

Geometrická morfometrie

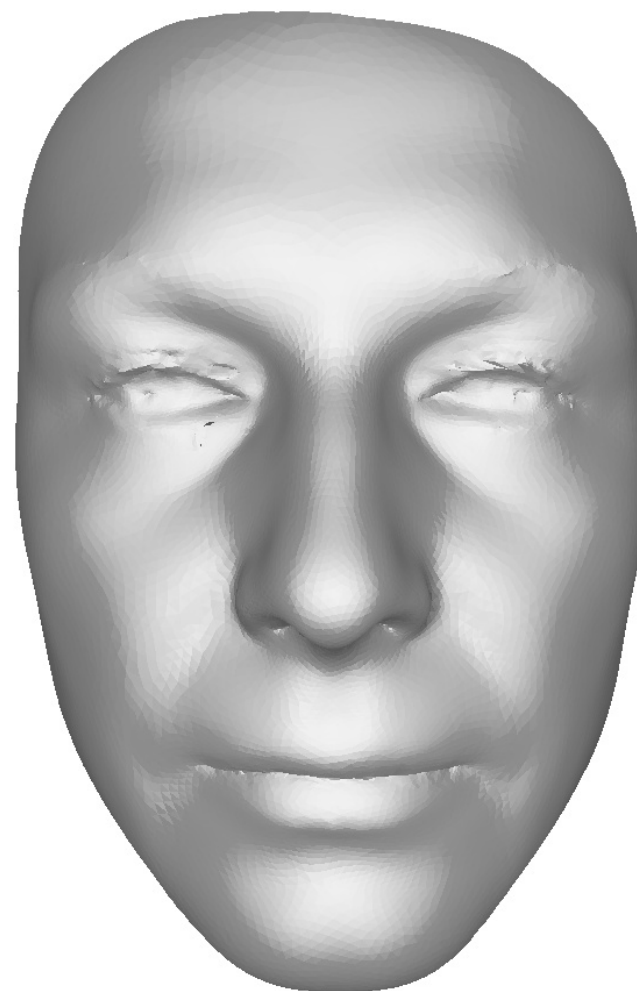
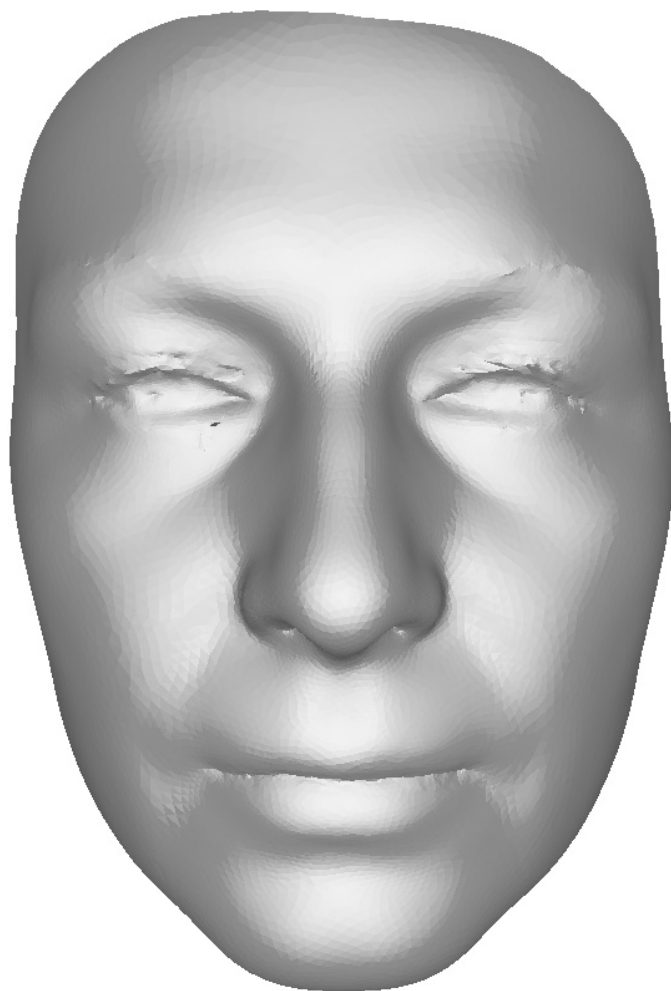
aneb „Co nám říká tvar obličeje?“

© 2011-15 Josef Pelikán, Ján Dupej, Václav Krajíček
CGG MFF UK Praha

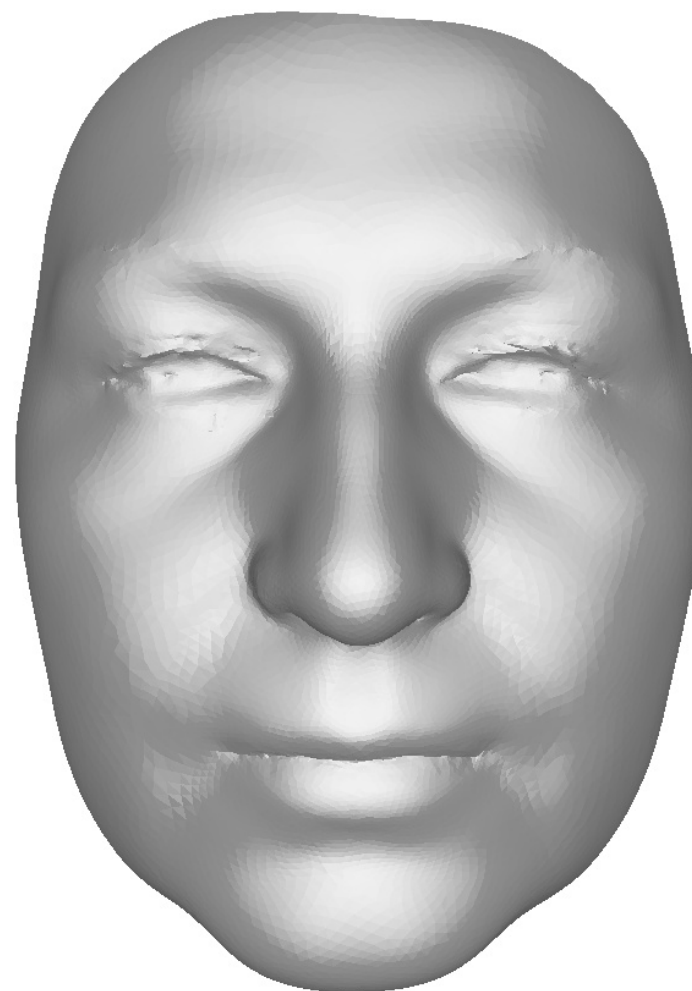
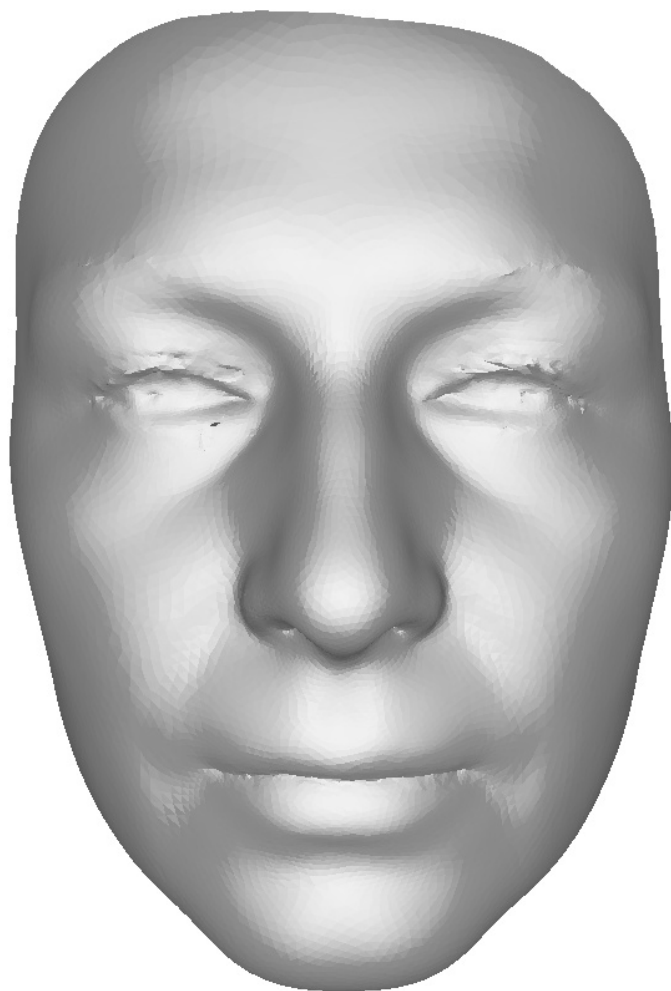
<http://cgg.mff.cuni.cz/>

{pepca|jdupej|vajicek}@cgg.mff.cuni.cz

Muž / žena ?



Mládí / stáří ?





matfyz



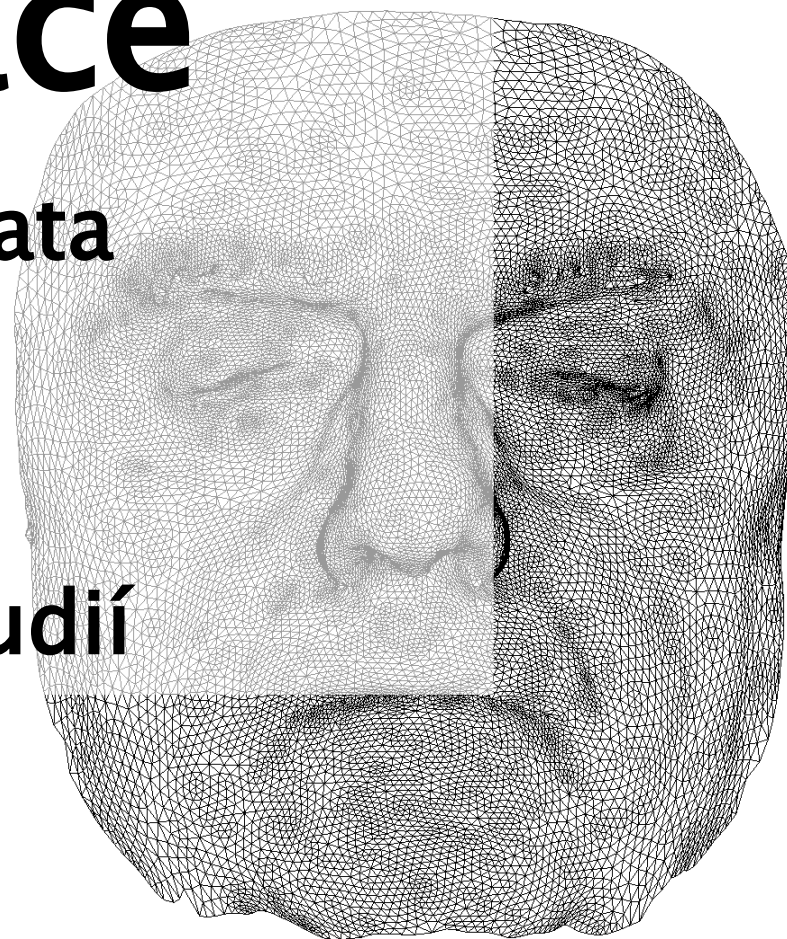
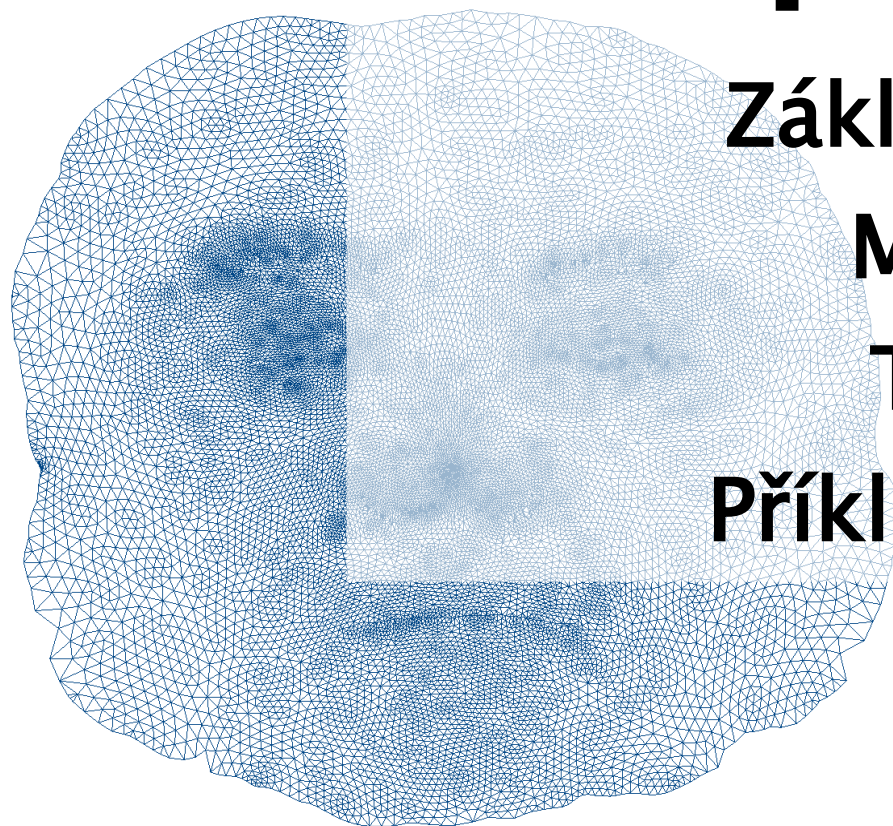
Aplikace

Základy, data

Metody

Teorie

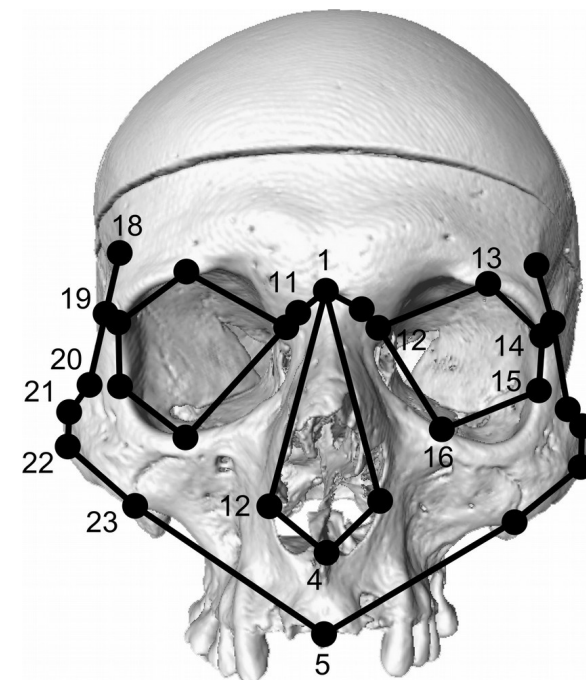
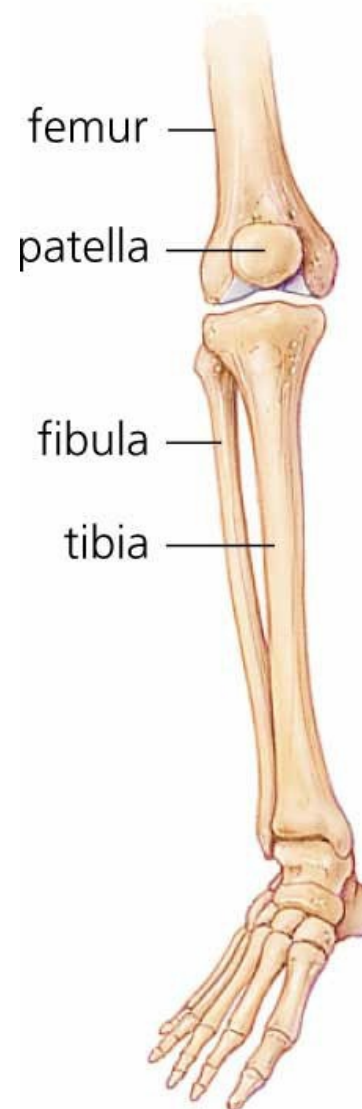
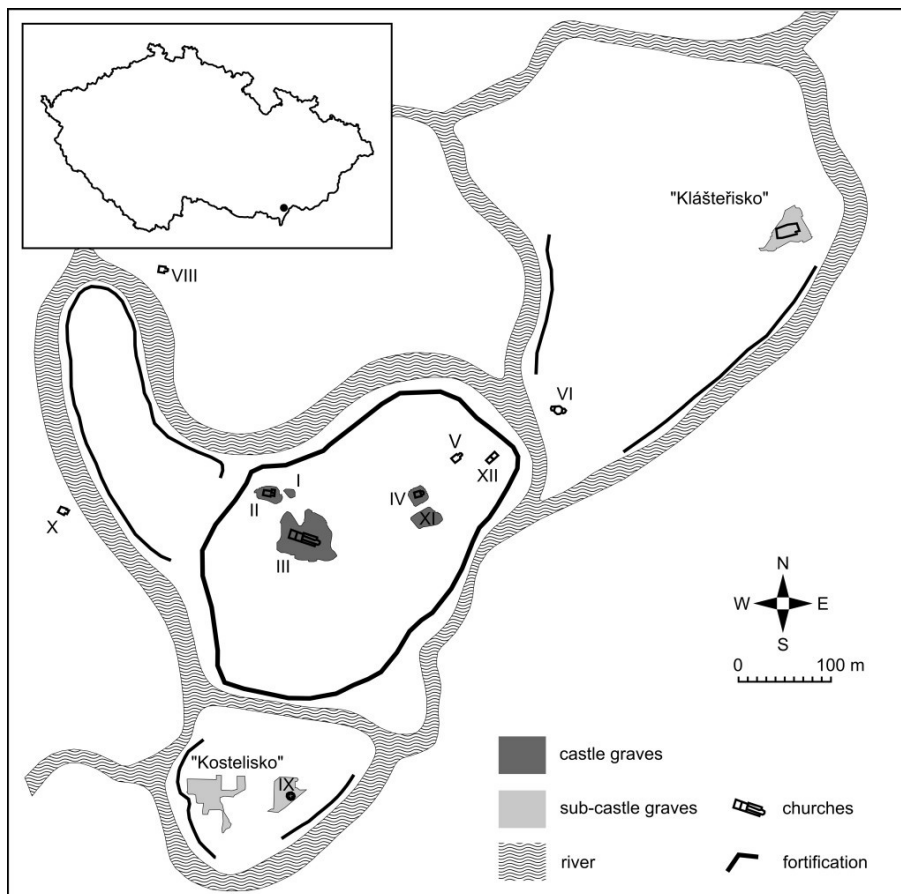
Příklady studií



