



1. ОСНОВНИ ПОНЯТИЯ. КЛАСИФИКАЦИЯ НА СИСТЕМИТЕ ЗА КОМПЮТЪРНО ЗРЕНИЕ.

1. Основни понятия. Класификация на системите за компютърно зрение.

Една система е с ИИ, ако отговаря на следните 5 критерия :

- 1) Изгражда ли системата модел на околната среда?
- 2) Използва ли системата този модел на околната среда за изграждане на план на действие в тази среда?
- 3) При изпълнението на плана за действие анализира ли системата алтернативни възможности?
- 4) Може ли системата да предефинира плана за действие, ако той води към непредсказуеми състояния за околната среда?
- 5) Използва ли системата минал опит за разширение и подобряване на модела на околната среда?

Определение.

Компютърното зрение (Computer Vision) е поле на компютърните науки, което работи за даване на възможност на компютрите да виждат, идентифицират и обработват изображения по същия начин, както прави човешкото зрение, и след това да осигурят подходящ изход. Това е като предаване на човешкия разум и инстинкти на компютър. В действителност обаче е трудна задача да се даде възможност на компютрите да разпознават изображения на различни обекти.

Компютърното зрение е тясно свързано с изкуствения интелект, тъй като компютърът трябва да интерпретира това, което вижда, и след това да извърши подходящ анализ или да действа съответно.

Целта на компютърното зрение е не само да вижда, но и да обработва и предоставя полезни резултати въз основа на наблюдението. Например компютър може да създаде 3D изображение от 2D изображение, като тези в автомобилите, и да предостави важни данни на автомобила и / или водача.

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Например автомобилите могат да бъдат оборудвани с компютърно зрение, което би било в състояние да разпознава и различава обекти на и около пътя, като светофари, пешеходци, пътни знаци и т.н., и да действа съответно. Интелигентното устройство може да осигури вход на водача или дори да накара колата да спре, ако има внезапна пречка на пътя.

Много големи компании като Microsoft продължават усилено да работят върху свои технологии (като Microsoft Research) за разпознаване и обработка на изображение.

Когато човек, който управлява автомобил, види, че някой внезапно се движи в пътя на колата, водачът трябва да реагира незабавно. В една секунда човешкото зрение е изпълнило сложна задача - тази да идентифицира обекта, да обработва данни и да решава какво да прави. Целта на компютърното зрение е да даде възможност на компютрите да изпълняват същия вид задачи като хората със същата ефективност.

----- www.eufunds.bg -----

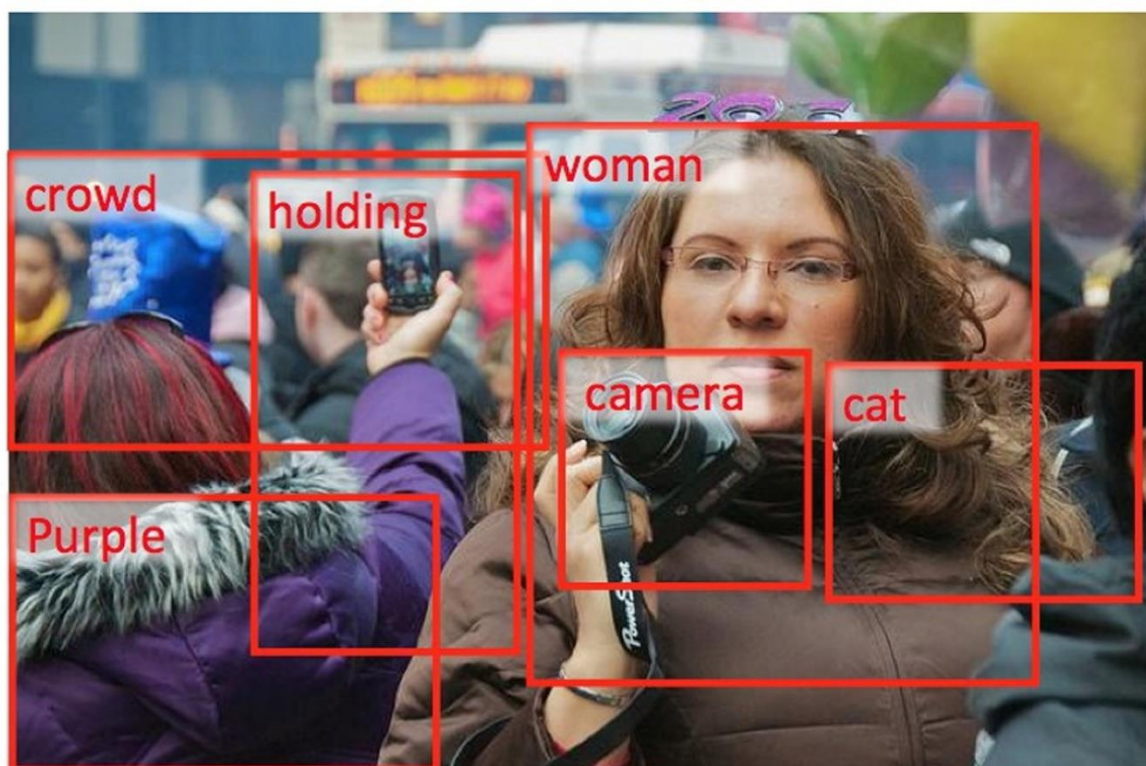
Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ



