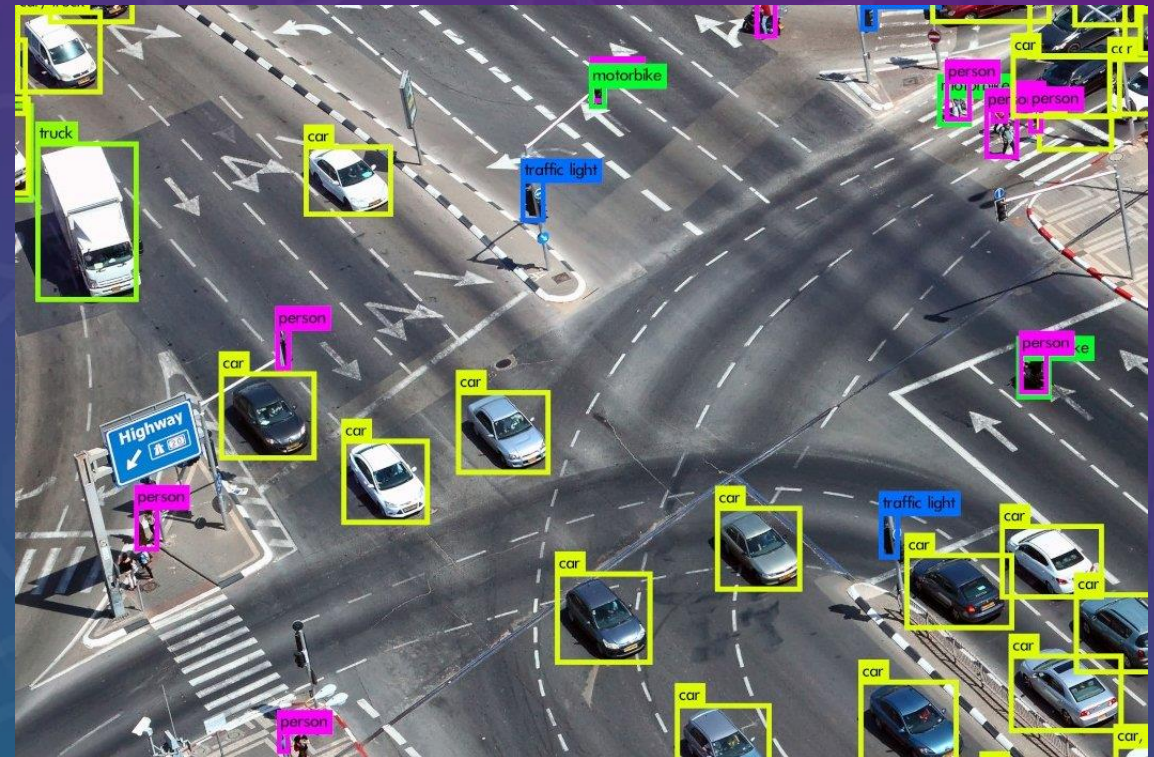


155YGEI GEOINFORMATIKA

DETEKCE VZORŮ V OBRAZECH
SEGMENTACE OBRAZU
STROJOVÉ UČENÍ

Jiří Cajthaml, Tomáš Janata
21. 11. 2023

ČVUT v Praze | Fakulta stavební | katedra geomatiky



POJMY

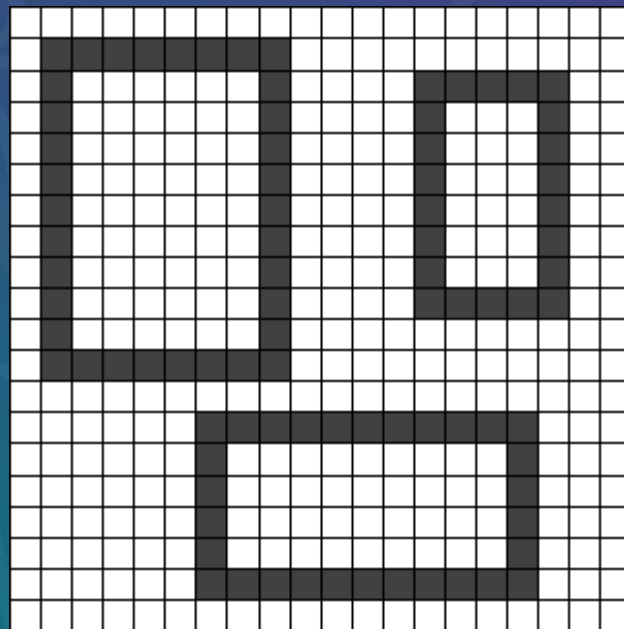
- **Strojové učení**
vědecká disciplína vytvářející a studující algoritmy, které se učí vytvářením statistických modelu z dat a používají se pro rozhodování a předvídání.
- **Rozpoznávání / klasifikace**
prirazuje skutečný objekt nebo událost do jedné nebo více předem stanovených tříd
- **Vzor**
je objekt nebo process, které lze pojmenovat a zařadit
- **Kategorie / třída vzorů**
je podmnožina vzoru, jejíž prvky sdílejí podobné rysy, tj. konečně rozpoznatelné vlastnosti (příznaky)

POJMY

- **Klasifikátor**
je stroj (algoritmus), který klasifikaci realizuje
- **Zde je vzorek (příklad, objekt)**
jedna instance vzoru, který míníme klasifikovat
- **Trénovací množina (data)**
je soubor příkladů, který se užívá k učení statistického modelu (klasifikátoru).
- **Testovací množina (data)**
je soubor příkladů, které jsou jiné než v trénovací množině
- **Príznaky**
jsou určující, podstatné vlastnosti dat pro uvažovanou klasifikaci

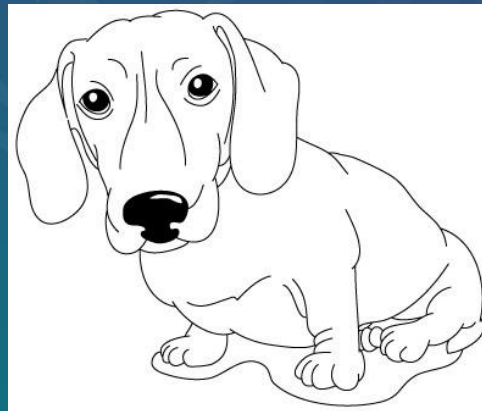
PŘÍKLADY TŘÍDY VZORŮ

- Třída syntakticky správných aritmetických výrazů, např. $2x(a + 3b) - 6y$
- Podmnožina množiny všech konečných řetězů (slov) nad nějakou abecedou
- Třída binárních obrazů obsahující nepřekrývající a sebe se nedotýkající obdélníkové rámečky s jednopixelovou tloušťkou. Je podmnožinou množiny všech pravoúhlých binárních obrazů.

