**sssssТ Е Х Н И Ч Е С К И У Н И В Е Р С И Т Е Т – С О Ф И Я**

**Ф а к у л т е т “Приложна математика и информатика”**

**Дипломна работа**

**Тема: „Система за лицево разпознаване и проследяване. Реализация във вградени системи.*“***

***на***

***Симеон Христов Лаплев***

**Факултетен номер: *181217011***

Изработил:............... Научен ръководител.................

/Симеон Лаплев/ / ас. Златко Захариев /

**София, 2021г.**

**Съдържание**

**ГЛАВА 1 - ВЪВЕДЕНИЕ**

* 1. **Обща информация**
     1. **Кратка история**
  2. **Цели и мотивация**

**ГЛАВА 2 - ФУНКЦИОНАЛНИ ИЗИСКВАНИЯ**

1. **Разпознаване на лице и локализация**
2. **Проследяване на лицето със сбор от сензори на вградената система**

**ГЛАВА 3 -** **ТЕХНОЛОГИИ ЗА ИЗГРАЖДАНЕ НА СИСТЕМАТА**

1. **Raspberry Pi 4**
2. **Python 3**
3. **OpenCV**

**ГЛАВА 4 - АРХИТЕКТУРА НА СИСТЕМАТА**

1. **Механична структура и базов набор от сензори**
2. **Софтуер**

**ГЛАВА 5 - ПОЯСНЕНИЯ НА КЛЮЧОВИ СПЕЦИФИКИ**

**ГЛАВА 6 - ОПИТНИ РЕЗУЛТАТИ**

**ГЛАВА 7 – ЗАКЛЮЧЕНИЕ**

**ИЗПОЛЗВАНА ЛИТЕРАТУРА**

Глава 1 – Въведение

# Обща информация

По дефиниция **система за лицево разпознаване** е технология пригодена за откриване на човешко лице от дигитално изображение (в частен случай кадър от видео). В зависимост от изискванията на системата, откритото лице може да се сравни с набор от записани лица в база данни и да се идентифицира – например с цел автентикация. В последно време системите за лицево разпознаване са станали изключително широко разпространени. Виждат огромно приложение в днешните смартфони и други типове технологии тясно свързани с роботиката. Тъй като компютърното лицево разпознаване включва измерването на човешки физиологични характеристики, то се категоризира като „биометрика“. Въпреки грубата прецизност на технологията в сравнение с други биометрични системи, лицевото разпознаване се прилага екстензивно в днешно време поради удобството на безконтактността, която е значително предимство на метода. В резултат, лицевото разпознаване се използва в области като: *сложни интеракции човек-компютър*, *видео наблюдение*, *автоматизирано индексиране и категоризация на изображения*; Използва се екстензивно и от органи на реда.

# Кратка история

Автоматизираното лицево разпознаване започва началото си през 60-те години на 20-ти век. Имена като Уди Бледсоу, Хелен Чан Волф и Чарлз Бисон работят по задачата на лицево разпознаване с компютри. Първите им опити за разработка на технологията е била значително опростено сравнение на дистанции между физиологични черти с помощта на компютър. Реалната идентификация на основните черти и техните позиции се е случвало с човешка помощ. Чертите на лицето са се посочвали от доброволци и координатите на съответните позиции са се записвали в база данни. След това компютъра минава през посочените черти в дадено изображение и сравнява дистанциите срещу записаните в базата данни. Това е ставало сравнително бавно, защото е зависело от механичен човешки труд.

След 90-те години на 20-ти век технологията започва да се проучва по-интензивно с подкрепата на американски военно-центрирани академични организации като DARPA и ARL. Интересът е завишен под консенсуса да се разработи „способност за автоматично лицево разпознаване“ за продуктивни приложения в истинския свят под претенцията: „подпомагане на охраняемостта, разузнаването, и органите на реда в изпълнение на техните длъжности.“ Не отнема дълго време разработените системи да видят използваемост, подпомагайки с индексиране на снимки върху лични документи с цел откриване на злоупотреби. До 1990-те системите за лицево разпознаване се разработват основно използвайки фото портрети на човешки лица. Проучването върху проблема с откриване на лица надеждно в образ, който съдържа и други обекти, се засилва през ранните години на същото десетилетие с така наречения *принципен компонентен анализ*. Метода се разработва от Матю Търк и Алекс Пентланд. Това решение приема идеята, че всяко едно уникално откриваемо лице може да се представи като тегловен процент на чертите от лицата записани в тренировъчния базис. Системата се оказва доста ефикасна със сравнително малък брой тренировъчни данни. Тренировъчните данни, въпреки това, изискват сравнително стриктна нормализация и осветление, което по това време не е било много изгодно за изчисление. Метода има и недостатъка, че е лабилен към не добре осветени примери. Промяната в емоционалното изражение на лицето също представлявало спънка за алгоритъма.

