

Objektové vnímání



Definice zrakového vnímání...

“procesem sestavování počitků do využitelné mentální reprezentace světa”
(Coon, 1983)

“dynamickým hledáním nejlepší interpretace dostupných údajů [...] vnímání jde za to, co nám bezprostředně říkají naše smysly” (Gregory, 1966)

“[vjemy nejsou] odhalením „toho, co je tam venku“, spíš mají povahu pravděpodobností a predikcí založených na dřívější zkušenosti” (Ittelson and Kilpatrick, 1951)

...obsahují pojmy jako reprezentace, interpretace,
porozumění

= schopnost vidět v podnětech
osobně smysluplné celky



Vnímání objektů

DETEKCE – určení výskytu objektu
v zorném poli

DISKRIMINACE – rozlišení mezi
větším počtem objektů

REKOGNICE – určení identity
objektu

Vnímání objektů

Vysoká přesnost a značná rychlost při detekci,
diskriminaci a rekognici...



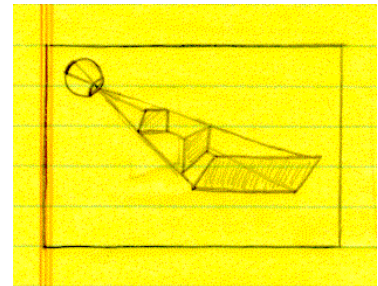
...nehledě na rozdíly v promítnuté podobě objektů
dané rozdíly v natočení, vzdálenosti, pozici vzhledem
k horizontu, vlastnostech osvětlení

Vnímání objektů

Vysoká citlivost lidského pozorovatele
kontrastuje s nízkou citlivostí počítače

- Schopnost vyrovnat se s omezeností
dostupné informace

- (i) Promítnutý obraz objektu je nejednoznačný – velikost a tvar projekce nekorespondují přímo s velikostí, vzdáleností a orientací objektu
- (ii) Promítnutý obraz objektu je nejednoznačný – světelné hodnoty v sítnicovém obraze přímo nekorespondují s barvou objektu; jsou výslednicí povrchových vlastností a parametrů osvětlení
- (iii) Objekty často nejsou vidět celé – jejich podobu je potřeba v mysli interpolovat



Percepční organizace

Nevnímáme izolovaná senzorká data, resp. jednotlivé prvky zorného pole, nýbrž objekty / jejich části / jejich skupiny

Počátek procesu zpracování podnětové informace = určení, které části sítnicového obrazu patří k sobě

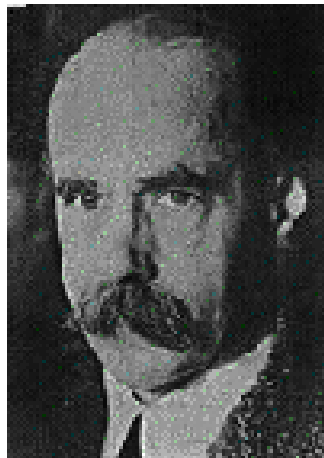
>>> teoreticky bezpočet variant dekompozice a integrace elementů

>>> reálně vybíráme jeden způsob

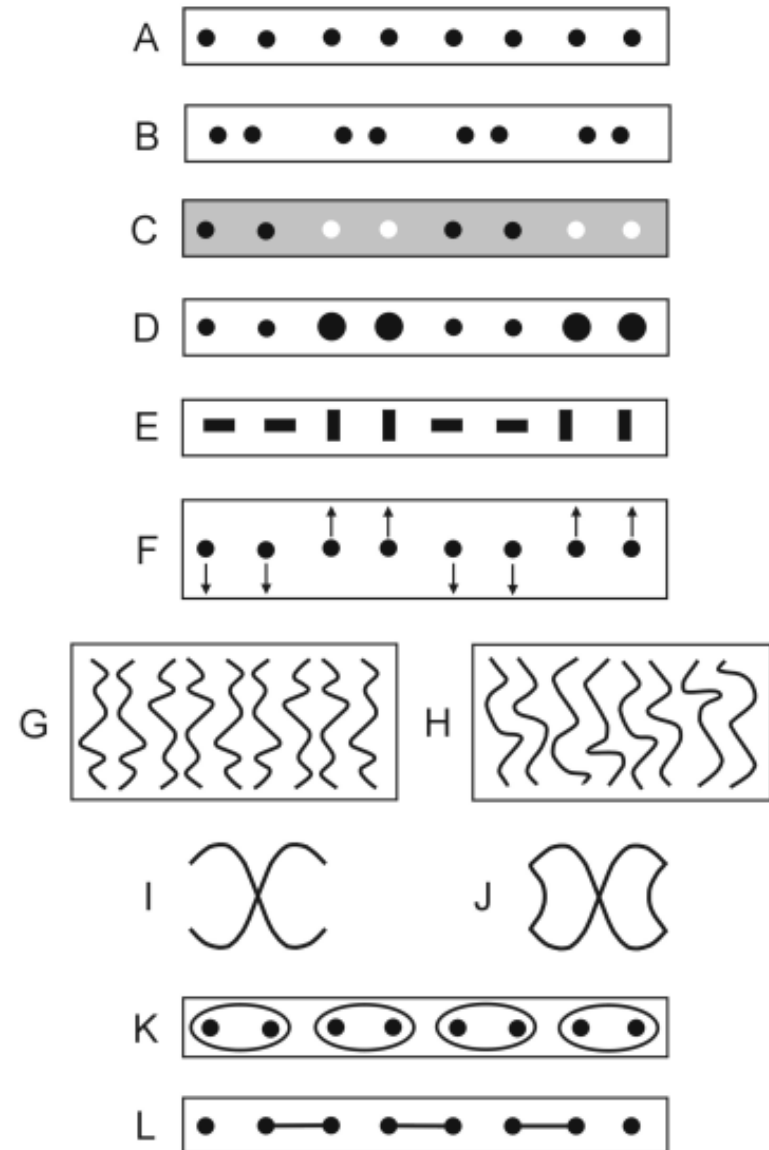
Pravděpodobnost seskupování určitých prvků k sobě

Percepční organizace

Max Wertheimer (Gestalt Ψ):
jaké vlastnosti podnětu
rozhodují o seskupování



Principy percepční organizace



Percepční organizace

Organizační principy prostupují naši každodenní
percepční zkušenost

Jejich následování se osvědčuje a je fylogeneticky správné
Výjimečně skutečnosti odpovídá „méně pěkné“ uskupení a
výjimečně malá rozmanitost zorného pole organizaci brání



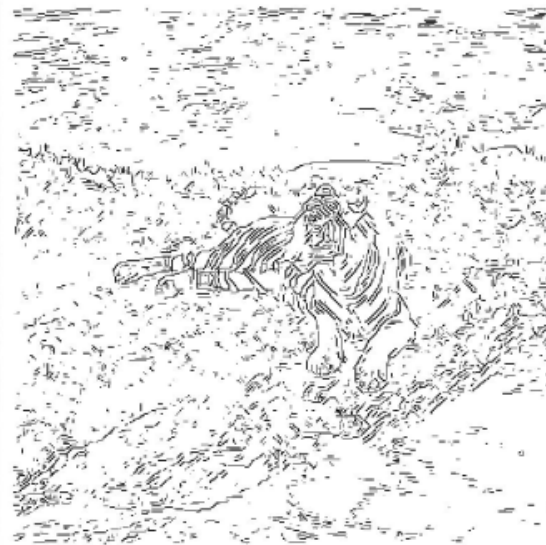
Percepční organizace

Síticový obraz je spojitý, neobsahuje „elementy“

Seskupování musí předcházet segmentace

Uniformní propojenost: rozčlenění obrazu do menších oblastí s uniformními vizuálními charakteristikami

Detekce hran: hrana v místě prudké změny jasové hodnoty
– segmentace do množství propojených oblastí



Percepční organizace

Proces oddělení figury od pozadí

Hrana = místo přechodu/setkání dvou oblastí sítnicového obrazu

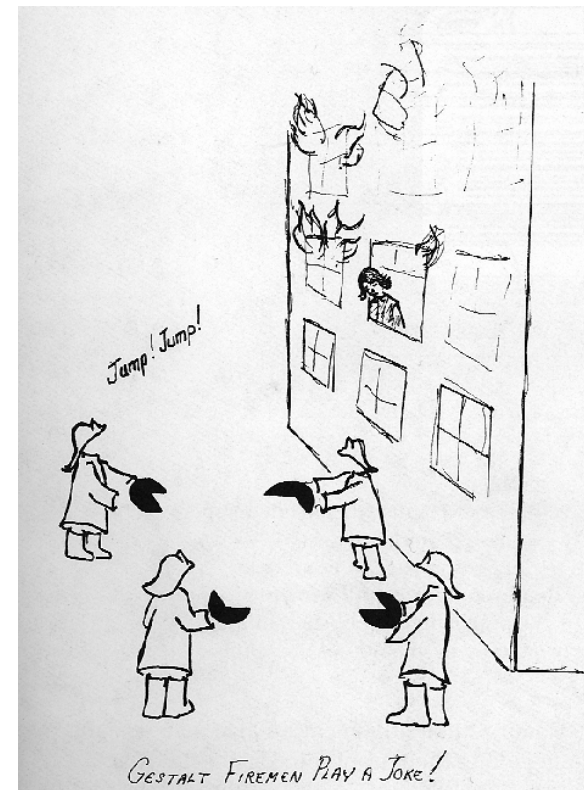
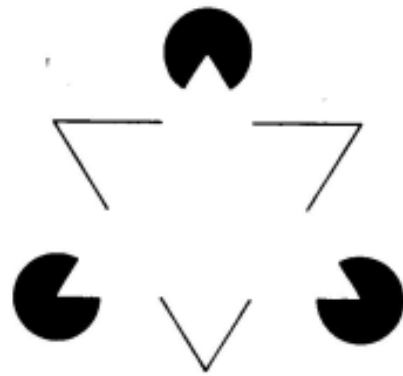
Faktory zvyšující či snižující P , že se daná oblast stane figurou, opticky ohraničeným Gestaltem:

Figurou se spíš stane oblast, která je (a) menší než sousední oblast, (b) symetrická, (c) konvexní a (d) uzavřená. Podle výsledků moderního výzkumu rovněž oblast, která je (e) více kontrastní (O'Shea, Blackburn & Ono, 1994), (f) oblast s vyšší prostorovou frekvencí (Brown & Weisstein, 1988) a (g) oblast tvořená oddělenými podoblastmi (Hoffman & Singh, 1997).