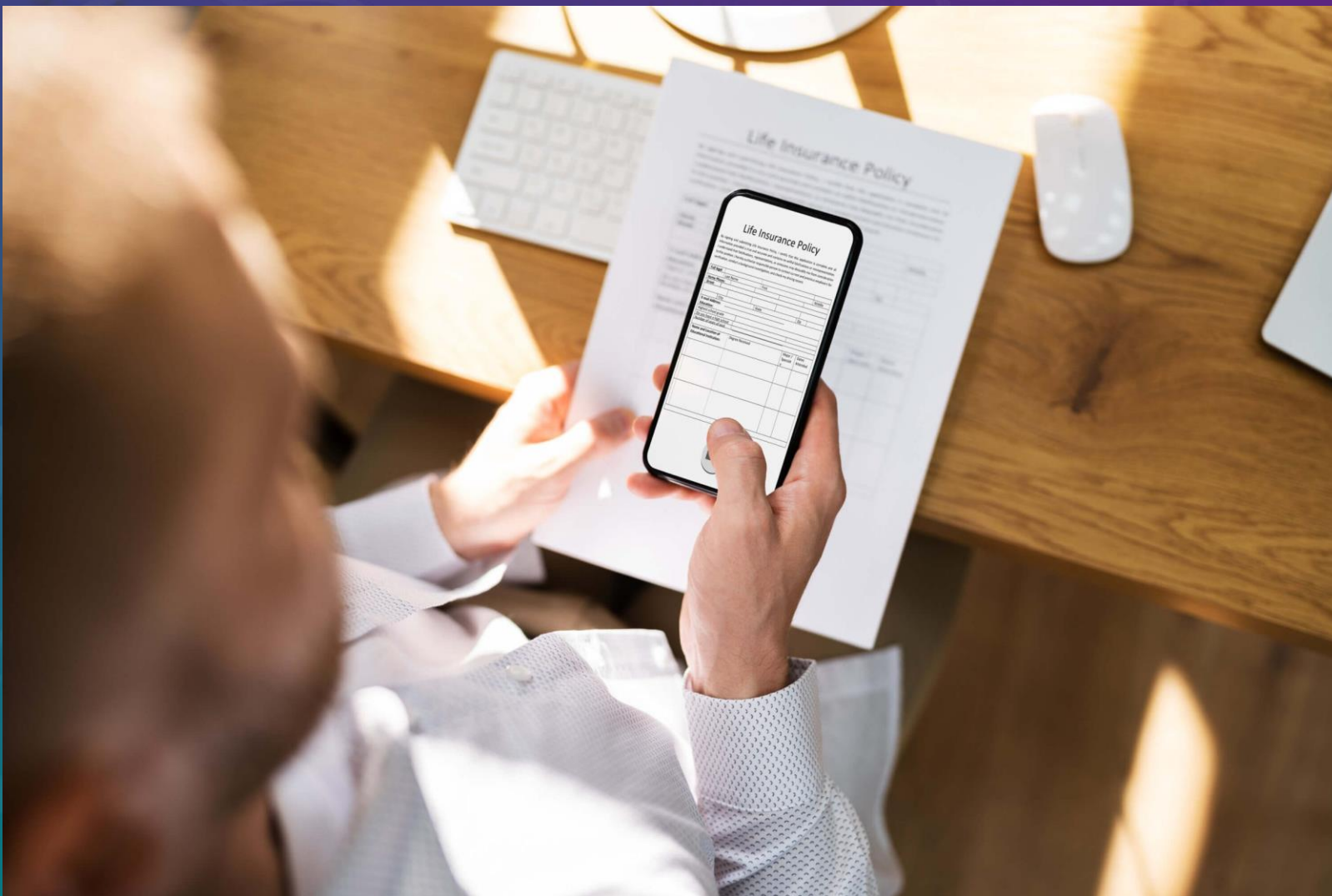


PŘÍKLADY TŘÍDY VZORŮ

- Množina všech psů v obrazech



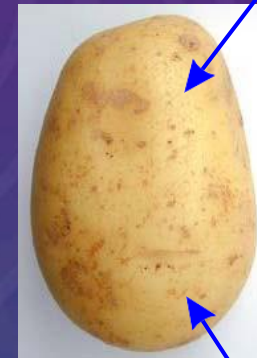
JEDNODUCHÝ A BĚŽNĚ ZNÁMÝ PŘÍPAD OCR



ZÁKLADNÍ POJMY, ILUSTRACE

- Studovaný vzor se analyzuje (například brambora, viz obrázek)
- Vektor příznaků $x \in X$ je vektor tvořený jednotlivými pozorováními (měřeními). Vektor x odpovídá jednomu bodu v prostoru příznaků X .
- Skrytý stav (ve obvyklém případě přímo značka třídy) $y \in Y$ není přímo pozorovatelný. Vzory se stejnými skrytými stavy vytvářejí jednu třídu.
- Úkolem je navrhnout klasifikátor (rozhodovací pravidlo) $q: X \rightarrow Y$, které přiřazuje pozorované instance vzoru ke skrytému stavu.