През 2001-ва година откриването на лица в истинско време посредством видео става възможно благодарение на „Виола-Джоунс методологията за откриване на обекти.“ Паул Виола и Майкъл Джоунс комбинират техния метод за откриване на лица с **подхода на Хаар-наподобяващи черти\*1** за откриване на обекти в дигитални изображения. Явява се така нареченият *AdaBoost*, първият фронтален лицев детектор, работещ в истинско време. До 2015-та година алгоритъма на Виола-Джоунс се имплементира във все по-малки и по-маломощни ръчни устройства и вградени системи. Съответно, Виола-Джоунс алгоритъма не само разширява практическото приложение на системите за лицево разпознаване, но и ги полага в центъра на вниманието, що се отнася до приложимост в преобладаващите технологии на модерния свят.

# Цели и мотивация

В днешно време достъпността на дребен, но същевременно способен хардуер е установена като факт. С един микроконтролер, човек има възможността да интегрира сложни концепции в система с най-различни приспособления. Целта на тази работа е да установи една базова интеграция на сравнително сложна концепция – именно лицево разпознаване и локализация. Да покаже, че приложността и разширяемостта на един подобен проект няма лесно начертана граница. Благодарение на технологиите, които предстои да се разгледат, ще се демонстрира достъпността на подобни разработки.

Когато се замислим каква е приложността на подобна интеграция, много примери излизат наяве. С оглед на простота на текста, ще дадем два конкретни примера.

*Пример 1.* Да кажем, че имаме определена личност под наблюдение с медицински цели. Въпросната личност се намира в помещение, в което е свободна да обикаля. Ключова информация се състои в образа и изражението на лицето на човека. Ключови кадри трябва да се събират с възможно най-голяма честота. Тази информация трябва да се снема по начин, по който ще е нужна възможно най-малко нормализация на данните. Проблемът е, че помещението е сравнително голямо и за това за възможно най-качествено изображение в далечина, фокусното разстояние на използваната леща е голямо. По този начин полезрението на камерата и/или другите сензори в конструкцията е доста ограничено. Не може лесно да се обхваща цялото помещение и да се очаква чист и постоянен образ на лицето на човека. Тази конструкция е важно да няма слепи точки и за това трябва да има поне две степени на свобода при насочване и да се намира на установена отворена позиция в близост до средата на помещението. Ако субекта е с лице към сензорите и се намира в полезрението на обектива, образа активно ще се центрира около лицето на въпросната личност. Ако лицето не е в полезрението на обектива, сборът от сензори ще почнат да преглеждат панорамно цялото помещение, докато лицето не бъде открито. По този начин се осигурява нужната ключова информация с възможно най-голяма честота в реално време.

*Пример 2*. Имаме нужда от система за домофон с видео връзка, която да има редица интеракции с потребителски разработен интерфейс състоящ се от много функционалности. За целта обаче ни трябва да се разработи сглобка, която да има свободно движеща се камера, която да центрира образа около всички лица в полезрението на домофона. Подобно на системата от *Пример 1* трябва да имаме механично проследяване и опцията за преглед на околностите със сглобката на въпросната камера. В случая, на разработчика му е нужно да състави интеграцията с налични технологии, за да има възможно най-голям контрол върху притока и нормализацията на данни от системата.

Мотивацията за проект с подобна интеграция може да варира драстично. От спонтанна идея за потребителски зададен характер на система за сигурност до нормализация на лицеви данни с академични или медицински цели. Границите на приложимостта се определят от въображението и/или работните изисквания на разработчика.

Глава 2 – Функционални изисквания

# Разпознаване на лице и локализация

Една от основните задачи на система за лицево разпознаване е да разграничи лицето в предоставения дигитален образ. В нашия случай изображението ще представлява кадър от видео, който се обработва в реално време. Трябва лицето да може да се разграничава при умерени светлинни условия с лека вариация в наклона и усукване на образа на лицето. С установена важност е това да се случва поне два пъти в секунда след всички допълнителни обработки върху кадъра.

В детайл, информацията която ни трябва от разграниченото лице е под формата на три различни стойности:

* **Двумерен вектор**  описващ координатите на пиксела върху изображението, от който може да се опише правоъгълния регион от образа, съдържащ лицето.
* Широчината на региона с лицето **w**
* Височината на региона с лицето **h**

В частност, като допълнение към задачата имаме възможността да предвидим притежателя на лицето ако имаме съществено трениран модел и предположението оправдава определена граница на сигурност. В такъв случай ще получим и **име на притежател** като текстови данни.

С получаването на горепосочената информация, имаме всичко нужно да правим допълнителна обработка на сегашния кадър от видеото, както и да манипулираме състоянието на сензорите прикачени към вградената система в зависимост от интерпретацията на тези данни.

Очевадно, искаме това да се случва с възможно най-малка вариация във времената за обработка, но това основно зависи от процесуалната мощ на компютъра, който ще се използва за проекта. Плюс някои очевидни оптимизации, които ще бъдат споменати в следващите глави, системата която ще използваме за изчисление ще оправдае няколкократно процесуалната мощ за тази задача.

# Проследяване на лицето със сбор от сензори на вградената система

След ключовата задача за разпознаване на лицето и неговата локализация, следва не по-малко важната задача за проследяване и центриране на лицата в образа ни.