Soustředění se na zajímavý, informačně nosný detail,
úspora energie

Percepční organizace

Paradoxní efekty



Percepční organizace

WHAT'S ON A MAN'S MIND



Rozpoznávání objektů



Různost témat a přístupů

- ❑ Popis procesu rozpoznávání
 - ❑ experimentální psychologie
 - ❑ počítačové vidění
- ❑ Zpracování v mozku
 - ❑ neurovědy
- ❑ Poruchy rozpoznávání
 - ❑ neuropsychologie

Rozpoznávání objektů

Podmínka úspěšného rozpoznávání: existence paměťového asociačního skladu, obsahujícího různé podoby objektu, s nimiž je aktuálně vnímaný podnět srovnávaný

Je to ale podmínka jediná a dostačující?

NE problémy umělých rekogničních systémů

NE schopnost rekognice u zvířat s nízkou procesuální kapacitou

=> Prosté poměrování shody promítnutého obrazu se zástupcem v paměti nemůže obsáhnout variabilitu promítnutých podob objektu

Strategie musí být komplexnější a robustnější !!!

Rozpoznávání objektů



Mnohost podob objektu při promítnutí \leq
extrinsické faktory (podmínky pozorování)

Nakolik pozorovatele při pozorování ovlivňuje

- Perspektiva
- Proměny tvaru
- Kontext
- Částečné zakrytí
- Světelné podmínky

Rozpoznávání objektů

Challenges 1: view point variation



Michelangelo 1475-1564

slide by Fei Fei,
Fergus & Torralba

Rozpoznávání objektů

Challenges 2: illumination



22

slide credit: S. Ullman

Rozpoznávání objektů

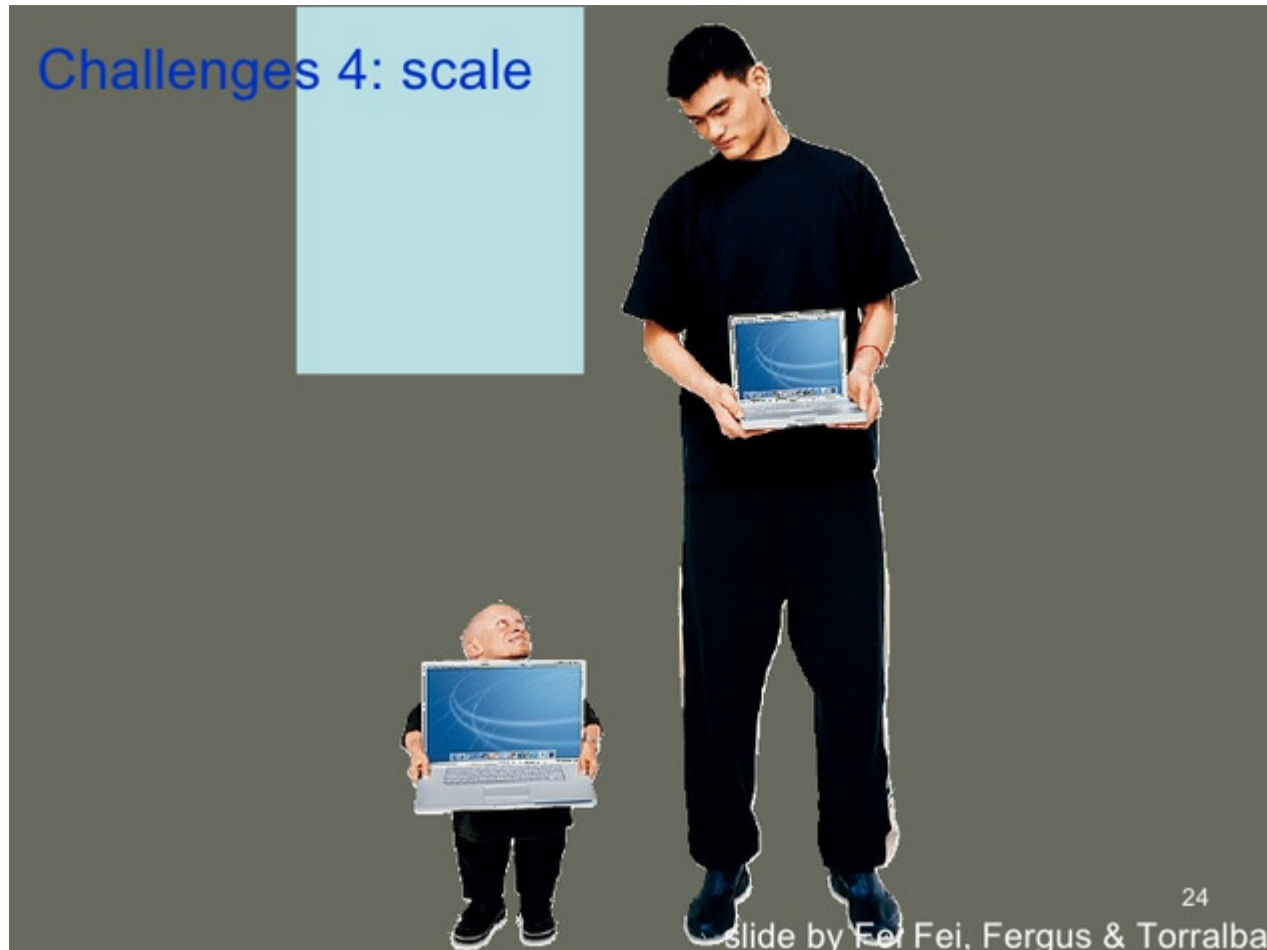
Challenges 3: occlusion

slide by Fei Fei, Fergus & Torralba

Magritte, 1957

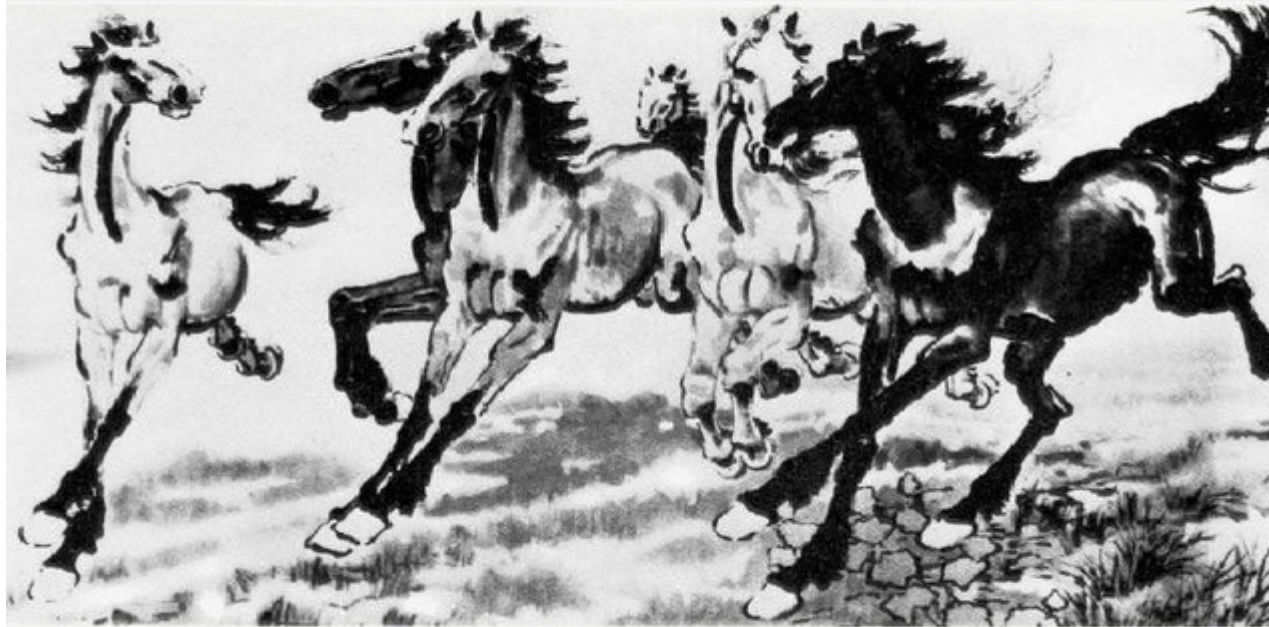


Rozpoznávání objektů



Rozpoznávání objektů

Challenges 5: deformation



Xu, Beihong 1943

25

slide by Fei Fei, Fergus & Torralba

Rozpoznávání objektů

Challenges 6: background clutter

slide by Fei Fei, Fergus & Torralba

Klimt, 1913



Rozpoznávání objektů

Challenges 7: object intra-class variation



slide by Fei-Fei, Fergus & Torralba

Rozpoznávání objektů

Challenges 8: local ambiguity

slide by Fei-Fei, Fergus & Torralba

Rozpoznávání objektů

Vliv perspektivy

Palmer, Roschová & Chase (1981):

skupina A: určit typičnost podoby, skupina B: klasifikace objektů; výsledek: lineární vztah mezi typičností a RČ; mohly být výsledky ovlivněné známostí kanonických podob?