skrytý stav (nebo značka třídy) y



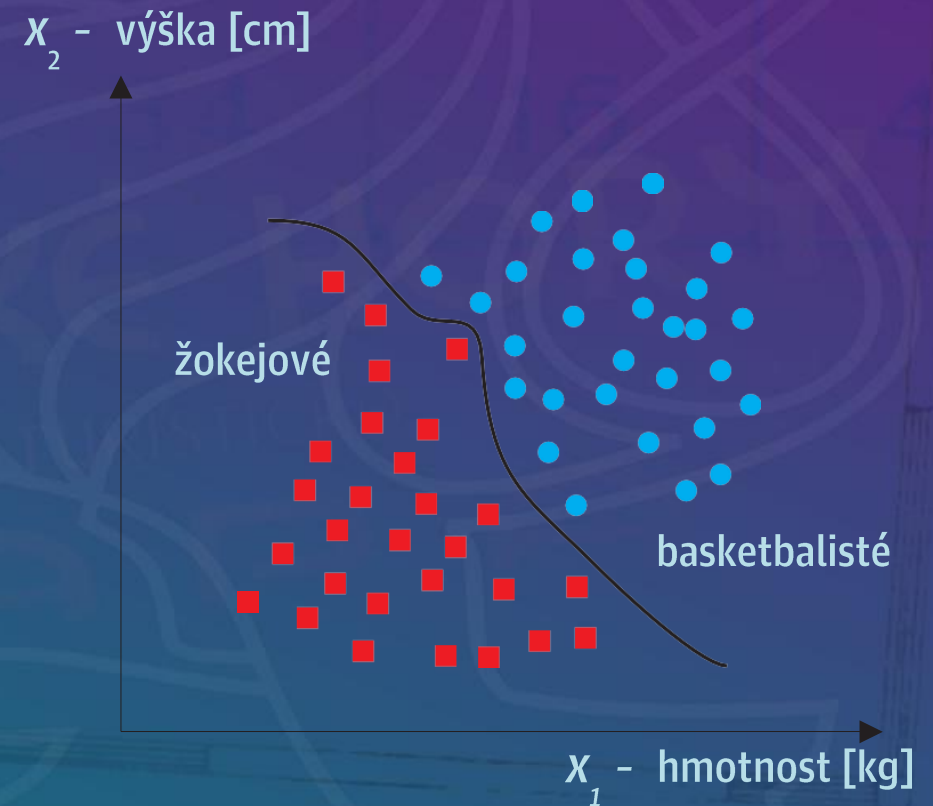
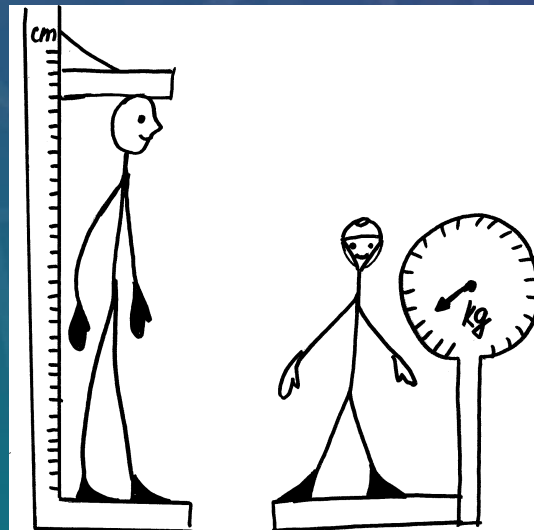
studovaný vzor

vektor příznaků X

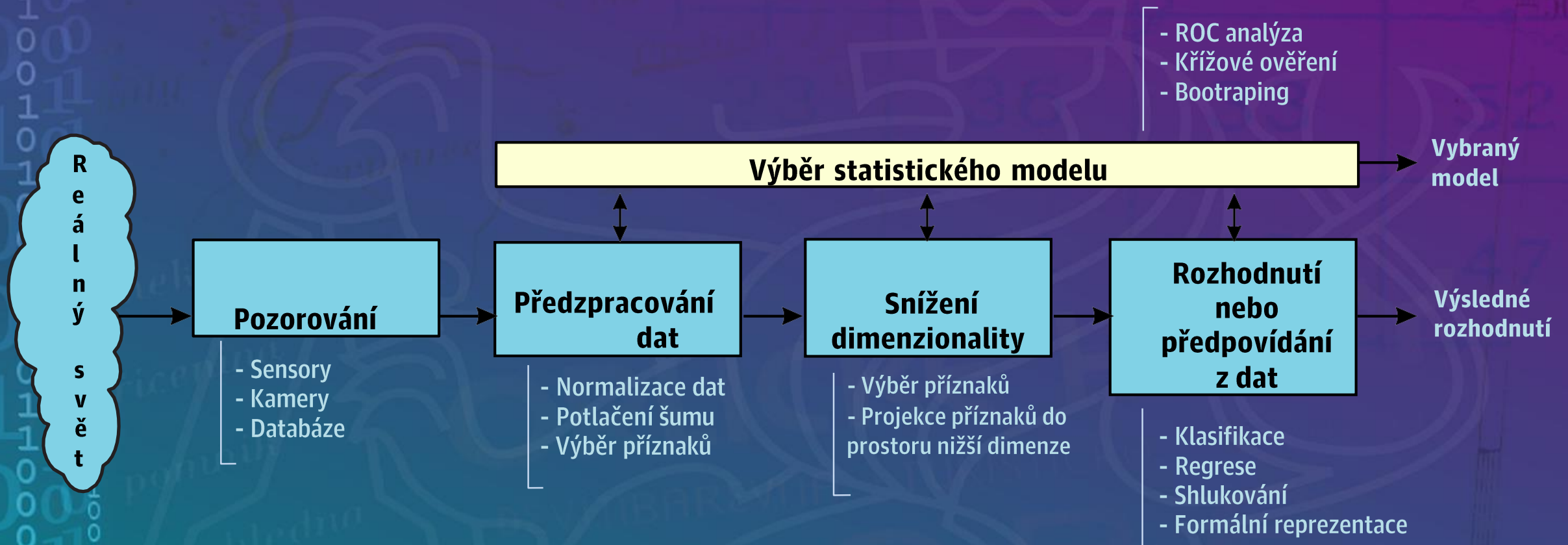
$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

ROZPOZNÁVÁNÍ, MOTIVAČNÍ PŘÍKLAD

- Objekt (situace) se popisuje dvěma parametry:
 - - pozorovatelný příznak (též pozorování);
 - - skrytý parametr (stav, speciální případ – klasifikační třída)
- Příklad:
statistické rozpoznávání:
žokejové a basketbalisté