Тук ключовите изисквания са тясно свързани със задвижването на сензорите прикачени към системата. В нашия случай имаме единия базов сензор – камерата. Моментално става ясно, механизма носещ сензорите трябва да има способността да се насочва ъглово с поне две степени на свобода. Обективът на камерата трябва да се намира в централната точка на въртене, за да няма перспективно изкривяване на образа при завъртане около двете си оси. Иначе казано, разстоянията между обектива и двете крепежни точки по съответните оси на ротация трябва да са равни.

Допълнително, ротацията на обектива трябва да се случва възможно най-плавно, за да не се прескачат ключови кадри и да не се представя ненужно лашкане.

Възниква и въпроса какво се случва при липса на открито лице във видео потока. Много подобно на споменатия *Пример 1* в първа глава, системата трябва да мине в режим на „сканиране“, при който камерата започва да се движи по предварително зададен ъглов маршрут с цел да се открие лице. В момента в който лице бъде открито, системата трябва пак да мине в режим „проследяване“ и да се стреми да центрира лицето подобаващо в кадър.

Конструкцията и крепежните точки трябва да са достатъчно здрави, за да елиминират излишни вибрации от ротацията по двете оси. Тъй като масата на товара, който се носи не е голяма (от порядъка на няколко грама за този проект), то има доста валидни варианти за материалите и крепежните елементи. Подхода, който е поет за тази конструкция ще се обсъди подробно в **Глава 4**.

Глава 3 – ТЕХНОЛОГИИ ЗА ИЗГРАЖДАНЕ НА СИСТЕМАТА

# Raspberry Pi 4

За сърцето на изчисленията бе избран един от най-обещаващите съвременни едноплаткови компютри с възможност за дигитално управление – Raspberry Pi 4. Този ARM базиран микроконтролер е повече от идеален за целите на тази система. Операционната система, която ще използваме е сравнително олекотен билд на Линукс. Съответната Линукс дистрибуция е базирана на Debian и предоставя нужните зависимости за ефективното използване на възможностите на платката.

*Фиг. 1 – Raspberry Pi 4*

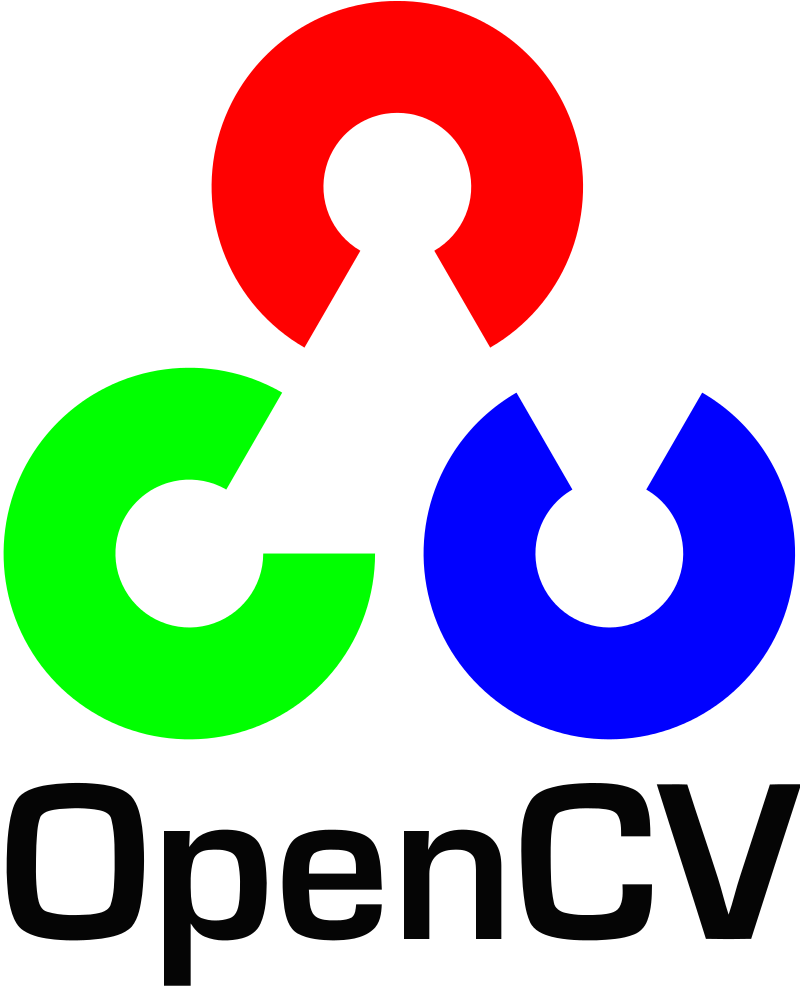
# Python 3

Raspberry Pi фондацията, предоставя инструменти за ползване на основния програмен език за платформата – Python. На пръв поглед този избор е спорен, тъй като основната цел на този компютър е да се използва във вградени системи. Python е език със сравнително голямо изпълнително време за повечето инструкции на високо ниво, тъй като в много от структурите на езика липсват педантичните компилаторски оптимизации, които можем да видим в някои по-общоприети езици използвани в тази сфера (пр. C, C++, Rust и пр.) По-общо погледнато обаче това съвсем не е лошо решение от гледна точка на достъпността за разработчици от всички инженерни и информатични сфери. Python е изключително достъпен и добре поддържан език с голям набор от библиотеки за всякакви цели. Основната идея на използването на подобен език е, че в повечето случаи, разработката е ускорена значително поради липсата на много инженерни специфики, които да касаят разработчика. В следствие, четимостта на софтуера е подобрена. Липсва нуждата от изграждане на посредствено скеле от код, което в крайна сметка е можело да го няма като цяло ако се крие зад нужните абстракции. Конкретиката не бива да е основно притеснение на разработчика, чието внимание следва да е центрирано изцяло върху приложните цели на системата.

