|  |  |
| --- | --- |
|  | **Министерство образования и науки Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ \_\_\_\_\_\_ИУ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

КАФЕДРА \_\_\_\_\_\_\_\_ИУ7\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**ОТЧЕТ ПО УЧЕБНОЙ ПРАКТИКЕ**

Студент\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Гаврилова Юлия Михайловна\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*фамилия, имя, отчество*

Группа\_\_\_\_ИУ7-37\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Тип практики \_\_\_\_учебная\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Название предприятия\_\_\_\_\_\_МГТУ им. Н.Э. Баумана\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Студент **\_**\_\_\_\_\_\_\_\_\_\_\_\_\_\_**\_\_\_ \_\_** Гаврилова Ю.М.**\_\_**

*подпись, дата фамилия, и.о.*

Руководитель практики **\_**\_\_\_\_\_\_\_\_\_\_\_\_\_\_**\_\_\_ \_\_** Филиппов М.В.**\_\_**

*подпись, дата фамилия, и.о.*

Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*2017 г.*

**Индивидуальное задание**

Провести анализ предметной области, выделить цель и основные задачи дипломной работы. Выполнить обзор существующих аналогов и инструментов для решения выделенных задач.

Оглавление

[Введение 4](#_Toc493775540)

[Применение сегментации изображений 6](#_Toc493775541)

[Особенности задачи определения движущегося объекта в видеопотоке 6](#_Toc493775542)

[Существующие решения 8](#_Toc493775543)

[1. Простейший алгоритм вычитания фона 8](#_Toc493775544)

[2. Кластеризация видеопотока 8](#_Toc493775545)

[3. Метод “разделить и объединить” (split and merge). 10](#_Toc493775546)

[4. Сегментация с помощью модели 10](#_Toc493775547)

[Заключение 12](#_Toc493775548)

[Доп. материалы для проработки: 13](#_Toc493775549)

# Введение

Тема дипломной работы: «Сегментация объекта в видеопотоке».

В рамках дипломной работы необходимо реализовать программное обеспечение, реализующее следующий функционал: загрузку видеопотока, разбиение входной информации на ключевые кадры, анализ гистограмм полученных изображений и сегментация ключевых кадров с последующим сохранением результата.

Данная работа может быть применена в таких сферах как: анализ медицинских изображений, выделение объектов на [спутниковых снимках](https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D1%83%D1%82%D0%BD%D0%B8%D0%BA%D0%BE%D0%B2%D0%B0%D1%8F_%D1%84%D0%BE%D1%82%D0%BE%D1%81%D1%8A%D1%91%D0%BC%D0%BA%D0%B0), р[аспознавание лиц](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D0%BF%D0%BE%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%BB%D0%B8%D1%86" \o "Распознавание лиц), системы управления дорожным движением и так далее.

Целью учебной практики является обзор предметной области и декомпозиция задачи. Для подзадач необходимо выделить существующие методы и инструменты их решения.

**Применение сегментации изображений**

Сегментация изображений — это процесс присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками имеют общие визуальные характеристики. Сегментация изображений обычно используется для того, чтобы выделить объекты и границы (линии, кривые, и т. д.) на изображениях. Наиболее распространенное применение сегментация изображений нашла в следующих сферах:

* Медицинские изображения
* Выделение объектов на [спутниковых снимках](https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D1%83%D1%82%D0%BD%D0%B8%D0%BA%D0%BE%D0%B2%D0%B0%D1%8F_%D1%84%D0%BE%D1%82%D0%BE%D1%81%D1%8A%D1%91%D0%BC%D0%BA%D0%B0)
* [Распознавание лиц](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D0%BF%D0%BE%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%BB%D0%B8%D1%86)
* Системы управления дорожным движением
* [Машинное зрение](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5)

Для большинства указанных выше сфер сегментация объектов в видеопотоке – следующий этап распознавания.