CELKOVÝ POHLED, ČÁSTI

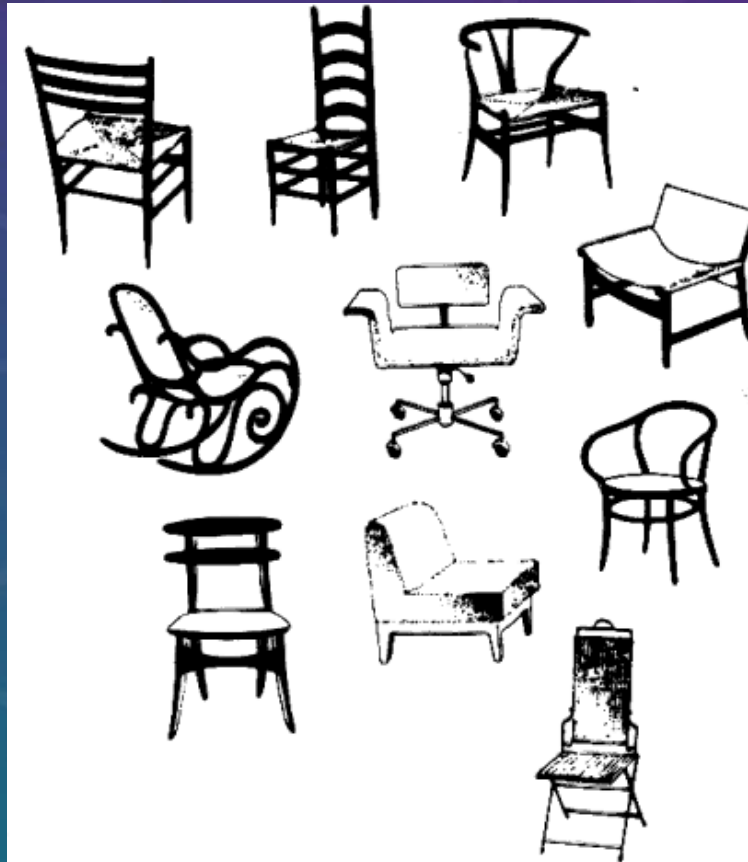


- Vstup: data, trénovací množina.
 - Statitické modely a jejich parametry se empiricky učí z trénovacích dat.
- Výstupy: různorodá rozhodnutí

DÁVNÁ VĚDECKÁ ÚLOHA, GNOZELOGIE

- Podstata klasifikace a rozhodování je hlavním tématem části filozofie, gnozeologie (teorii poznání), která studuje podstatu znalosti
- Základy rozpoznávání lze tedy odkázat až k Platónovi a jeho žákovi Aristotelovi
- Oba rozlišovali mezi:
 - základními vlastnostmi sdílenými všemi příslušníky třídy;
 - nahodilými vlastnostmi, kterými se mohou jednotliví příslušníci jedné třídy lišit

KLASIFIKACE/KATEGORIZACE (NEBO POPIS PODLE TOHO, K ČEMU OBJEKT SLOUŽÍ)



TYPY MOŽNÝCH ROZHODNUTÍ / PREDIKČNÍCH ÚLOH

- Klasifikace - přiřazuje pozorování do jedné z malé množiny tříd. Výstupem je identifikátor třídy, její značka. Např. značka označující kvalitu jablka jako A, B, C a odmítnutí (zmetek)
- Regrese - předpovídá hodnotu z pozorování. Zobecňuje klasifikaci. Např. výstupem může být reálné číslo odhadující příští hodnotu akcie na burze podle předchozích hodnot a dalších indikátorů chování akciového trhu
- Učení bez učitele (shlukování) - uspořádává pozorování do smysluplných tříd podle jejich vzájemných podobností. Např. v genetice hledá skupiny genů s podobnými vzory exprese
- Reprezentace strukturních vztahů se opírá o primitiva, např. vyjádřuje člověka pozorovaného dohlížecí kamerou pomocí předem definovaných poloh těla a s nimi spojených prototypů aktivit

DALŠÍ OBORY SDÍLEJÍCÍ PODOBNÉ HLAVNÍ MYŠLENKY

- Statistické modelování - hledá (generativní) model popisující objekt zájmu, např. pravděpodobnostní rozdělení a ohodnocuje kvalitu modelu statistickými metodami
- Strojové učení (což je dnes častější název pro rozpoznávání) - při dané trénovací množině se má rozhodovací pravidlo naučit automaticky. Člověkem zadaná (subjektivní) pravidla nejsou použita. Různé úlohy potřebují různé trénovací množiny
- Data mining - hledání explicitních, předem neznámých a potenciálně užitečných znalostí v datech
- Vizualizace ve vědě - vysokedimenzionální úloha se člověku zobrazuje v pro něj přirozeném 2D obrázku nebo 3D scéně. My lidé více dimenzí nevidíme
- Neuronové sítě - jeden z matematických formalismů řešící rozhodovací bez nutnosti vytvářet generativní model skutečného biologického systému

BIOLOGICKÁ MOTIVACE

- Člověk je na špičce pomyslné pyramidy živočichů i proto, že je schopen přemýšlet o postupech, jakými sám uvažuje.
- Panuje všeobecný zájem o strojové napodobení biologického vnímání s cílem napodobit inteligentní chování v nepříliš známém prostředí.
- Základním atributem inteligentního chování je schopnost učit se na základě vnímání okolního prostředí.
- Klíčová je otázka reprezentace znalosti. Přirozený jazyk je nejdokonalejší nástroj lidí pro vyjádření pozorování, pro popis jevu, formulaci úloh, jejich řešení a pro související otázky učení.

SLOŽITÉ JEVY A SYSTÉMOVÉ MYŠLENÍ

- Potřeba porozumět složitým jevům například v biologii, technice nebo sociálních vědách vede k nutnosti zkoumat jevy komplexně v mnoha souvislostech.
- Přístup je nazýván systémovým myšlením, aby se odlišil od newtonovské snahy zredukovat každý jev na vztahy mezi základními prvky a jejich vlastnostmi.

POJMY Z TEORIE SYSTÉMŮ

- Při zkoumání složitého jevu se omezujeme na část, která nás zajímá, a říkáme jí objekt (nebo systém)
- Vše ostatní, co nám z daného pohledu připadá nezajímavé, nazýváme pozadí
- Objekty většinou nezkoumáme v celé jejich složitosti. Při jednom zkoumání pozorujeme nebo měříme jen určité vlastnosti, které nám právě připadají zajímavé. Teorie systémů zde používá pojem rozlišovací úroveň
- Popis a chápání objektu se přirozeně může vyvíjet s měnící se rozlišovací úrovní. Jde o metapohled hledající kvalitativní změnu v popisu objektu

GENERATIVNÍ × DISKRIMINATIVNÍ REPREZENTACE OBJEKTU

- Snaha o exaktní popis objektů (složitých dějů) matematickými nástroji vyústila (zhruba řečeno) ve dva možné přístupy:
- **Generativní ⇒ modelování**
- Snaží se o porozumění fyzikálním/jiným principům a jejich vyjádření modelem. Tento model umí generovat data dosti se podobající empirickým pozorováním
- Příkladem je matematické modelování fyzikálního/technického děje (v newtonovském smyslu)
- **Diskriminativní ⇒ klasifikace**
- Snaží se porozumět vnějšímu chování bez detailní znalosti dílčích principů (která u složitých objektů/dějů ani většinou není)
- Výstupem jsou rozhodnutí/předvídání ve smyslu regrese
- Příkladem je rozpoznávání, např. stanovení diagnózy lékařem / počítačovým programem

MATEMATICKÉ MODELOVÁNÍ

- Podstatné rysy objektu se napodobují formou matematických rovnic. Často se hledá relace mezi vstupem a výstupem.
- Má blíže k newtonovskému pojetí. Snaha o co nejpodrobnější a deterministické vysvětlení.
- Příklad: dobrý matematický model elektrárenského kotle v teorii řízení předpovídá téměř stejné odezvy na vstupní signály jako kotel sám.
- Protipříklad 1: V mnoha případech nejsme schopni matematický model vůbec vytvořit (např. model fungování lidského těla).
- Protipříklad 2: Počítačové vidění. Inverzní úloha k fyzikálnímu postupu vzniku obrazu je příliš složitá, a tudíž prakticky nepoužitelná.