Настрана от това отстъпление, Python е изключително способен и изчистен език, центриран върху производителност, макар и не от най-бързите. Компютърът който ще използваме в тази работа има повече от нужния капацитет да изпълнява код на високо ниво, без това да се отрази значително върху крайния резултат.

Редица Python библиотеки и зависимости ще бъдат използвани за програмирането на системата. Тяхното инсталиране ще бъде тривиално благодарение на вградения мениджър на пакети, който идва с Python 3 на име **Pip**.

# OpenCV



Основната библиотека в центъра на цялата система е OpenCV.

OpenCV (Open Source Computer Vision Library) е библиотека основно насочена към областта на компютърно зрение. Тя бива разработена от Интел и те остават основен поддръжник на проекта и до днес. От 2011 насам, OpenCV поддържа ускорение с графичен хардуер за по-представителна работа в реално време.

Софтуера е разработен за свободно ползване, лицензиран под общите условия на Apache лиценза. Всеки, който се интересува от сферата на компютърно зрение и/или машинно обучение може свободно да използва тази мощна библиотека, да я модифицира и преустановява както пожелае.

OpenCV има много възможности вградени в себе си, но за целта на сегашната задача, ние ще се интересуваме главно от модула за разпознаване на лица.

Благодарение на екстензивната ѝ функционалност, можем сравнително лесно да изградим основата на проекта като прозореца на програмата за демонстрация, заедно с по-голямата част от обработката на образа.

