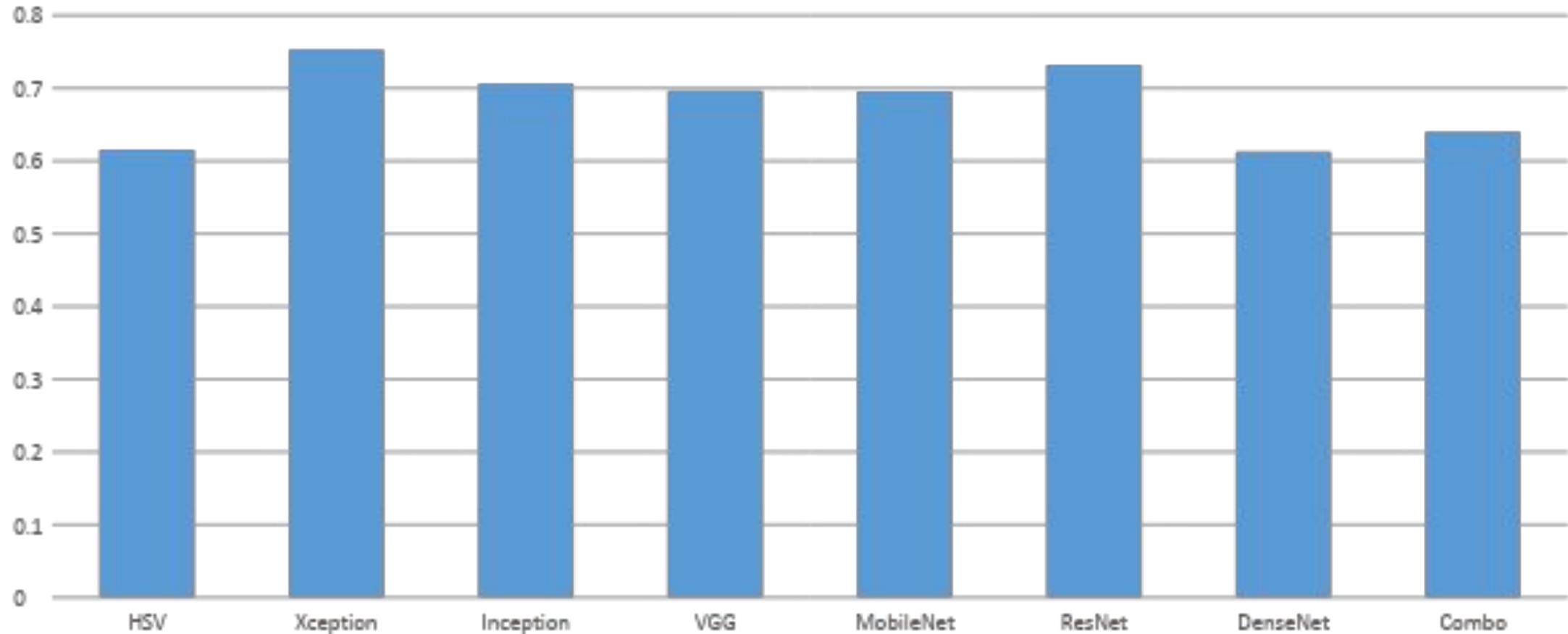
Coverage
("Покрытие")



$$O(L_{new}, L_{orig}) = 1 - \frac{\sum^n overflow(L_{new}, L_{orig_i})}{\sum^n len(L_{orig_i})}$$

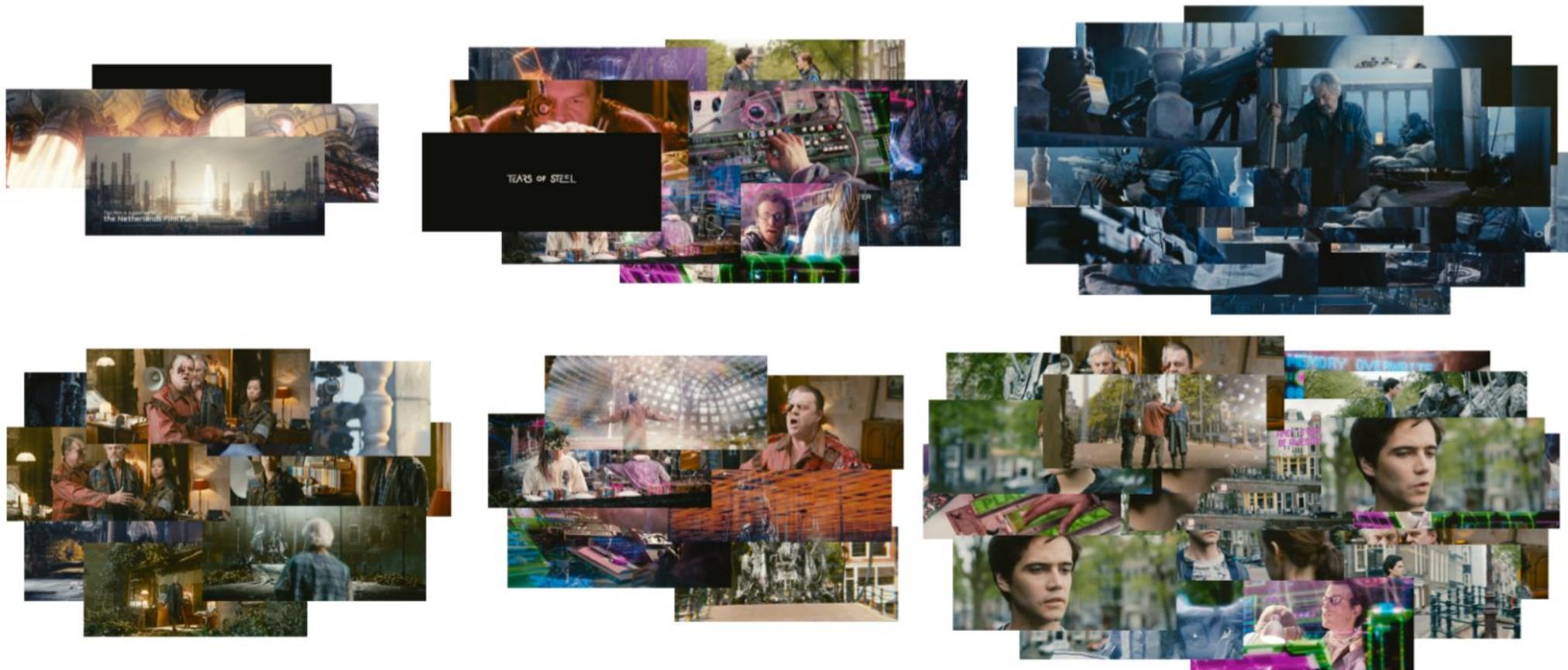
Overflow("Переполнение")

Зависимость работы метода от извлеченных особенностей



Значение F-меры Coverage+Overflow для набора данных OVSD

Результаты создания СВЯЗНЫХ СЦЕН



Заключение

В результате выполнения работы были решены поставленные задачи:

- проведён анализ предметной области, выделены основные подходы к созданию связных сцен;
- разработан алгоритм решающий поставленную задачу;
- описаны специфические особенности предложенного метода;
- разработано программное обеспечение для создания связных сцен;
- исследованы наиболее важные аспекты разработанного программного продукта.

Направления развития

- повышение точности предсказания количества сцен;
- извлечение высокоуровневых особенностей планов;
- повышение точности сегментации;
- замена аддитивной функции стоимости на функцию менее зависящую от номера плана ;
- использование сиамских нейронных сетей для определения расстояния между планами