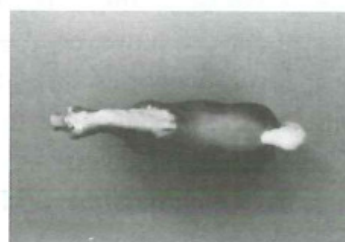
BEST (1.60)



FRONT-SIDE-TOP (2.80)



FRONT-TOP (4.80)



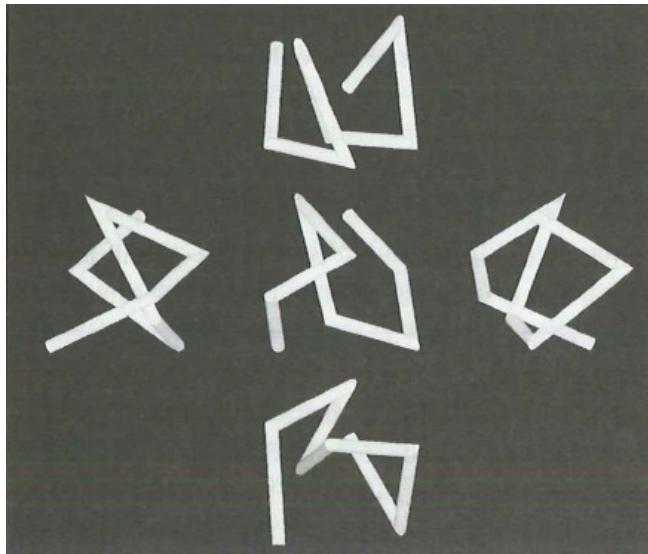
TOP (6.36)

Rozpoznávání objektů

Vliv perspektivy

Edelman & Bühlhoff (1992):

Neznámé objekty (zprohýbané kancelářské spony);
1. učení + 2. rekognice; výsledek: nárůst RČ s
odchýlením sledované podoby od naučené

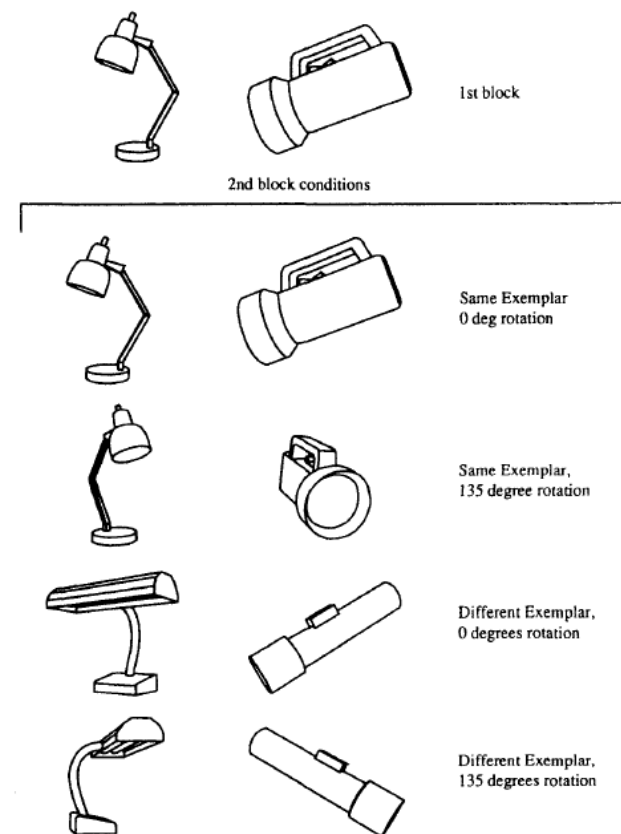


Rozpoznávání objektů

Vliv perspektivy

Biederman & Gerhardstein (1993):

Metoda primingu; výsledek: nárůst RČ pouze u objektů, u nichž při změně perspektivy mizí kritické, nenáhodné znaky (geony)

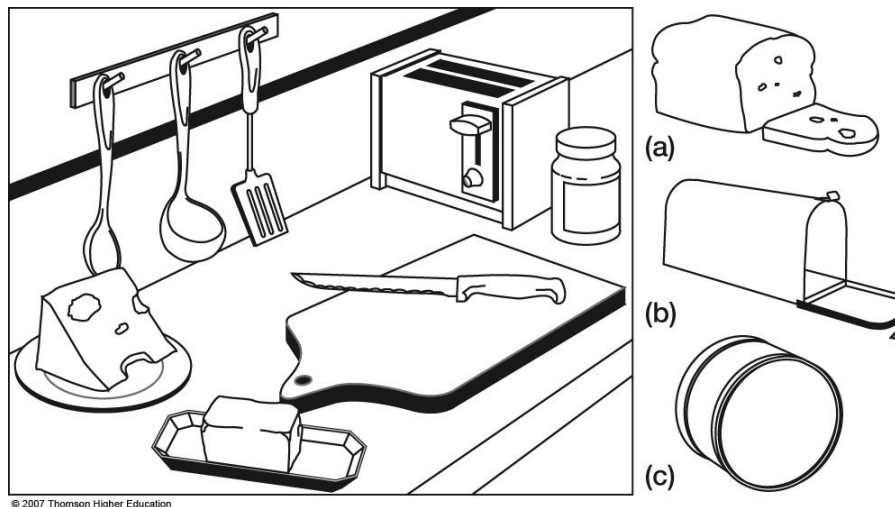


Rozpoznávání objektů

Vliv kontextu

Palmer (1975):

1. tématická scéna + 2. cílový objekt; ten vzhledem ke scéně buď konzistentní, nebo ne leč tvarově podobný (paradox!), nebo ne a tvarově odlišný; pomalejší rekognice a vyšší chybovost při paradoxních podmínkách



Rozpoznávání objektů

Vliv kontextu

Davenportová a Potterová (2004):

Fotografie scény s figurou buď adekvátní nebo nepatřičnou (doklíčovanou); prezentace 80 msec; úkol: rozpoznání (a) F, (b) G, (c) F+G; výsledek: i při krátké prezentaci je F zpracovávána ve vztahu k G



Teorie rozpoznávání objektů

- Možnost porovnání sledované/promítnutého objektu s předobrazem v mysli předpokládá odlišení vlastní (intrinsické) informace o objektu od ne-vlastních (extrinsické)
- Jak toho dosáhne? 2 skupiny teorií:
 - ✓ Teorie popisu struktury (Marr, Nishihara, Biederman) – předpoklad nezávislosti procesu rozpoznání na perspektivě
 - ✓ Teorie rozpoznání založeného na sítnicovém obrazu (Tarr, Bulthoff, Edelman, Lawsonová) – předpoklad závislost na perspektivě

Teorie rozpoznávání objektů

Teorie popisu struktury

Jakou má objekt strukturu a jak ji nejlépe (z hlediska operacionalizace) popsat

Rozpoznávání probíhá v posloupnosti (1) detekce hran (lokálních sítnicových znaků) --- (2) kombinace do složitějších a komplexnějších struktur --- (3) srovnání těchto struktur (definujících podnět) s reprezentací (v podobě struktur)

Vychází z pojetí v počítačovém vidění (např. Binford, 1971)

Teorie rozpoznávání objektů

Teorie popisu struktury

Model Marra a Nishihary (1978; Marr, 1982)

Předpoklady:

1) Informace nutná k r

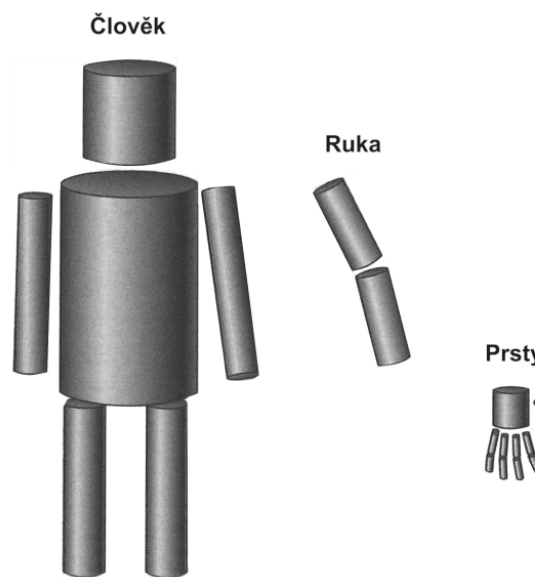
2) I. je extrahovatelné]