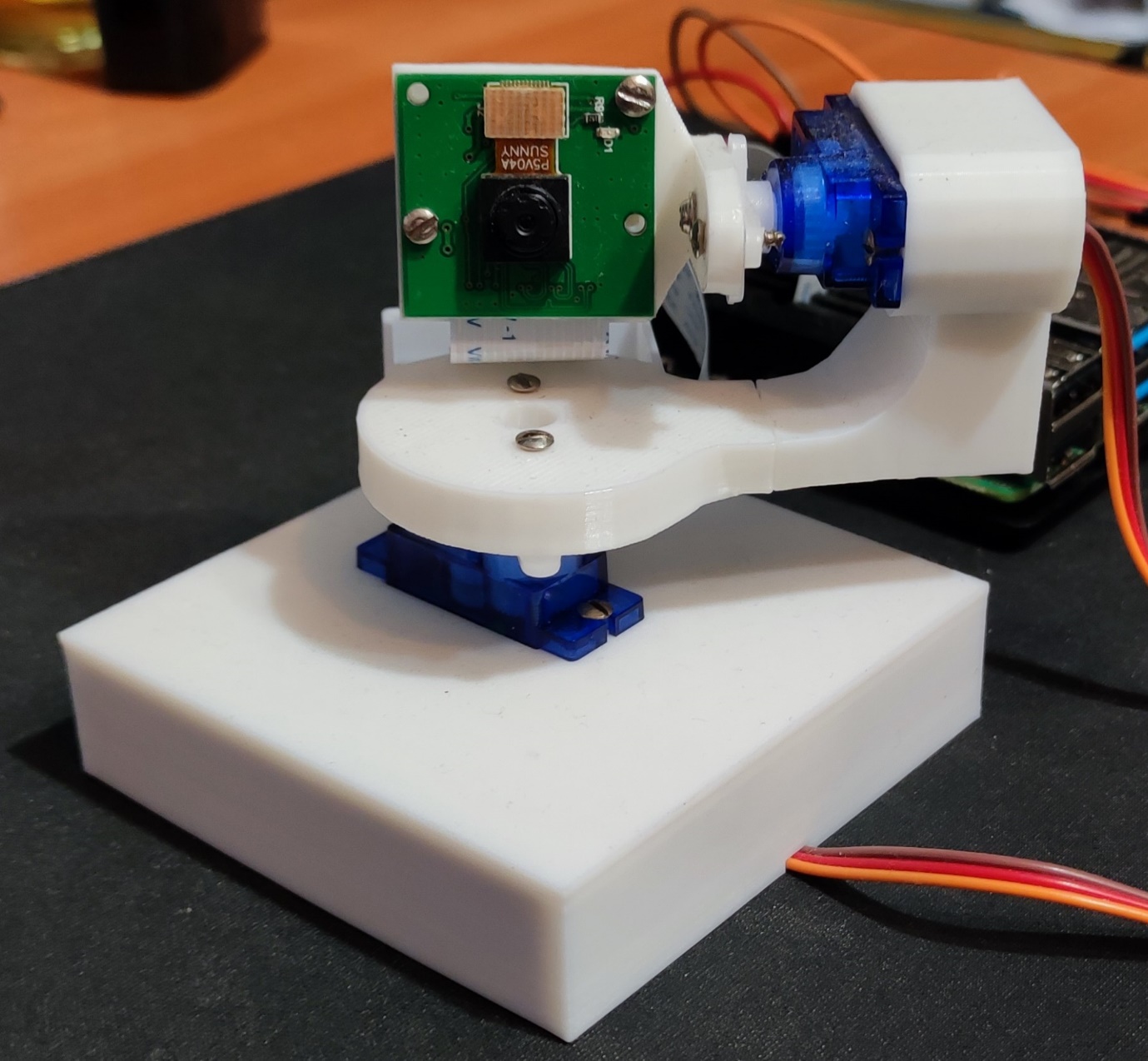
Глава 4 – Архитектура на системата

# Механична структура и базов набор от сензори

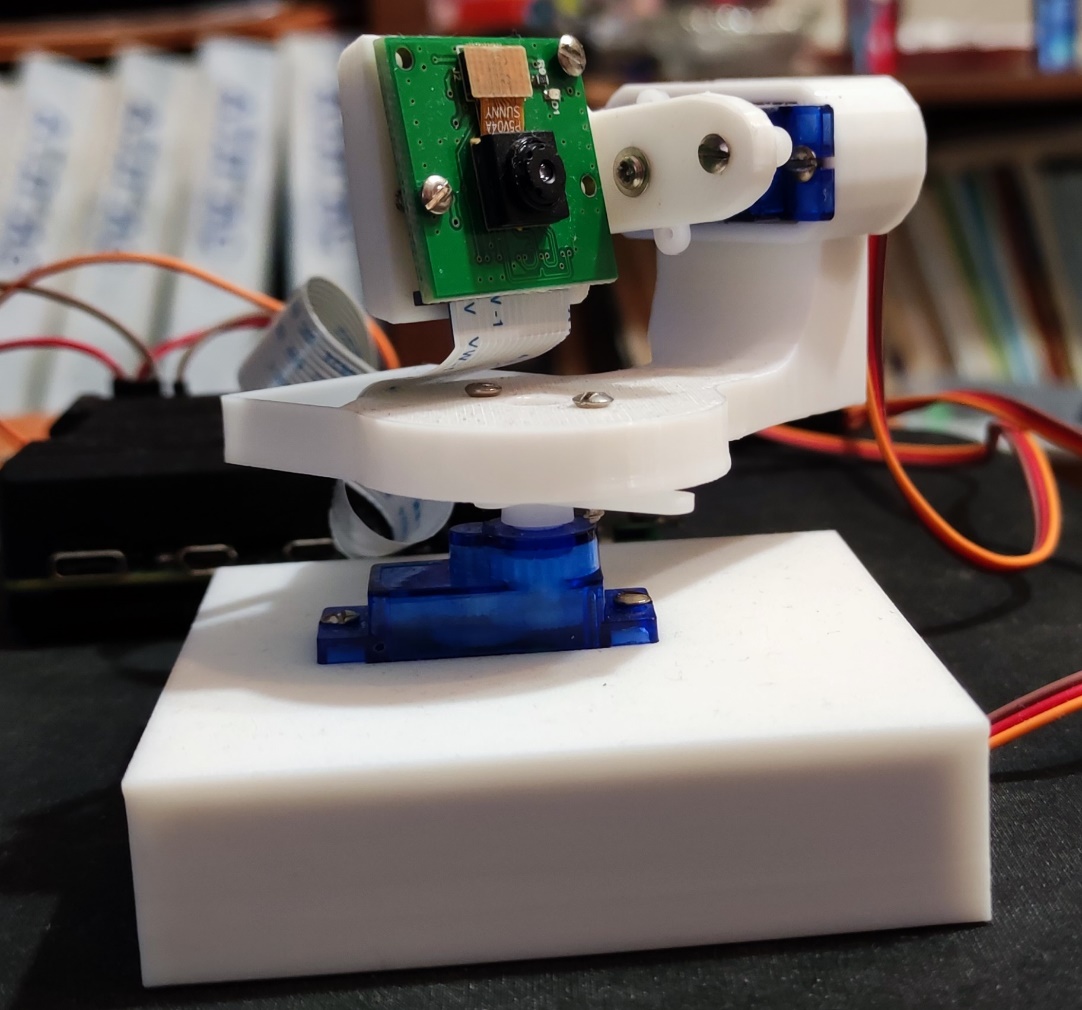
Конструкцията на механизма трябва да спазва функционалните изисквания подчертани в Раздел 2.2

За пълна яснота нека разгледаме подробно готовата конструкция и процеса на произвеждане на механизма.