Фиг. 1. Анализ и разпознаване на околната среда

Microsoft Research сега не само разпознава основните елементи във фотографиите, но автоматично може да опише в текстов вид, обектите във фотографиите (фиг. 2.). Това е поредната стъпка в създаването на по-съвършен изкуствен интелект.

www.eufunds.bg

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ



GT: spotlight
1: grand piano
2: folding chair
3: rocking chair
4: dining table
5: upright piano



GT: spotlight
1: acoustic guitar
2: stage
3: microphone
4: electric guitar
5: banjo



GT: spotlight
1: altar
2: candle
3: perfume
4: restaurant
5: confectionery



GT: restaurant
1: wine bottle
2: candle
3: red wine
4: French loaf
5: wooden spoon



GT: restaurant
1: goblet
2: plate
3: candle
4: red wine
5: dining table



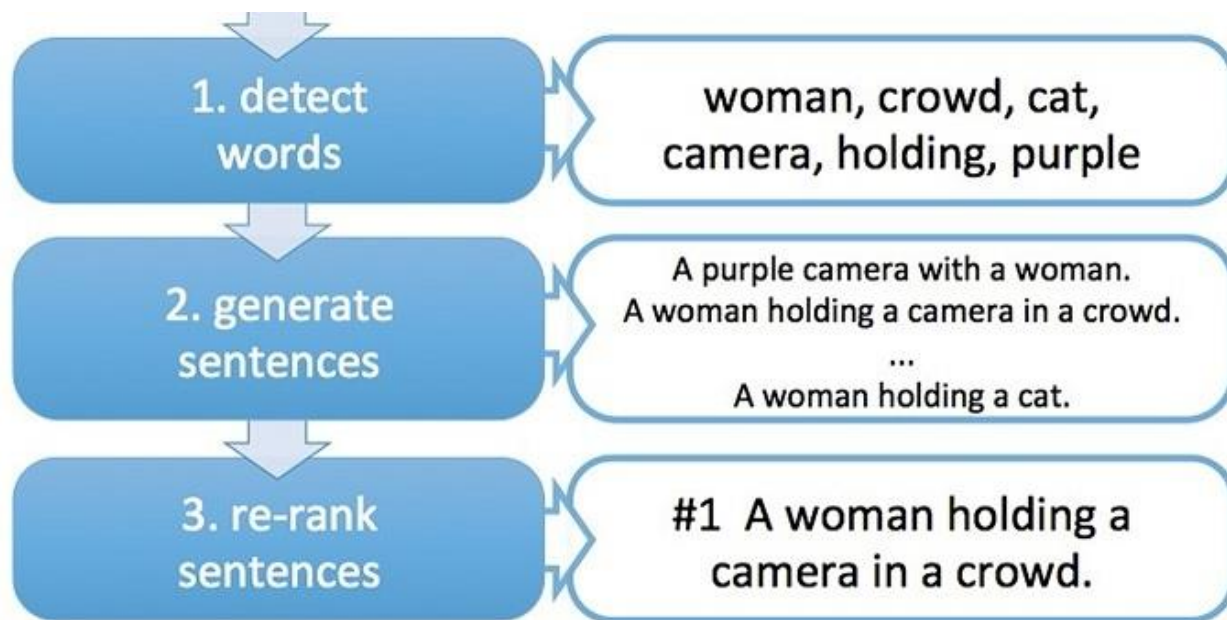
GT: restaurant
1: plate
2: meat loaf
3: ice cream
4: chocolate sauce
5: potpie