3) Jedna reprezentace]

Jednoznačná, neměnná i

Komponenty = objemov
vzájemných orientací

Rozklad obrazu, složení



ecifikovaná v obraze

orování

různé reprezentace se liší

pení komponent

a různých

Teorie rozpoznávání objektů

Teorie popisu struktury

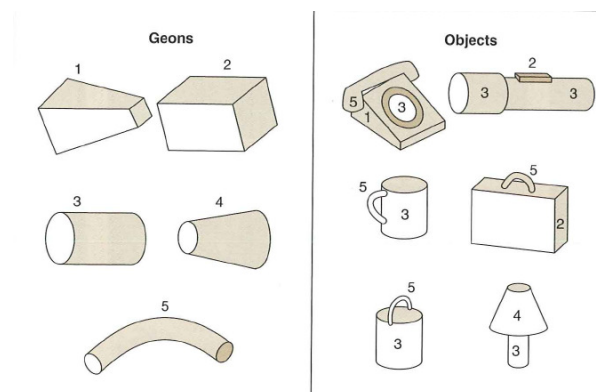
Model Biedermana (1985, 1987)

Návaznost + empirická podpora

Rozpoznání postavené na srovnání kompozice jednoduchých komponent (geonů) s reprezentací (=porovnání strukturního popisu)

Komponenty nemají jednotný tvar – klasifikace: rovné vs. zakřivené hrany, rovná vs. zakřivená hlavní osa, šířka ve směru hlavní osy, osová symetrie => 36 tvarově jedinečných geonů

Dvojice, trojice geonů „skládají“ širokou varietu možných tvarů



The set of geons is generated by variations in the production function for generalized cylinders that produce viewpoint-invariant (= nonaccidental) shape differences

1. Cross Section: Straight vs. Curved

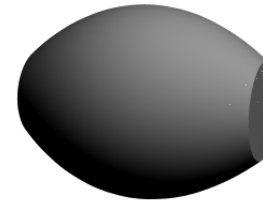
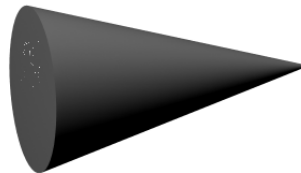


2. Axis: Straight vs. Curved

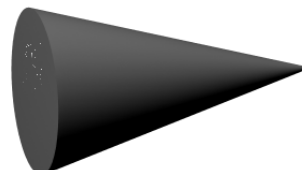


3. Size of Cross Section:

Constant (parallel sides) vs. Expand vs. Expand & Contract vs Contract & Expand



4. Termination of Geon when Nonparallel: Truncated vs. Pointed vs. Rounded



Teorie rozpoznávání objektů

Teorie rozpoznání založeného na podobě obrazu

NE jediná abstraktní reprezentace popisující strukturu,
ale sada reprezentací ve formátu sítnicového obrazu,
informace o povrchových vlastnostech

Více různých reprezentací jednoho objektu

Vznik v návaznosti na...

- experimentální výsledky potvrzující závislost úrovně rozpoznání na podmínkách pozorování, především úhlu pohledu
- úspěchy rekogničního systému Poggia a Edelmana (1990)

Teorie rozpoznávání objektů

Teorie rozpoznání založeného na podobě obrazu

Kolik uchovávaných podob? Jakým způsobem k nim vnímané podoby přiřazujeme?

Tarr & Pinker (1989): v mysli jen několik typických podob, známých ze života; mentální rotace vnímané podoby; čím delší nutná rotace, tím delší rekognice

Poggio & Edelman (1990) aj.: určování podobnosti vybraných znaků obrazu; k úspěšné interpolaci stačí disponovat třemi podobami

Teorie rozpoznávání objektů

Informace popisující strukturu

NEBO

Informace obsažená v sítnicovém obrazu?

Aktuální otázka v 90. letech, v poslední době pokusy o integraci

Východisko: v některých podnětových situacích je vnímání více a v jiných méně ovlivněné úhlem pohledu

Hummel & Stankiewicz (1998, 2002): při možnosti možnosti porovnání 2D obrazů přímé a rychlé rozpoznání, při nemožnosti nutnost důkladného a pomalého popisu struktury

Rozpoznávání objektů v mozku

Projekce ventrálního zrakového proudu – spodní část spánkového laloku – inferotemporální kortex (IT)

IT je klíčový pro rozpoznávání:

- ❑ Poškození či odstranění vede ke zhoršení až neschopnosti rozpoznávat objekty
- ❑ Elektrofyzilogická měření a zobrazovací techniky zjistili silnou aktivaci v IT při prezentaci komplexních podnětů (jednoduché ve V1 a V2)

Rozpoznávání objektů v mozku

Od poloviny 90. let využívání zobrazovacích technik

- hlavně fMRI (měření rozsáhlejších oblastí při zachování vysokého prostorového rozlišení)

1. etapa výzkumů: lokalizace oblastí v IT výběrově reagujících na objekty a tváře + určení jejich funkční specializace – tváře, části těla, písmena, místa

Novější výzkum: sledování dalších determinant rozpoznávání

Rozpoznávání objektů v mozku

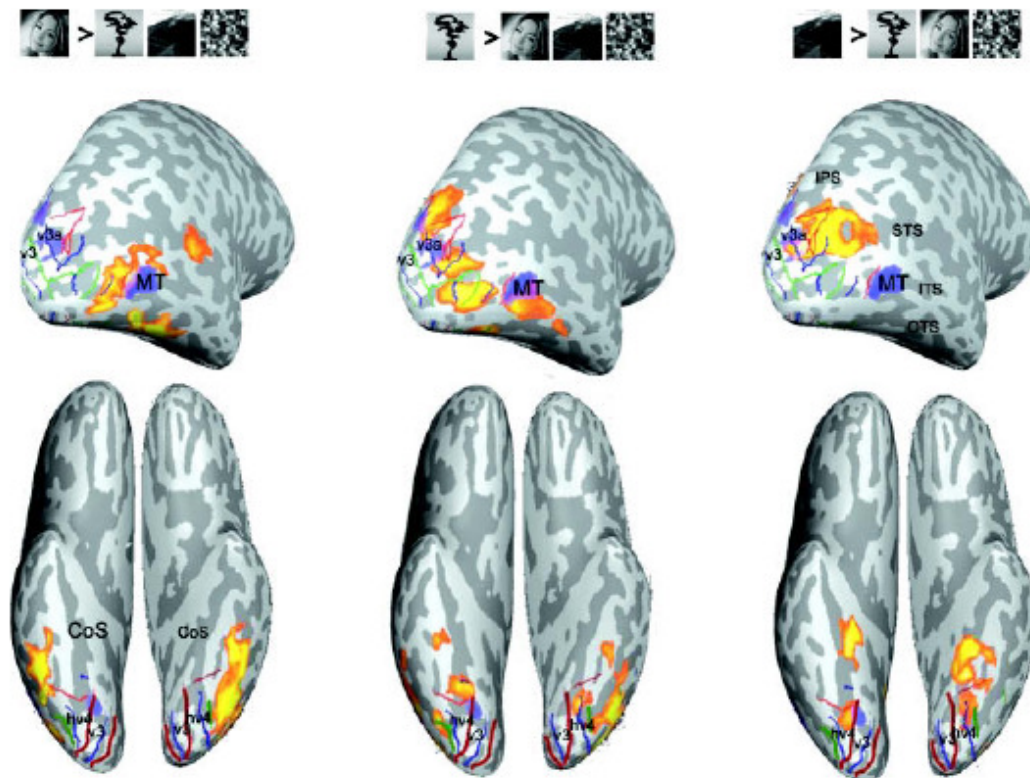
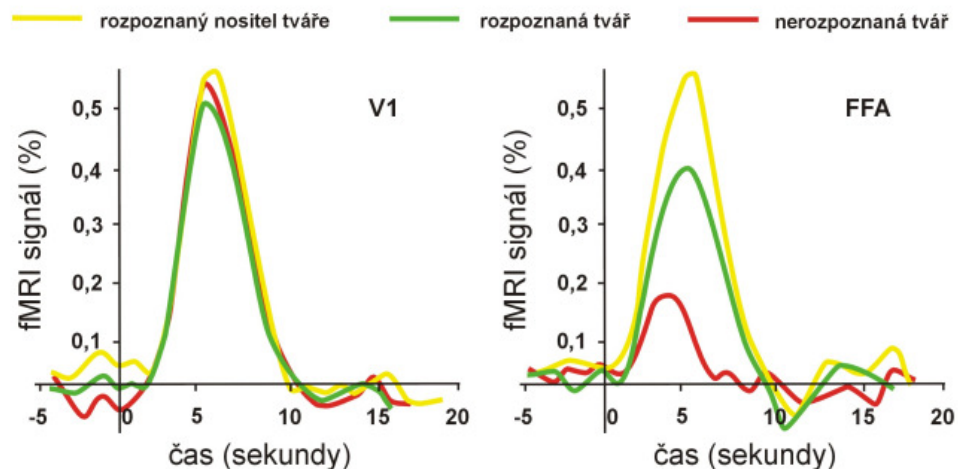


Figure 6 Face-, object-, and place-selective regions in the human brain displayed on an inflated surface representation of subject JHT. Icons indicate the comparison done in the statistical tests. *Left*: areas responding more strongly to faces than objects, places, or textures. *Center*: areas responding more strongly to objects than faces, places, or textures. *Right*: areas responding more strongly to places (scenes) than faces, objects, or textures. Yellow and orange indicate statistical significance: $p < 10^{-12} < p < 10^{-6}$. Colored lines indicate borders of retinotopic visual areas. Blue indicates area hMT+, defined as a region in the posterior bank of the inferotemporal sulcus that responds more strongly to moving versus stationary low-contrast gratings (with $p < 10^{-6}$).

