Заглавна страница

Съдържание

# УВОД

Медът е сладък продукт, който се получава от медоносните пчели от вида *Apis mellifera*. Те могат да го произвеждат като събират и преработват нектара на растенията, или секретите(сладки сокове) отделяни от растенията, или от екстременти на смучещи насекоми по растенията. Медът е един от най – ценните натурални продукти, които човечеството познава и използва от древни времена, както заради високите си хранителни качества, така и за доказаните полезни свойства за здравето на човека. Широко залегнал е в редица традиционни медицини по цял свят.

Пчелният мед се дели основно на три вида: манов, нектарен и смесен. Мáновият се получава от сладките сокове – мàна, които се отделят от растенията под формата на секрети или от дребни насекоми като екстременти. Това разделя мàната на два вида – животинска и растителна. Нектарният мед се произвежда от нектарния сок на растенията. Бива два вида: монофлорен – когато в съдържанието му преобладава нектарен сок от определен вид растение и полифлорен – когато в състава му влизат нектарите на две или повече растения, без ясно изразено водещо съдържание. Смесеният се получава, когато се смесят двата вида мед.

В състава си медът съдържа различни захари, основно фруктоза и глюкоза, както и други вещества като органични киселини, ензими и чужди примеси, получени от добива на мед. Цветът на меда варира от почти безцветен до тъмнокафяв. Консистенцията му може да бъде течна и вискозна или от частично до напълно кристализирал. Вкусът и ароматът варират в зависимост от ботаническия произход на меда.

Имайки предвид сложността и времето необходимо за добиване на качествени натурален пчелен мед, то „натурален“ е термин, с който много често боравят търговците, но реално малко от количеството, предлагано в търговската мрежа, действително отговаря на изискванията за „натурален пчелен мед“. За натурален се смята медът, които не е смесван, не е загряван, няма примеси от други храни и подсладители (глюкоза, фруктоза, инвертна захар глюкозо-фруктозен сироп) и не е получен, чрез преработката на захарен сироп, с който са захранвани пчелите по време на медосбор.

Медоносните пчели са първенците по качество и обем на опрашване на ентомофилни растения, затова съхраняването на традициите в България и по света за производство на качествен, натурален пчелен мед е от изключително важно значение за човечеството и природата. Осигуряването на качествени и обективни методи за анализ и оценка на пчелния мед помага за развитието на производството, разпространението и консумацията на натурален пчелен мед.

# Глава първа Анализиране на качеството и автентичността на пчелния мед

## Европейско и местно законодателство относно пчелния мед

Стандартизацията на пчелните продукти има за цел определянето и регламентирането в стандартизационни документи на качествените изисквания, показателите и нормите, на които те трябва да отговарят, както и на стандартни методи за анализ. Окачествяването на пчелните продукти представлява определяне на степента на съответствие на продуктите към стандартните изисквания и норми, което дава възможност да се правят изводи за техния произход и автентичност, наличието на фалшификати или остатъчни количества от други вещества, здравословността и безопасността им като храна, условията на преработка и съхранение и т.н. (Иванов, 2004).

Съгласно действащата Наредба за изискванията към пчелния мед, предназначен за консумация от човека, приета с постановление №3/06.01.2023г., се определят изискванията към наименованията, състава, характеристиките, етикетирането, вземането на проби и лабораторното изпитване на пчелния мед, предназначен за консумация от човека (България). Пчелният мед, който се предлага на пазара, трябва да отговаря на определени физико-химични норми и изисквания, посочени в таблицата по-долу:

Таблица 1.0‑1

Изисквания към състава и характеристиките на пчелния мед



Съгласно Наредба №9/22.06.2005г. за условията и реда за одобряване и регистрация на предприятията за преработка на восък и производство на восъчни основи, както и на предприятията за производство и търговия с пчелен мед и пчелни продукт и Приложение № 1 към чл. 4, ал. 9 са въведени изисквания към органолептичните и физико-химичните показатели на пчелния мед – някои видове монофлорен (акациев, липов, лавандулов и слънчогледов като най – разпространени в България), полифлорен, манов и смесен пчелен мед. Посочените изисквания Наредбата отговарят на изискванията на Директива 2001/110/ЕО на Съвета от 20 декември, 2001 година относно меда, както и Директива 2014/63/ЕС на Европейския парламент и на съвета от 15 май, 2014. Тези изисквания са в съответствие със стандарта Кодекс алиментариус за меда (Codex Stan 12-1981) (Europe Union).

В Директивата на ЕС и нашата Наредба е регламентирано съществуването и предлагането в търговията на монофлорни видове мед с определен растителен произход, които трябва да притежават "органолептични, физико-химични и микроскопски характеристики на съответния източник". Същевременно, освен посочените изключения за някои видове монофлорен мед, тези характеристики не са специфицирани. За да се отстрани този недостатък, през 1998 г. в рамките на Международната комисия по меда беше създадена работна група с българско участие със задача да разработи качествени критерии за по-важните монофлорни видове мед, произвеждани в Европа. В проучването бяха включени 16 вида монофлорен мед, като бяха обобщени и обработени математически литературните данни за техните органолептични свойства, микроскопска характеристика (поленов спектър) и химичен състав. Обработени са данните за 5 000 проби мед от 20 страни по редица показатели (общо над 60 000 анализа). За всеки вид мед е направена пълна характеристика, която може да се използва при определянето на растителния произход на меда. Необходимо е да се отбележи, че като метод за определяне на растителния произход на меда в БДС е включен поленовият анализ. За да се приеме меда за монофлорен /от определено растение/, поленът от това растение трябва да е не по-малко от 40% от общата количество полен. За акациевия и липовия мед долната граница е 30%, а за лавандуловия - 15%. (Иванов, 2004)