Aplikace



Archeologie, antropologie



Carlyn Iverson

Aplikace



◆ Forenzní antropologie





matfyz



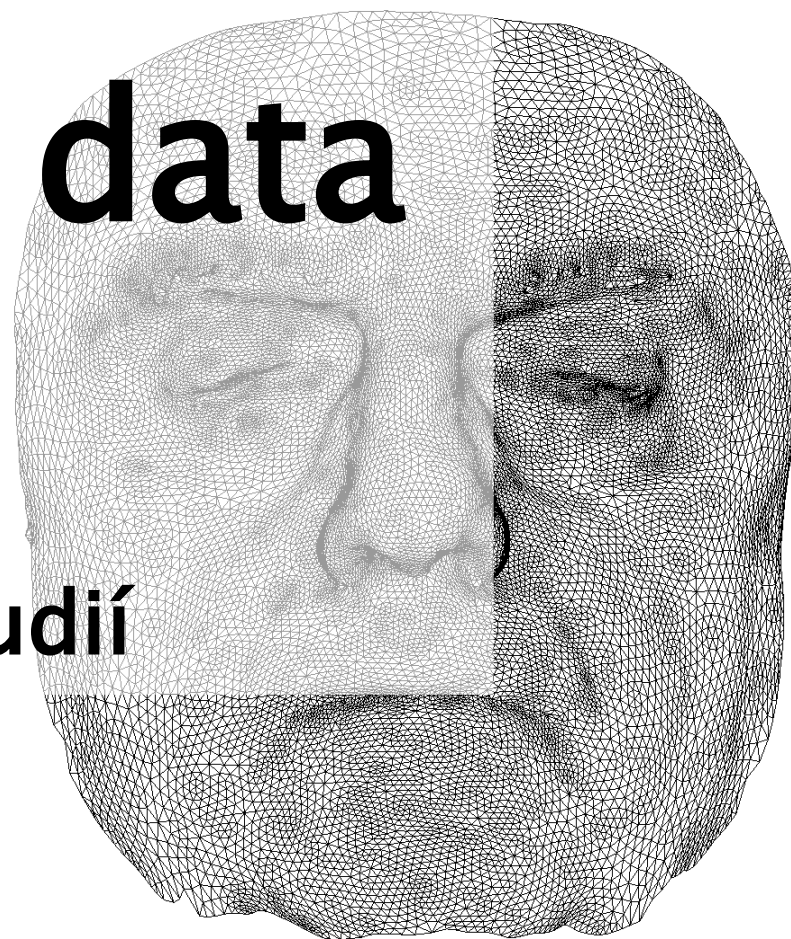
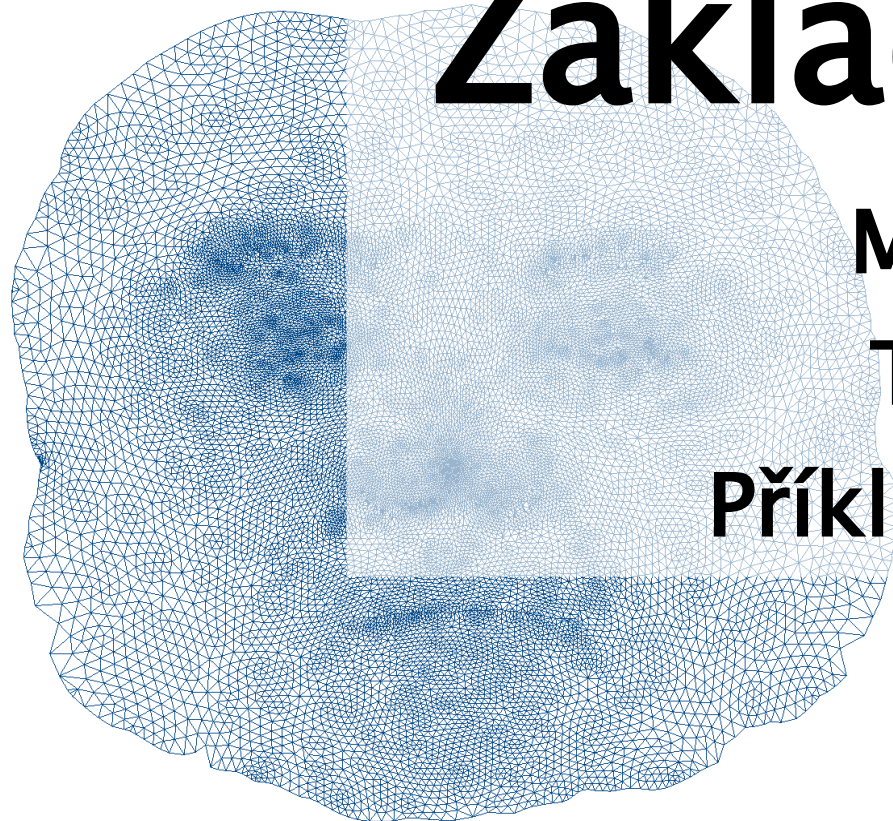
Aplikace

Základy, data

Metody

Teorie

Příklady studií

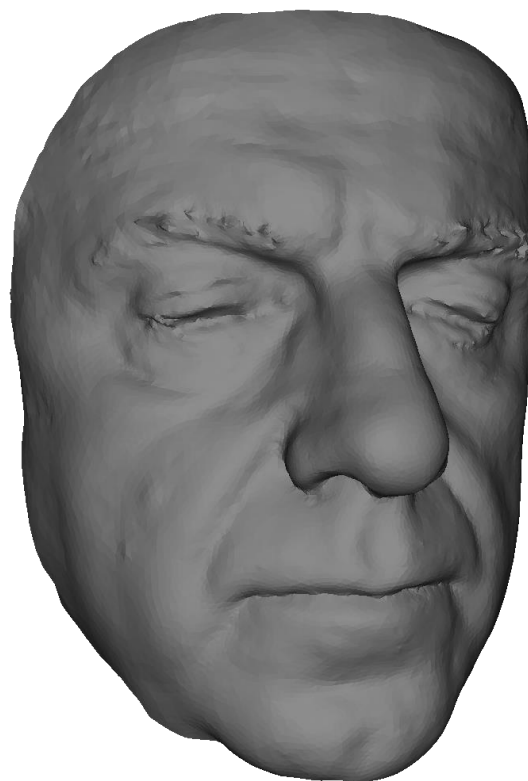
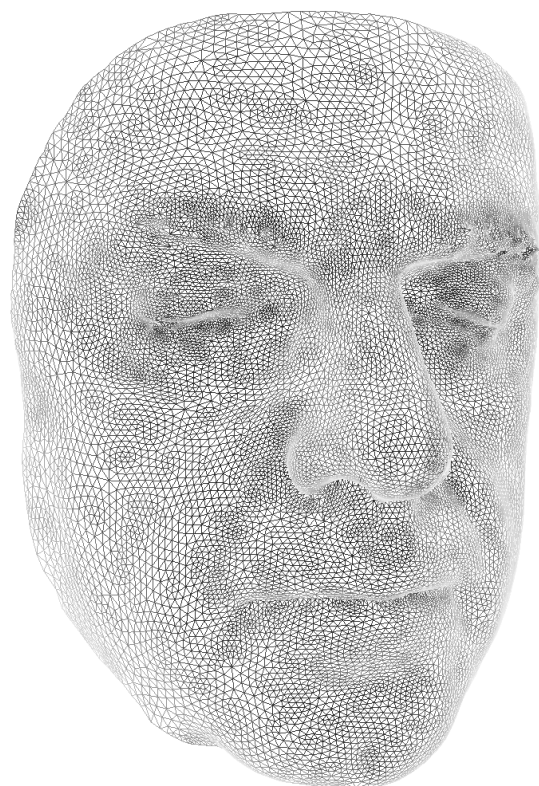


Geometrická morfometrie

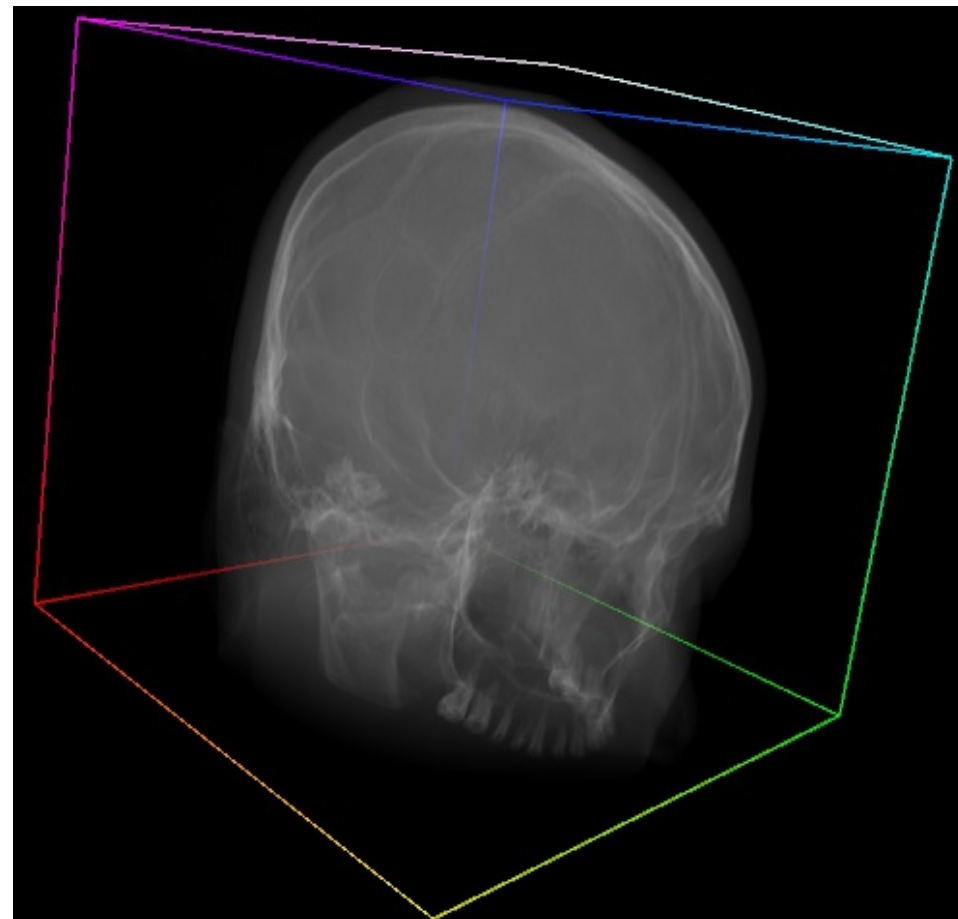
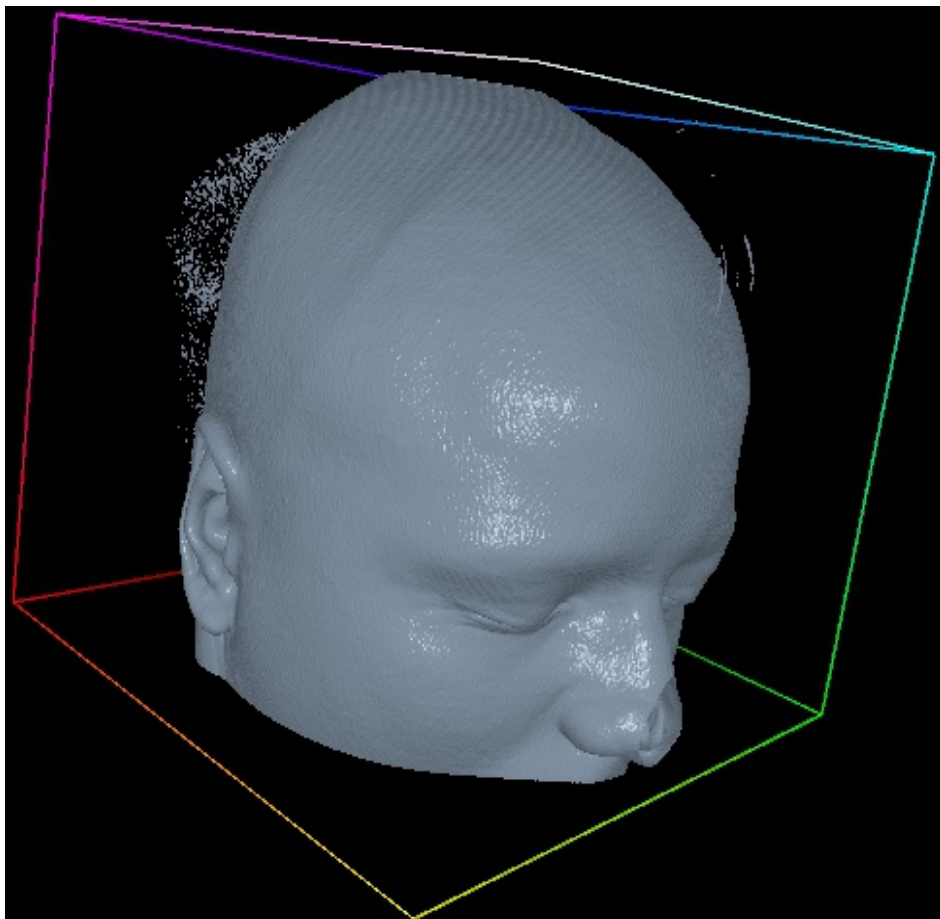


- ◆ **morfometrie:** disciplína zabývající se zkoumáním změn tvarů objektů, jejich variabilitou a klasifikací
 - ◆ biologie, archeologie, forenzní vědy
- ◆ **geometrická morfometrie (GMM):** geometrický popis tvaru vychází z tzv. landmarků
 - ◆ 2D nebo 3D souřadnice biologicky významných bodů
- ◆ **moderní GMM:** další metody, jak popsat tvar
 - ◆ trojúhelníkové sítě, analytické reprezentace křivek a ploch, objemová data

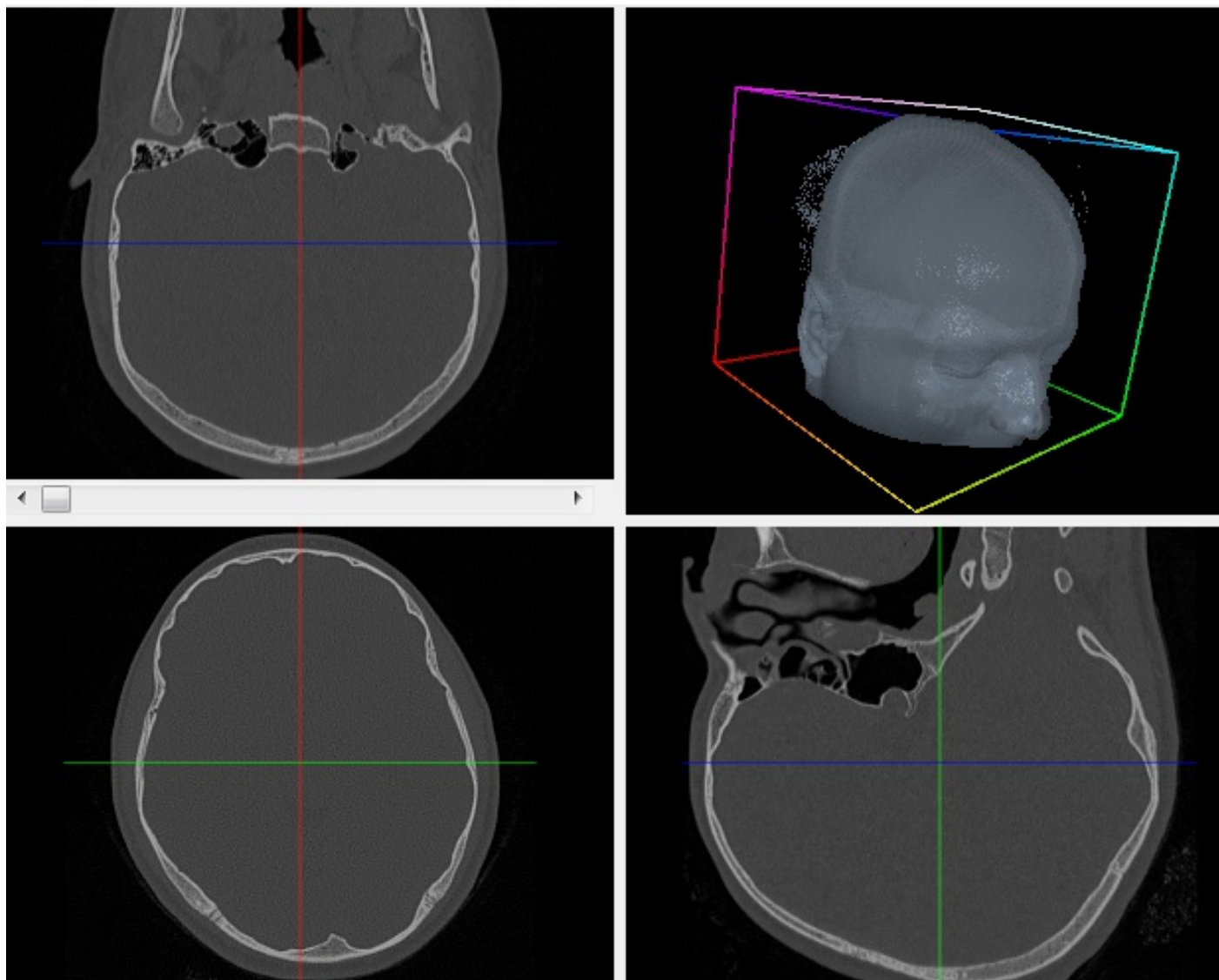
Data – 3D obličejový skener



Objemová data



Zdroj – počítačový tomograf (CT)