Rozpoznávání objektů v mozku

Grill-Spectorová a kol. (2004):

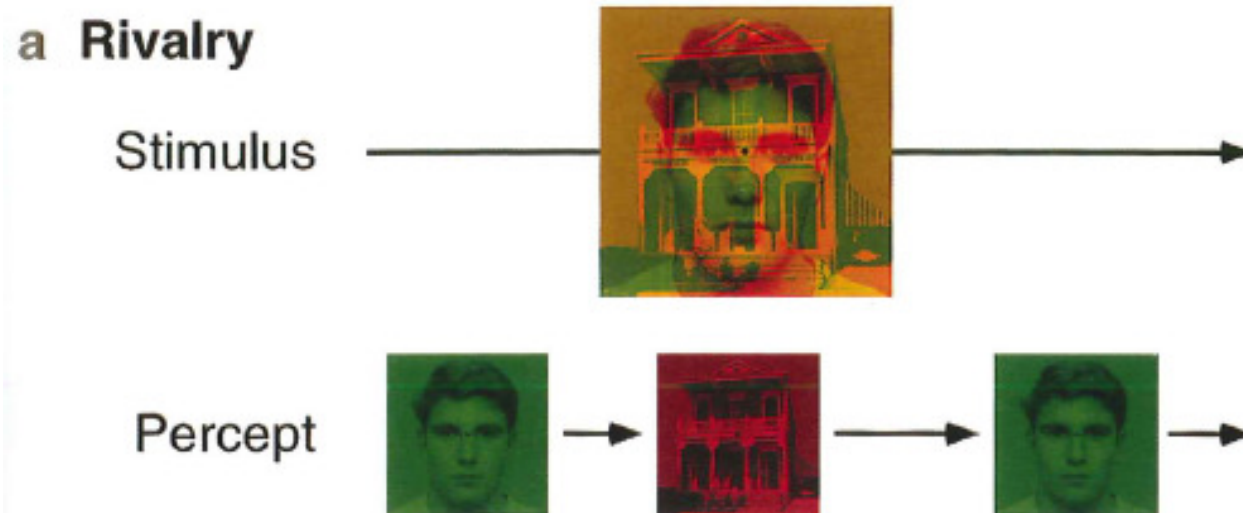
Vztah aktivity a úspěšnosti rozpoznávání; aktivita ve V1 a FFA; prezentace známé tváře, neznámé tváře nebo vzoru; 50 msec (—> vysoká chybovost); výsledek: nejsilnější reakce při rozpoznání, slabší při poznání alespoň tváře, nejslabší při selhání



Rozpoznávání objektů v mozku

Tong a kol. (1998):

Nejednoznačné podněty (binokulární rivalita) – změna vjemu bez změny podnětu; jedno oko: tvář, druhé oko: dům; aktivita ve FFA a PPA; změny aktivace (nárůst a pokles) sladěné s překlápěním vjemu



Poruchy rozpoznávání

Neuropsychologický výzkum zrakové agnozie

Následek otravy oxidem uhelnatým, rtutí, srdeční zástavy, úderu do hlavy, kortikální atrofie

Snížená schopnost rozpoznávání se projevuje nesprávným pojmenováním předkládaných objektů, chybami při rozřazování do kategorií, potížemi při překreslování

Zrakové funkce (ostrost vidění, citlivost ke kontrastu, výpadky zorného pole) nejsou postižené. Podobně ani reprezentace v mysli.

Poruchy rozpoznávání



Zraková agnozie je „selhání rekognice, které nemůže být přisouzené elementárním senzoryckým deficitům, zhoršení mentálního stavu, narušení pozornosti, nesprávnému pojmenovávání v důsledku afázie nebo neznalosti prezentovaného podnětu“ (Bauer, 1993).

Poruchy rozpoznávání

Rozlišení na **apercepční** a **asociační** agnozii
(Lissauer, 1890)

Apercepční: podnět není správně vnímán

Asociační: vjem není propojený s relevantní
reprezentací

Pochybnosti o smysluplnosti uvedeného dělení

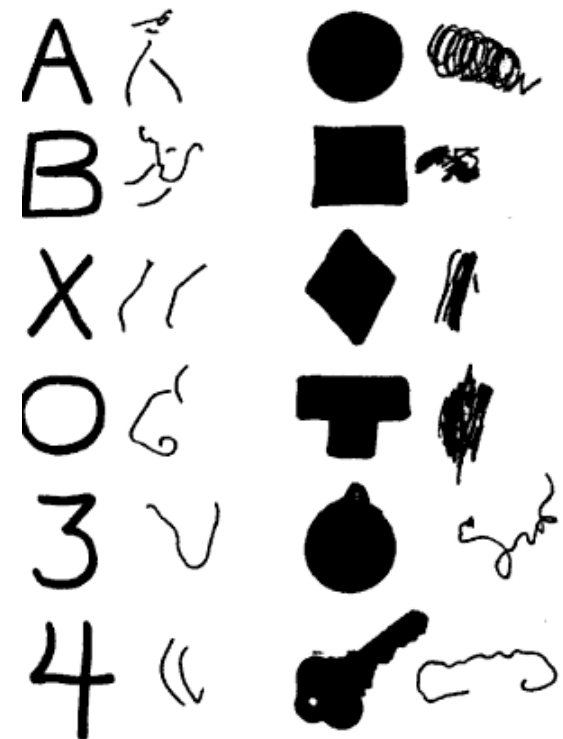
Poruchy rozpoznávání

Apercepční agnozie

Poškození v raných fázích zpracování zrakového podnětu

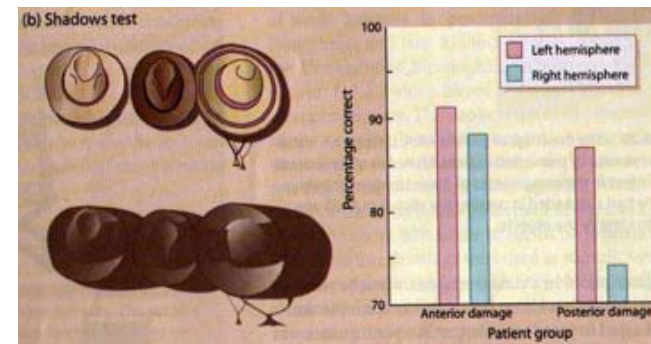
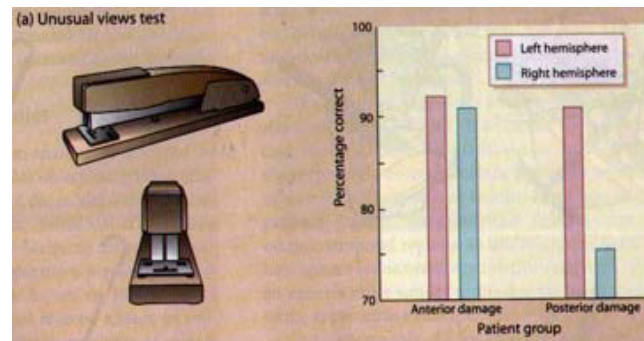
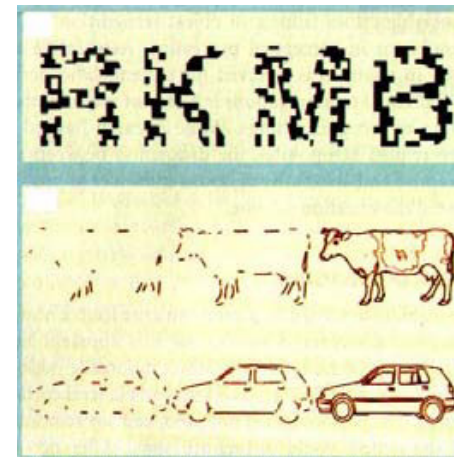
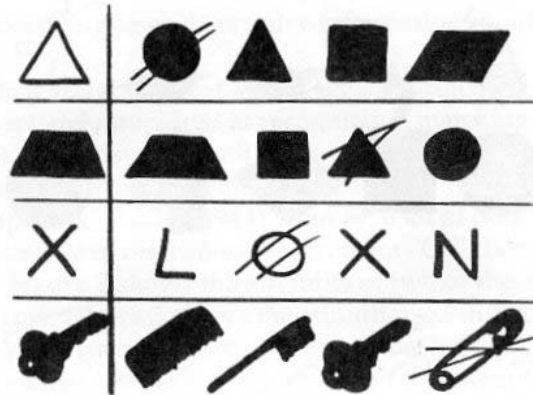
Vyšetření: problémy při rozpoznání, porovnání a překreslení předložené kresby