Фиг. 2. Автоматично описване в текстов вид на обекти във фотографии.

Алгоритмите за разпознаване и интерпретиране на изображенията се опитват да имитират работата на човешкия мозък. Първоначално се определят всички основни елементи в изображението, а след това се генерират предложения, базирани на взаимодействието между обектите (фиг.3).

www.eufunds.bg

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



Фиг. 3. Алгоритми за разпознаване и интерпретиране на изображенията.

Като пример се дава фотографията, показана по-горе, за която ИИ разпознава обектите „жена, блъсканица, котка, камера, държи, виолетово“ и генерира предложенията: „Виолетова камера с жена“, „Жена с фотоапарат в тълпата“ и „Жена, която държи котка“. В крайна сметка ИИ се спира на „Жена с фотоапарат в тълпата“ (фиг. 1.).

Основното тук е изграждане на модела на околната среда. Целта на компютърното зрение е да се изгради модела за наблюдаваната околна среда. Трябва да извлече информация за тази околна среда, за да се постигне целта.

Околна среда: съвкупност от обекти, попадащи в полето на наблюдение (възприемане) от съответната видеовъзприемаща апаратура.

Визуална сцена: двумерна картина на съвкупността от двумерни или тримерни обекти попадащи в полето на наблюдение на видеозаснемащата апаратура.

Класификация на системите за компютърно зрение може да бъде направена според:

----- www.eufunds.bg -----



- 1) Област на приложение: визуален контрол, медицина, авионаблюдение, космически наблюдения.
- 2) Сложност: прости, със средни възможности, с големи възможности.
- 3) Размерност: двумерни и тримерни.
- 4) Тип на изображението: двоично, полутоново, цветно.
- 5) Сложност на задачата: центрирани детайли, детайли в разширен безпорядък, детайли в бункер.
- 6) Използвани средства: телевизионни камери (видикони и твърдотелни камери), мултиспектрални скенери, компютърни системи.
- 7) Начин на обработка на изображението: последователна, паралелна.

Двумерни системи за компютърно зрение: когато става дума за обработка на двумерни обекти.

Класификация според типа на изображението :

- 1) двоично изображение: използват се само 2 нива - черно (1)/бяло (0).
- 2) полутоново изображение: използват се повече от 2 нива. Обикновено 256 нива на сивото. Това изображение дава по-добра представа за формата на обектите във визуалната сцена.
- 3) цветно: използват се 3 цветови канала (син, зелен и червен).

Степента на сложност нараства както с преминаване от двоично, през полутоново до цветно изображение, така и при преминаване от двумерни към тримерни сцени. Разделянето е на малка сложност, средна сложност и голяма сложност (цветно тримерно изображение).

Класификация според сложността на задачата:

- 1) Центрирани детайли: имаме конвейер с детайли. Детайлите не се препокриват.
- 2) Детайли в равнинен безпорядък: На конвейера детайлите могат да се закриват един друг.
- 3) Детайли в бункер: Системата трябва да вземе конкретен детайл от бункера.

----- www.eufunds.bg -----



Класификация според използваните средства:

- 1) Телевизионните камери могат да бъдат черно-бели, полутонови и цветни.
- 2) Мултиплексорни скенери: измерват дължината на електромагнитното колебание.

Зрение: процеса на определяне какво се съдържа в наблюдаваната околна среда и къде се разполагат обектите в тази наблюдавана среда.