## Медоносни растения в България

Медоносни растения са култури, от които пчелите събират цветен нектар и цветен прашец. Има различни видове класификации на медоносните растения и в зависимост от това могат да се групират в различни категории. В зависимост от времето на цъфтеж различаваме основно: пролетни, летни и есенни. Някои растения цъфтят през повече от един сезон, дори и целогодишно. Според вида паша разграничаваме: растения, отдаващи само цветен прашец, растения отдаващи цветен нектар и цветен прашец и растения отдаващи само цветен нектар. По местоположение растенията се разделят на: медоносни горски дървета, медоносни горски храсти, медоносни горски тревни и храстовидни растения, ливадни медоносни растения, полски медоносни растения, всестранно разпространени медоносни растения и специални медоносни растения. Несериозно е да се определи точният брой на медоносните растения в България, но със сигурност наброява няколко хиляди вида. Има, обаче, растения, които дават най-високи добиви на мед (кг/дка) и те се считат за основни медоносни растения. В най-голям брой публикации е потвърдено доминиращото значение и продуктивност на: *Robinia pseudoacacia L. (Бяла акация), Amorpha fruticosa L. (Храстовидна аморфа, Черна акция), Phacelia tanacetifolia Benth. (Калифорнийски див зюмбюл), Cotoneaster integerrimus Medicus. (Целокраен котонеастер), Cornus mas L. (Обикновен дрян), Corylus avellana L. (Обикновена леска), Tilia tomentosa Moench. (Сребролистна липа), Acer campestre L. (Полски клен), A. platanoides L. (Шестил), Crataegus monogyna Jacq. (Глог), Galanthus elwesii Hook. (Елвезиево кокиче), Primula veris L. (Иглика), Anemone ranunculoides L. (Лютиковидна съсънка), Viola odorata L. (Теменужка), Paeonia peregrinа Mill. (Червен божур), Fritillaria pontica Wahl. (Черноморска ведрица), Pulmonaria officinalis L. (Медуница), Geranium macrorrhizum L. (Здравец),* и от видовете от родовете *Thymus (Мащерка), Teucrium (Поддъбиче), Salvia (Салвия), Sedum (Седум), Marrubium (Пчелинок), Ajuga (Срещниче), Еchium (Усойниче), Echinops (Кръглоглав челядник), Carduus (Магарешки бодил), Genista (Жълтуга) и Lathyrus (Секирче).* В България, едни от най-високодобивните видове са разпространени повсеместно като: *Robinia pseudoacacia L.(Бяла акация)*, видовете от род *Tilia(Липи), Trifolium repens L (Детелина)*., *Vicia sativa L*. *(Пролетен фий),* *V. villosa Roth (Вълнеста глушина)*, *Fagus sylvatica L*. *(Бук)*, видовете от род *Quercus (Летен дъб), Erodium cicutarium (L.)* *L'Her. (Цикутово часовниче)*, *Juglans regia L*. *(Орех)*, видовете от род *Lamium (Мъртва коприва), Mentha (Мента), Plantago (Теснолистен живовляк), Origanum vulgare L. (Риган), Salvia verticillata L. (Прешленест конски босилек), Teucrium chamaedrys L. (Червен поддъбиче), Clematis vitalba L. (Обикновен повет), Crataegus monogyna Janq. (Глог), Fragaria vesca L. (Горска ягода), Prunus avium L. (Череша)* и др.. (ВЕЛИНОВА, 2018).

Таблица 1‑2

|  |  |  |  |
| --- | --- | --- | --- |
| бяла акация  добив: 11…30кг/дка | дипа  добив: 10…32кг/дка | дрян  съдържание захар в ha: 36 кг/hа | глог  съдържание захар в ha: 16 кг/hа |
| мента  добив: 20кг/дка | слънчоглед  добив: 3.5…6кг/дка | лавандула  добив: 14кг/дка | люцерна  добив: 8…10кг/дка |
| салвия  добив: 5…34кг/дка | череша  добив: 3…16кг/дка | ябълка  добив: 1…10кг/дка | вишна  добив:1.4…18кг/дка |

Предвид огромното разнообразие на медоносни растения става ясно и факта, че 100% монофлорен мед е почти невъзможно да се получи. Някои от основните видове монофлорен мед, налични на пазара в България са: акациев, липов, лавандулов, рапицов, слънчогледов, от магарешки бодил и др. За да бъде признат за монофлорен даден мед, той трябва да съдържа минимум 40% полени от растението, за което се сертифицира, като при акацията и липата (30%), както и при лавандулата (15%) се допускат намалени количества. Монофлорният мед е по – скъп и като такъв, неговият произход трябва да бъде лабораторно доказан, чрез някой от методите за анализ на ботаническия произход на пчелния мед.

## Видове анализи за изследване на ботаническия произход на меда

Пчелният мед, произвеждан на територията на България е напълно достатъчен за задоволяване на вътрешните ни нужди. Дори нещо повече – голяма част от продукцията се изнася към Европейския съюз и извън него. По – голямата част от пчелния мед, произвеждан в България е монофлорен (например от рапица, слънчоглед, магарешки бодил, липа, акация или кориандър) (Dinkov, 2015).В сегашните пазарни условия клиентите искат сигурност, обективност и надеждност относно истинността на ботаническия произход, които те консумират. Това поставя учените, изследователите и лабораториите пред предизвикателство – намирането на подходящите методи за идентификация, позволяващи надеждно разграничаване на фалшификати и некачествен мед.

Традиционно, ботаническия произход на монофлорния мед се доказва чрез мелисопалинологичен анализ – съдържанието на полени се определя чрез микроскопско изследване, което е скъпо, времеемко и изисква обучен персонал. Съществуват и други методи, използвани за доказване на ботаническия произход на пчелния мед.

### Поленов анализ

Широко използван метод за определяне на ботаническия произход на пчелния мед. Методът включва микроскопско изследване на пчелен мед и идентифициране на поленовите зърна открити в него. Всеки вид растение има уникална морфология на полените, което се използва за определяне на произхода на меда. Поленовите зърна се идентифицират и броят от опитен лаборант-специалист. Специалистът сравнява наблюдаваните поленови зърна с референтна колекция от поленови зърна с ясен, определен произход и определя относителното им количество в изследваната мостра. Веднъж определил и преброил поленовите зърна в дадена мостра, лаборантът може да определи ботаническия произход и вида на пчелния мед.

Поленовия анализ има няколко преимущества като метод за определяне на ботаническия произход на пчелен мед. Характеризира се с висока точност и надеждност, поради уникалната морфология на полените и способността да се разграничат едни от други. Методът сам по себе си не е деструктивен и мострата мед може да бъде използвана за храна отново впоследствие. Като недостатък може да се посочи необходимостта от наличие на скъпа апаратура (микроскоп), както и обучен кадър с експертиза в областта. Изисква доста време и труд. Методът не е приложим за мед, подложен на загряване или ултрафилтриране, тъй като тези процеси могат да премахнат или унищожат полените.

### Ядрено-магнитен резонанс (ЯМР)

ЯМР спектроскопия е техника, която е често използвана за определяне на ботаническия произход на меда. Пчелният мед съдържа различни химични компоненти, включително аминокиселини, органични киселини и захари, които са специфични за различните източници на цветен нектар. ЯМР спектроскопията може да бъде използвана да се за откриване на тези специфични компоненти и определянето на ботаническия произход на меда.