ALTERNATIVOU K MODELOVÁNÍ JE ROZPOZNÁVÁNÍ

- Rozpoznávání zařazuje pozorování podle nějakého rozhodovacího pravidla do předem známých tříd.
- Třídy ekvivalence (relace ekvivalence: reflexivní, symetrická, tranzitivní). Uvnitř těchto tříd jsou si objekty podobnější než mezi třídami navzájem.
- V rozpoznávání bývá porozumění objektu méně podrobné než v modelování.

ROLE UČENÍ V ROZPOZNÁVÁNÍ

- Výhodou rozpoznávání je, že člověk vytvářející rozhodovací pravidlo (strategii, klasifikátor) nemusí rozumět složité podstatě jevu, o kterém se má rozhodovat
- Rozhodovací pravidlo může být naučeno empiricky z mnoha pozorovaných příkladů
- Paradox znalostního inženýrství: Pro člověka je snazší poskytnout příklady správné klasifikace než explicitně vyjádřit pravidlo, podle kterého rozhoduje
- Tři hlavní přístupy k učení:
 - Učení s učitelem na základě trénovací multimnožiny zahrnující pozorování a informaci o třídě, kterou přisoudil učitel (znalec).
 - Učení bez učitele na základě hledání podobnosti mezi pozorováními, aniž by byla k dispozici znalcová klasifikace.
 - Podporované učení (reinforcement learning) místo informace od učitele odhaduje odměny nebo pokuty z prostředí. Maximalizuje se kumulativní odměna

METODY ROZPOZNÁVÁNÍ A APLIKACE

- Velkou výhodou rozpoznávání je, že je lze oddělit od aplikačních disciplín.



HLAVNÍ PŘÍSTUPY K ROZPOZNÁVÁNÍ

- **Statistické (příznakové) rozpoznávání**

Předpokládá se existence statistického modelu jednotlivých vzorů a tříd vzorů. Souřadné osy prostoru odpovídají jednotlivým číselně vyjádřeným pozorováním, tedy příznakům

Objekty jsou reprezentovány jako body ve vektorovém prostoru

- **Strukturní rozpoznávání**

Mezi pozorováními existuje struktura a ta je reprezentována. Nejrozvinutější a nejstarší je reprezentace struktury gramatikami

- **Umělé neuronové sítě**

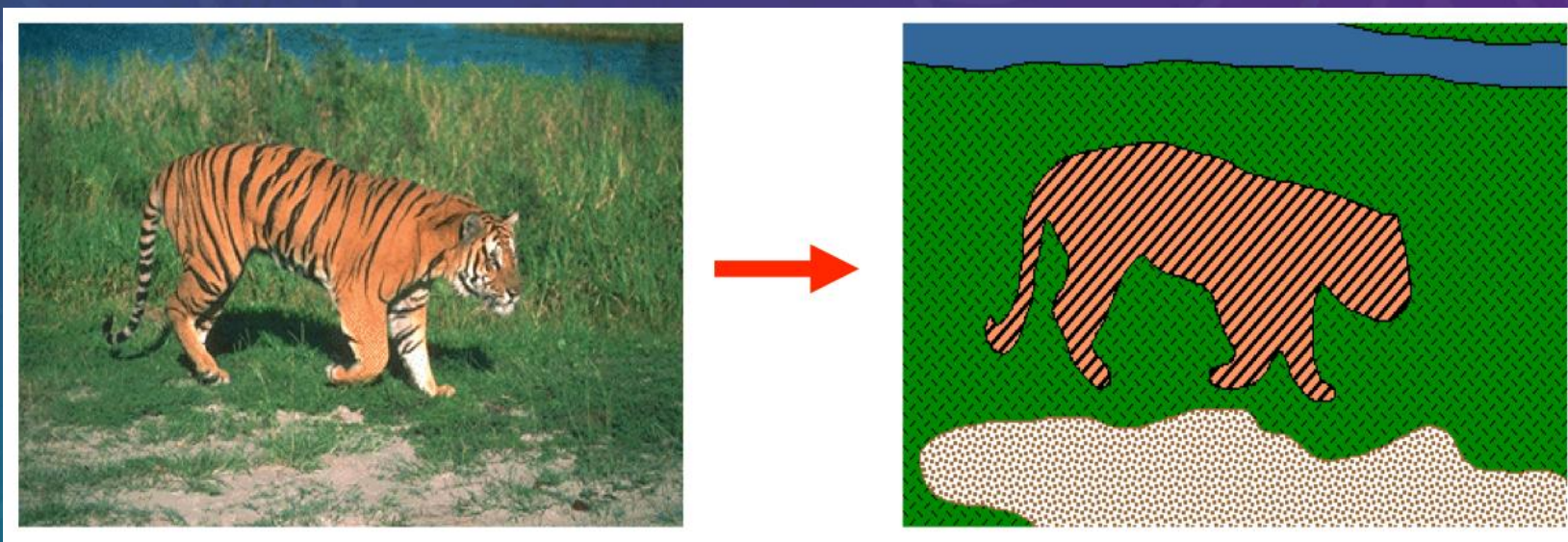
Klasifikátor je realizován sítí navzájem propojených uzlů, které modelují neurony v lidském mozku (přístup konekcionisty, např. model dopředné neuronové sítě)

OBECNOST FORMULACE ROZHODOVÁNÍ

- Pozorování x může být číslo, symbol, funkce dvou proměnných (např. obrázek), graf, algebraická struktura, atd.

Aplikace	Pozorování	Rozhodnutí
hodnota mince v mincovním automatu	$x \in \mathbb{R}^n$	hodnota
rozpoznávání znaků (OCR)	2D šedotónový obrázek	znaky, slova
rozpoznávání RZ aut	2D šedotónový obrázek	písmena, číslice
rozpoznávání otisků prstů	2D šedotónový obrázek	identita člověka
rozpoznávání řeči	signál z mikrofону $x(t)$	slova
analýza EEG, EKG signálů	$x(t)$	diagnóza
identifikace zásilek	binární obrázek	{ano, ne}
ověření řečníka	signál z mikrofonu $x(t)$	identita člověka

SEGMENTACE OBRAZU



CO JE SEGMENTACE OBRAZU ?