Трябва да се получи изображение на тази среда. То трябва да се представи по определен начин в системата. На базата на това изображение се определят обектите. Системата трябва да бъде обучена да разпознава обекти.

Трябва да се определят отношенията между обектите. Определя се центъра на тежестта на всеки разпознат обект и се определят техните отношения. Може да има неразпознати обекти. Целият този процес изисква съхраняването на голям обем информация.

Система за компютърно зрение (СКЗ): представлява съвкупност от специализирани устройства за формиране на изображението на околната среда.

От друга страна за СКЗ са необходими методи, алгоритми и програми за обработка на изображения; определяне на признаци на обектите; построяване на формално описание на наблюдаваната сцена; разпознаване на обектите и изграждане на околната среда.

2. Физически основи на компютърното зрение

Всяко тяло, което има температура над абсолютната нула генерира електромагнитно излъчване. Излъчването се движи със скоростта на светлината. С увеличаване на вътрешната температура на тялото се увеличава и количеството на излъчваната енергия, както и излъчването преминава от по-дълги към по-къси вълни.

Ако се вземе по абсциса дължината на вълната то имаме следните области (започва се от най-късите вълни):

- Гама лъчи;
- Х лъчи (рентгеново излъчване);
- Ултравиолетови лъчи;

----- www.eufunds.bg -----



- Видима светлина;
- Инфрачервена (отразена и термална) светлина;
- Микровълни;
- Радиовълни.

Видимата област на електромагнитния спектър е в рамките 400-700 нанометри. За хората основен източник на електромагнитно излъчване е Слънцето. То излъчва в много честотни диапазони : от 1 пикометър до 30 микрометра. Максимум на излъчване е около 500 нанометри. Температурата на Слънцето е 6000K.

Земята също излъчва електромагнитна енергия при средна температура 300K. Нейният максимум обаче не е в във видимата част на спектъра, а е на 9,7nm.

Видимата област се разделя на 3 подобласти :

- синьо (blue) - 400-500nm;
- зелено (green) - 500-600nm;
- червено (red) - 600-700nm.

Човешкото зрение може да се разглежда като дистанционно наблюдение (измерване от разстояние). Ако при него имаме естествен източник на електромагнитни колебания, то тогава системата е пасивна. Ако източника на светлина е изкуствен - системата е активна.

Светлината се дели на :

1) ахроматична: не съдържа цвят, изменя се от черно до бяло през нива на сивото. Основна нейна характеристика е интензитета на отразената светлина.

2) хроматична: характеризира се със цвят. Анализа на цветни изображения става на базата на трите основни цвята червен, зелен и син - Red, Green и Blue (RGB). Всеки цвят може да се получи като комбинация от компоненти по всеки един от тези основни цветове.

Трите основни цвята образуват тримерно пространство на цветовете. Ако стойностите са дискретни и ограничени, пространството е куб. Началото на координатната система е (0,0,0). Колинеарните вектори в това пространство представят един и същ цвят, но се различават по яркостта.

----- www.eufunds.bg -----



За да получим всеки един от тези 3 цвята от едно хроматично изображение се използват филтри. Интегралните уравнения за реализиране на филтрите са:

$$\begin{aligned} R &= \int E(\lambda) R^I(\lambda) d\lambda, \\ G &= \int E(\lambda) G^I(\lambda) d\lambda, \\ B &= \int E(\lambda) B^I(\lambda) d\lambda, \end{aligned}$$

където $R^I(\lambda)$, $G^I(\lambda)$, $B^I(\lambda)$ са функции на разпределение за съответните основни цветове, а $E(\lambda)$ - спектралния състав на цвета в дадена точка.

Във всяка от трите части на видимата част на спектъра, кривата на спектралния състав има изпъкнала форма. Максимумите в отделните части са определени да са: 435,8 nm (blue), 546,1 nm (green), 700 nm (red).

Има и други системи за анализ на цветове и изображения.

Начина на възприемане на човешкото око се доближава повече до системата HIS (Hue, Saturation, Intensity);

- цветови отенък (Hue);
- наситеност (Saturation);
- яркост (Intensity) - пропорционална на интензитета на отразената светлина.