Problémy především v sub-optimálních podmínkách = méně obvyklé natočení, snížená kvalita osvětlení, dopad stínu na povrch, neúplné figury



Poruchy rozpoznávání

Apercepční agnozie



Poruchy rozpoznávání

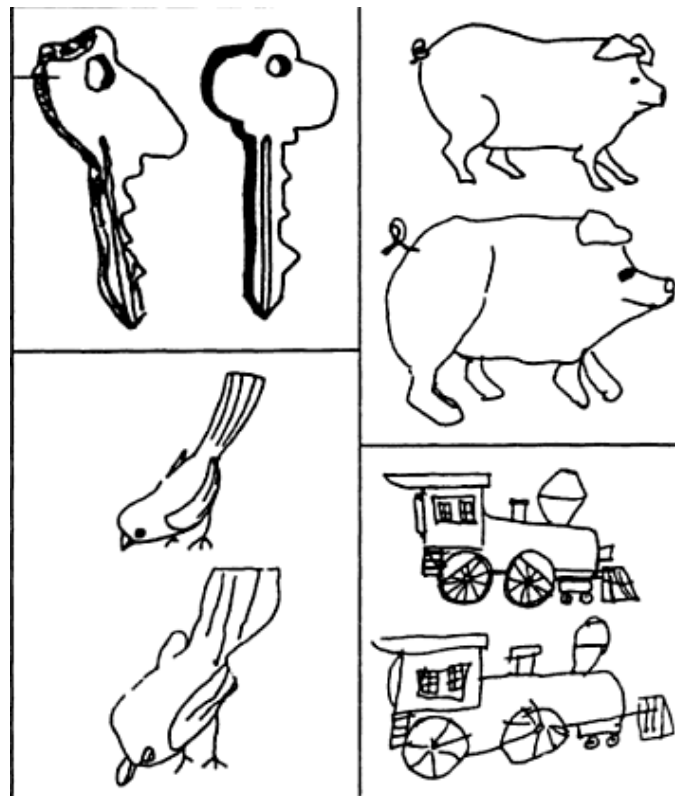
Asociační agnozie

Neschopnost asociovat
vjem s pamětí

Nenarušená schopnost
popisu vzhledu,
překreslení, porovnání
dvou obrázků

Problémy při rozřazování

Farahová: alternativní
výklad



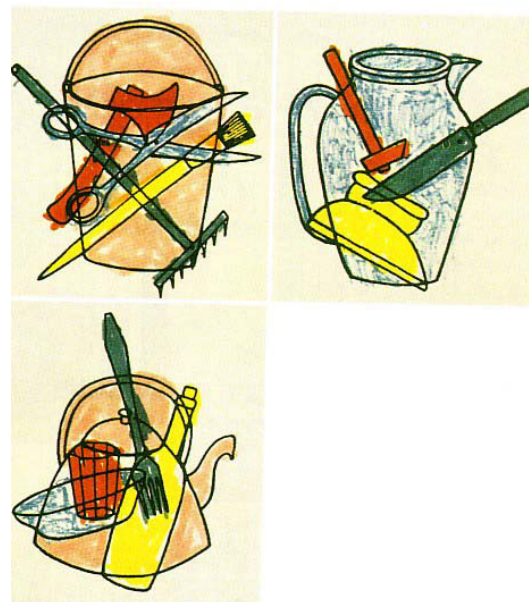
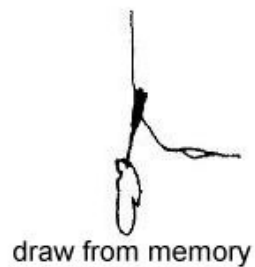
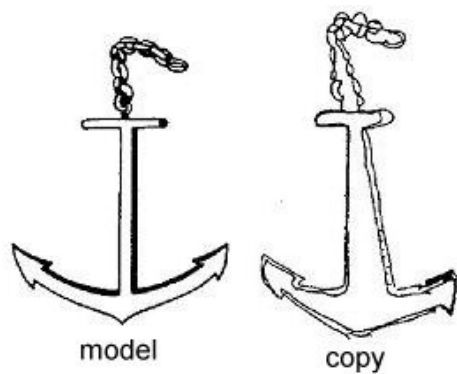
Poruchy rozpoznávání

Asociační agnozie



Poruchy rozpoznávání

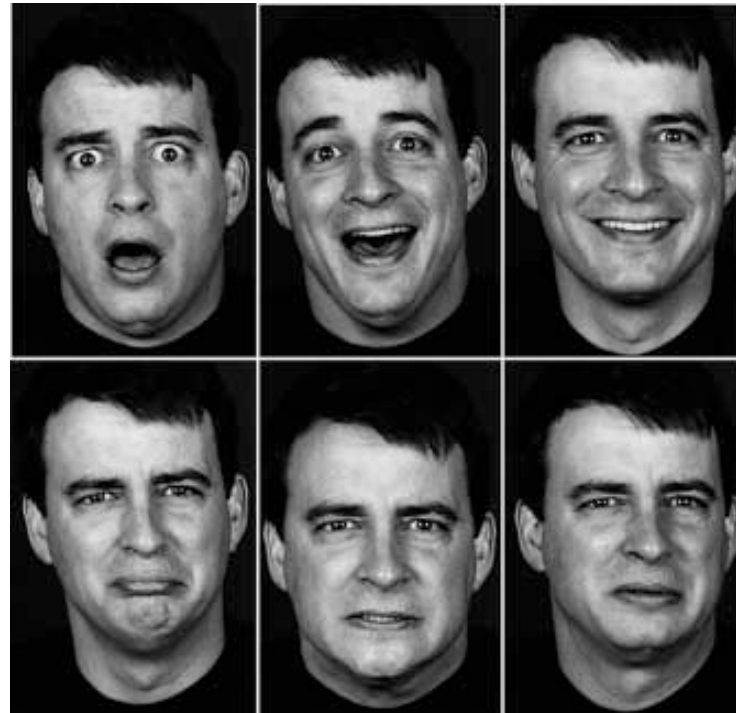
Asociační agnozie



Vnímání tváří



- Tvář je z informačního – i jiných – hlediska výjimečná
- Stačí letmý pohled na tvář a dokážeme určit identitu, odhadnout věk, pohlaví, rasu, zdravotní stav, z tváře usoudit na osobnost, momentální náladu a záměry nositele
- Více v Blažek a Trnka: Lidská tvář (2010)



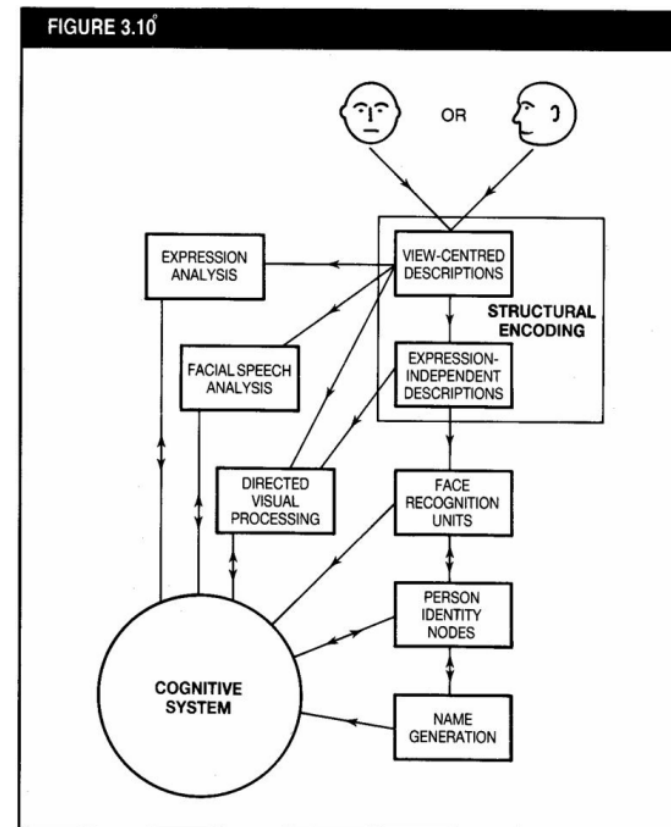
Kognitivní kapacita

- Přesnost a detailnost informací z tváře je překvapující:
 - (i) Lidské tváře jsou si navzájem velmi podobné (více než jiné typy objektů)
 - (ii) Žádná tvář ve dvou situacích nepromítá stejný obraz
- Proč tedy nemáme s rozpoznáváním a získáváním informací v životě výraznější problémy?