- Segmentace je soubor metod, které umožňují interpretovat prostorově blízké části obrazu jako objekt.
- Oblasti (tj. kompaktní množiny) přirozeně reprezentují prostorovou blízkost, a jsou tak důležitým stavebním krokem k segmentaci. Objekty ve 2D obraze velmi často odpovídají rozlišitelným oblastem.
- Objektem je vše, co je na snímku zajímavé (z hlediska konkrétní aplikace). Zbytek obrazu je pozadí.
- Tento přístup je podobný přístupu používanému při rozpoznávání vzorů, tj. rozdělení obrazu na množinu tříd ekvivalence.



SEGMENTACE MŮŽE BÝT OBTÍŽNÁ, NAPŘÍKLAD



- V uvedené oblasti je obtížné najít hranici mezi hrnečkem a pozadím, protože se v místním pohledu neliší
- Hádanku může vyřešit pouze znalost sémantiky hrnečku

SEGMENTACE OBRAZU, TROCHA MAGIE

- Často neexistuje jediná odpověď, jak segmentovat. Segmentace je většinou založena spíše na ad hoc metodách
- Neexistuje žádná obsáhlá teorie segmentace. Několik nedávných teoreticky podložených přístupů však segmentaci formulovalo jako optimalizační úlohu
- Často se setkáváme se speciálním případem segmentace popředí vs. pozadí. Segmentace má obvykle smysl v rámci konkrétní aplikace

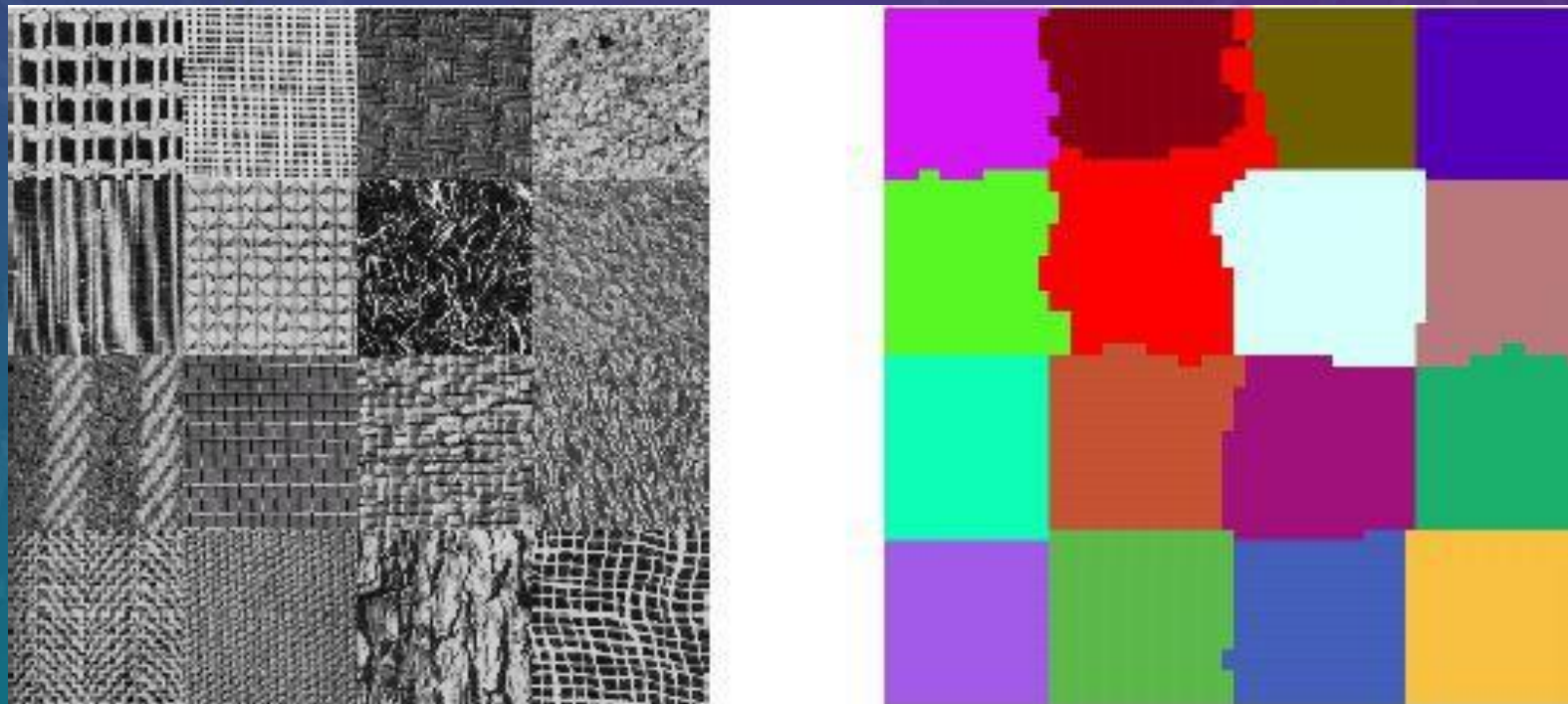


SEGMENTACE ZÁVISÍ NA APLIKACI

- Segmentace závisí na aplikaci, její sémantice. Metody nejsou univerzálně použitelné pro všechny obrazy.
- Přímá segmentace vstupního obrazu pragmaticky zohledňuje sémantické informace o konkrétní aplikaci.
- Tímto způsobem lze obejít špatně řešené problémy v inverzních úlohách souvisejících s fyzikou tvorby obrazu.
- Zásadní role apriorních informací:
 - Nízká úroveň: např. jas, prostorová koherence, barva, textura, pohyb, ...
 - Střední úroveň: symetrie objektů, blízkost ve větším měřítku, ...