Преминаването от едната система в другата става лесно чрез използването на следните формули:

$$I = R + G + B; H = (G - B) / (I - 3B); S = (I - 3B) / I$$

Описание на различните части от електромагнитният спектър:

1) Гама лъчи (< 0.03 nm): Поглъща се напълно от горните слоеве на атмосферата.

2) Рентгенови лъчи (0.03 - 30 nm): Напълно се поглъща от горните слоеве на атмосферата.

3) Ултравиолетови лъчи (0.003 - 0.4 μ m):

Вълните с дължина < 0.3 μ m напълно се поглъщат от озона в горния слой на атмосферата.

----- www.eufunds.bg -----



4) Фотографска УВ област ($0.3 - 0.4 \mu\text{m}$): Минава през атмосферата. Улавя се със фоточувствителен филм. Атмосферното разсейване е много силно.

5) Видима област ($0.4 - 0.7 \mu\text{m}$): Изобразяване с филми и фото детектори.

6) Инфрачервена област ($0.7 - 100 \mu\text{m}$): При взаимодействие с материята (отражение) променя дължината на вълната си. Съществуват атмосферни прозорци на пропускане.

7) Отражена ИЧ ($0.7 - 3 \mu\text{m}$): Съдържа информация за топлинните свойства на материята. Лентата ($0.7 - 0.9 \mu\text{m}$) се улавя с филм и се нарича фотографска ИЧ област.

8) Термална ИЧ ($3 - 5 \mu\text{m}$), ($8 - 14 \mu\text{m}$): Принципи атмосферни прозорци.

9) Микровълни ($0.1 - 30 \text{ cm}$): По-дългите вълни проникват през облаците. Изображение може да бъде получено по активен начин.

10) Радио област ($> 30 \text{ cm}$): Най-дългите вълни в електромагнитният спектър.

Значението на яркостите на изображенията се определя от четири фактора:

- Геометрични свойства на обектите;
- Коефициент на отражение на видимите повърхности;
- Осветеност на сцената;
- Позиция на наблюдателя.

Задачата за предварителна обработка на зрителна информация се заключава в сортиране на изменението на яркостта в зависимост от причината, така че да получим представяне, в което четирите фактора се отчитат независимо един от друг.

Тази цел се постига на два етапа:

1. Съответстващите представления се базират на основата на изменение на яркостта и структурата. В резултат на тази обработка се получава така наречения първоначален ескиз.

----- www.eufunds.bg -----



2. Геометричните преобразувания върху видимите повърхности. Така се получава 2,5 мерен екскиз.

И двата екскиза се построяват в равнината (координатната система) на наблюдателя.

3. Информационен модел на система за компютърно зрение.

Информационният модел дава етапите, през които преминава входното изображение, за да се получи модел на околната среда. Етапите за обработка са следните.

- 1) Квантуване. Филтрация. (вход: $f_1(x,y)$, изход: $g(x,y)$)
- 2) Отделяне на ръбове и граници (вход: $g(x,y)$, изход: $b(x,y)$)
- 3) Отделяне на признаци. Сегментация.
- 4) Построяване на описание.
- 5) Интерпретация. Разпознаване.
- 6) Изграждане на модел на околната среда.

Тези стъпки могат да се разделят по два начина на нива. Едната система е с 2 нива (ниско/високо), другата система е с 3 нива (ниско/средно/високо). При първата, ниското ниво са стъпки 1,2 и 3, останалите са във високото. При втората система разделянето е на по две стъпки в ниво в същата посока както при първата система.

На входа на информационният модел постъпва функция $f_1(x,y)$, която се квантува до функция $g(x,y)$. Върху тази функция (дискретизирана) се прилага филтрация с цел подобряване на първоначалното изображение.

Сегментация: определяне на формата на различните обекти в сцената. $b(x,y)$ е двоично изображение, т.е. стойностите на функцията са само 1 или 0. Границите на ръбовете са означени с 1, а всичко останало е 0 (процес на отделяне на ръбове).