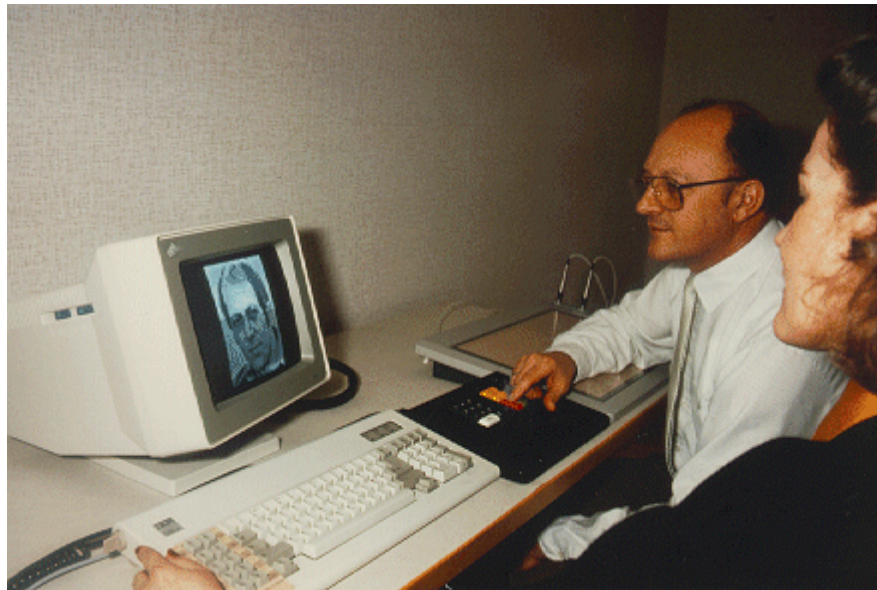
Proces rozpoznávání tváří

Model Bruceové a Younga
(1986):

Nižší (promítnutá podoba tváře)
a vyšší (dílčí informace vazující
se ke tváři) úroveň zpracování



Krátký experiment (Warringtonová, 1984)



- Kapacita se váže k optimálním (informačním) podmínkám; v suboptimálních rapidně klesá.

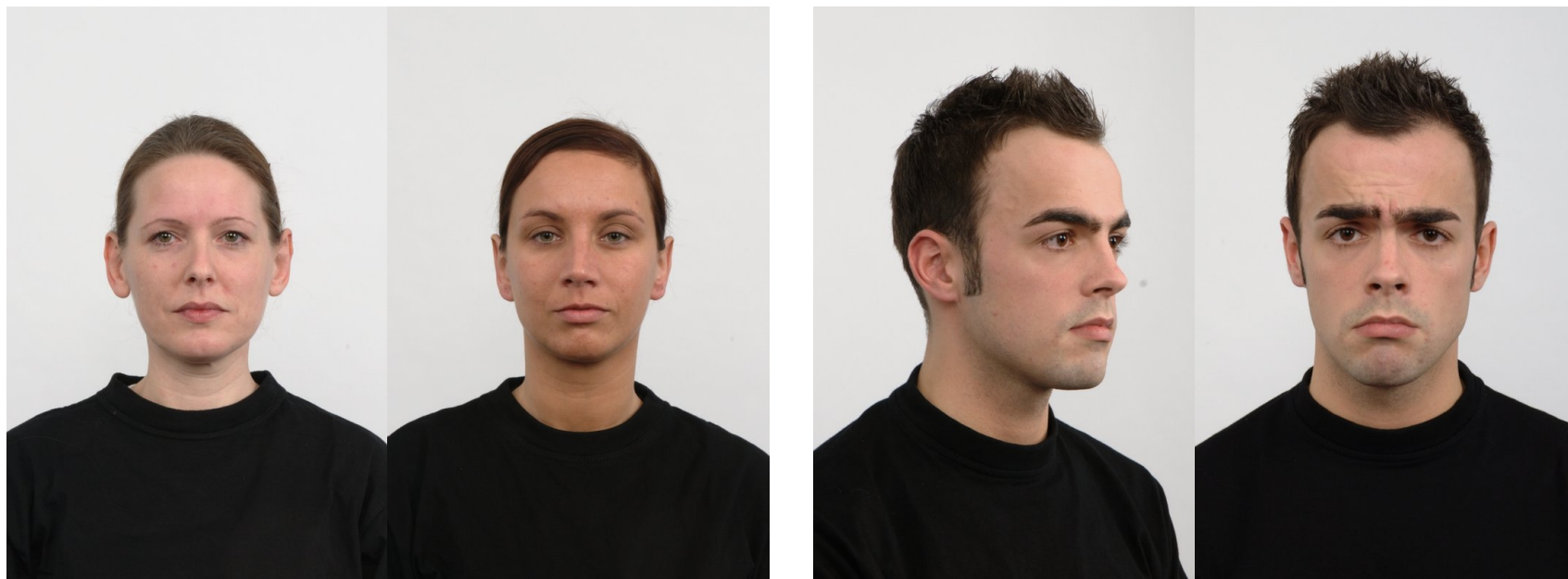
... i když ...



- Ztížení podmínek pro rozpoznávání (přirozená změna, cílená manipulace)

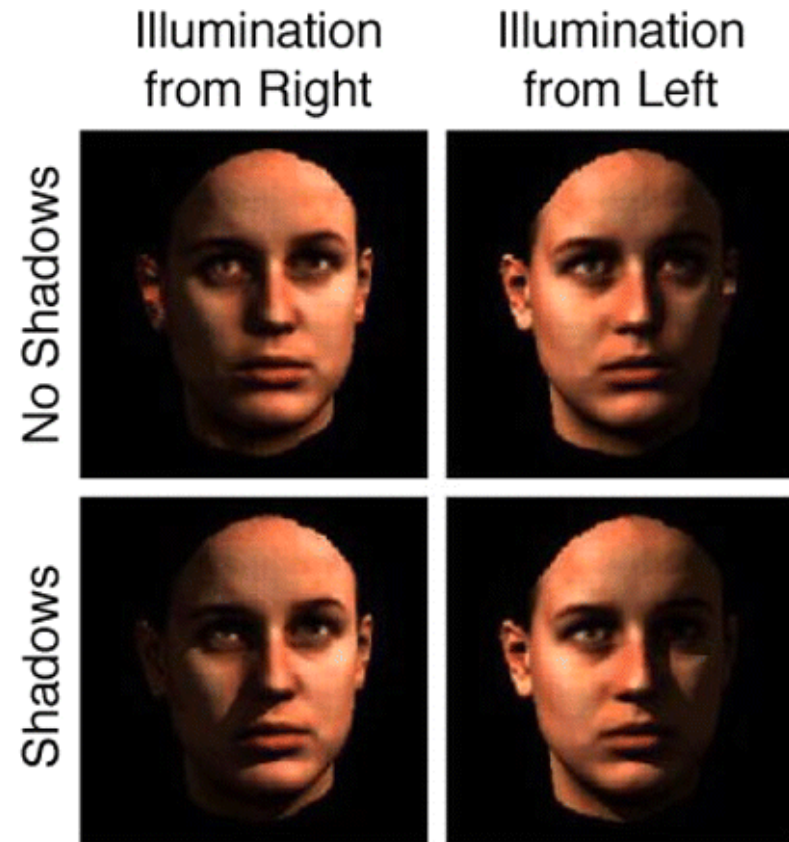
Změna natočení, změna emoce

Bruceová (1982): rekogniční úloha – shodná nebo odlišná tvář (2AFC); >90% při shodě úhlu i emoce, 76% při jiném úhlu, 61% při jiném úhlu i emoci



Změna osvětlení

Brajeová, Kersten, Tarr & Troje (1998): rekogniční úloha při prezentaci dvojice počítačově vytvořených tváří, jedna osvětlená zleva, druhá zprava; výrazný pokles



Negativ, otočení o 180°

Neschopnost zužitkovat dílčí nápovědi (stíny, pigmentace, konfigurace tvářových znaků)

Podnět v mozku zpracováváný jako ne-tvář

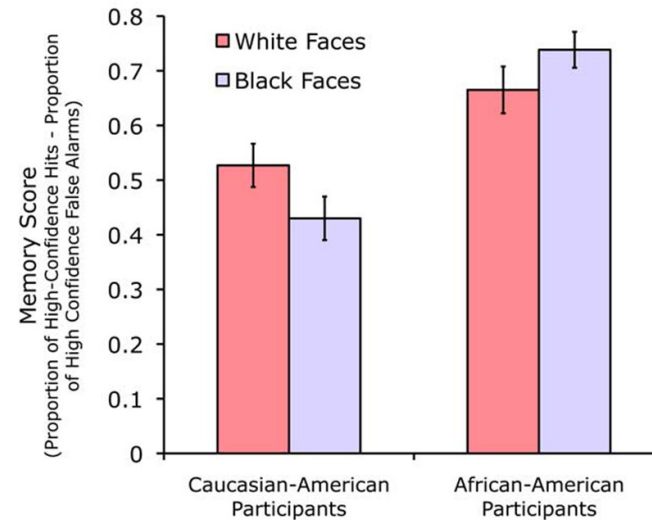
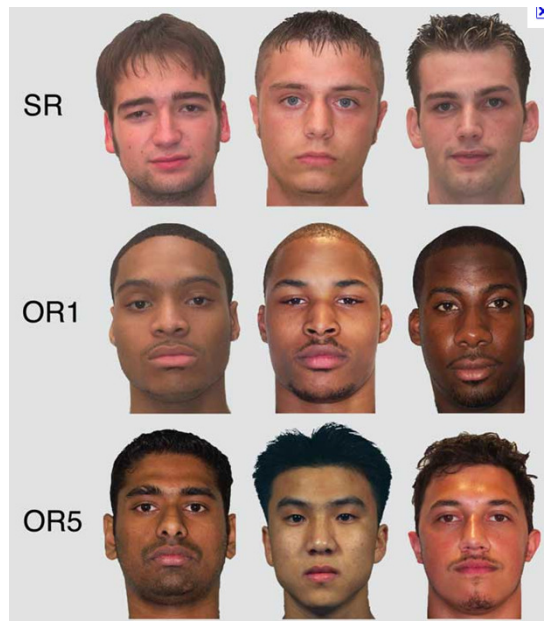




o Výkonové rozdíly nejen v závislost na podmínkách,
ale i na nositeli tváře

- stejnost/rozdílnost rasy, pohlaví, věku

o Efekt jiné rasy: příslušníky stejné rasy podle tváře
rozpoznáváme snáze a rychleji než jiné rasy