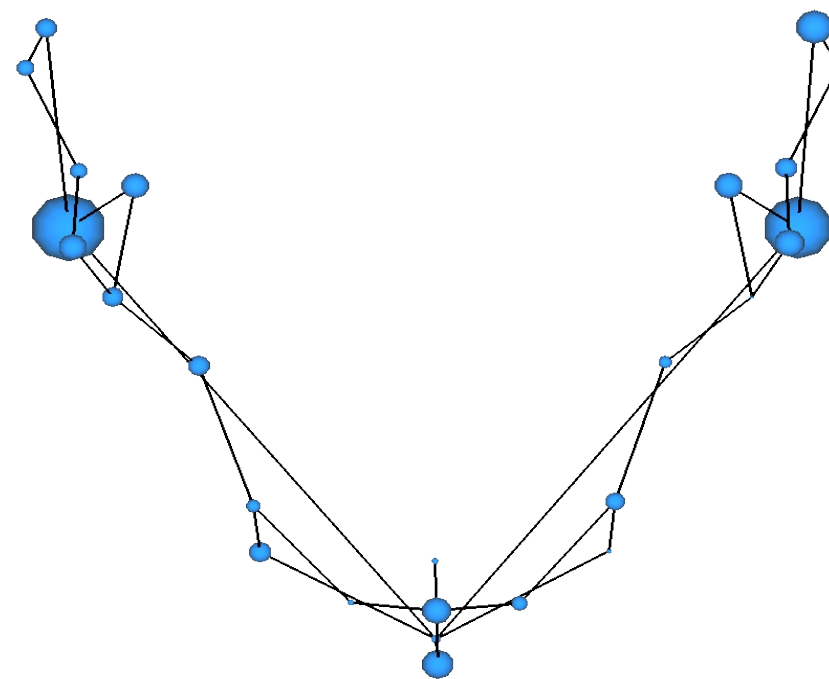


Přímé snímání souřadnic



MicroScribe: 3D data s přesností zlomku mm

Landmarky na čelisti

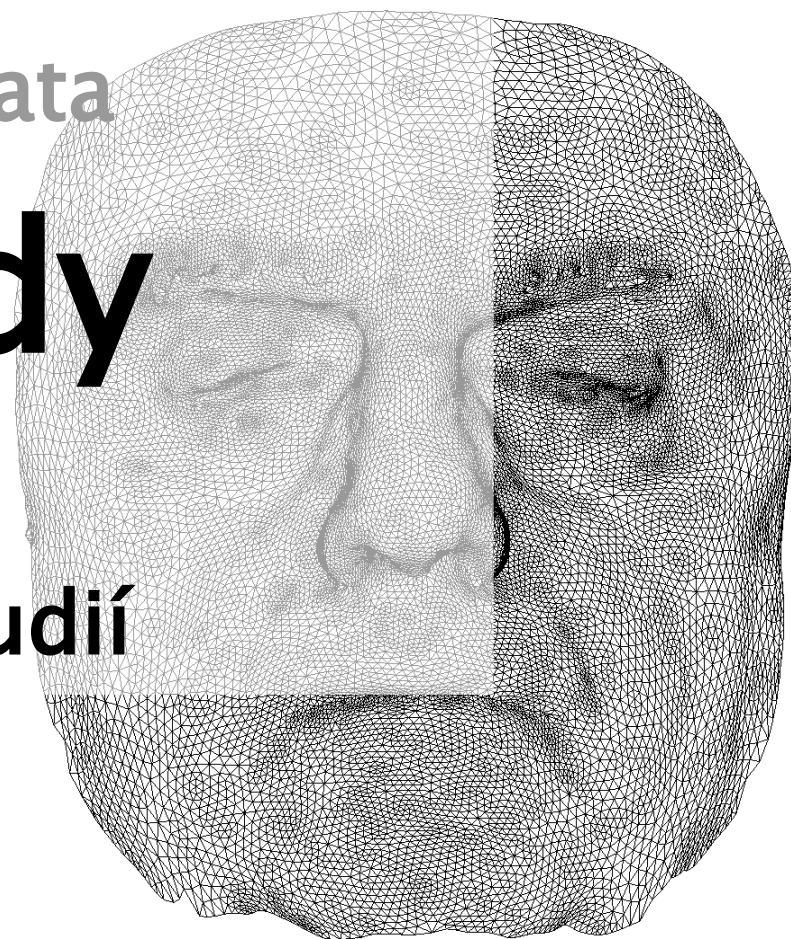
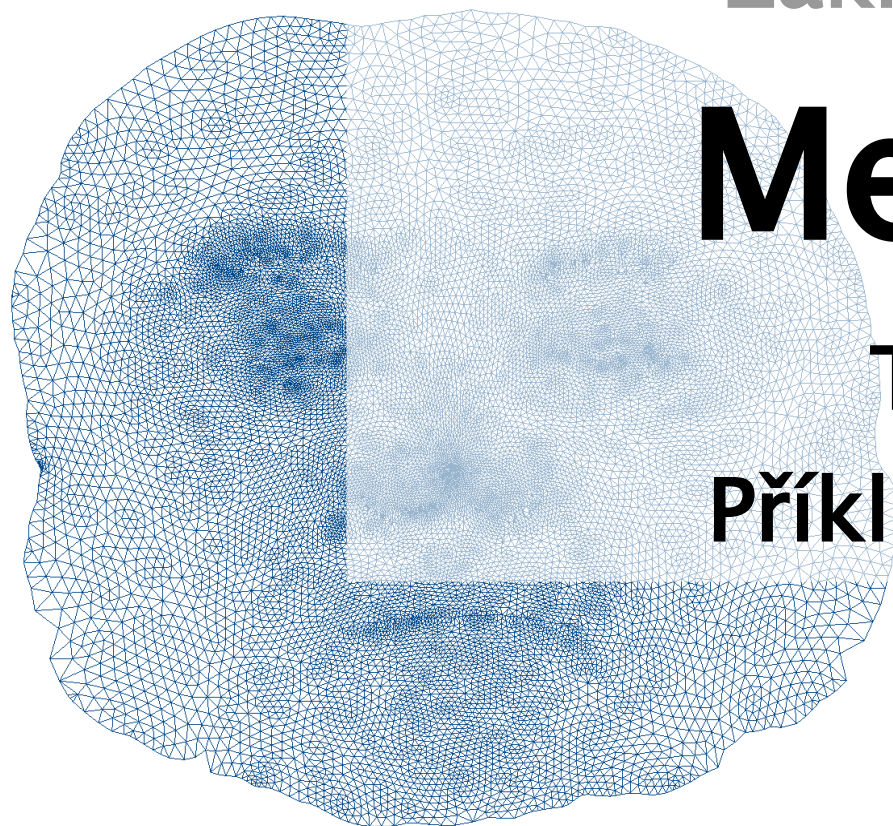


Aplikace
Základy, data

Metody

Teorie

Příklady studií



Cíle GMM



- **odhalit skrytá fakta a souvislosti**
- **než začneme** se vztahy → popis exempláře
 - landmarkové metody, křivky, trojúhelníkové sítě
- **nic nevíme** → získávání nových poznatků studiem tvaru u skupiny exemplářů
 - PCA, shluková analýza, vizualizace
- **něco tušíme** → ověření nebo vyvrácení hypotézy
 - statistické testy, korelace, regresní analýza
- **víme co, nevíme kolik** → stanovení míry odlišnosti
 - metriky

Příklady konkrétních úloh GMM

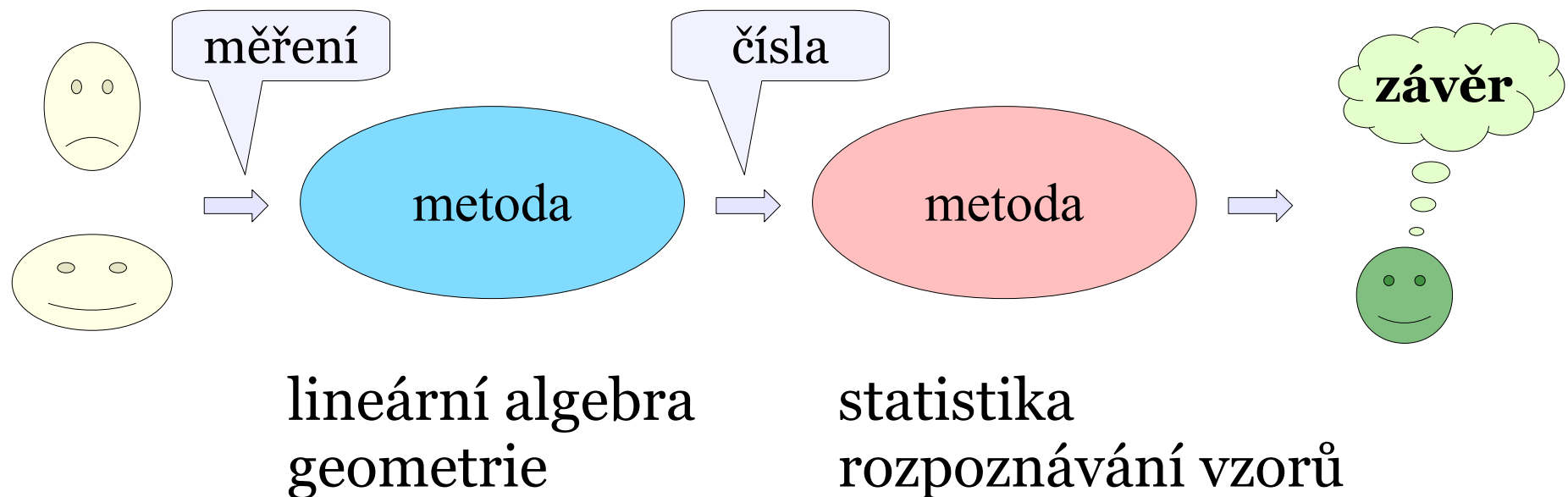


- ◆ porovnávání dvou (mnoha) tvarů
- ◆ hledání odlišností mezi dvěma či více skupinami
- ◆ hledání průměrného jedince v daném souboru
- ◆ studium symetrií (různé druhy asymetrie)
- ◆ hledání trendu vývoje jedince (ontogeneze) nebo druhu (fylogeneze)
- ◆ klasifikace tvarů
- ◆ interpolace a extrapolace (predikce)

Obecný postup GMM



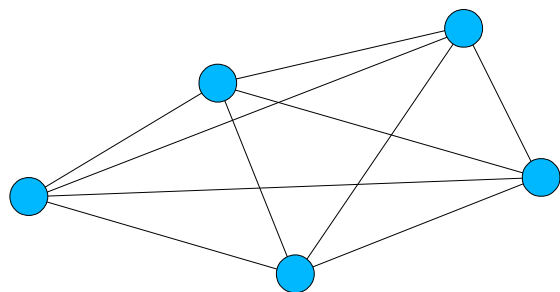
1. popsat tvar čísla
2. čísla statisticky vyhodnotit



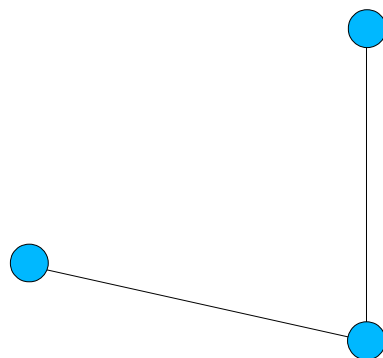
Tradiční (stará) morfometrie



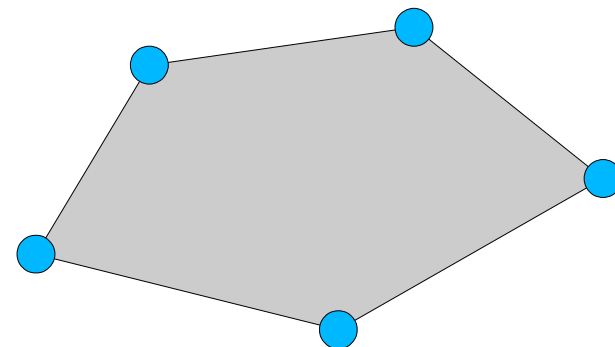
- poloha „bodů“ nehraje roli (ani se nedala dobře měřit)
- pracovalo se:
 - ♦ s vzdálenostmi mezi dvěma „body“, poměry vzdáleností
 - ♦ s úhly mezi třemi „body“
 - ♦ s plochou oblasti (polygonu)



$$\binom{p}{2}$$



$$p \binom{p-1}{2}$$

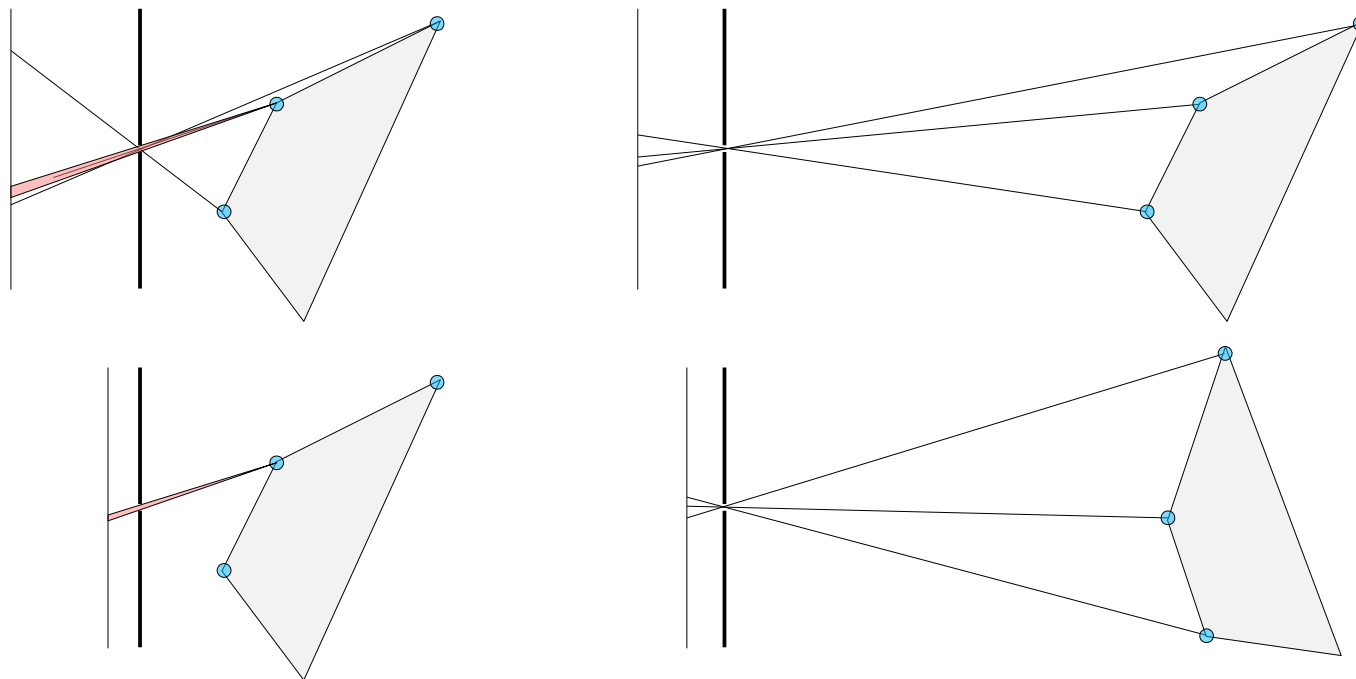


Plocha složena z ploch nepřekrývajících se trojúhelníků

Pořizování dat



- ◆ **2D**: fotografie, video
 - ◆ perspektivní projekce – potenciální zdroj chyb měření
- ◆ **3D**: objemová data, povrchové skenery, polohovací zařízení

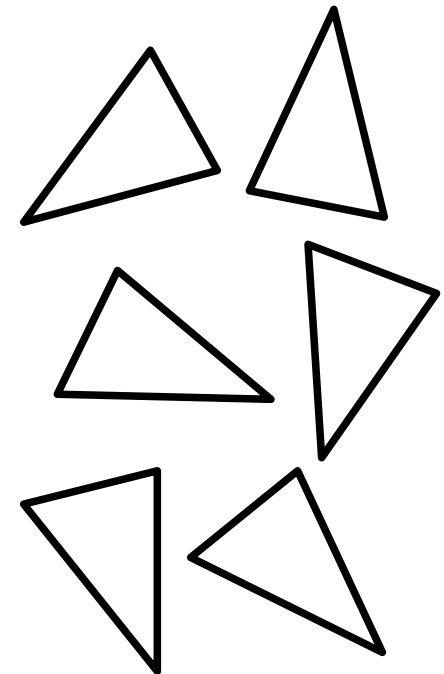
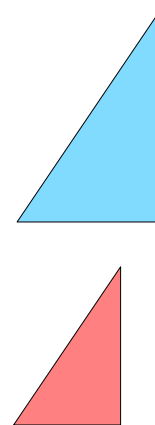
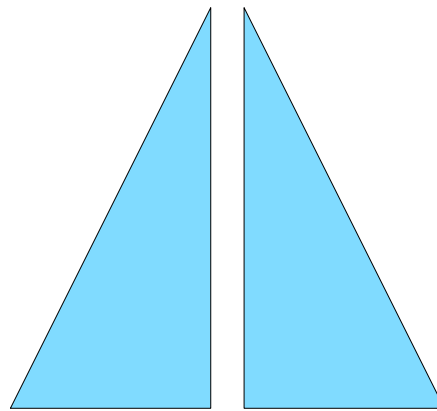
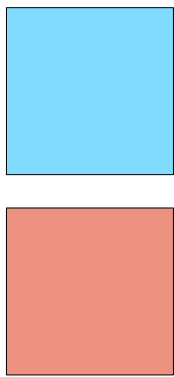


Moderní metody GMM



- „Metody měření, zpracování, analýzy a zobrazení s cílem studia tvarů a jejich změn“
- **tvarové proměnné** – popisují tvar, nemění se s velikostí

forma = velikost + tvar





- ◆ Definice (Kendall):

Tvar je veškerá geometrická informace, která objektu zůstane po odstranění vlastností polohy, měřítka a otočení

- ◆ „tradiční“ morfometrie

- ◆ pracuje přímo se vzdálenostmi (vylepšení – poměry vzdáleností)
- ◆ samotné vzdálenosti stěží odrážejí tvar
- ◆ dva exempláře jsou tak tvarem neporovnatelné

- + **oddělení velikosti a tvaru** v datech

- ◆ jedna z hlavních myšlenek geometrické morfometrie

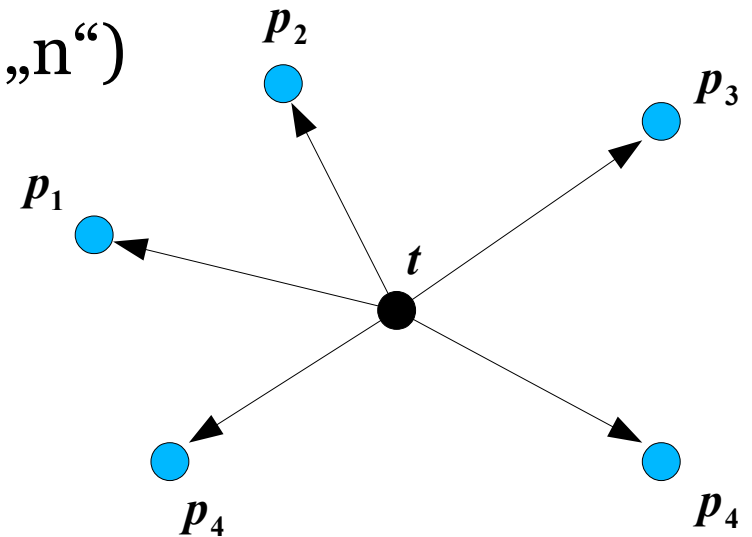
Velikost



- ◆ nezáporný atribut, linearita
- ◆ **středová velikost** (Centroid Size)
 - ◆ matematicky je „nezávislá na tvaru“
 - ◆ výpočet z landmarků (pozor na „n“)

$$\mathbf{t} = \frac{1}{n} \sum_{i=1}^n \mathbf{p}_i$$

$$CS = \sqrt{\sum_{i=1}^n (\mathbf{p}_i - \mathbf{t})^T (\mathbf{p}_i - \mathbf{t})}$$



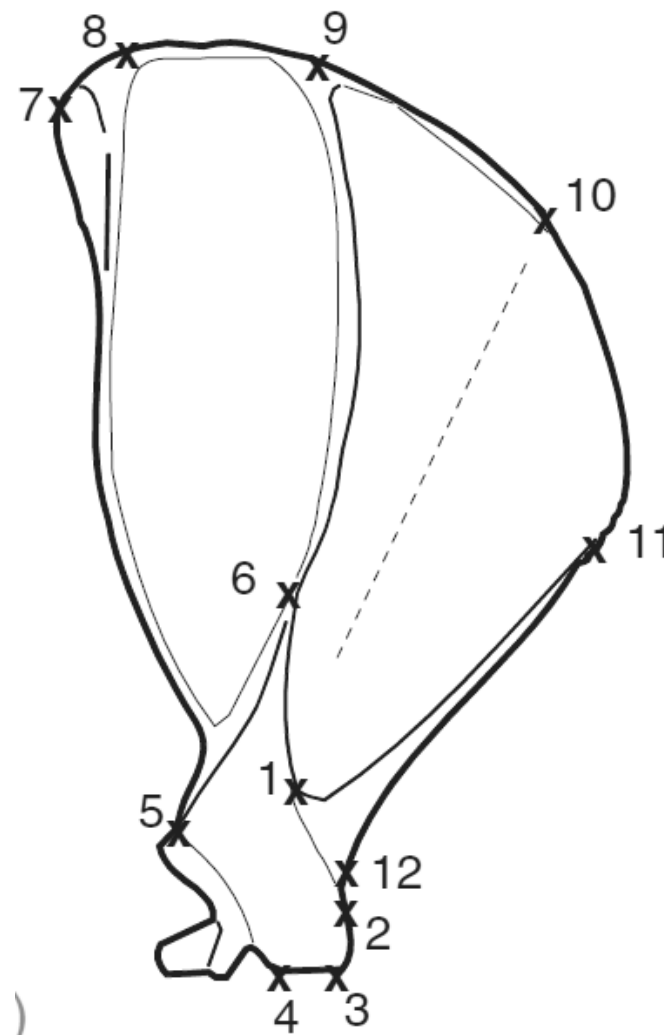
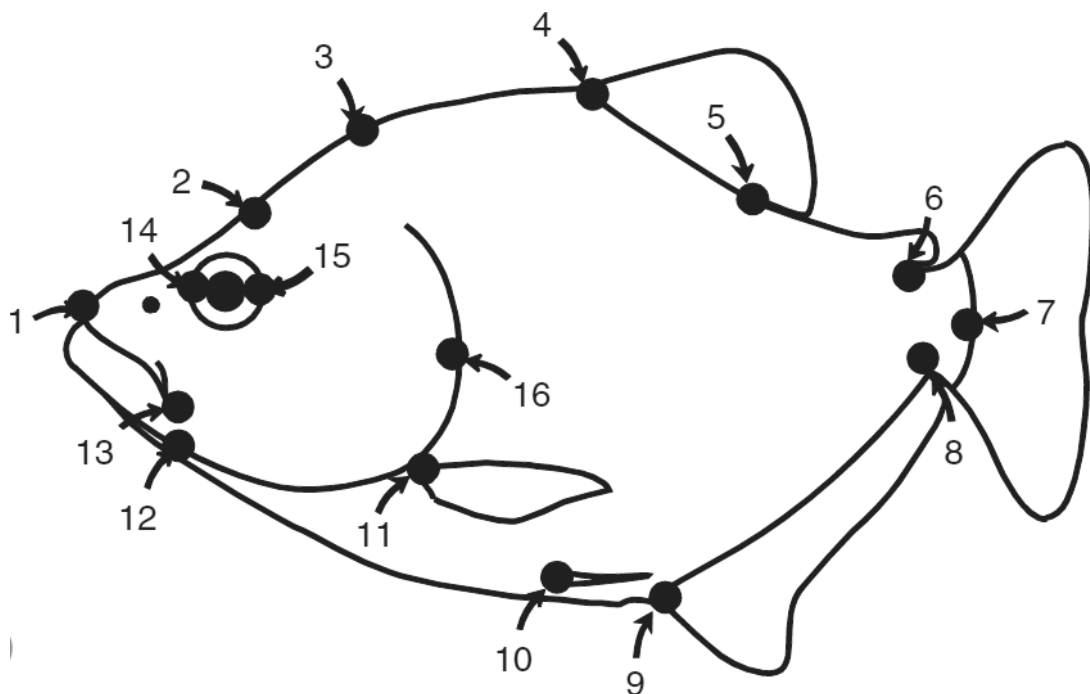
- ◆ často se hledá konkrétní souvislost středové velikosti s tvarem