ЯМР спектроскопията е базирана на принципа на ядрено-магнитния резонанс, което е свойство на атомните ядра, поставени в магнитно поле. Когато тестваната проба е поставена в магнитно поле и облъчена с радио-вълни, атомните ядра от пробата абсорбират енергия и повторно излъчват като резонантен сигнал. Честотата и интензитета на резонантния сигнал са характеристика на атомните ядра в пробата и могат да бъдат използвани за идентифициране на химичния състав на меда.

Спектралната характеристика на пчелен мед, най – често съдържа пикове кореспондиращи на различните химични съставки като захари, органични киселини и аминокиселини. Преходът от един пик към друг, както и техните интензитети, са специфични за всеки източник на цветен нектар (вид растение), тъй като различните видове растения произвеждат различни по вид и количество химични компоненти.

Веднъж получени, тези спектрални характеристики могат да бъдат анализирани, използвайки разнообразни статистически методи, като „Метод на главните компоненти“ или „Линеен дискриминантен анализ“, за да се определи ботаническия произход на пчелния мед. Тези методи сравняват спектралната характеристика на тествания мед с база данни от референтни спектрални характеристики на различни видове монофлорен мед и търсят най-доброто съвпадение.

Като преимущества на ЯМР спектроскопия могат да се посочат високата чувствителност и точност, тъй като могат да се засекат ниски концентрации на различни химични компоненти в меда. Недеструктивен метод е и пробата може да бъде използвана отново след анализа. Като недостатък може да се изтъкне необходимостта от силно специализирано оборудване и обучен персонал с експертиза в областта за извършване на анализа.

### Инфрачервена спектроскопия с Фурие трансформация

Инфрачервената спектроскопия с Фурие трансформация е мощна аналитична техника, която често се използва за определянето на ботаническия произход на пчелен мед. Този метод е базиран на принципа, че различните източници на цветен прашец, съдържат уникални химични компоненти, които могат да бъдат идентифицирани по абсорбционен спектър в инфрачервената област на електромагнитния спектър. Извършва се като малка проба мед се поставя на статив и после се облъчва с източник на инфрачервена светлина. Пробата абсорбира част от светлината и останалата светлина е приета от детектор. Резултатния спектър се нарича Инфрачервен спектър с трансформация на Фурие и съдържа информация относно химическия състав на меда. Всяко вещество в меда абсорбира светлина с различна честота, което се отразява като характерни пикове в получения Инфрачервен спектър с трансформация на Фурие(ИСТФ). Тези пикове могат да бъдат използвани за идентификация на различните вещества в меда и да определят неговия ботаничен произход. Получените спектрални характеристики се анализират с различни статистически методи и получените резултати се сравняват с база данни от референтни спектрални характеристики, търсейки най – добро попадение. Например медът от лавандула ще съдържа висока концентрация на линалол, което може да бъде идентифицирано от характерен пик в ИСТФ. Като преимущество на метода може да се посочи бързината на метода и високата му чувствителност. Не е деструктивен, изисква малко количество проба мед за изследване и не уврежда пробата. Може да бъде използван за определяне на фалшификати или замърсители в меда, което е важно за подсигуряването на автентичността и качеството на продуктите. Като недостатък може да се посочи отново необходимостта от специализирано оборудване и персонал.

### Органолептичен (сензорен) анализ

Органолептичният анализ е метод, използващ оценката на вкусът, ароматът, текстурата, консистенцията и външния вид на пчелния мед, като това дава информация за качеството и ботаническия произход на меда. Някои видове мед(акациев, липов, кестенов и др.) имат характерни органолептични показатели. Този тип анализи се извършват от опитни дегустатори, които са способно да окачествят пробите използвайки сетивата си. Дегустаторите опитват вкуса на пчелния мед за сладост, киселост, горчивина и други вкусови показатели. Вкусът на меда силно се влияе от неговия ботанически произход, тъй като различните растения произвеждат различни видове нектар с уникални вкусови качества. Ароматът и цветът също са силно повлияни от източникът на цветния нектар, от които е произведен меда. Текстурата се оценя на база на вискозитета, кристализацията и други физични показатели.

Органолептичният анализ се може да бъде дескриптивен и дискриминантен. Дескриптивните методи използват дегустатори, които трябва да опишат органолептичните показатели на меда, използвайки стандартизирана терминология. Дискриминантните методи използват специалисти, които трябва да разграничат различните проби мед и да определят техния ботанически произход.

Сензорният анализ е ценен инструмент при оценката на качеството и произхода на меда, предоставяйки информация за характеристиките вкус, аромат и текстура, които иначе трудно биха могли да бъдат измерени с технически средства. Основният недостатък на метода е неговата субективност. Резултатите могат да бъдат повлияни от лични предпочитания, моментно състояние и др. на дегустатора.

### Физико-химичен анализ

Физико-химичният анализ на пчелния мед включва набор от изследвания за да се определят физичните и химичните характеристики на меда. Тези методи могат да предоставят информация за качеството и автентичността на меда, както и за неговите географски и ботанически произход. Най – важните и информативни физико-химични показатели са:

1. Влага – съдържанието на вода в меда е важен параметър по отношение на качеството на продукта, тъй като високото съдържание на вода може да доведе до ферментация и разваляне на меда. Максимално допустимото ниво на влага в меда е 20%.
2. Водородният показател pH – pH-то на меда е обикновено в границите от 3.4 до 6.1 и дава информация за киселинността на пчелния мед. Стойностите на pH на пчелния мед се влияят до голяма степен от източника на цветен нектар, но и от други фактори като температура и условия на съхранение на меда.
3. Електрическа проводимост – електрическата проводимост на меда е свързана с минералното съдържание на меда и може да даде информация за географския му произход.
4. Общо захарно съдържание – общото съдържание на захари се определя с измерване на съдържанието на глюкоза и фруктоза, които се влияят от източника на цветен нектар и други фактори (например допълнително изхранване на пчелите и вида храна, с която се извършва то).
5. Хидроксиметилфурфурал (HMF) – ХМФ-то е вторичен продукт, получаващ се при загряване на пчелния мед или продължителното му съхранение. Количеството му в пробата може да достави информация относно свежестта и качеството на меда.
6. Цвят – Цветът на пчелния мед силно варира в зависимост от вида на цветния нектар, географския регион, времето за което е бил съхраняван меда и др.

Физико-химичните анализи на меда могат да се извършат, използвайки разнообразни методи като рефрактометрия, pH измерване, измерване на електрическа проводимост, колориметрия. Тези методи обикновено са недеструктивни и изискват малко количество продукт за проба. Общо казано физико-химичните методи за оценка на меда са полезен инструмент за осигуряването на качеството и автентичността на меда, както и неговия ботанически и географски произход.