**Особенности задачи определения движущегося объекта в видеопотоке**

Основной задачей сегментации объекта в видеопотоке является выделение фона и самого объекта. В случае, когда фон также имеет тенденцию к изменению (например, движение листьев на дереве, перемещение облаков и т.п.), появляются ограничения на используемые алгоритмы и входные значения.

Существует 2 основных вида ошибок:

* переднеплановый пиксель был классифицирован как заднеплановый (например в следствии шума или из-за того, что цвет переднепланового объекта совпадает с цветом заднего плана)
* заднеплановый пиксель был классифицирован как переднеплановый (например из-за шума).

Оба вида ошибок можно связать с изменением следующих параметров:

* Изменения заднего плана или динамический задний план

Например, видеокамера установлена на улице и в поле ее видимости находится машина, через некоторое время приходит хозяин машины садится в нее и уезжает. После того, как машина уедет, камере откроется часть сцены, которую машина собой загораживала. Если в заднем плане сцены подобного рода изменения будут происходить достаточно часто, то в итоге большая часть сцены будет сегментироваться в передний план и данные сегментации будут совершенно непригодны для анализа и обработки.

* Изменение освещения

Изменения освещения практически полностью меняют цветовые характеристики сцены. В итоге из-за сильного изменения освещения алгоритмы, чувствительные к цвету, будут сегментировать в передний план большую часть сцены, что разумеется совершенно неприемлемо. Стоит разбить изменения освещения на 2 группы: быстрые (такие, например, как включение света в комнате) и медленные (например изменение освещения в течение дня). С медленными и быстрыми изменениями освещения можно бороться по-разному.

* Движущиеся тени

Автоматическое нахождение теней - довольно сложная задача, особенно при движущихся тенях, т.е. от движущихся объектов сцены. Дело в том, что падение тени на объект довольно сильно меняет цвет объекта.

* Шум камеры

Разные камеры шумят по разному и это также может быть проблемой для вычисления фона. Неадаптивный порог приходится настраивать вручную для разных камер и разных настроек камеры по разному, что разумеется нежелательно для автоматических систем.

**Существующие решения**

На практике – наиболее часто используются следующие алгоритмы:

1. **Простейший алгоритм вычитания фона**

Будем считать, что с видеокамеры поступают полутоновые изображения (если с камеры поступают цветные изображения то их очевидно можно привести к полутоновым, например используя формулуhttp://masters.donntu.org/2011/fknt/galiakberow/images/substraction/image006.gif). Алгоритм сохраняет первый кадр видеопоследовательности, а потом для каждого следующего кадра применяет порог к модулю разности текущего и сохраненного изображения по каждому пикселю

1. **Кластеризация видеопотока**

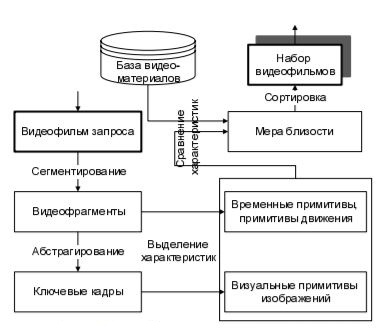
Одним из наиболее известных и распространенных методом кластеризации является [k-средних](https://ru.wikipedia.org/wiki/K-means) [2]

Это [итеративный](https://ru.wikipedia.org/wiki/%D0%98%D1%82%D0%B5%D1%80%D0%B0%D1%86%D0%B8%D1%8F) метод, который используется, чтобы [разделить изображение](https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%82%D0%B5%D1%80%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) на K кластеров. Базовый [алгоритм](https://ru.wikipedia.org/wiki/%D0%90%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC) приведён ниже:

1. Выбрать K центров кластеров, случайно или на основании некоторой [эвристики](https://ru.wikipedia.org/wiki/%D0%AD%D0%B2%D1%80%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC)
2. Поместить каждый пиксель изображения в кластер, центр которого ближе всего к этому пикселю
3. Заново вычислить центры кластеров, [усредняя](https://ru.wikipedia.org/wiki/%D0%A1%D1%80%D0%B5%D0%B4%D0%BD%D0%B5%D0%B5_%D0%B0%D1%80%D0%B8%D1%84%D0%BC%D0%B5%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5) все пиксели в кластере
4. Повторять шаги 2 и 3 до сходимости (например, когда пиксели будут оставаться в том же кластере)

В настоящее время интерес представляют именно методы кластеризации, позволяющие работать с большими наборами данных. Наиболее известные – CLARANS, DBSCAN и X-средних.

На вход системе распознавания видеоданных подается видеофайл, который в результате сегментирования разбивается на последовательность видеофрагментов. Для каждого полученного видеофрагмента выделяется ключевой кадр (абстрагирование), для которого считаются различные характеристики: визуальные примитивы (характеристики цвета, формы, текстуры объектов изображения ключевого кадра), временные примитивы и примитивы движения. В случае, если на вход системе подается отдельный видеокадр, то для него считаются только визуальные примитивы, которые затем сравниваются с визуальными примитивами ключевых кадров видеофрагментов анализируемой коллекции видеоматериалов.



1. **Метод “разделить и объединить” (split and merge).**

Данные метод применяется по аналогии кластеризации видеопотока выделением ключевых кадров и последовательным применением Split&Merge [3] к выбранным файлам**.** Метод “разделить и объединить” требует, чтобы все данные были организованы в пирамидальную решеточную структуру из областей, где каждая область организована в группы по восемь (для трехмерного случая). Любая область может быть разбита на восемь частей, и, соответственно, восемь частей могут быть объединены в одну большую область. Как и в случае алгоритма наращивания области, критерием для объединения может быть все, что угодно: интенсивность вокселя (как самый простой вариант), или некоторое условие, основанное на результатах предварительной сегментации. Допустим, что таким условием является С. Тогда алгоритм можно записать следующим образом: 1) выбрать область R в пирамидальной решеточной структуре; если C(R) = ложно, тогда разбить область на восемь частей; если для восьми областей R1, R2,…,R8 C(R1, R2,…,R8) = верно, о бъединить их в одну область; когда не останется областей, которые можно объединить, остановиться; 2) если все соседние области Ri и Rj таковы, что C(Ri∪Rj) = верно, то объединить эти области. Большим преимуществом данного метода перед алгоритмом наращивания области является отсутствие необходимости вмешательства оператора для выбора начальной точки. С другой 15 стороны, метод требует организации исходных данных в пирамидальную решеточную структуру, что может быть нежелательно для больших объемов данных

1. **Сегментация с помощью модели**

Основное предположение этого подхода — то, что интересующие структуры или органы имеют повторяющиеся геометрические формы. [4] Следовательно, можно найти вероятностную модель для объяснения изменений формы органа и затем, сегментируя изображение, накладывать ограничения, используя эту модель как априорную. Такое задание включает в себя (i) приведение тренировочных примеров к общей позе, (ii) вероятностное представление изменений приведённых образцов и (iii) статистический вывод для модели и изображения. Современные методы в литературе для сегментации, основанной на знании, содержат активные модели формы и внешности, активные контуры, деформируемые шаблоны и методы установления уровня.

# Заключение

В результате выполнения учебной практики были выделены основные направления развития современной сегментации видеопотоков. Проведено подробное изучение существующих методов и реализация прототипа программного комплекса.

**Доп. материалы для проработки:**

1. Jianbo Shi, Jitendra Malik. Normalized Cuts and Image Segmentation
2. Богучарский С. И. Методы и моедели толерантноой кластеризации в коллекциях изображений
3. https://compscicenter.ru/courses/images-and-video-2/2014-spring/classes/289/
4. http://dic.academic.ru/dic.nsf/ruwiki/1623093
5. http://www.jprr.org/index.php/jprr/article/view/54/16