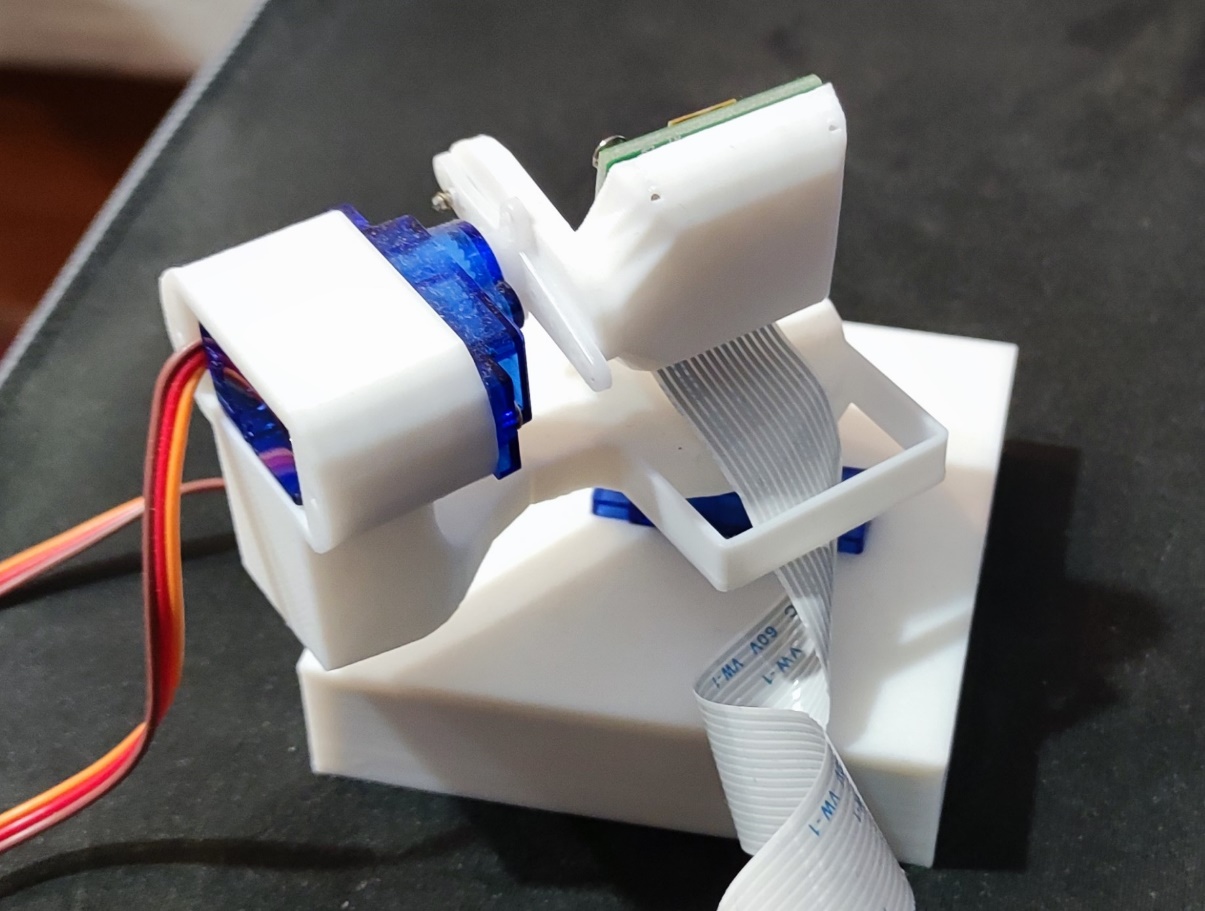
*Фиг. 2.1 – Готовият механизъм (отпред)*



*Фиг. 2.2 – Готовият механизъм (отстрани)*



*Фиг. 2.3 – Готовият механизъм (отзад)*

**

Първият компонент, който ще вземем под внимание са именно двете сервота, които ще задвижват конструкцията. Конкретно, нека опишем продукта и да минем през някои негови характеристики. Също така предстои обяснение на приложението му в контекста на тази задача.