## Компютърно зрение

Компютърното зрение е дял от компютърните науки, които се причислява към сферата на изкуствения интелект и позволява на компютрите да добиват смислена и полезна информация от дигитални изображения, снимки и други визуални източници и да предприемат действия или да правят препоръки и становища, базирани на входящата информация. Ако може да се каже, че изкуствения интелект кара компютрите да „мислят“, то компютърното зрение им позволява да „гледат, наблюдават и разбират“ заобикалящият ги свят.

Системата за компютърно зрение може да се определи като: съвкупност от специализирани устройства за формиране на изображение на околната среда, компютърна система и конкретно програмно осигуряване. Програмното осигуряване се разработва на основата на методи и алгоритми за обработка и анализ на изображението, обучение и разпознаване на обектите от визуалната сцена и изграждането на модел за околната среда. Проблемът за вътрешното представяне и съхраняване на информацията изисква колкото се може повече тя да бъде намалена в сравнение с входното изображение. За тази цел се търсят характерни свойства(признаци), на основата на които адекватно да могат да се опишат обектите. Анализът и разпознаването се правят на основата на съхраняваните описания. (Гочев, 1995).

Учените и инженерите отдавна се опитват да създават нови начини машините да „виждат“ и „разбират“ видима информация от около 65г. Експериментите започват през 1959г., когато неврофизиолози показват на котка група изображения, опитвайки се да направят взаимовръзка между изображенията и реакцията на мозъка на котката (електрически импулси). Те открили, че мозъкът първо реагира на силно изразените ръбове, ъгли и линии, което научно доказва, че обработката на изображенията започва с прости форми като прави ъгли и др.

Почти по същото време, първата технология за сканиране на изображения била изобретена, позволявайки на компютрите да дигитализират и запазват дигитални изображения. Друг крайъгълен камък е достигнат през 1963г., когато компютрите вече са били способни да трансформират двуизмерни изображения в триизмерни форми. През 1960г., Изкуствения интелект се появява като академична дисциплина, което също маркира началото на надпреварата в сферата на компютърното зрение.

През 2000г. на фокус за научни изследвания е разпознаването на обекти, а първата компютърна програма за лицево разпознаване в реално време се появява през 2001г. През 2010г. е представена и става достъпна базата данни ImageNet - <https://www.image-net.org/>. Съдържа са качени милиони изображения на хиляди обекти класове и дава основата за обучение и развитие на невронните мрежи и методи за „дълбоко обучение“ налични днес.

Няколко примера за решени ежедневни проблеми, с помощта на компютърното зрение са:

* система класификацията и идентификация на обекти. Системите за разпознаване на обекти виждат изображение и могат да го класифицират, стига да са обучени за целта. Намира приложение в изключително широк спектър от области – здравеопазване, охрана, индустриална автоматизация, медицина, опазване на обществения ред, туризъм и много други;
* система засичане на обекти. Използва класификацията на обекти за да „засече“, намери присъствието на даден обект в изображение и после да предприеме действие или друго, в зависимост от логиката и структурата системата. Например откриването на повреди на линия за сглобяване, или откриването на гнили ябълки, на поточна линия, или откриването на залепен накриво или наобратно етикет на бутилка и др.;
* система за следене на обекти. Веднъж засечен даден обект, той започва да бъде проследяван. Това най – често намира приложение в поредица от заснети снимки в кратък период или видео в реално време. Използва се, например, в автономните превозни средства (следване на линия, засичане на пешеходци или други превозни средства, които се движат);

## Видове алгоритми за разпознаване на изображения

В машинното обучение, компютърното зрение и разпознаването на образи, терминът *признак* е индивидуална, измерима характеристика или феномен. Избирането на информативни, описателни и независими признаци е от фундаментално значение за ефективността и правилната работа на алгоритмите. Признаците обикновено са числени или бинарни (двоични), но може да са и структурни като стрингове или графи. Концепцията признак може още да се нарече и информативна променлива, използвана в статистическите техники (Wikipedia).

Синтезът на признаци е една от най-важните задачи при разработването на разпознаващи системи. Формални правила за формиране на признаци не съществуват, затова в много случаи процесът е основан на знанията, уменията и интуицията на специалиста в дадената област. Основните условия, на които е необходимо да отговарят признаците са да бъдат информативни и да има възможност за формално описание, т.е. да се изразяват количествено. Процесът на формиране на признаково пространство има за цел получаването на подходящо представяне на обектите, от което съществено зависи последващия процес на синтез на класификатор и неговата ефективност. Този процес обикновено се разделя на две части:

- Синтез на признаци *(“Feature Extraction”*). Редуциране на размерността чрез линейна или нелинейна трансформация от  -мерно пространство в  размерен вектор, като .

- Селекция на най-информативните от тях (*“Feature Selection”*). Допълнително редуциране на размерността чрез селекция на подмножество от информативни признаци. (Начев)

Избирането на значими, информативни признаци е от голямо значение за работоспособността и верификацията на резултатите. Повечето признаци не винаги водят до по – добри резултати (overfitting).

Както вече стана ясно, разпознаването на изображения и класификация на обекти стои в основата на компютърното зрение. Класифицирането на изображения се дели на два вида: супервайзорно(с учител) и несупервайзорно(без учител).

Несупервайзорното обучение за класификация на изображения е напълно автоматизиран метод, който не разчита на обучаващи данни. Това означава, че този метод за машинно обучение се използва за анализ и клъстеризация на данни в т.нар. *суров* вид. При този тип данни няма предварително пояснение като клас на принадлежност, свойство експертна оценка. Този тип алгоритми трябва да открият скрити шаблони или да групират данните в отделни групи, без необходимостта от човешка намеса.

Супервайзорното обучение за класификация на изображения използва предварително класифицирани данни за да обучи класификатора и впоследствие да има възможността да класифицира нови, непознати за него данни към един от предварително определените класове. Този тип обучение може да разчита на разстояние между отделните обекти за класификация (Евклидово разстояние, Манхатъново разстояние), на предварително избрани признаци, на предварително зададени условия на класифициране – брой групи(клъстери), сложност на разделителния инструмент (линия, равнина) и др.