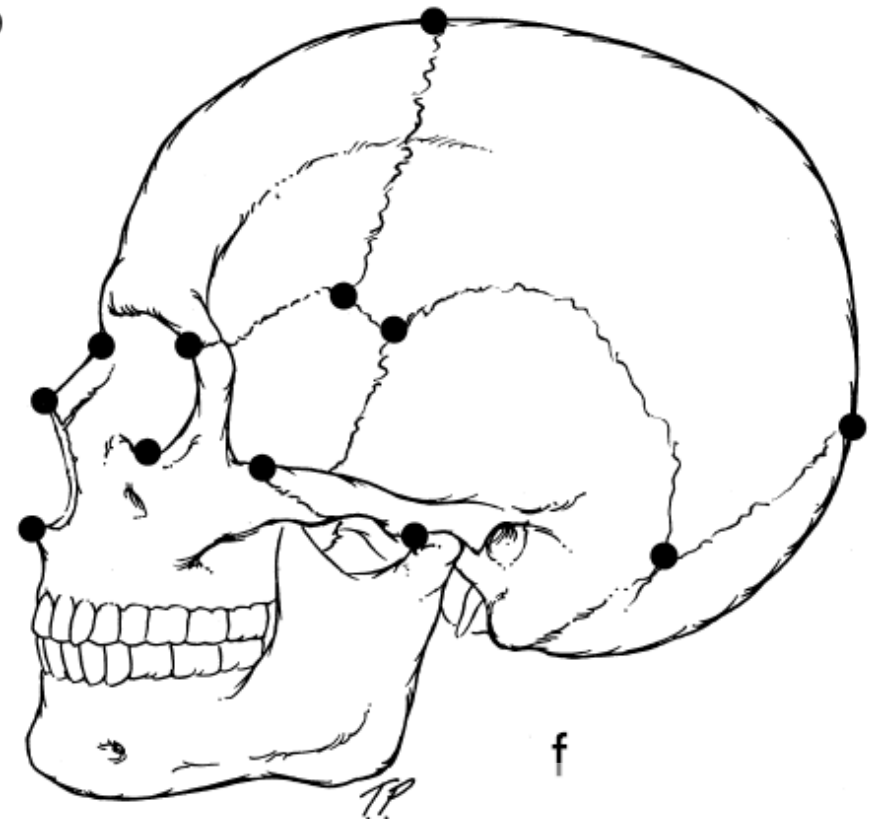
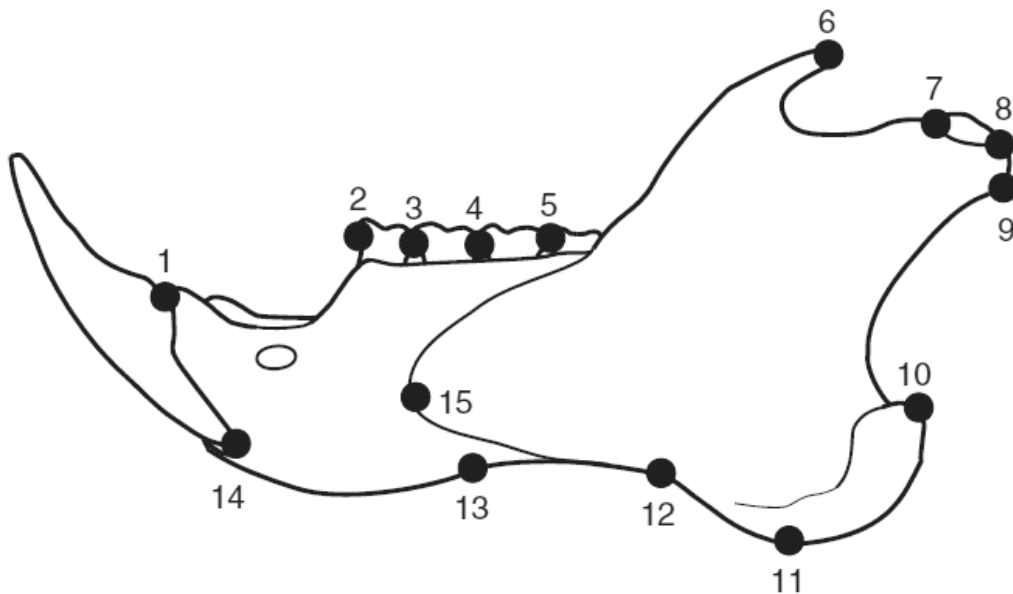
Landmarky

- ▶ **významné body** umístěné na objektu
 - ♦ biologický význam (pojmenování)
 - ♦ souřadnice v 2D nebo 3D prostoru
- ▶ **typické umístění**, snadno reprodukovatelné:
 - ♦ extrém tvaru (roh)
 - ♦ průsečík dvou linií
 - ♦ extrém křivosti, ohyb, příp. inflexní bod
- ▶ **semi-landmarky**
 - ♦ pomocné body dopočítané nějakým matematickým postupem

Landmarky – příklady



Landmarky – příklady

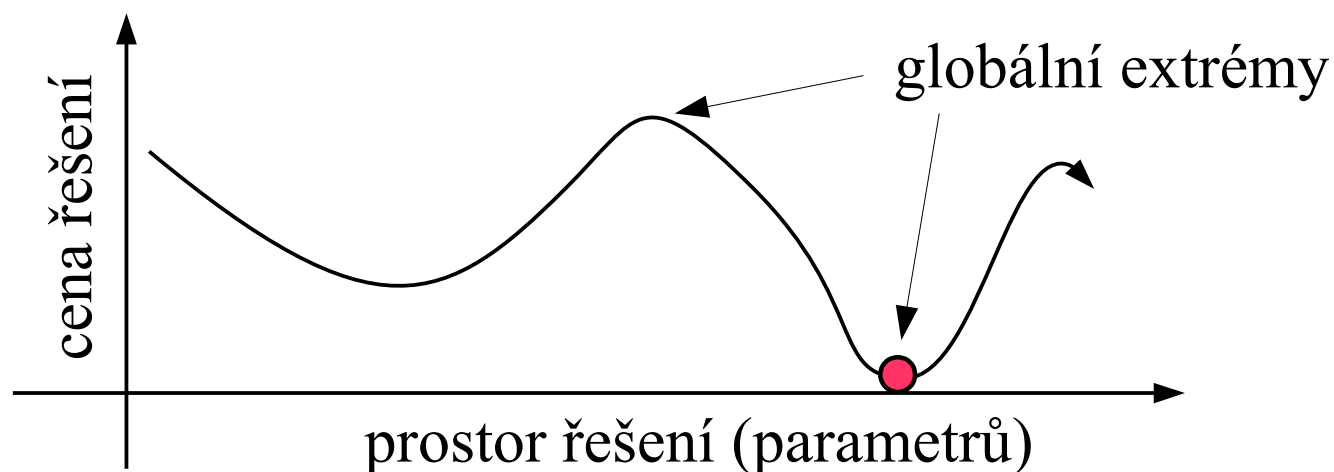


(převzato ze Zeldich et al: Geometric morphometrics for biologists: the primer)

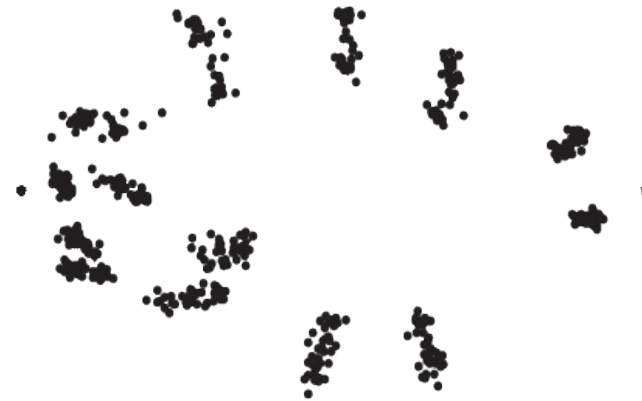
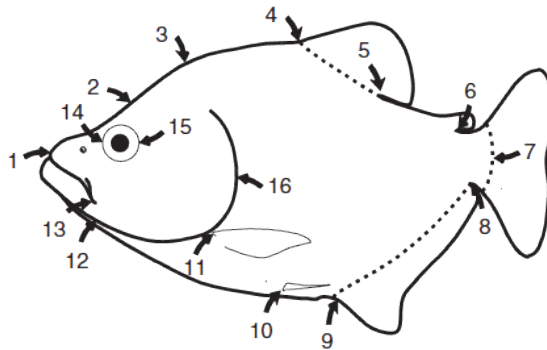


Prokrustovská analýza (GPA)

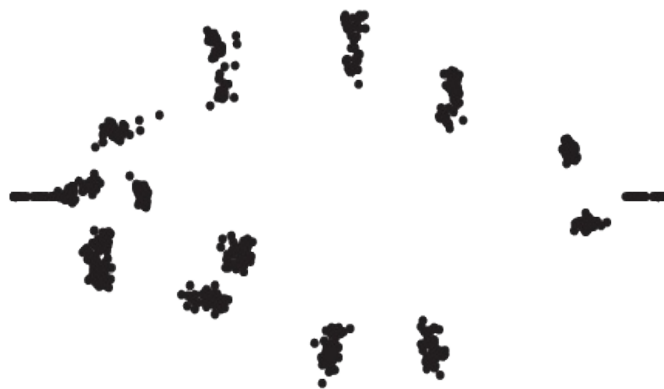
- ◆ populární metoda registrace podle landmarků:
 - ◆ funguje i v 3D, není ovlivněna výběrem žádné základny
- ◆ hledání nejlepších parametrů transformace (posunutí, škálování, rotace) ve smyslu **nejmenší vzdálenosti** odpovídajících si landmarků (nejmenší čtverce)
- ➔ optimalizační problém



Registrace podle landmarků



Booksteinova registrace



Klouzavá základna



GPA

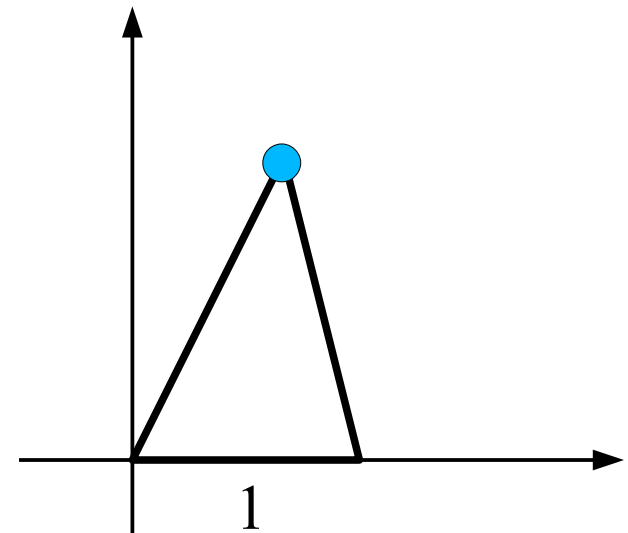
Prostor tvarů – trojúhelník



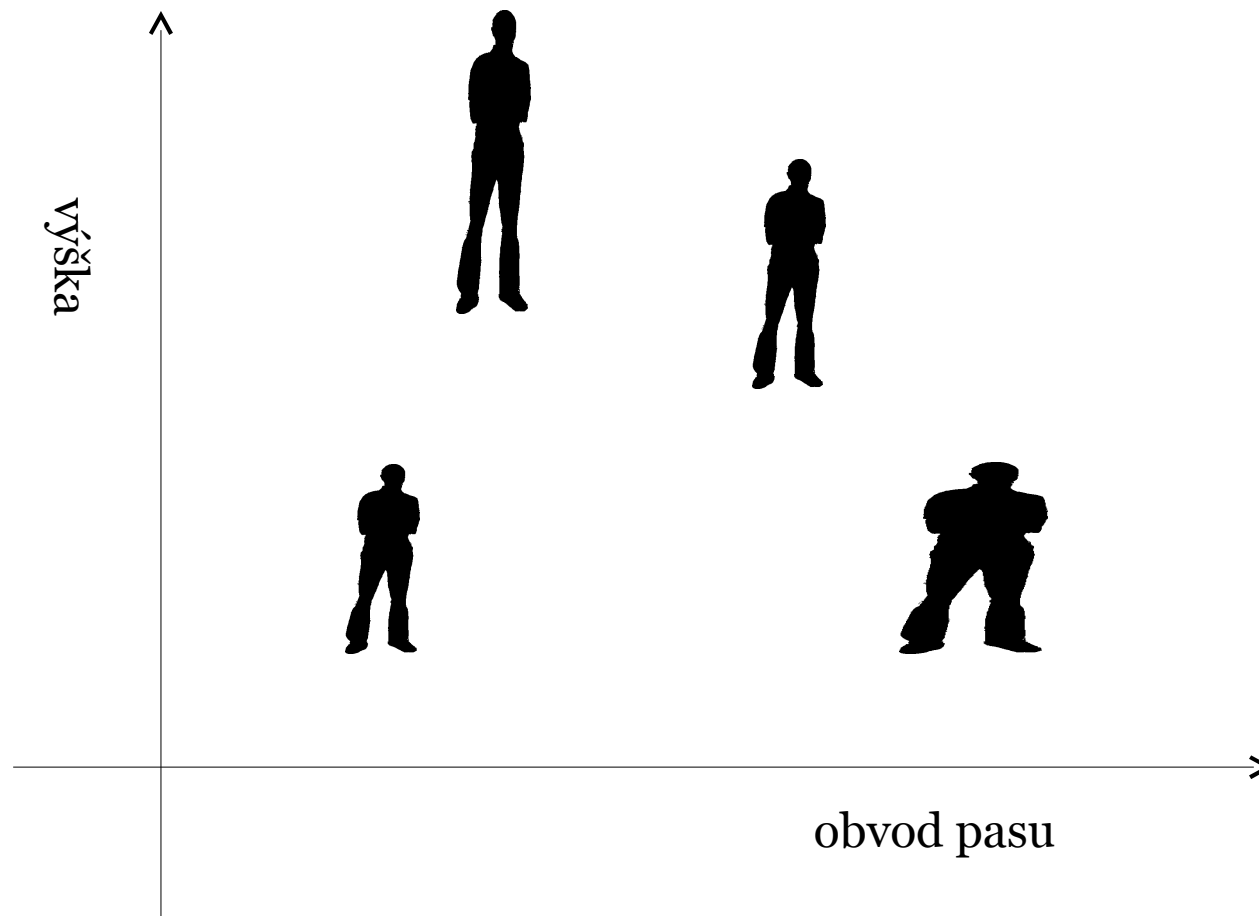
- nejjednodušší geometrický objekt, u kterého se dá hovořit o tvaru
- úplný popis tvaru trojúhelníku v rovině: dvojice čísel
- **stupeň volnosti tvaru**: počet proměnných po odstranění velikosti, otočení a posunutí

- 2D: $2p - 4$
- 3D: $3p - 7$
- k-dim: $pk - k - k(k - 1)/2 - 1$

p – počet bodů



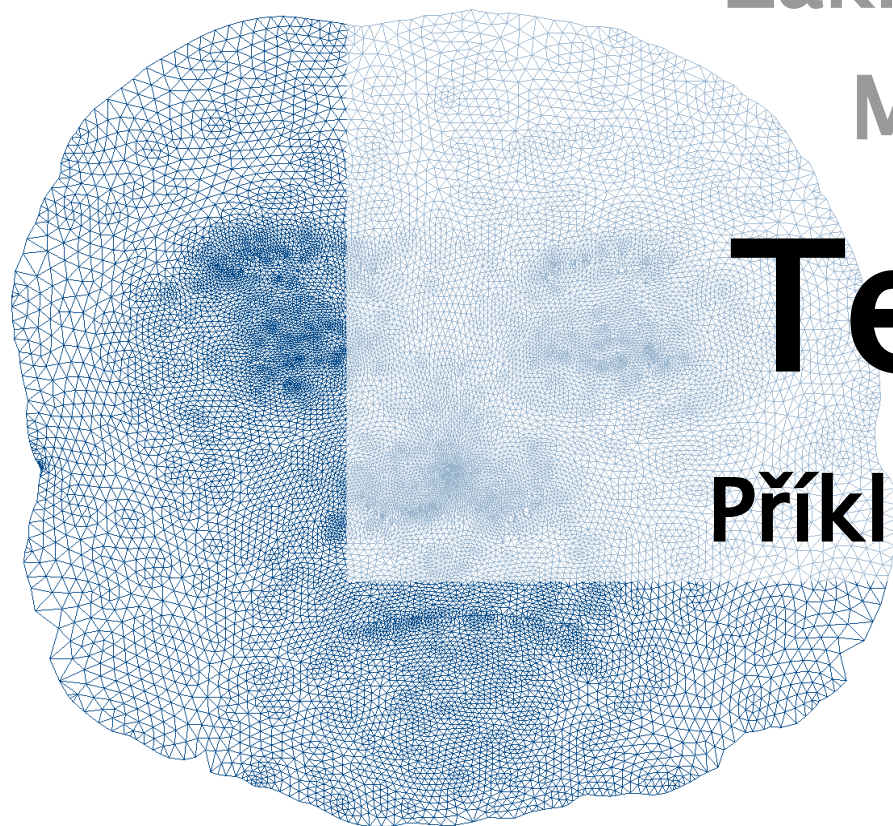
Prostor tvarů – lidská postava



Aplikace
Základy, data
Metody

Teorie

Příklady studií



Teorie GMM – statistické metody



- ◆ klasické přístupy
 - ◆ testování hypotéz (stejná či odlišná distribuce..)
 - ◆ rozhodování na zvolené hladině významnosti
- ◆ N-rozměrná analýza
 - ◆ jedinec ... vektor čísel (souřadnice landmarků, koeficienty, vzdálenosti, úhly, povrchy, objemy..)
- ◆ jsou dvě množiny statisticky významně oddělené?
 - ◆ pro klasifikaci v terénu, forenzní aplikace..
 - ◆ diskriminační analýza (často LDA), další klasifikační postupy (SVM)