*Фиг. 3 – Micro servo SG90*

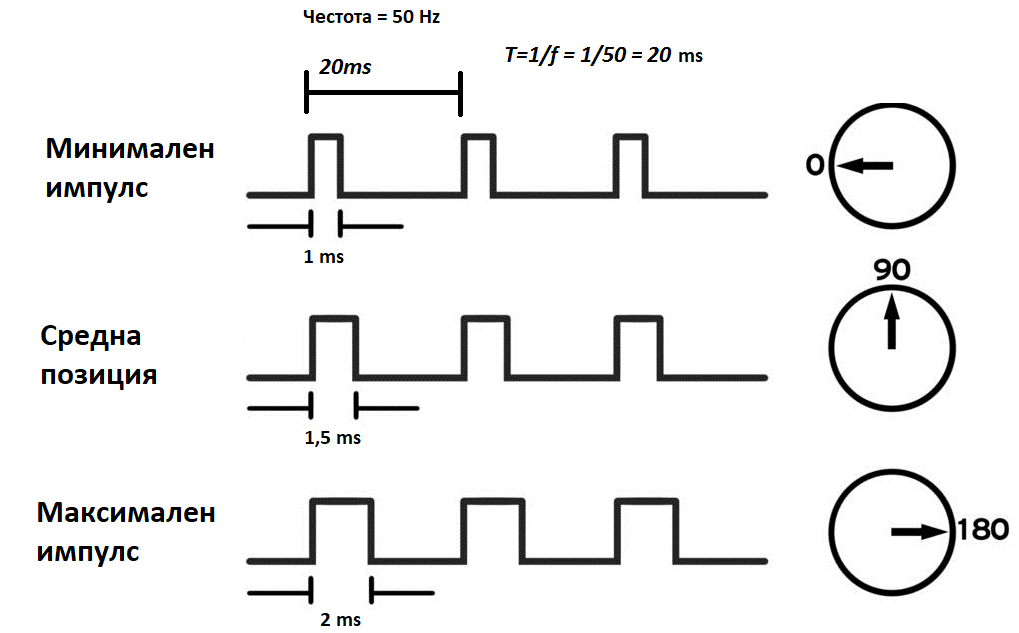


Базовите характеристики на сервото са следните:

* Оперативен волтаж – Между 3 и 6 волта
* Оперативна скорост – От **0,12 сек / 60 градуса** при 4.8 волта захранващо напрежение до **0,1 сек / 60 градуса** при 6.0 волта захранващо напрежение (без прикачен товар)
* Максимален поддържан въртящ момент – Около 1 кг / см
* Ъглов диапазон – 180 градуса
* Импулсна мъртва точка – 7 микросекунди

Сигналът нужен за задаване на ъгъл на ротация е модулиран дигитален сигнал (импулсна модулация) с период от 20 милисекунди като ъгъла се определя от широчината на импулсите в този период. Широчината на импулса може да варира в диапазона 1-2 милисекунди. Съответно ако широчината е 1 милисекунда, ъгъла е минималния за диапазона на въртене. При импулс с широчина 2 милисекунди, ъгъла е максималния за диапазона. При 1,5 милисекунди, зададения ъгъл е в средата на диапазона.

*Фиг. 4 – Форма на управляващия сигнал*

**

Ключово е да се отбележи, че ъгловата резолюция на сервото зависи както от импулсната мъртва точка, така и от предавателния механизъм, заедно в всякакви загуби, които той може да ни представи. С други думи, факта че на теория имаме резолюция от порядъка на части от микросекунди когато боравим със сигнала, това не задължително означава, че такава малка промяна в широчината на импулса ще се отрази във физическо движение на моторната ръка. Всяко серво има характерна „импулсна мъртва точка“, това е минималната времева промяна в импулса на която мотора всъщност отговаря с физическо движение. В нашия случай се използва хоби серво с не много добра отзивчивост. Както е посочено в спецификациите, импулсната мъртва точка е 7 микросекунди в нашия случай. Сигналния диапазон с който работим е 1000 микросекунди и имаме общ ъгъл с големина 180 градуса в този диапазон. Вземайки на предвид и минималната отзивчивост от 7 микросекунди, то след проста сметка можем да намерим частта от целия ротационен диапазон, която отговаря на минималната отзивчивост на сервото:

градуса

За нашите цели, тази ъглова резолюция е напълно достатъчна, макар и не идеална ако целим абсолютно плавно въртене.

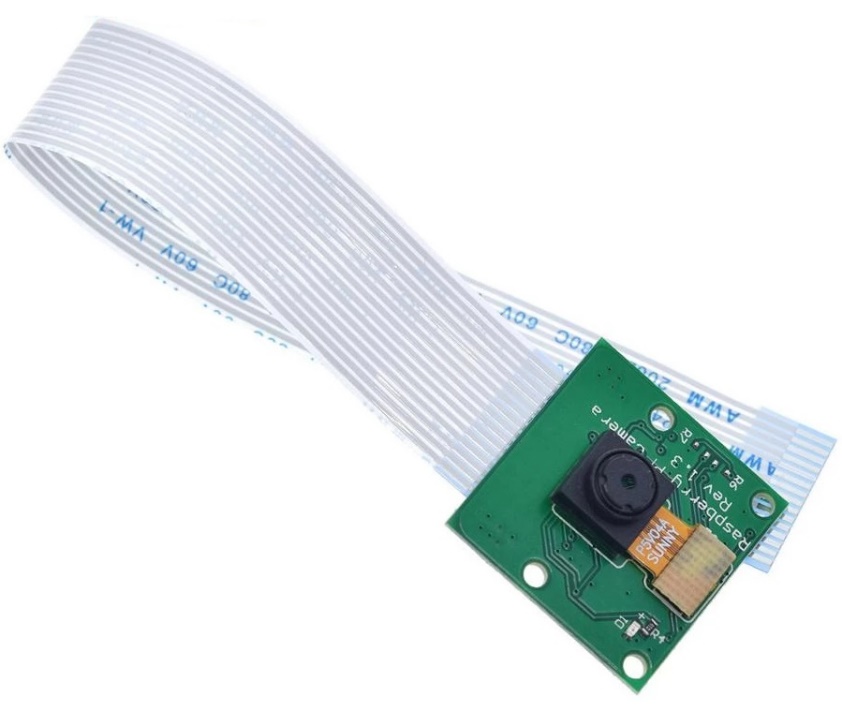
Самият механизъм в цялост се дели на две категории части: **активни** и **пасивни**.