В днешно време с развитието на компютърната наука и технологии са известни много видове подходи и алгоритми при разпознаването и класификацията на изображения. Някои по – известни са:

* Метод на главните компоненти(МГК) е популярна техника за анализиране на големи бази данни, с висока размерност или признаково простра, която повишава тълкуваемостта на данните, запазвайки максимално количество информация и позволявайки визуализацията на многоизмерни данни. Накратко МГК е статистическа техника за редукция на размерността съвкупност от данни. Това се постига, чрез линейната трансформация на данните в нова координатна система, където вариацията в данните може да се опише с по-малко размерни данни, сравнени с първоначалните. Много изследвания използват първите две компоненти, за визуализация на данни в две измерения и визуално идентифицират клъстери в двуизмерното пространство. МГК се използва в множество сфери като генетика, микробиомни изследвания, разпознаване на образи, изследвания на атмосферата и др. (Wikipedia)
* Метод на опорните вектори (SVM, Support Vector Machines) представя обучаващите примери като точки в n-мерно пространство. Примерите са проектират в пространството по такъв начин, че да бъдат линейно разделими. При работа с два класа се търси начин да се начертае линия, която да разделя данните от двата класа. Линията, която разделя данните, се нарича разделителна хиперравнина. Тази хиперравнина трябва да се избере по такъв начин, че да се намира възможно най-далеч от примерите и на двата класа. (Технически университет гр.Варна)
* Метод на К най-близкия съсед за класификация на обекти, основан на оценка на сходство на обекти. Класифицираният обект се отнася към този клас, към който принадлежат най-близките до него обекти на обучаващата извадка (КNN - k nearest neighbours). Алгоритъмът на работа е следния: 1.Определяне на К броя обекти, които ще участват в класификацията; 2.Изчисляване на дистанцията между всеки два обекта от обучаващата извадка, чрез използване на подходяща функция за измерване на разстояние между две точки;   
  3. Избор на К обекти от обучаващата извадка, разстоянието до което е минимално; 4.Класа на класифицирания обект – това е класа на новия обект, намерен въз основа на заключенията, направени по отношение на К най-близки съседи (Технически университет гр.Варна)
* Изкуствени невронни мрежи - симулация на невронна мрежа, изкуствена невронна мрежа, разговорно наричана само невронна мрежа, е изчислителен модел за обработка на информация, вдъхновен от изучаването на биоелектричните мрежи в мозъка на човека и животните, образувани от неврони и техните синапси. Математическият аналог на биологичната невронна мрежа представлява множество от взаимосвързани прости изчислителни елементи (възли или изкуствени неврони). Всеки от тях приема сигнали от другите (под формата на числа), сумира ги, като сумата минава през активационна функция, и така определя своята активация (степен на възбуда), която се предава по изходящите връзки към другите неврони. Всяка връзка има тегло, което, умножавайки се със сигнала, определя неговата значимост (сила). Теглата на връзките са аналогични на силата на синаптичните импулси, предавани между биологичните неврони. Отрицателна стойност на теглото съответства на потискащ импулс, а положителната – на възбуждащ (Wikipedia)

## Интелигентни класификатори на полени заснети на микроскопски снимки

Има редица разработки на учени по цял свят, целящи синтез на интелигентни системи, разпознаващи с висока точност различни видове полени. Различните разработки използват различни признаци за разпознаване на полените – геометрия, морфология, контури, текстура, цвят или комбинации от няколко, използвайки различни подходи при класификацията.

Група учени – Amar Daood, Eraldo Ribeiro и Mark Bush от Флоридският институт по технологии работят над метод за разпознаване на микроскопски изображения на полени, използвайки многослойна признакова декомпозиция. Първо нормализират изображенията, после декомпозират изображението на няколко слоя, използвайки клъстеризация, извличайки признаци от всеки слой. За класификация използват Методът на опорните вектори (SVM), когато се касае за бинарна класификация. Когато има повече от два вида полени за разпознаване използват техниката ECOC (error-correcting output code), което разширява възможностите на Методът на опорните вектори, позволявайки му да класифицира множество класове полени (30 класа полени). Резултатът, получен в този научен труд дава 86.94% успеваемост. (Amar Daood)

Celeste Chudyk и Frank Boochs от Университета по приложни науки, Майнц, Германия и Hugo Castaneda, Romain Leger и Islem Yahiaoui от Дижонският технологичен институт, Бургундски университет, Франция разработват система за автоматична класификация на полени. Съставят си база данни от около 600бр. микроскопски снимки на 5 различни вида полени. Извличат признаци за формата като периметър, площ, овалност, диаметър и др. и признаци на текстурата на полените, използвайки похвати като Габор филтри, трансформации на Фурие и др. Получените признаци използват за класификация, използвайки метод на опорните вектори и метод на случайни дървета за решение(random decision forests не знам как му е точното наименование на български) като класификатори. Методът дава успеваемост от 87% ±2% (Chudyk, et al.).

Димитър Николов и проф. Диана Цанкова от Университета по Хранителни технологии, катедра Автоматика, информационна и управляваща техника правят разработка на тема „Извличане на признаци за разпознаване на полени в пчелен мед използвайки филтри на Габор“. Те събират база данни от изображения на няколко вида полени и обработват изображенията с цел нормализацията им. Нормализираните изображения биват филтрирани с филтри на Габор за получаването на характерни признаци. Размерността на получените данни бива намалена, чрез използването на метода на главните компоненти. Същинският класификатор се синтезира на принципа на Линейния дискриминантен анализ(LDA). Успеваемостта на метода е 90.32%. (Diana Tsankova, 2018)

## Цел на дипломната работа и поставени задачи

Целта на дипломната работа е да се разработят алгоритми за класификация на полени на медоносни растения, като се използват фотоизображения от оптичен микроскоп и хемометрия върху предварително обработени изображения върху предварително обработени изображения.

В съответствие с поставената цел се решават следните задачи:

1. Събиране на база данни, включваща фотоизображения на микроскопски слайдове на полени на медоносни растения, широко разпространени в България;
2. Разработване на алгоритми и програми за предварителна обработка на микроизображения на базата на: филтър на Габор, матрица за съвместна поява на нива на сивото (GLCM), изчисляване на размери и форма на полените;
3. Разработване на алгоритми и програми за класификация на полени на базата на хемометрия (Линеен дискриминантен анализ, нелинеен дискриминантен анализ, к-най-близки съседи) върху предварително обработени изображения;
4. Генериране на подобрения на разглежданите класификатори чрез създаване на хибридни структури;
5. Разработване на симулационна среда (на езика MATLAB) за тестване на разглежданите алгоритми(базови и подобрени варианти);
6. Сравнителен анализ на предложените класификатори;