Statistický model tvaru

- **model:** matematický popis jevu z reálného světa
- příklady:
 - ♦ model tvaru lebky může být „vhodným“ reprezentantem
 - splňuje dokonale pouze on sám, extrémny jsou daleko
 - ♦ průměrný reprezentant ze zkoumané populace
 - nesplňuje ho nikdo, ale všichni jsou mu dostatečně blízko
 - ♦ průměr + odchylka
 - může popisovat i nemožné exempláře
- parametrický model
 - ♦ každý konkrétní exemplář odpovídá číselné konfiguraci (např. PCA souřadnice)



Hledání průměrného jedince

- ◆ průměr je důležitý pro další analýzu
 - měří se proti němu nový jedinec
 - průměr reprezentuje skupinu (ženy, muži)
- ◆ **N-tice** popisující jedince (triviální)
 - jenom musí být model odchylek symetrický
- ◆ plochy v 3D
 - ◆ obtížný úkol
 - ◆ parametrizace ploch (PCA) a potom průměr hodnot parametrů..
 - ◆ použití zavedených (klasických) landmarků



Statistický model tvaru II

- zohledňuje tendence vyskytující se v pozorované **množině exemplářů (jedinců)**, vzorku

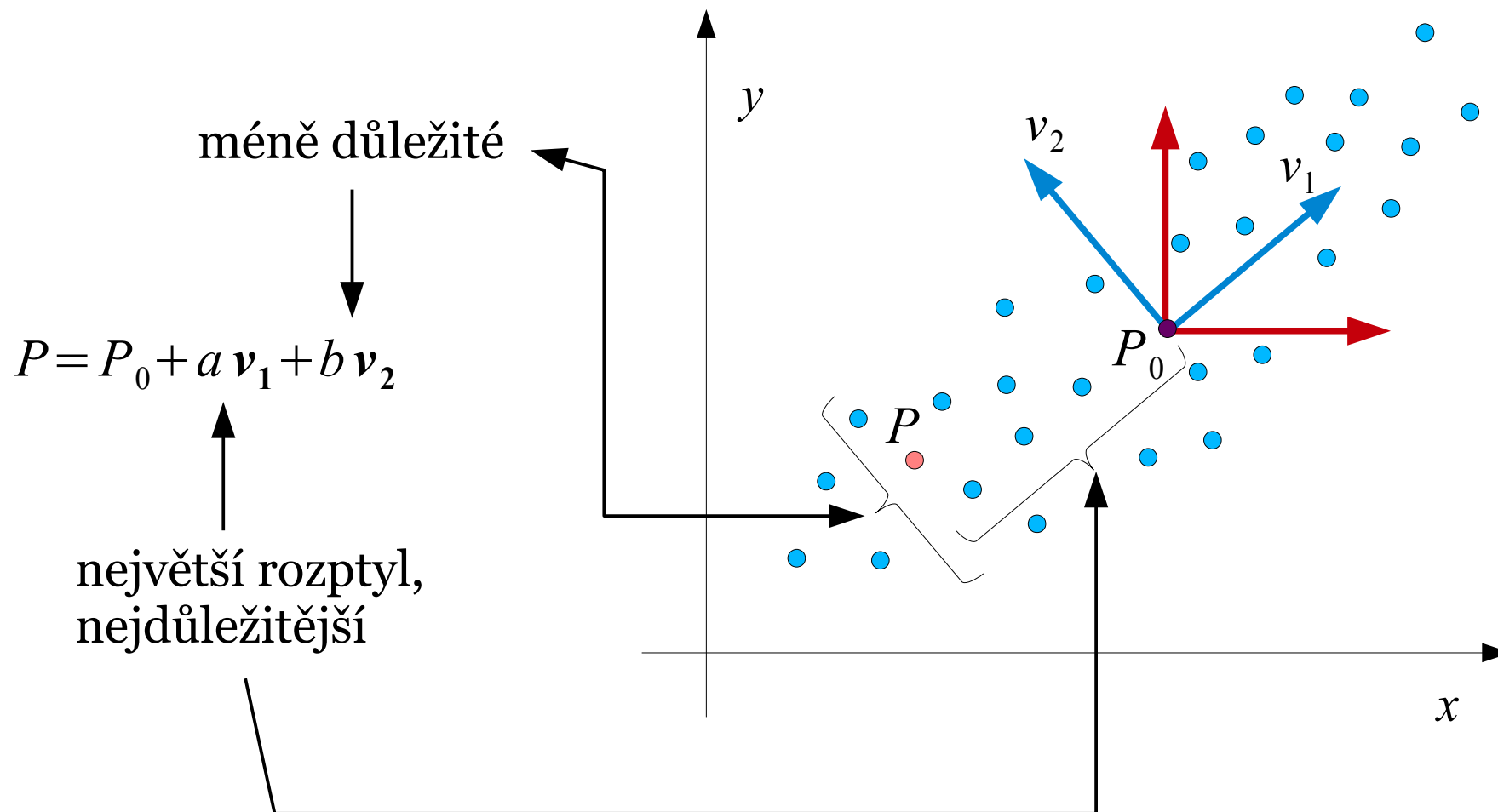
$$\{P_1, P_2, \dots, P_n\} \rightarrow F(k_1, \dots, k_m)$$

- dokáže popsat nekonečně velkou populaci ... generování exemplářů
- nástroj: **Analýza hlavních komponent** („Principal Component Analysis“ = PCA)
 - exemplář se skládá z příspěvků různých komponent
 - hledáme komponenty, které nejvíc přispívají do každého exempláře ve vzorku (nesou nejvíce informace)

Analýza hlavních komponent (PCA)

- ◆ máme **N-tici parametrů** popisující exemplář
- ◆ potřebujeme **oddělit významnější složky** od méně významných
 - ◆ např. je potřeba zredukovat dimenzi prostoru kvůli objemu dat, náročnosti měření, apod.
 - ◆ nebo hledáme minimální dimenzi parametrů popisující dobře strukturu nějaké populace (postačující pro separaci druhů..)
 - ◆ pro (interaktivní) vizualizaci a animace je přehledné pohybovat se jen v 2D prostoru parametrů..

Analýza hlavních komponent





PCA – matematika

Matematicky je to jen
lineární transformace souřadnic (změna baze)

1. kovarianční matice S:

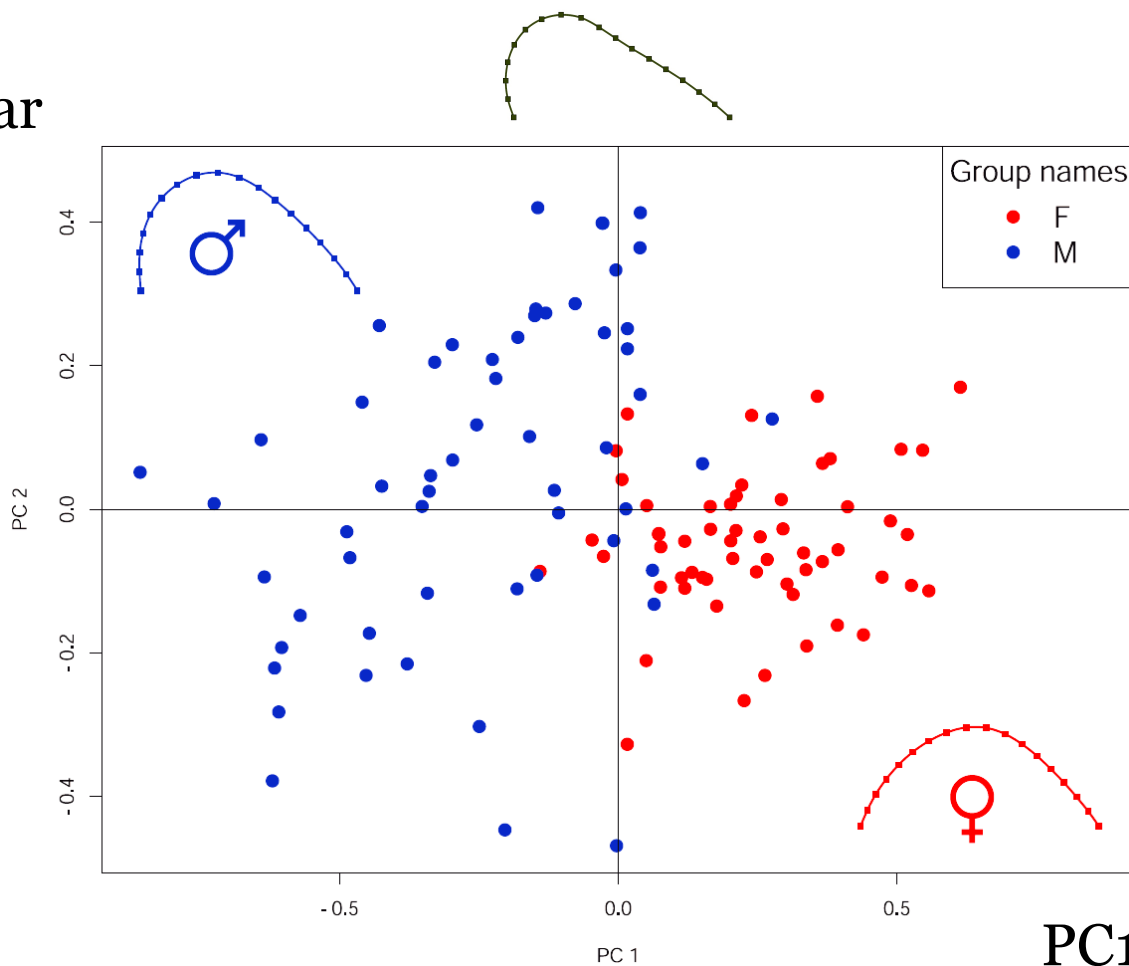
$$s_{ij}^2 = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i) \cdot (x_{kj} - \bar{x}_j)$$

- 2.** vlastní vektory matice S tvoří PCA bazi
- 3.** komponenty se setřídí podle velikosti příslušných vlastních čísel
- 4.** suma velikostí vlastních čísel ... celková informace (může se počítat podíl informace PCA složek)



PCA příklad – incisury

PC2 ... 27% var



PC1 ... 50% var

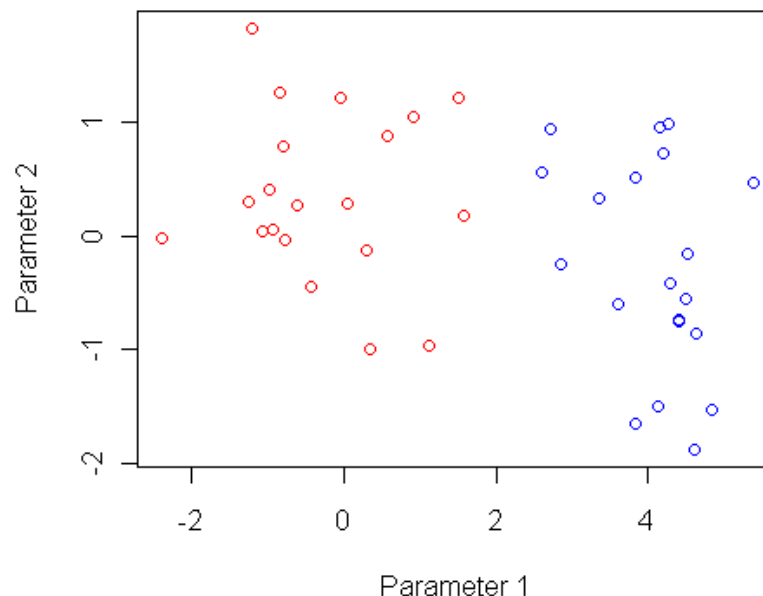


Diskriminační analýza

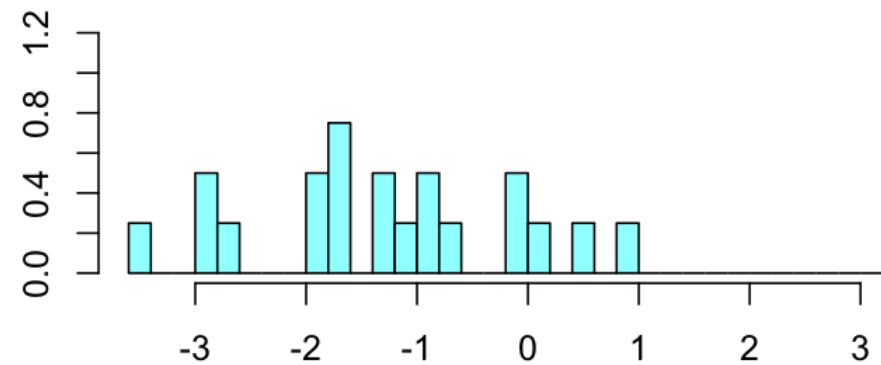
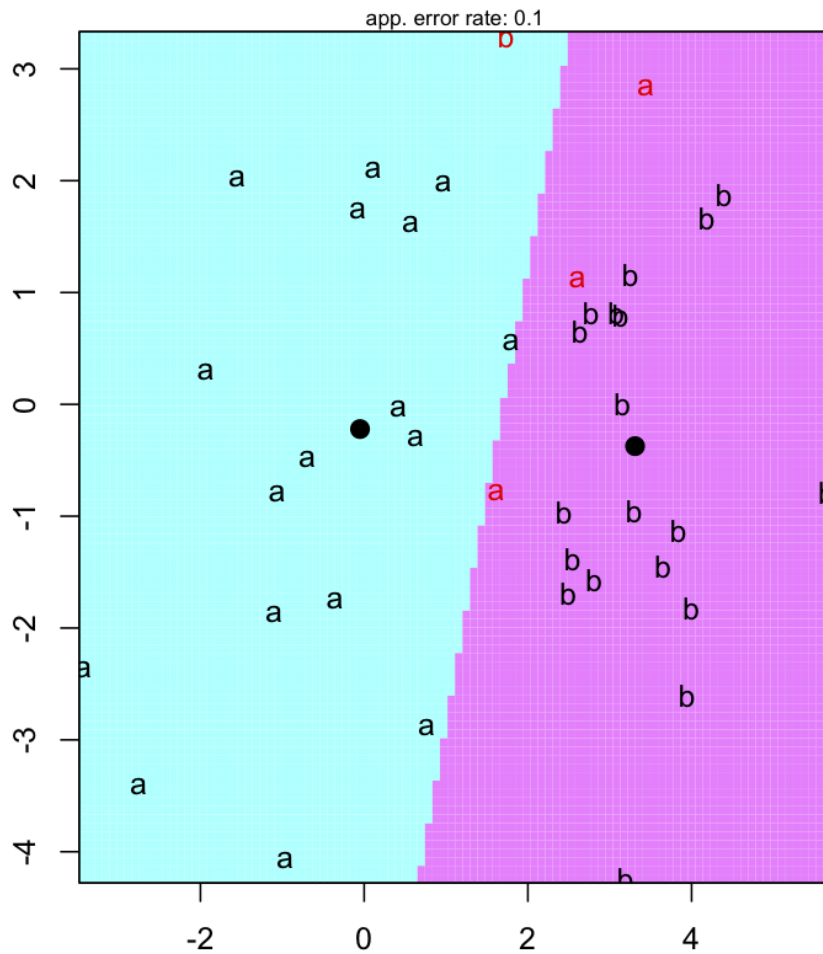
- ◆ dichotomie – rozdělitelnost souboru na dvě skupiny
- ◆ hledání **diskriminační funkce f** , která jedince \mathbf{x}
 - $f(\mathbf{x}) > 0$ přiřadí do první skupiny
 - $f(\mathbf{x}) < 0$ přiřadí do druhé skupiny
- ◆ lineární diskriminační analýza (**LDA**) → f je lineární

$$f(\mathbf{x}) = \mathbf{w} \mathbf{x} + c$$

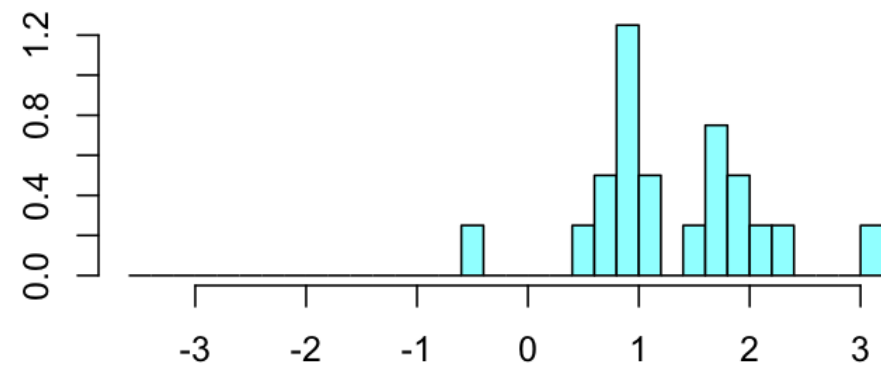
- ◆ hledání vektoru \mathbf{w} a skaláru c
- ◆ dimenze dat je libovolná
- ◆ rozšíření pro více skupin