Интерпретация, разпознаване: Системата трябва да е предварително обучена, за да може да прави разпознаване.

Накрая, ако не са разпознати някои обекти, те ще участват в модела, но ще трябва да се дообучи системата за да ги разпознава следващ път. Моделът се създава на база отношението между обектите в сцената.



На основание на обработката се получават Базис от данни (БД) и Базис от знания (БЗ).

Информацията, която се получава в компютъра се дели на :

- декларативна: това са данните, които се обработват;
- процедурна: множество програми, по които става обработката.

Особености на една система за компютърно зрение по отношение на данните:

1) Обемът данни, които се обработва в една система за компютърно зрение е много голям. Следователно капацитета на БД е много голям.

2) Съществува йерархичност на данните.

Ако информационните единици (ИЕ) в БД придобият следните свойства те стават БЗ:

1) Вътрешна интерпретируемост: към информационните единици можем да се обърнем с ключ. С вторични ключове можем да се обърнем към определени полета на ИЕ.

2) Структурност: едни ИЕ да могат да се включват в други ИЕ. (част - цяло, обект - клас)

3) Свързаност на ИЕ: определя конкретни отношения за информационните единици ("пред", "над", "причина - следствие", "аргумент - функция"). Тези отношения се отнасят към декларативната информация. "аргумент - функция" се отнася към процедурната информация.

4) Семантична свързаност: отнася се за конкретни предметни области. Близки информационни единици се анализират с цел формиране на някакво понятие.

5) Активност: анализирайки, конкретните данни и отношенията между тях, могат да се предизвикват конкретни действия.

Представянето на знанията в системата за компютърно зрение трябва да отговаря на следните условия:

1) Трябва да можем да представяме в системата различни видове знания.

2) Системата трябва да води до ефективно съхранение на данните.

3) Лесна манипулация с елементите на знанието.



4) Знанието трябва да бъде така представено, че да бъде лесно разбираемо за човека.

Начини за представяне на знанието :

1) Използват се портретни модели (іsonіc) - изходното полутоново или цветно изображение. Тези портретни модели са много чувствителни по отношение на точката на наблюдение и осветеността. Тези портретни модели изискват голям обем памет.

2) Графови модели: възли, които съответстват на отделни елементи на сцената (стени, обекти) и клонове (ребра) свързващи възлите. Ребрата указват връзките между елементите в сцената. Графовите модели са по-малка чувствителност към осветеността и ориентацията на обектите в сцената.

3) Семантични мрежи: развитие на граф модела. Това е ориентиран граф, възлите на който показват елементи на сцената, а ребрата именувани (смислови) отношения.

Фреймите са друг начин за представяне на декларативно процедурно значение. Фрейма представлява структура от данни за определяне на стереотипни ситуации. Като под ситуация се разбира визуален образ, събитие. Фреймовото представяне на знанието представлява ориентиран граф, като възлите на този граф представляват елементите на наблюдаваната сцена. Ребрата на графа дават отношенията между елементите. Всички възли на ориентираният граф на горни нива са определени, а на най-ниското ниво възлите са празни. Празните възли се запълват в зависимост от конкретната ситуация. Всеки възел може да бъде обръщане към подфрейм, при което обръщане могат да се активират процедури за изчисление на информация за конкретният възел или за даване на конкретна информация за него.

От всички представяния на знанието, най-универсално е фреймовото.

Управление на обработката на информацията :

1) “отдолу - нагоре” (управление на данните) - на входа на всеки етап постъпват данни, на изхода излизат резултати, които активират изпълняването на следващият етап и т.н.

2) “отгоре - надолу “ (управление на моделите) - прави се хипотеза, че при дадена обработка като краен резултат ще се получи резултат

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

съответстващ на заложеният. Недостатък е възможността да се направят голямо количество тестове преди да се достигне желаният резултат. Използва се често за анализ на тримерни сцени.

3) Обратна връзка - анализират се резултатите на изхода на всеки етап. Ако те са удовлетворителни се минава към следващият етап. Ако не са удовлетворителни се връщаме на входа на този етап и модифицираме обработката в някаква посока с цел получаване на различен резултат.

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.