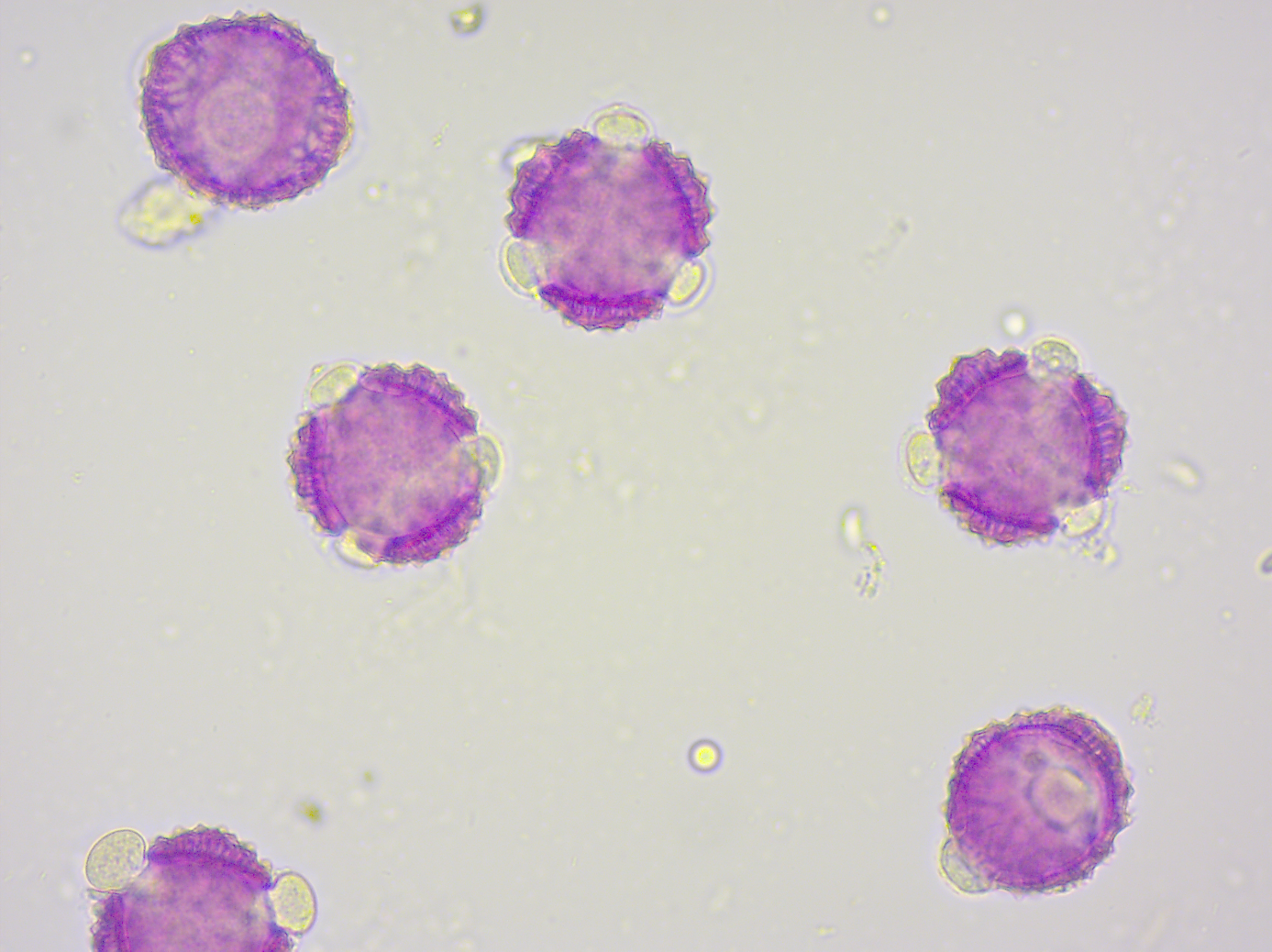
# Глава втора методи и алгоритми за събиране и обработка на изображенията

## 2.1. Събиране на база данни от фотоизображения

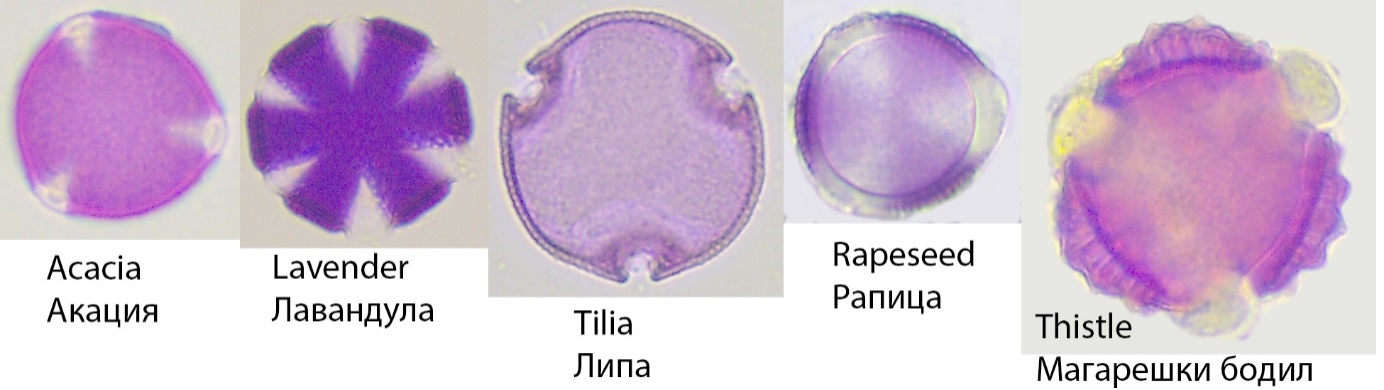
В глава I са споменати някои от основните видове медоносни растения. За съставянето на база данни, включваща фотоизображения на микроскопски слайдове на полени на медоносни растения, широко разпространени в България се използват следните пет вида растения: акация, липа, рапица, магарешки бодил и лавандула.

За преброяването на поленовите зърна, съгласно БДС 3050-80, е необходимо предварително да се подготвят еталонни препарати от цветен прашец на различни медоносни растения. Това става, като се вземе прашец от съответното растение и се постави на предметно стъкло. Намокря се с капка дестилирана вода и след подсушаване се фиксира с капка спирт, оцветен с фуксин ( за да се придаде цвят). След това препаратът се покрива с капка подгрят глицерин-желатин и се поставя внимателно покривно стъкло. След 3-4 дни краищата на покривното стъкло се уплътняват с канадски балсам. Препаратът е траен дълго време. Глицерин-желатиновата смес се подготвя като 7g желатин се разбъркват в продължение на два часа в 42 cm3 дестилирана вода. Прибавят се 50g глицерин и 0.5g фенол(карболова киселина). Нагрява се 15 min и се филтрува на топло. Преди употреба се нагрява до втечняване (Динков, 2013). Описаната методология на работа се използва и при подготовката на базата данни от изображения в настоящата дипломна работа.