LDA – příklad



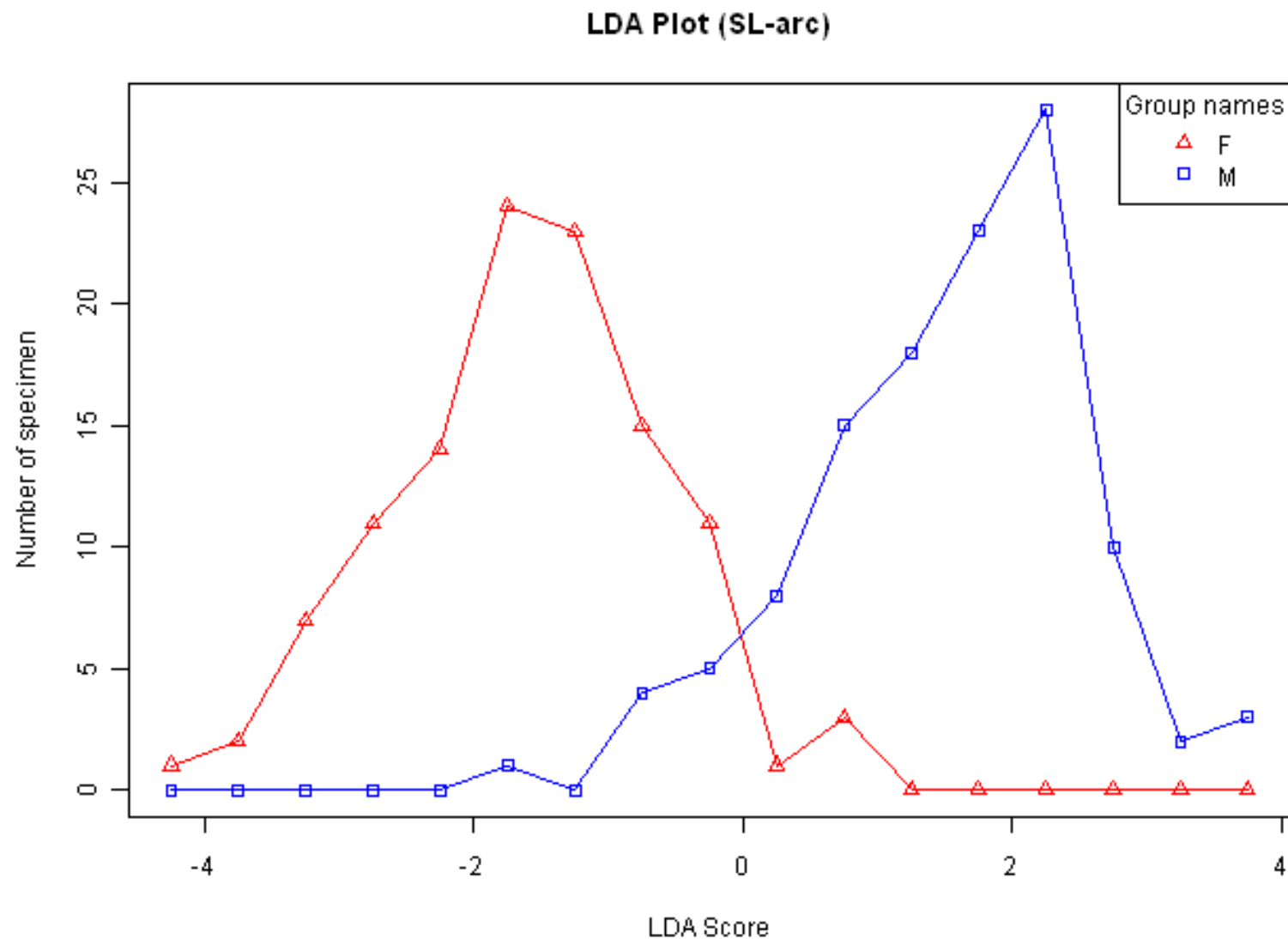
group a



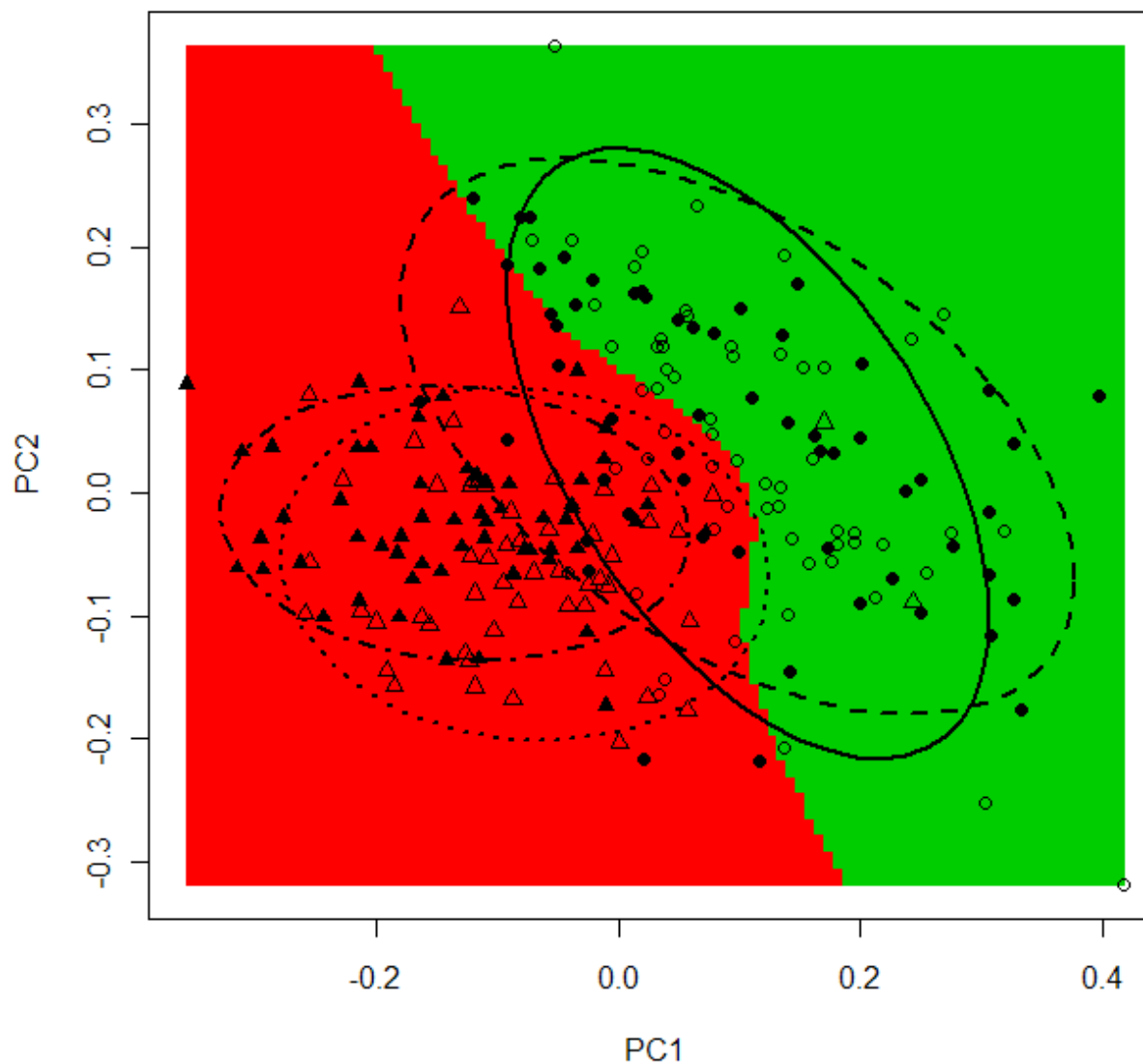
group b



LDA – příklad II (incisury)



Složitější diskriminační analýza



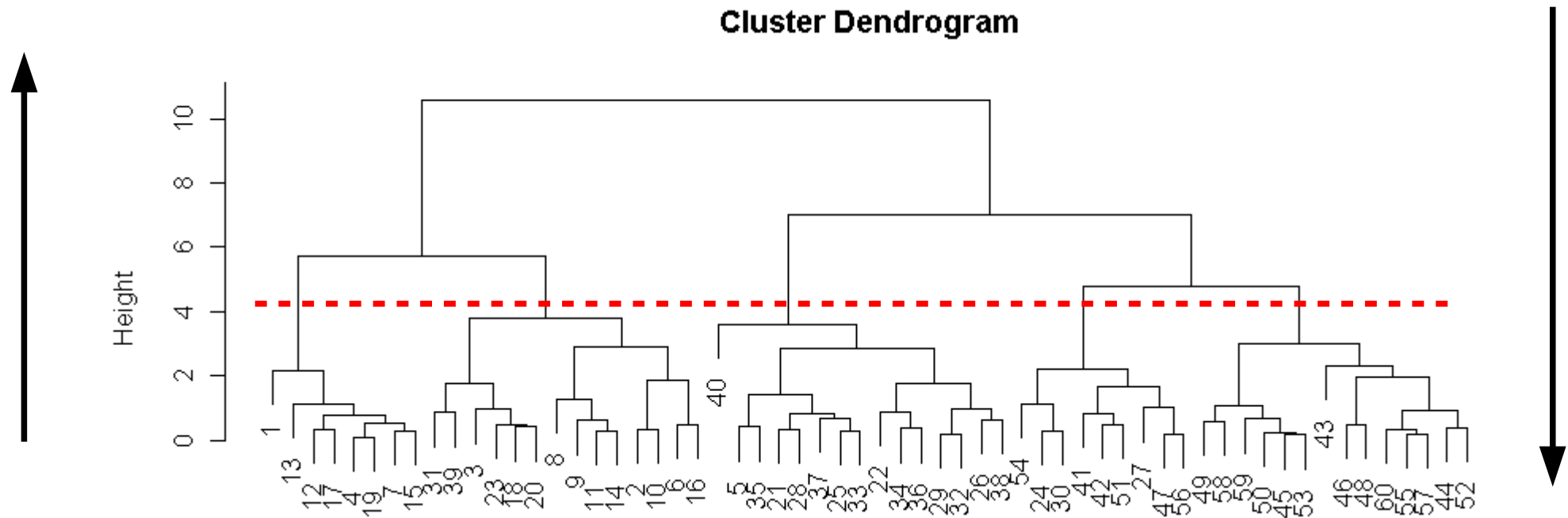
Shluková analýza (Cluster Analysis)

- ◆ **vstup:** soubor k N -rozměrných vektorů popisujících k jedinců
- ◆ **výstup:** dendrogram (hierarchie shlukování)
 - ◆ postupné slévání / oddělování skupin na různých hladinách významnosti
- ◆ problémy:
 - ◆ jakou zvolit metriku (diameter, minimum, mean dist.)
 - ◆ neuniformní distribuce entropie v jednotlivých dimenzích (každá složka N -tice má jinou váhu / důležitost)

Shluková analýza – příklad

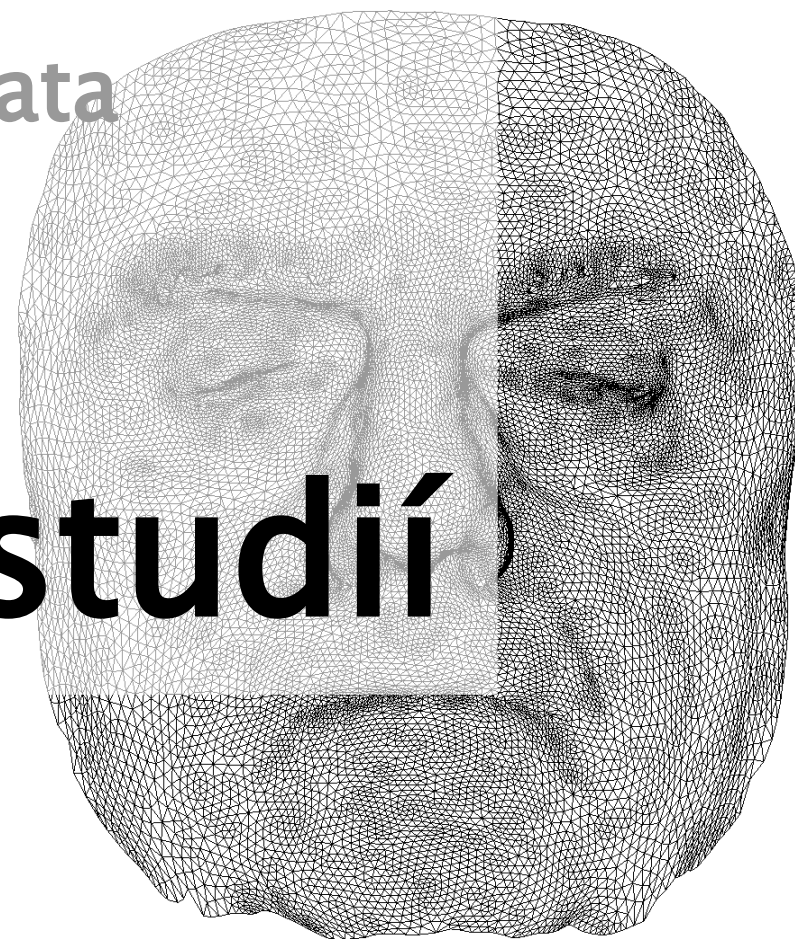
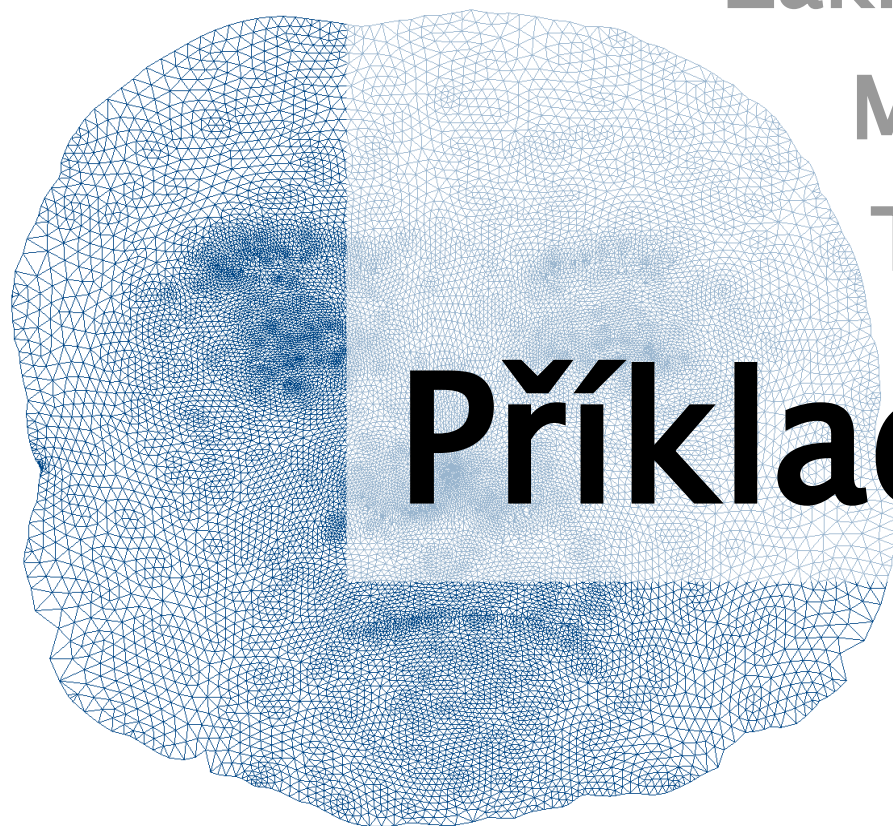


- hierarchické shlukování ... výsledný dendrogram:



Aplikace
Základy, data
Metody
Teorie

Příklady studií



Studie 1: longitudální změny obličeje

- opakování stejných měření v delším časovém úseku
 - studium růstu jedince (ontogeneze)
 - 3D skeny obličejů



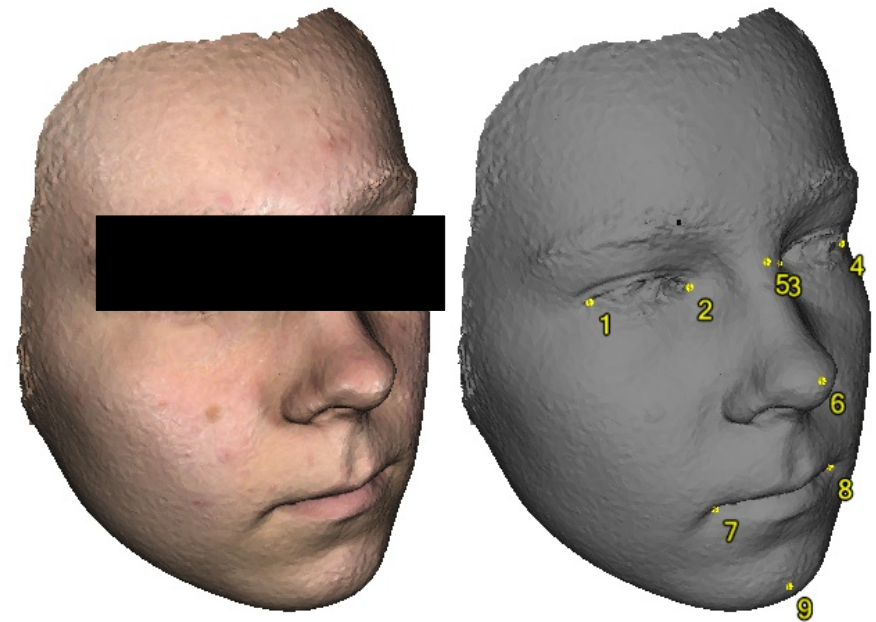
Data pořízená v 12 až 15 letech věku

3D obličejový skener



◆ Canfield Vectra 3D

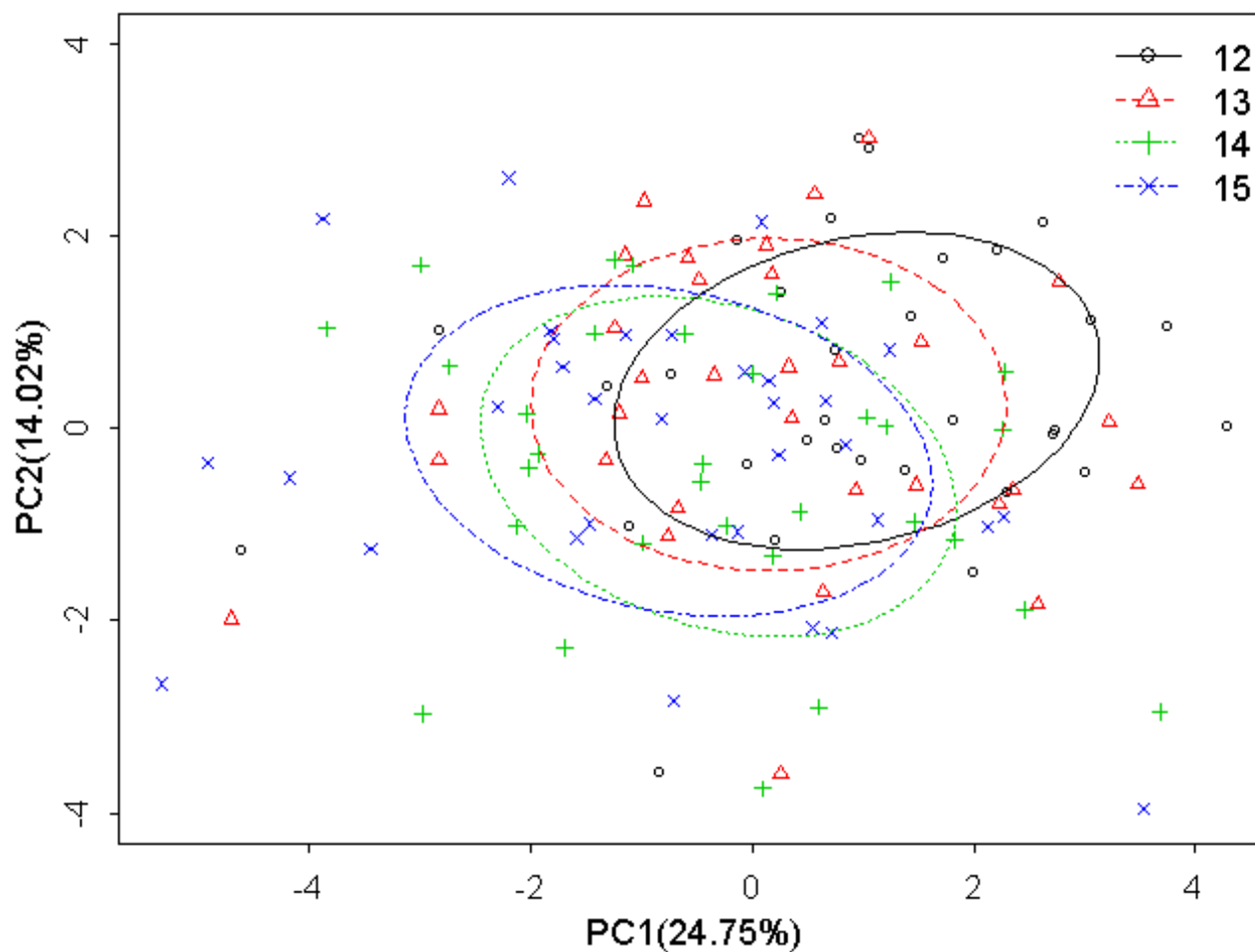
- ◆ optický skener (stereo-fotogrammetrie)
- ◆ projekce umělého vzorku (podpora 3D registrace)