Efekt jiné rasy

Výzkumné designy: paměťové úlohy (odstup několik minut až několik dní), tachistoskopická prezentace, očitě svědectví

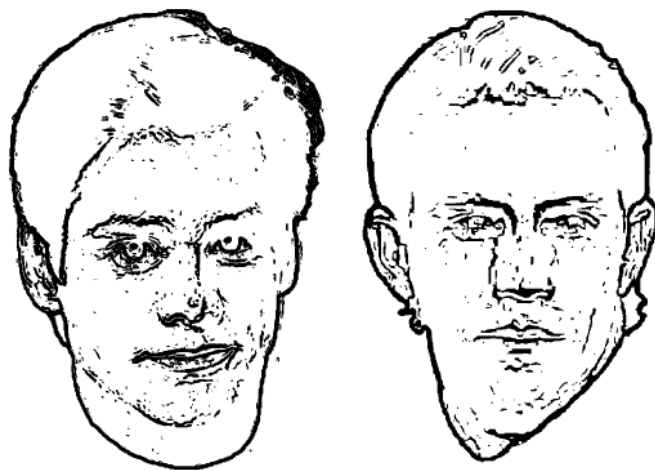
Vysvětlení: Nikoliv rozdíly v podobě, ale spíš v rozsahu a intenzitě zkušenosti (lepší diferenciaci u vlastní, homogenita u jiné). Nejednoznačné empirické doklady.
Alternativní hypotéza: rozdíl ve směřování pozornosti

Nápovědi při rozpoznávání tváře

Mnohost zdrojů informací o tváři

- tvar (umístění tvářových znaků, vzdálenosti mezi nimi, tvarová specifika)
- pigmentace (barva, jas, textura, průsvitnost tváře)

Ad object rec.: tvar > barva, ad face rec.: tvar = barva



Nápovědi při rozpoznávání tváře

O'Tooleová, Vetter a Blanz (1999)



Shape



Pigmentation



Shape +
Pigmentation



bflmpsvz

Jak jsou při rozpoznávání zapojené dílčí nápovědi (diskrétní údaje)?

Analytický přístup: vnímání = extrakce informací ze struktur tváře, které jsou diagnostické (tvář vnímáme podle toho, jaký kdo má nos, ústa atd.)

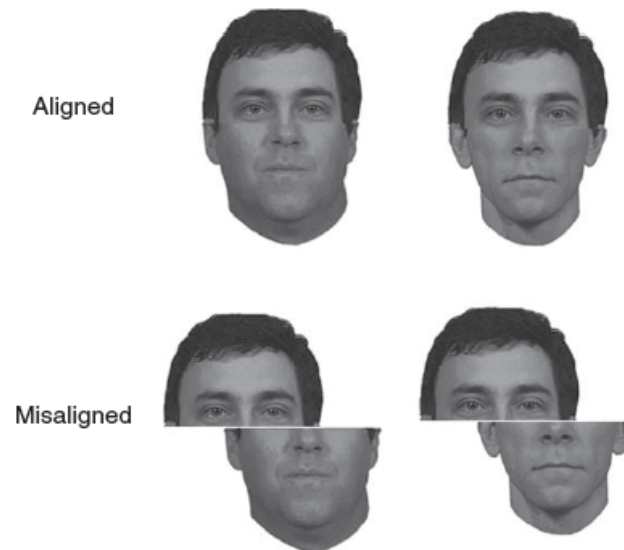
- PROS: (i) oční pohyby, (ii) schopnost identifikace z mála

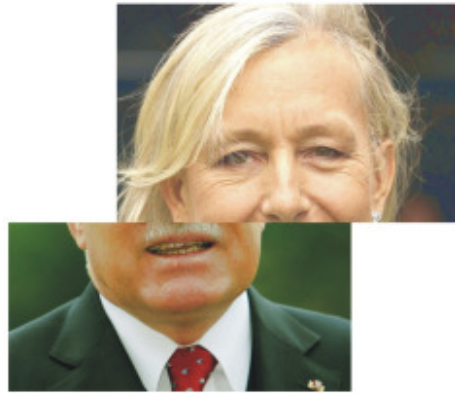
Holistický přístup: struktury vnímáme jako integrovaný celek; určující vliv kontextu obličeje



Holistický přístup

Young, Hellawellová & Hay (1987): rozpoznání patřičnosti či nepatřičnosti dvou polovin kompozitní fotografie; při lícování nižší přesnost, spolehlivost a delší RČ než při odsunutí; interpretace: líčující tvář vnímaná holisticky, posunutá tvář analyticky





Holistický přístup

Hole, Georgeová, Eavesová & Rasek (2002): komprese výšková, stranová, shear; změna podoby dílčích struktur i vzdálenosti, zachovaná konfigurace; žádný problém; problém při deformaci jen části obličeje





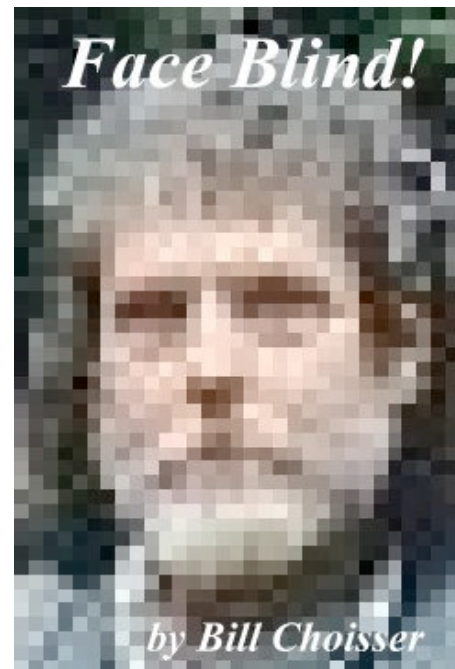
Holistický přístup

Sinha... (????): schopnost identifikace tváře navzdory nemožnosti vyhodnotit jednotlivé tvářové znaky



Prosopagnózie

Výrazně narušená schopnost rozpoznávat tváře



Prosopagnózie

Strategie rozpoznávání

Skoro vždy kombinovaná

Implicitní schopnost

Testy

Vrozená vs. získaná

Vnímání scén

Mary Potter (1976)

Mary Potter (1975, 1976) demonstrated that during a rapid sequential visual presentation (100 msec per image), a novel picture is instantly **understood** and observers seem to comprehend a lot of visual information



Objects and Scenes

Stimuli from Hock, Romanski, Galie, and Williams (1978).



TYPE I



TYPE II



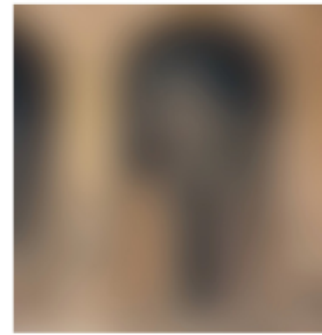
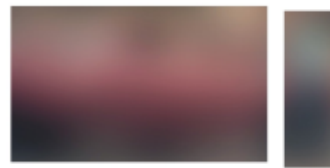
TYPE III



TYPE IV

Biederman's violations (1981):

1. *Support* (e.g., a floating fire hydrant). The object does not appear to be resting on a surface.
2. *Interposition* (e.g., the background appearing through the hydrant). The objects undergoing this violation appear to be transparent or passing through another object.
3. *Probability* (e.g., the hydrant in a kitchen). The object is unlikely to appear in the scene.
4. *Position* (e.g., the fire hydrant on top of a mailbox in a street scene). The object is likely to occur in that scene, but it is unlikely to be in that particular position.
5. *Size* (e.g., the fire hydrant appearing larger than a building). The object appears to be too large or too small relative to the other objects in the scene.



Scene perception: change blindness

failure to notice changes in natural scenes

<http://www.sinauer.com/wolfe/chap8/flickerF.htm>

<http://viscog.beckman.uiuc.edu/grafs/demos/12.html>



<http://viscog.beckman.uiuc.edu/grafs/demos/11.html>

suggests attention acts to solidify only part of our representation
of the world

rather than filtering out unattended material,
attention gives life to attended