Използва се оптичен микроскоп Carl Zeiss Primo Star, с обектив с четиридесеткратно увеличение. За заснемането на фотоизображенията се използва професионална дигитална камера за микроскоп Zeiss AxioCam ERc 5s. Изображенията са с резолюция 2560х1920 пиксела като размерът на всеки пиксел е 0.110µm x 0.110 µm. Получените изображения могат да съдържат един или няколко полена, в зависимост от това колко на брой попадат в заснетия участък.  
  
От така получените фотоизображения се изрязват отделните полени, така че да се получи индивидуално фотоизображение, съдържащо само един полен.



Изображение 1. Магарешки бодил



Изображение 2. Акация, лавандула, липа, рапица и магарешки бодил

На изображение 2. са показани по един пример от петте вида, от които е съставена базата данни с полени. При обработката на снимките, те не са преоразмерявани за да се запазят в реални размери. Вижда се, че са с различна форма и големина. Цветът на повечето е с близки нюанси, което се дължи на оцветяването от фуксина. Съставена е база данни, съдържаща по 35бр. изображения от всеки от петте избрани вида полени. Всички изображения са отрязани и обработени по един и същи начин за максимална обективност.

## 2.2. Алгоритми и програми за предварителна обработка на изображенията

### 2.2.1. Филтър на Габор

Филтрите на Габор са широко използвани в сферата на обработка на снимки и компютърно зрение за задачи като засичане на контури, анализ на текстура на изображение и извличане на признаци. По същество представлява линеен филтър, който анализира наличието на някаква определена честота в дадено изображение, насочена в определена посока в конкретен регион или точка от изображението. Намирането на определена честота в изображение и неговата ориентация наподобява зрителната система на бозайниците, според някои учени. (Bruno A. Olshausen, 1996).

Филтърът на Габор притежава две важни характеристики: способност за пространствена локализация и честотна селективност (избираемост). Способността за пространствена локализация дава възможност на филтъра да извлича признаци само от определени места от изображението, а честотната селективност съдържа в себе си възможността да извлича информация за специфични честоти. Това е и причината, когато се извличат признаци от фотоизображение, чрез филтър на Габор, да се използва набор от филтри с различни честоти и ориентации. Математическото описание на филтъра на Габор е следното:

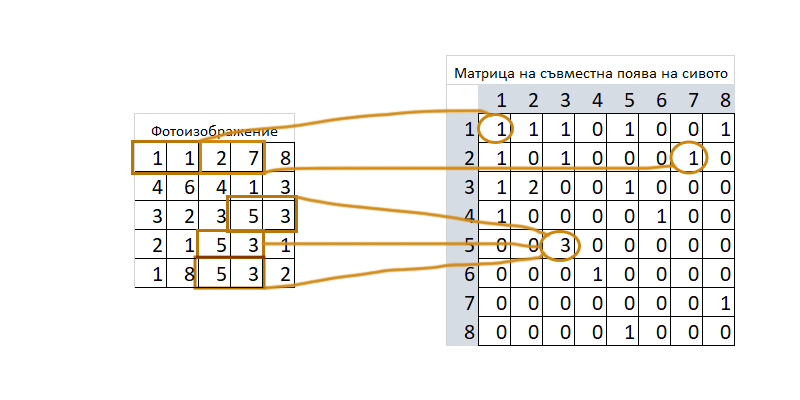
където е *g(x, y)* е Гаусовата функция, определяща се по следния начин:

*λ* е дължината на вълната на филтъра на Габор, *θ* е ъгълът на ориентацията на филтъра на Габор, *σx, σy* е стандартното отклонение по *x* и *y* координатите. Ако приемем, че *σx=σy=σ,* се получава следното:

Тук може би трябва да пиша още, но ми е малко объркано

### 2.2.2. Матрица за съвместна поява на сивото (Gray level co-occurrence matrix)

През 1973 Робърт Харалик представя матрицата за съвместна поява на сивото. Тази матрица се създава, изчислявайки колко често пиксел (с координати x, y) с интензитет на сивото със стойност *i,* се появява на предварително определено разстояние (*Δx, Δy*) от пиксел (с координати *x + Δx, y + Δy*) с интензитет на сивото със стойност *j.* Това определено разстояние (*Δx, Δy*) се нарича пространствена взаимовръзка между два пиксела. Всеки елемент *i, j* в така получената матрица за съвместна поява на сивото е сумата от броят пъти, които пиксел с интензитет на сивото *i,* се появява на предварително определено разстояние (*Δx, Δy*) до пиксел с интензитет на сивото със стойност *j.* Броят на нивата на сивото във монохромното фотоизображение определя и размера на матрицата, например при 8 нива на сивото, размерът на матрицата е 8x8. На изображение 3. е илюстриран пример за това как се конструира матрицата. Математическото уравнение на матрица за съвместна поява на сивото е :



Изображение 3. Конструкция на матрица за съвместна поява на сивото

където:

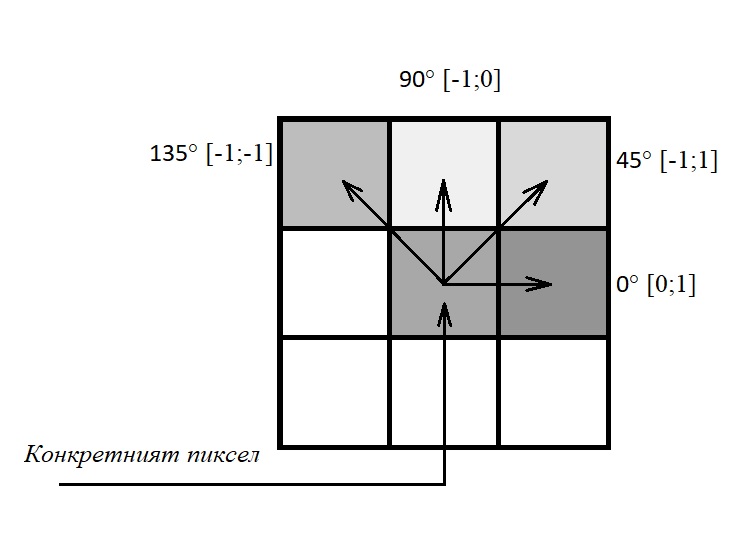
*i, j* – координатите (пикселите) в матрицата за съвместна поява на сивото

*x, y* – са координатите (пикселите) в изображението I, с размери (n x m)

*Δx, Δy* – избраната пространствена взаимовръзка между два пиксела

– стойността на пиксела с координати (x, y)

*I(x + Δx, y + Δy)* – стойността на пиксела с координати (x + *Δx*, y + *Δy*)

**Матрицата за съвместна поява на сивото се използва често за анализ на текстурата на изображения и последваща им класификация. Анализът на текстурата на дадено изображение по същество трябва да е инвариантен по отношение на ротацията на изображението. Заради това е обичайна практика е да се използват 4 броя пространствени взаимовръзки [0 1;-1 1; -1 0; -1 -1], които да сравнят конкретния пиксел на позиция (*x, y*) и по отделно следните: съседния отдясно на 0°, съседния отдясно и отгоре по диагонал­­­­­­­­­­­­­­­­­­­­­­­­­­ на 45°, съседния отгоре 90° и съседния отляво и отгоре по диагонал 135°. След изчислението се получават 4бр. матрици за съвместна поява на сивото, всяка с по една от отделните пространствени взаимовръзки (*Δx, Δy).* Така получените матрици, след използване на някои от известните математически подходи (Метод на главната компонента, Факторен анализ и др.) за редуциране на размерността, могат да бъдат използвани за обучение и класификация.