Variabilita tvaru, vývoj

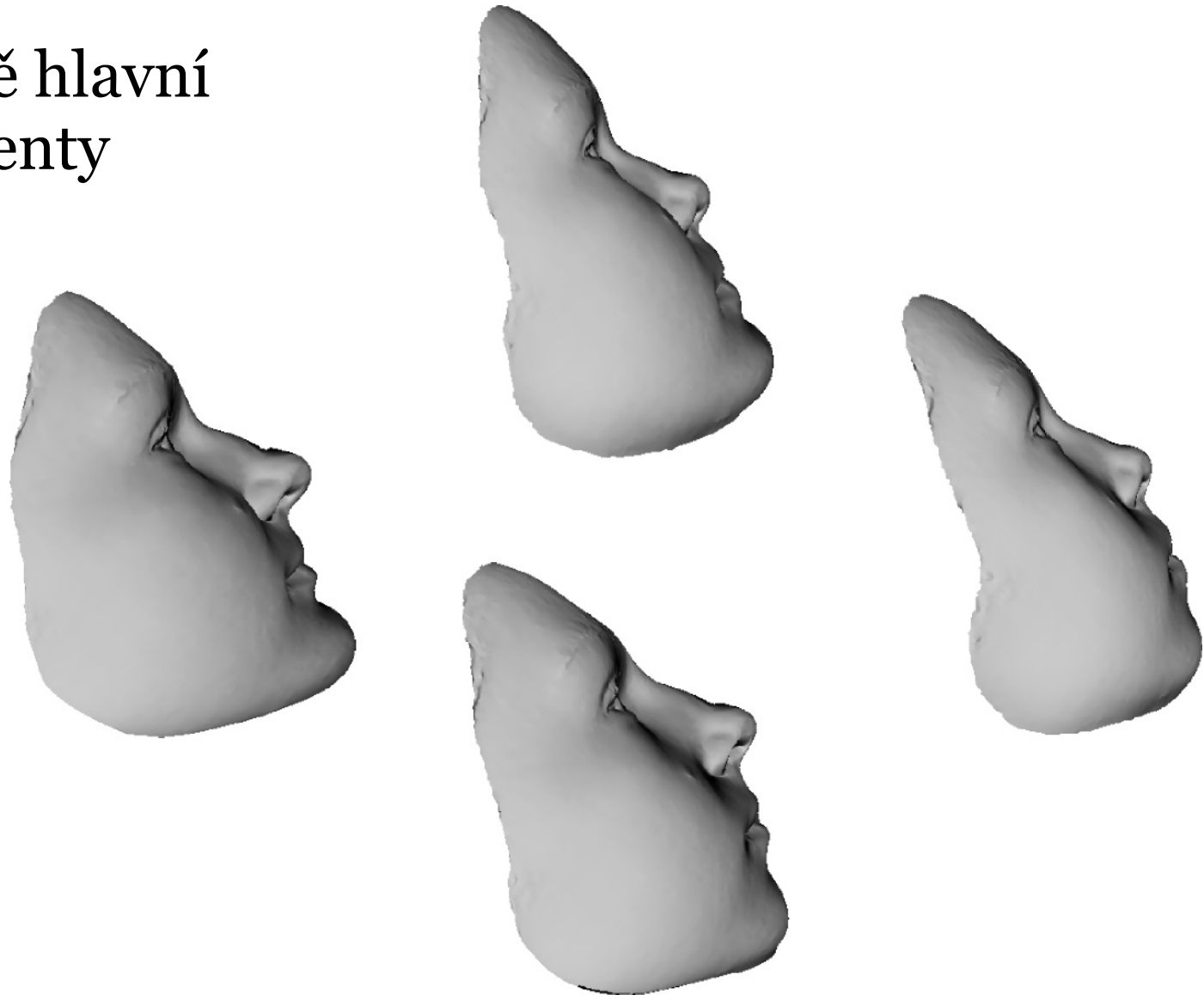
- ◆ celkový soubor, variabilita se mění s věkem



Variabilita tvaru, interpretace



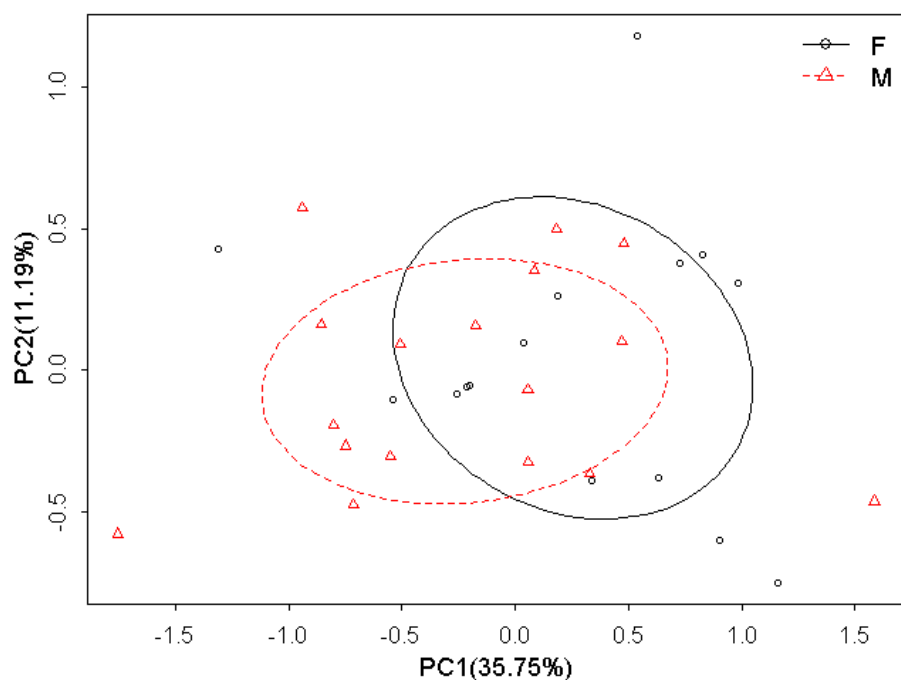
- ◆ první dvě hlavní komponenty



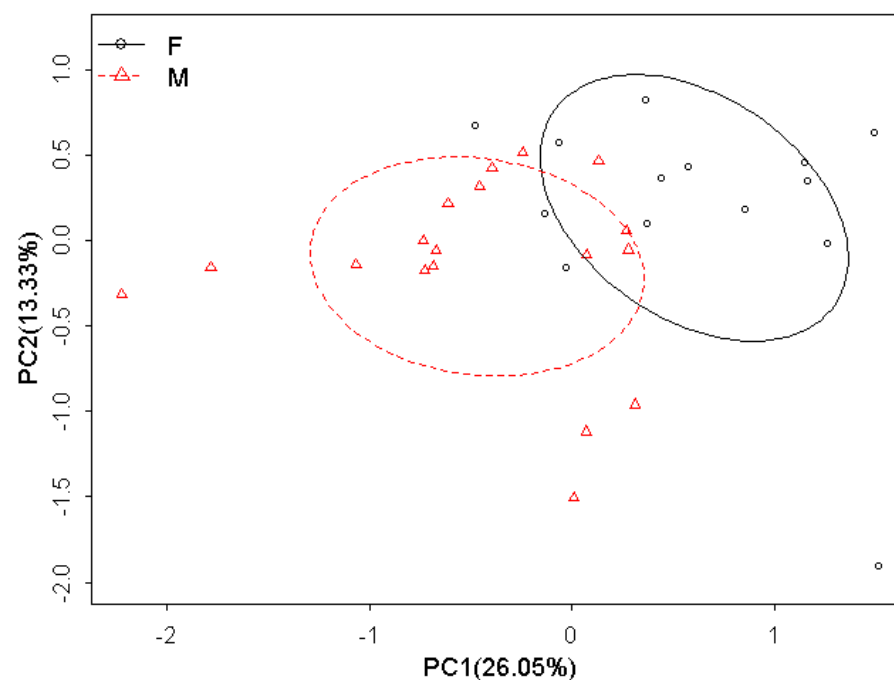


Pohlavní dimorfismus

- ▶ chlapci se od dívek odlišují výrazněji až ve vyšším věku



12 až 13 let

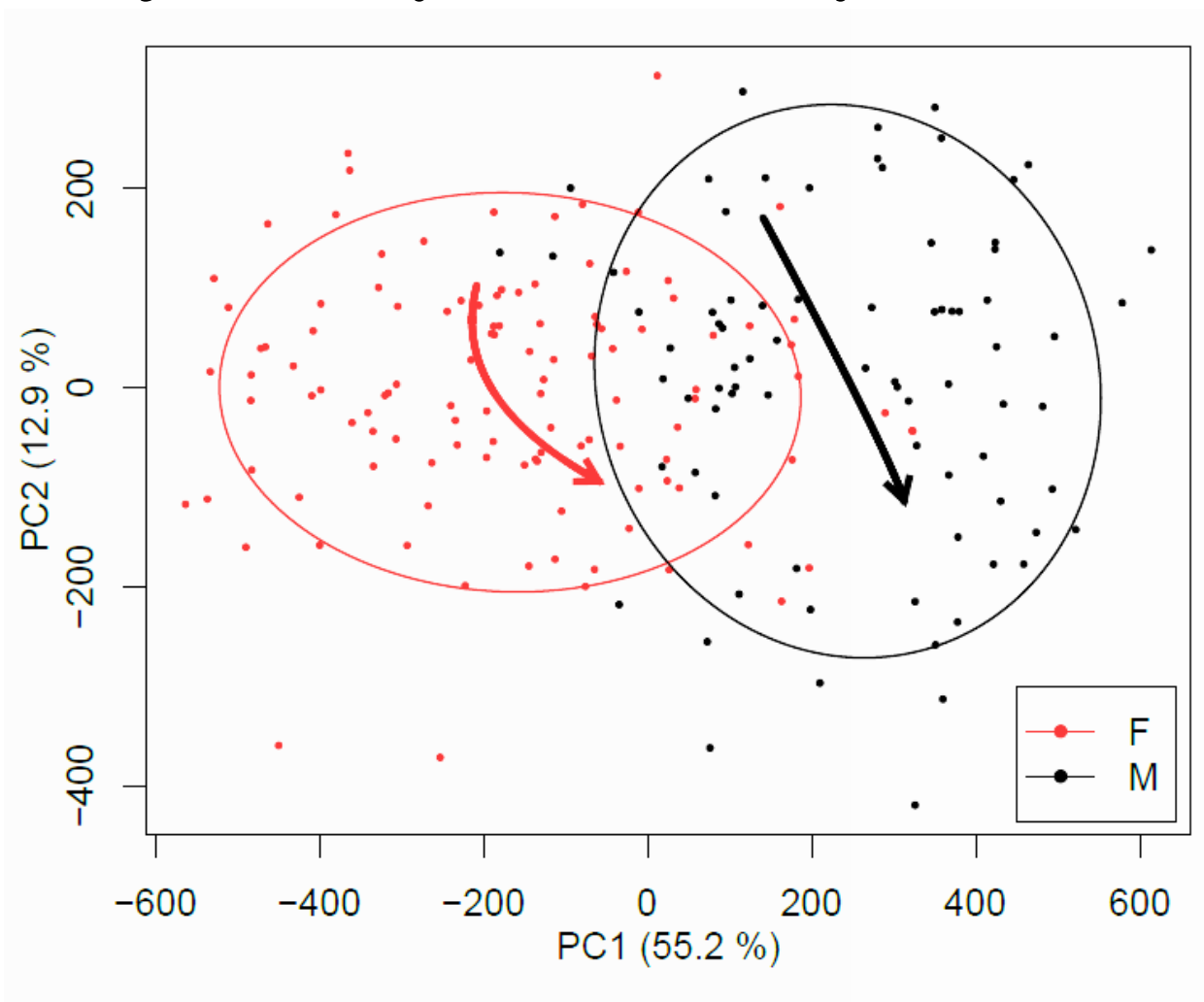


12 až 15 let

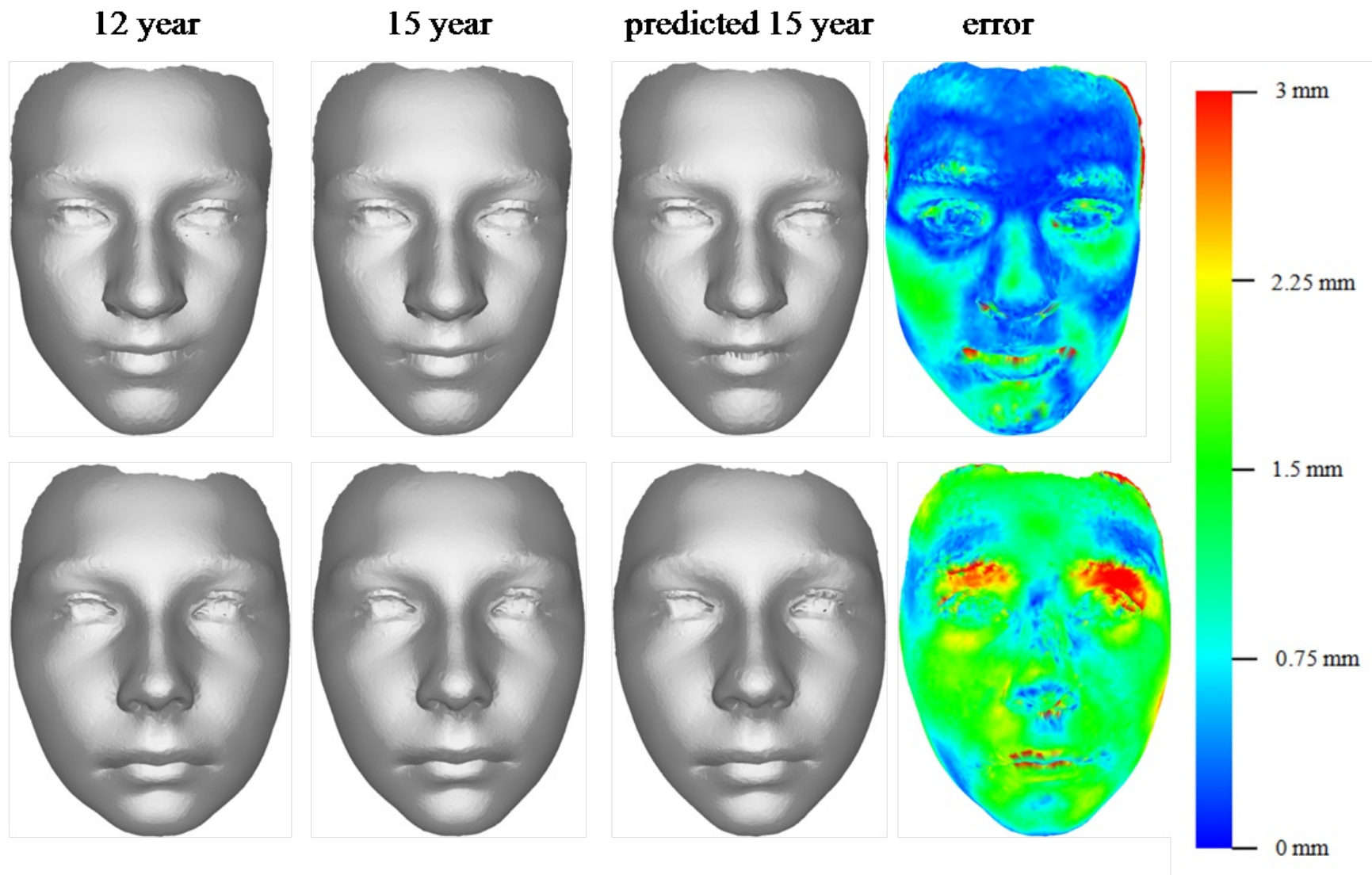


Longitudální studie 20–80 let

- variabilnější datový soubor, ženy i muži

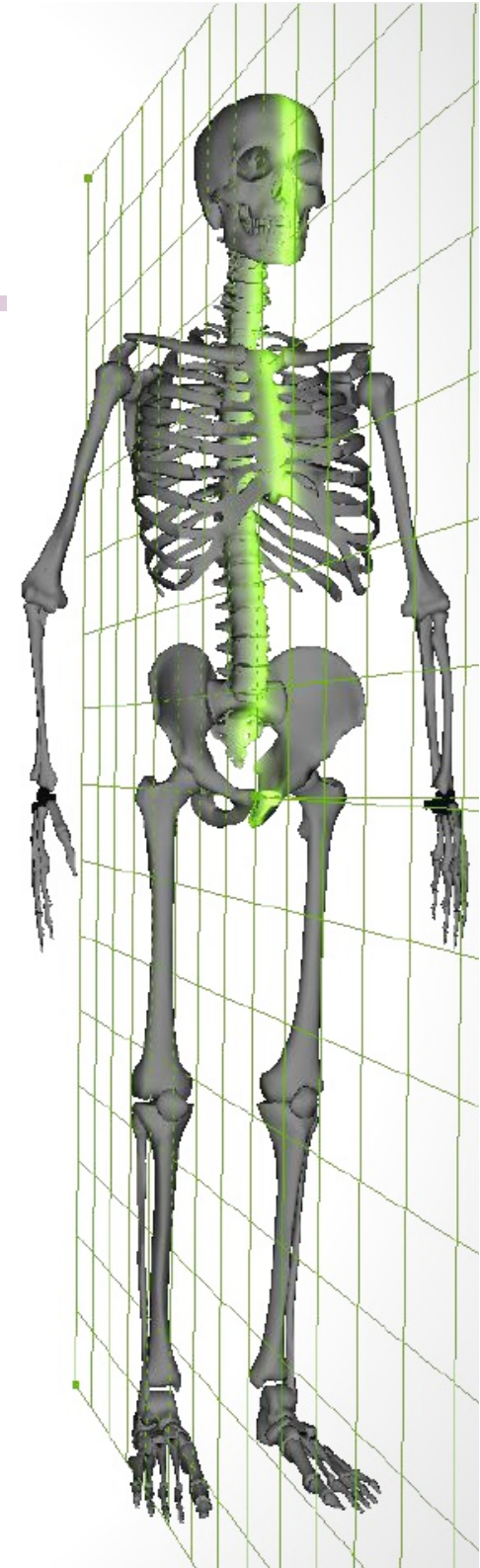


Predikce stárnutí (12–15 let)

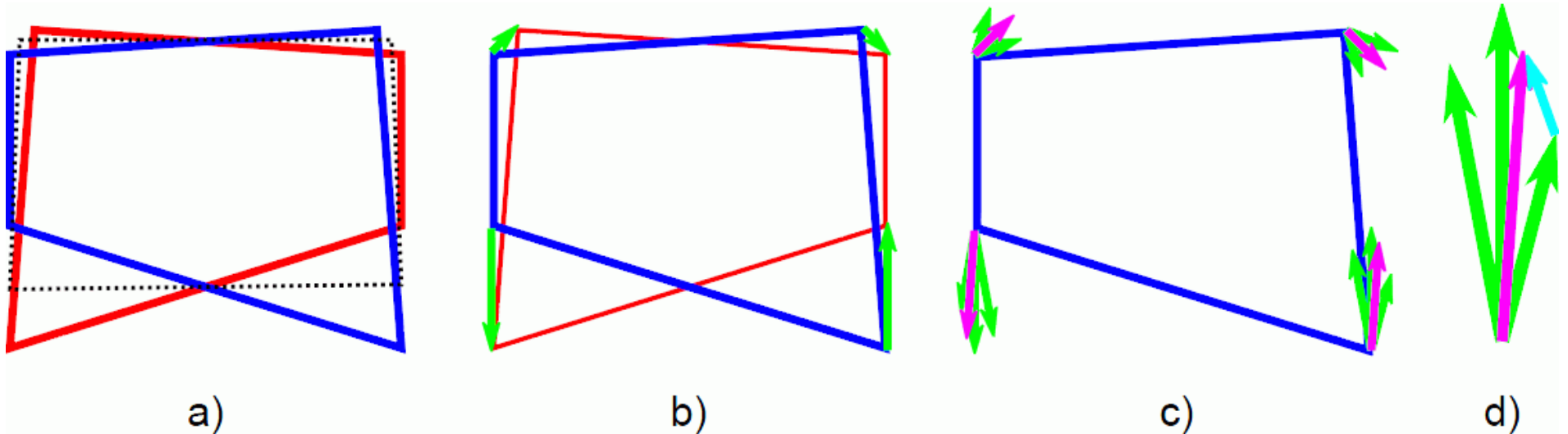


Studie 2: symetrie

- ◆ **symetrie**
- ◆ **asymetrie** (nejčastěji se vyskytuje trend jedním směrem)
- ◆ **anti-symetrie** (žádná symetrická data)
- ◆ **hledání roviny symetrie**
 - ◆ nemusí být jednoznačně definovaná (např. povrch obličeje člověka)
 - ◆ matematický přístup: minimalizace kvadratické vzdálenosti plochy a jejího zrcadlového obrazu