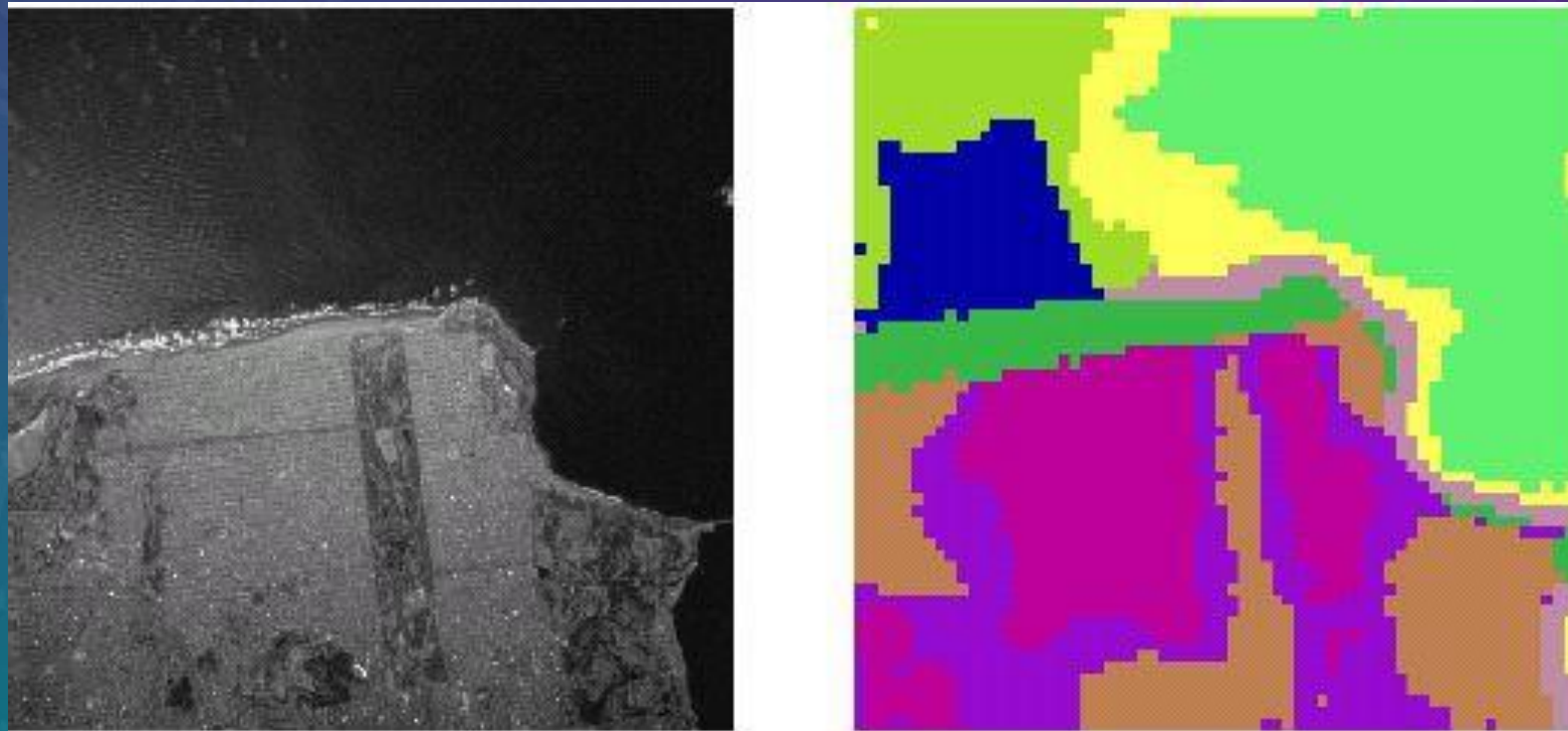
SEGMENTACE, PŘÍKLAD 1

- Segmentace podle textury.



SEGMENTACE, PŘÍKLAD 2

- Segmentace leteckého snímku mořského pobřeží.



ÚPLNÁ VS. ČÁSTEČNÁ SEGMENTACE (1)

- **Úplná segmentace** - rozdělí obraz na nepřekrývající se oblasti, které odpovídají objektům reálného světa.
- Úplná segmentace rozdělí obraz R na konečný počet S oblastí R_1, \dots, R_S
- **Částečná segmentace** - v obraze je možné najít pouze části se sémantickým významem (např. oblasti, soubory hran), které povedou k interpretaci při pozdější analýze.

ÚPLNÁ VS. ČÁSTEČNÁ SEGMENTACE (2)

- Pro přechod od částečné k úplné segmentaci je třeba prozkoumat vyšší úroveň zpracování informací.
- To lze provádět iterativně ve zpětné smyčce.
- Využívají se informace o sémantice konkrétní aplikace. Částečná segmentace snižuje množství dat, která je třeba zpracovat.
- Příklady kompletní 2D segmentace

Vyhledávání kontrastních objektů na homogenním pozadí.

Prahování intenzity poskytuje siluetu odpovídající objektům, např. tištěným znakům, buněčným jádrům, detailům osvětleným protisvětlem kontrolovaným v průmyslu.

POZOROVÁNÍ (NEBO RYSY) A SEGMENTACE

- Je tu důležitá otázka: Která pozorování (v terminologii rozpoznávání vzorů rysy) rozlišují oblasti odpovídající různým objektům na obraze?
- Příklady rysů: intenzita, barva, tvar oblasti, textura, pohyb ve videu, disparita ve stereoskopickém zobrazení.
- Primární funkce poskytuje přímo senzor (např. intenzita, barva, hloubka u dálkoměrné kamery, teplota u dálkové infračervené (termální) kamery).
- Z primárních znaků je třeba vypočítat složitější sekundární znaky, jako jsou: parametry textury, tvarové parametry oblasti, vzájemné vztahy mezi oblastmi, parametry pohybu ve videu, stereo disparita atd.

POUŽITÍ APRIORNÍCH INFORMACÍ

- Čím více apriorních informací, tím lépe. Ilustrativní příklady:

Požadovaný tvar oblasti. Požadovaná poloha, orientace

Známy počáteční a koncový bod hranice (např. v aplikaci analyzující tvar kapky polymeru, kde vzorek polymeru vychází z trubice o známé poloze)

Vztah uvažované oblasti k jiným oblastem s požadovanými vlastnostmi (např. nad, uvnitř apod.)

- Příklady ze dvou oblastí použití:

Dálkový průzkum Země: Hledejte lodě ve vodě. Typické vlastnosti železničních tratí, dálnic (minimální zakřivení). Řeky se nekříží.

Lékařské: Krevní cévy jsou přibližně rovnoběžné. Vztah k anatomickému atlasu (modelový přístup).