**Активните части** в този случай са камерата прикачена към системата и двете нужни сервота за въртене по двете определени оси.

**Пасивните части** са всички поддържащи и корпусни части, заедно с крепежните елементи за извършване на сглобката.

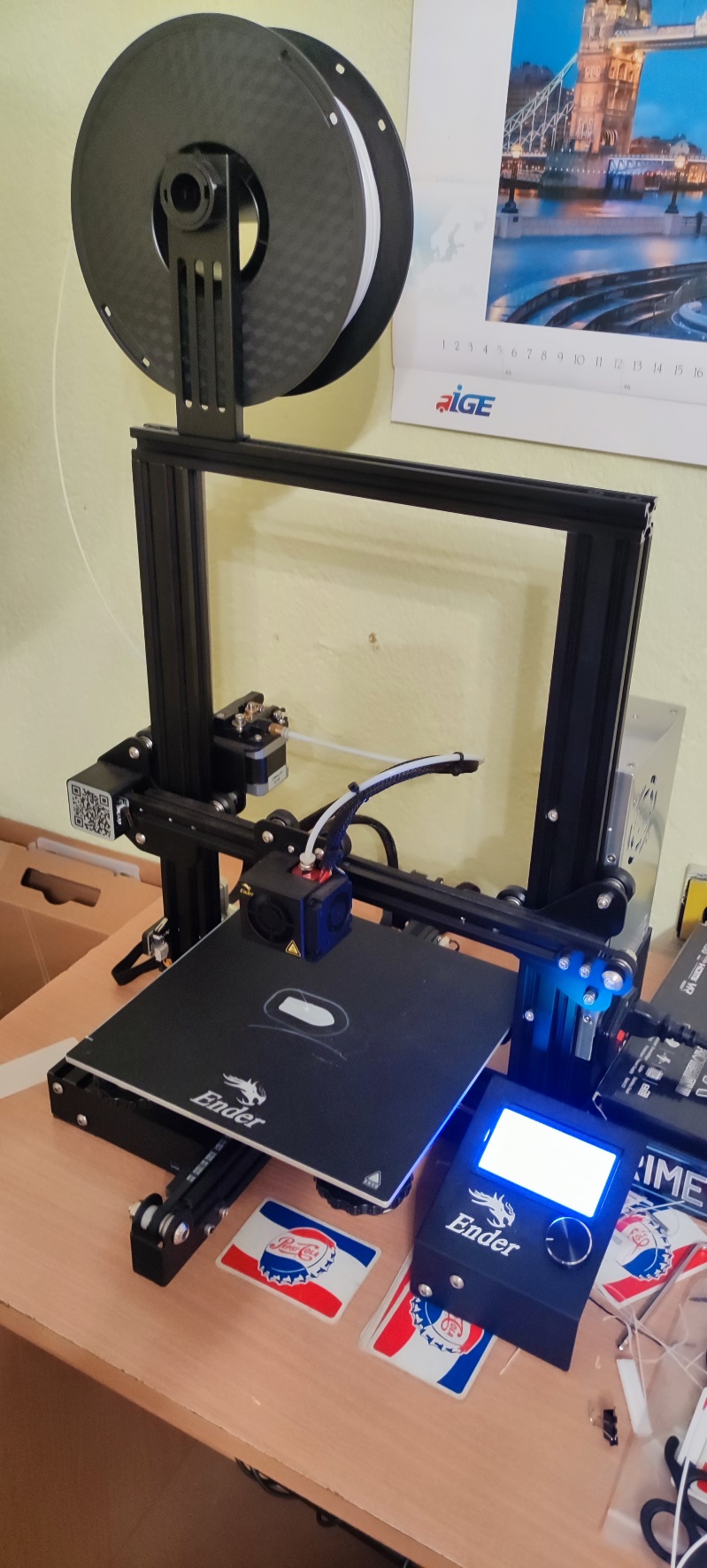
Що се отнася до активните части, вече разгледахме серво моторите. Следва да споменем камерата, в чийто специфики е ненужно да се спускаме в дълбочина. Само е редно да се отбележи, че камерата е с 5 мегапиксела разделителна способност и сензора поддържа автоматично регулиране на експозицията на образа. Това са двете базови изисквания, за по-ефикасна работа на системата. Приемлива разделителна способност, и да не се губят прекалено много детайли при лека промяна в светлинните условия.

*Фиг. 5 – Камерата*



По отношение на пасивните части, предприетия подход за целта на тази задача е 3D принтиране на нужните поддържащи части. Всяка една поддържаща част от механизма е проектирана на CAD програма, след което е принтирана на Creality Ender 3™.

*Фиг. 5 – 3D Принтера*



Механизма се състои от 3 поддържащи части. Следва да обсъдим проектирането им без да се спускаме в излишни подробности.

*Фиг. 6 – Поддържащите части в CAD*