Изображение 4. Ротация при сравняване на пикселите

Ако се използва оригиналното изображение с всички нива на сивото от 0 до 255, матрицата за съвместна поява на сивото става с изключително големи размери - 256х256, което води до необходимост от голяма изчислителна мощ. Друг проблем е, че матрицата има много ниско ниво на запълненост и малка част от позициите в нея биха били различни от нула. За статистически по – достоверни резултати е по – добре матрицата да има по – високо ниво на запълненост. Това се постига основно по два начина:

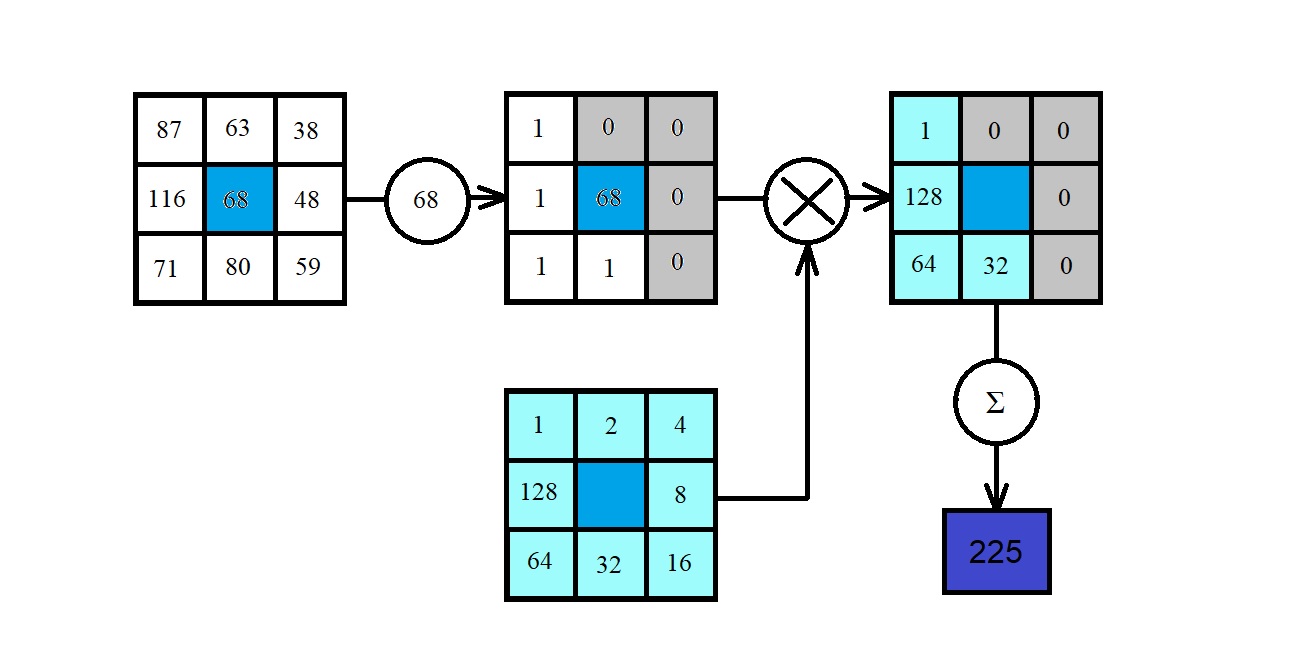
а) ограничаване броя нивата на сивото, чрез интерполацията им от оригиналните нива на сивото ( от 0 до 255) до диапазон с с по - малко на брой нива ( напр. от 0 до 8, от 0 до 16, от 0 до 32). Това води значителна редукция на броя изчисления и значително повишаване на нивото на запълненост на матрицата за съвместна поява на сивото.

б) избирането на характерна част от изображението, при която текстурата е по – ясно изразена и локалното определяне на матрица за съвместна поява на сивото.

От получените матрици също така могат да бъдат извлечени определени признаци за класификация като контраст, ентропия, хомогенност и корелация. Изчисляват се математически по определени формули. Те могат да бъдат използвани като признаци за обучение и класификация на даден класификатор.

### 2.2.3. Локални двоични модели (Local binary patterns)

Локалният двоичен модел е известен подход в сферата на машинното обучение, обработката на изображения и компютърното зрение. За пръв път е представен през 1994г. и оттогава е сред най – използваните подходи за определяне и класификация на текстура в изображенията. Методът е оператор за описание на текстура, базиран на знаците на разликите между съседни пиксели и един централен. На изображение 5 са показани етапите при изчисление на LBP  
стойностите.   
За всеки отделен пиксел във фотоизображението се изпълнява алгоритъм, при които той се сравнява със съседните на него 8 пиксела. Ако конкретният пиксел е по – голям от съседа му, на мястото на съседа се записва 0, ако е по – малък се записва 1. Получената 3х3 матрица се умножава с друга 3х3 матрица. Тази друга матрица съдържа степени на 2-ката от 0 до 7 – общо 8 на брой, като те се записват в крайните клетки по посока на часовниковата стрелка, започвайки от позиция 0,0 (или 1,1). Стойностите в получената резултанта матрица се сумират и резултатът се записва на мястото на конкретния пиксел – в случая 68 - > 225. По този алгоритъм се минава през цялото изображение.   
  
Математическото представяне на алгоритъма е следното:



Изображение 5 Етапи при изчисление на LBP

Това е класическият вариант алгоритъма, но съществуват и други вариации. След като алгоритъмът премине през всички пиксели се изчислява хистограмата на изображението. В някои случаи се нормализира. Получените вектор-ред (или вектор-стълб) се използват като признаци за обучение на някакъв вид класификатор.

### 2.2.4. Изчисляване на геометрична форма и размери