HLAVNÍ PŘÍSTUPY K SEGMENTACI

- Prahové, podle globální vlastnosti, obvykle intenzity, kde globální znalost je reprezentován histogramem intenzity.
- Prostorová koherence (\approx shlukování "tokenů").
- Připojení např. hran, protože hrany jsou často důležitou informací o objektech
- Slučování/rozdělování oblastí. Oblasti vznikají spojením pixelů s podobnými vlastnostmi (kritérium homogeneity).
- Shoda se šablonou - detekce a přiřazení tokenů v obraze k předem známé šabloně.
- Detekce parametrického modelu, např. přímka, kružnice, elipsa, ... na základě neobvyklých jevů.
- Detekce maskování na základě neobvyklé textury.
- Segmentace oblastí pro kompresi obrazu.

PŘÍKLAD, PRAHOVÝ VLIV



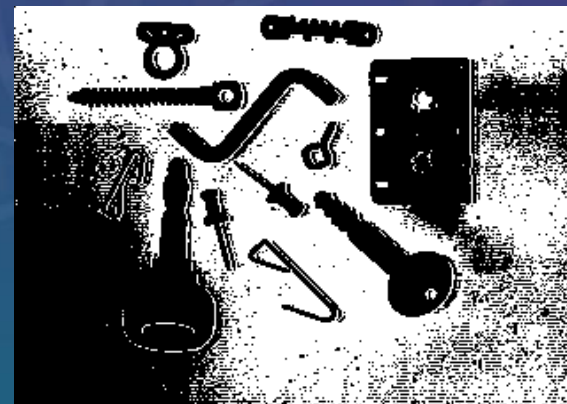
- Původní obrázek.



- Prahová segmentace.



- Příliš nízká prahová hodnota.

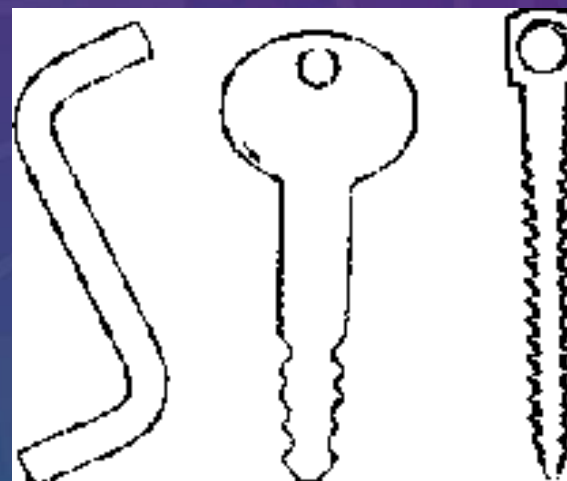


- Příliš vysoká prahová hodnota.

PŘÍKLAD. HRANIČNÍ OBLASTI POMOCÍ PRAHOVÁNÍ PÁSEM

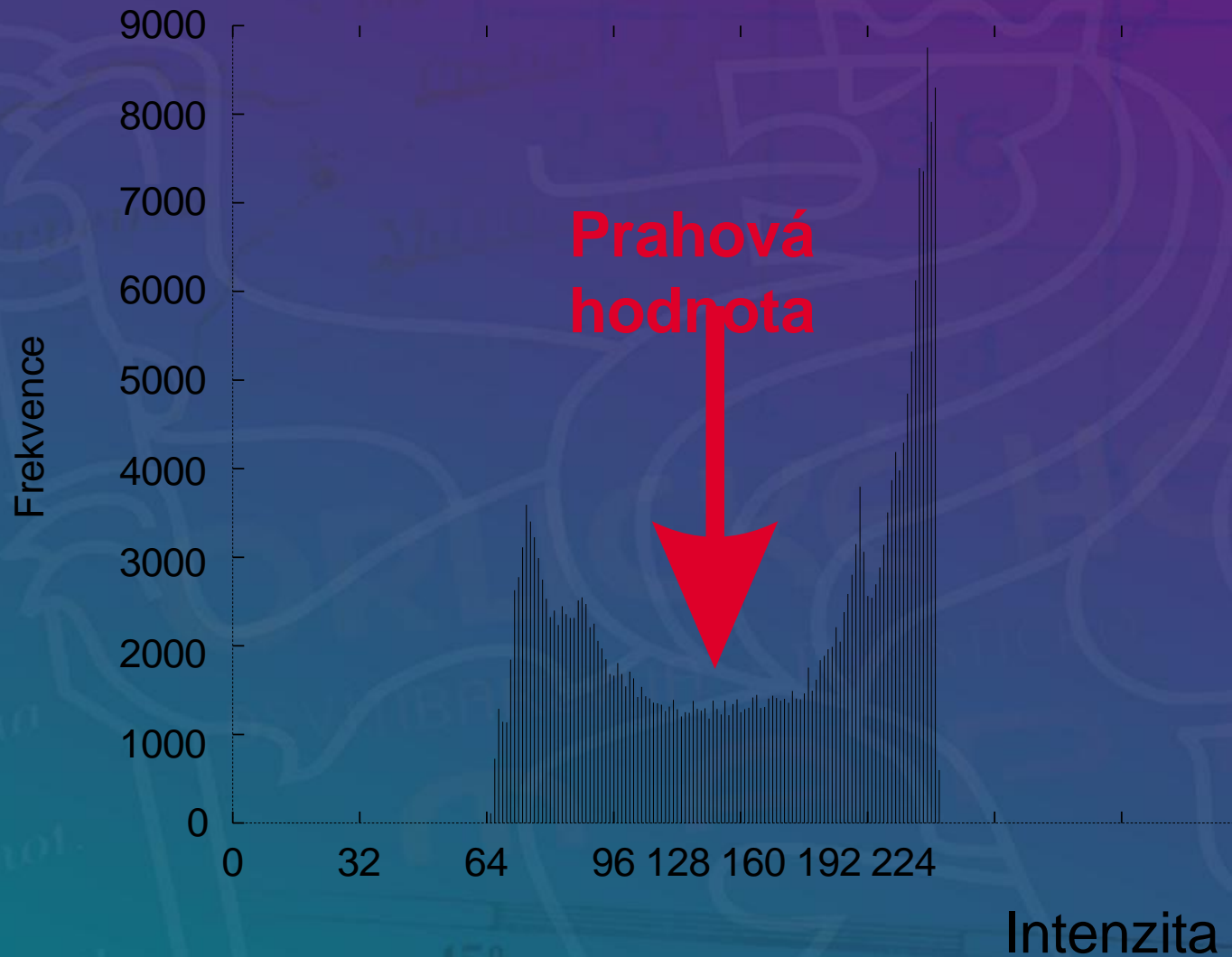


- Původní obrázek.



- Zjištěné hraniční oblasti.

Automaticky nalezený práh podle bimodálního histogramu



Optimální prahování pomocí směsi Gaussů