Analýzy asymetrie

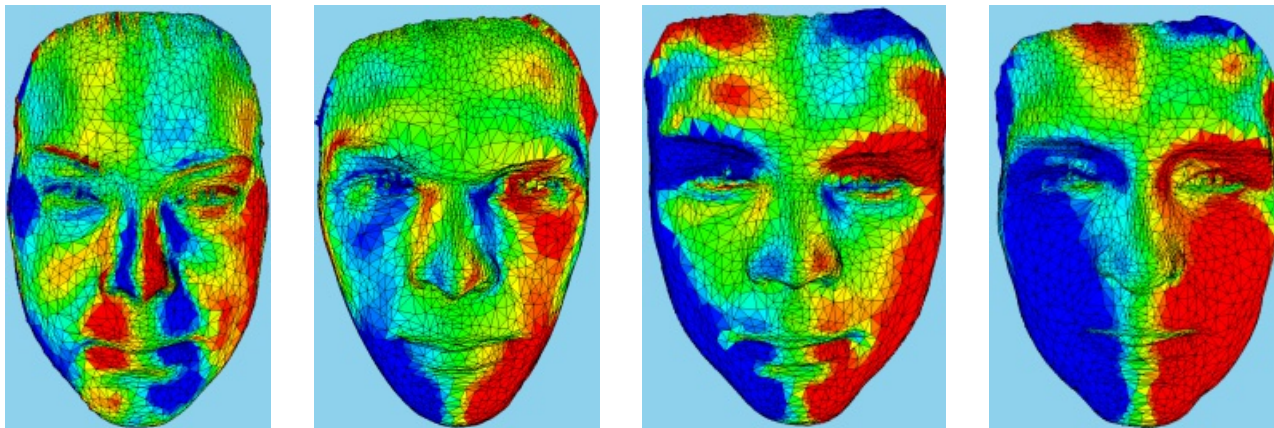
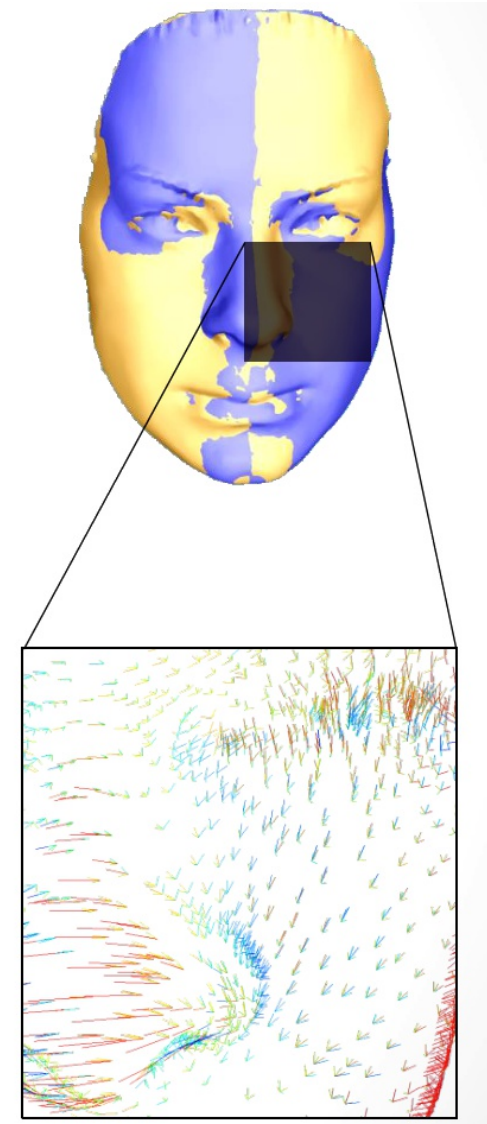
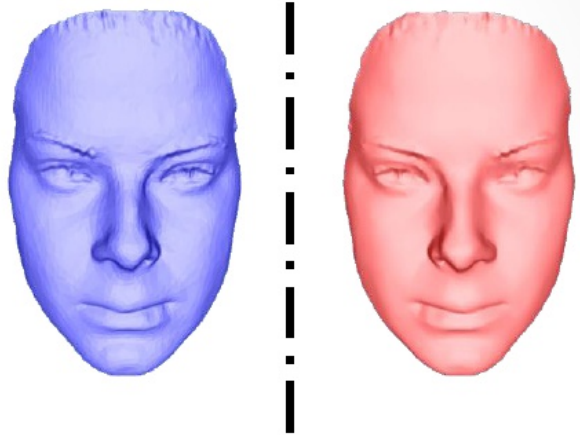


- a) konstrukce zrcadlového a dokonale symetrického jedince
- b) individuální asymetrie (rozdíly mezi landmarky pův. a zrcadl.)
- c) direkcionální asymetrie (zprůměrované individuální asymetrie)
- d) fluktuační asymetrie (rozdíl individuální a direkcionální asymetrie)

Analýzy asymetrie obličeje



- ◆ individuální asymetrie
 - ◆ zarovnání sítě se zrcadlovým obrazem

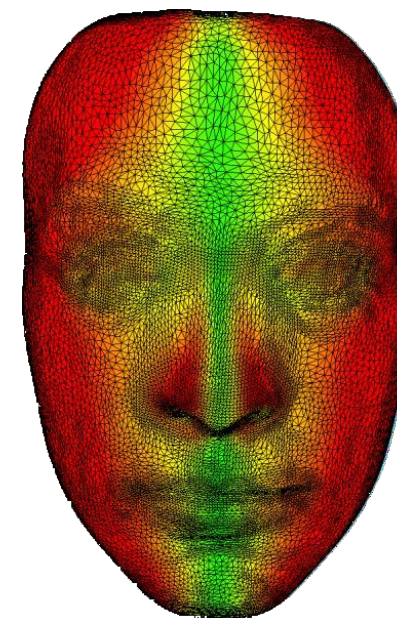
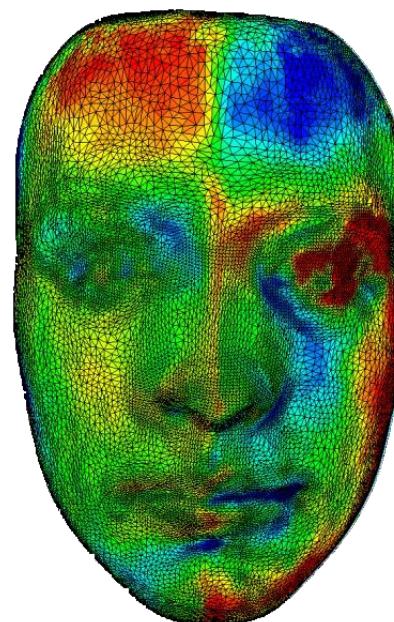




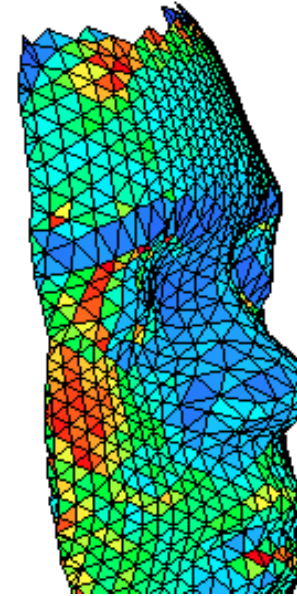
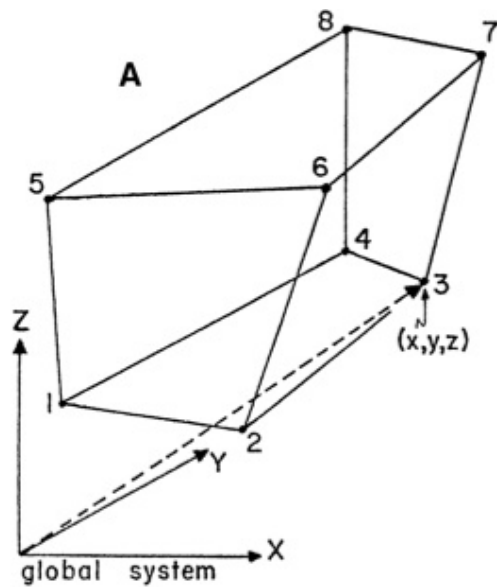
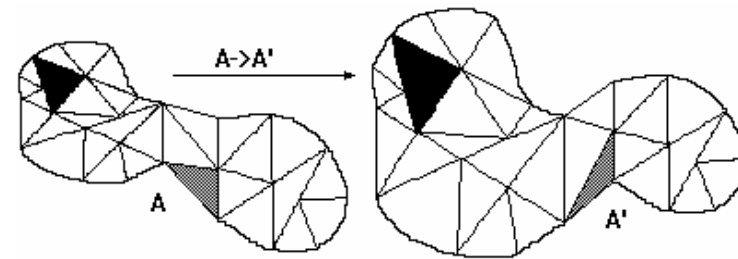
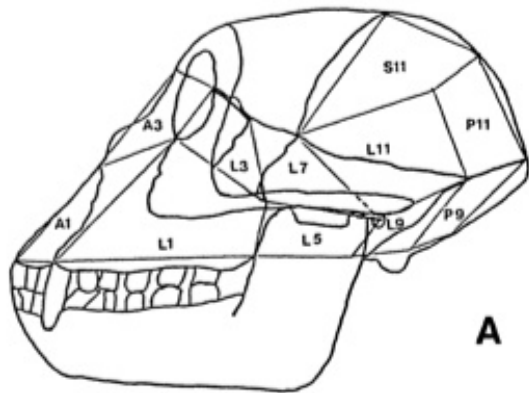
Zkoumané asymetrie

- ◆ direkcionální asymetrie
 - ◆ systematická asymetrie
 - ◆ soubor 71 jedinců

- ◆ fluktuační asymetrie
 - ◆ asymetrie jedince po odečtení systematické asymetrie
 - ◆ ukázka: průměrné absolutní hodnoty



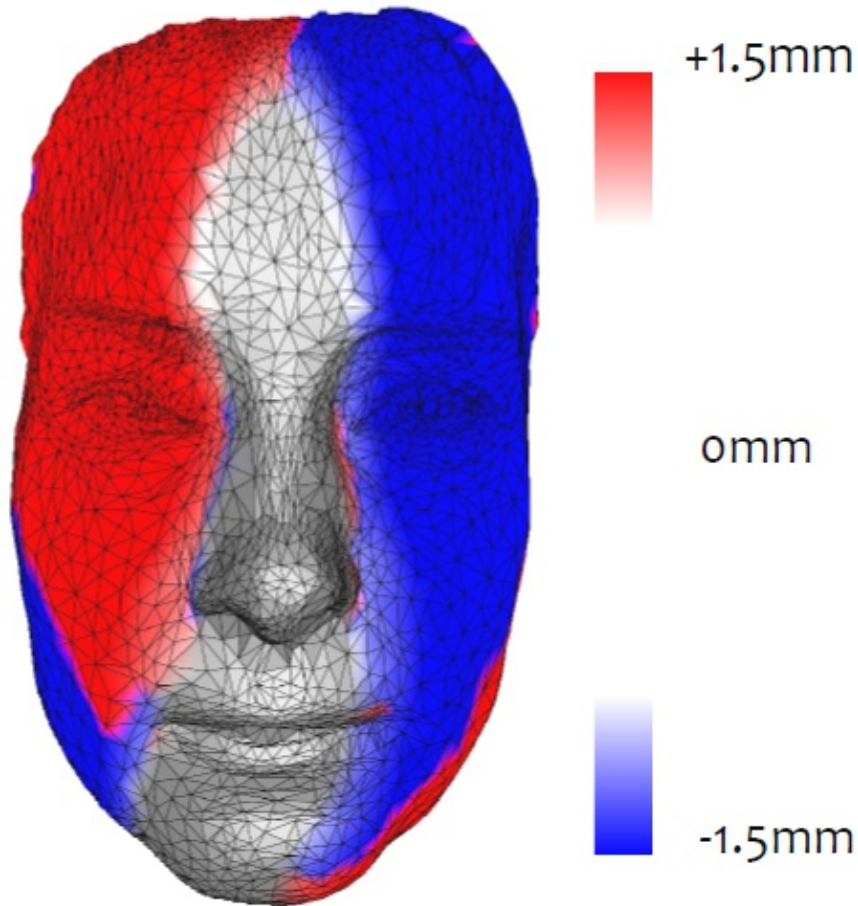
Metody konečných prvků (FEM)



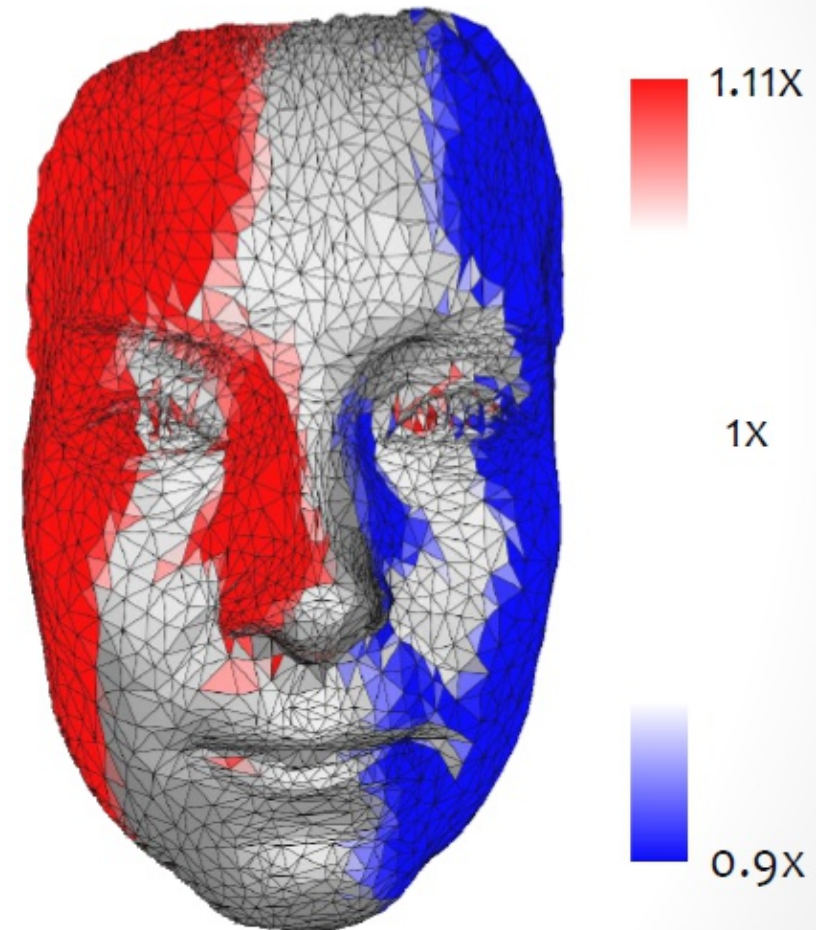
FESA (analýza růstu)



- lokální zvětšení/zmenšení sítě



Orientované posunutí



FESA

Studie 3: analýza křivky (kontury)



- ◆ hladká hranice, kontura
 - ◆ neobsahuje přímo landmarky (jen koncové body)
- ◆ segmentace kontury

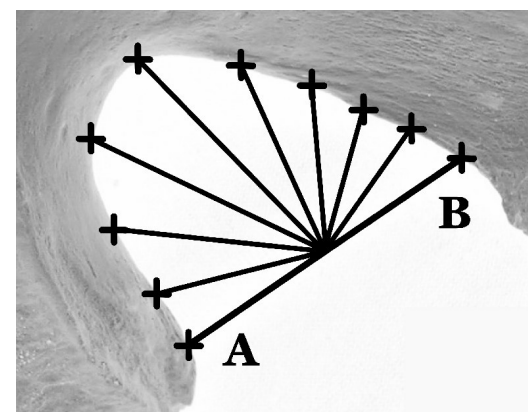
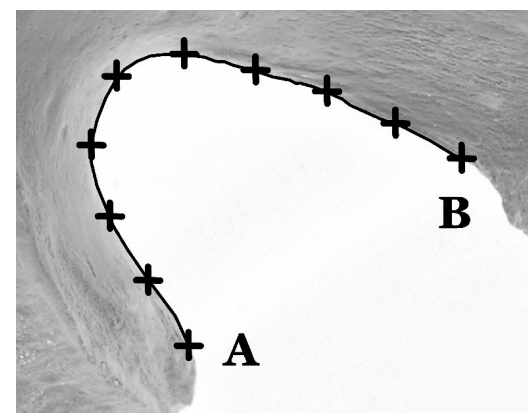
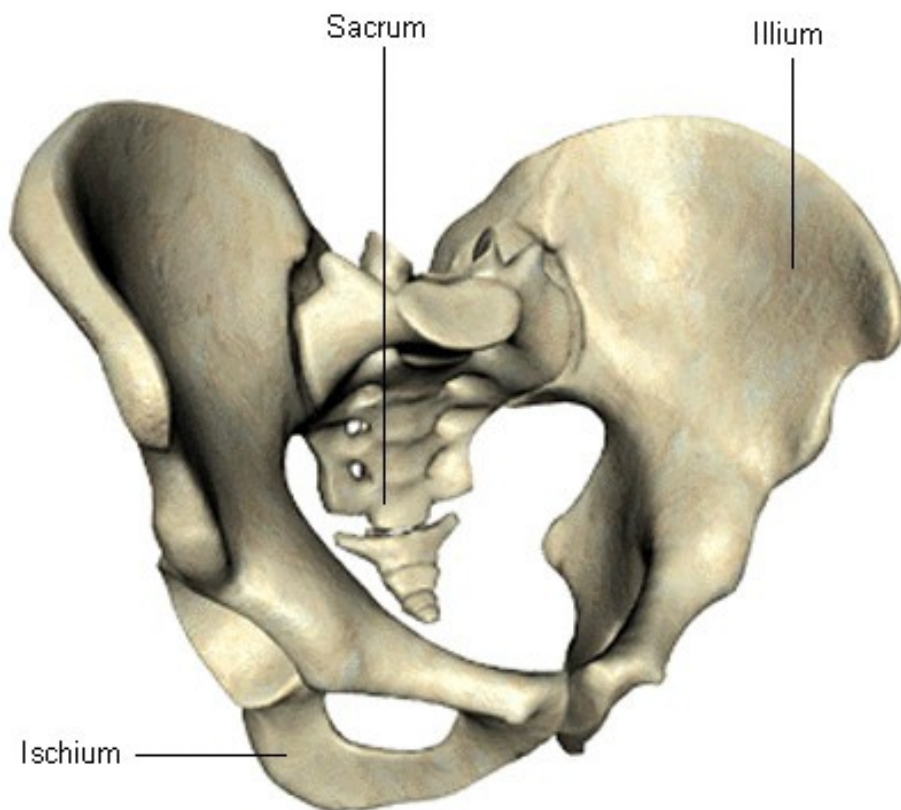
Jak konturu reprezentovat čísly?

1. semilandmarky - body
 - ◆ rozdělení kontury na úseky podle délky, nebo úhlu
2. transformace na soubor koeficientů
 - ◆ aproximace nějakou křivkou a práce s koeficienty
 - ◆ Waveletová (vlnková) / Fourierova transformace navzorkované křivky



Konkrétní studie – incisury

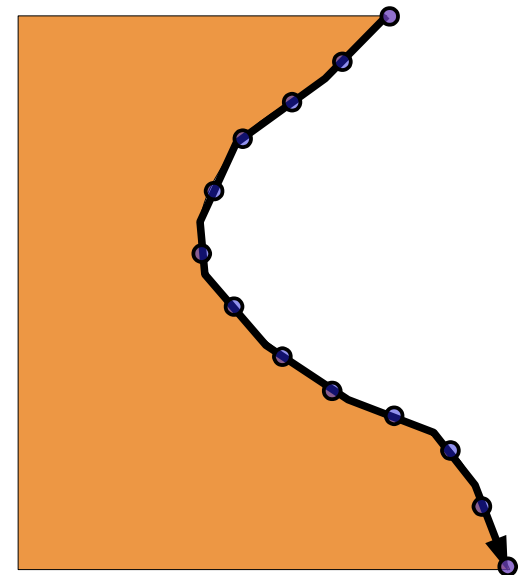
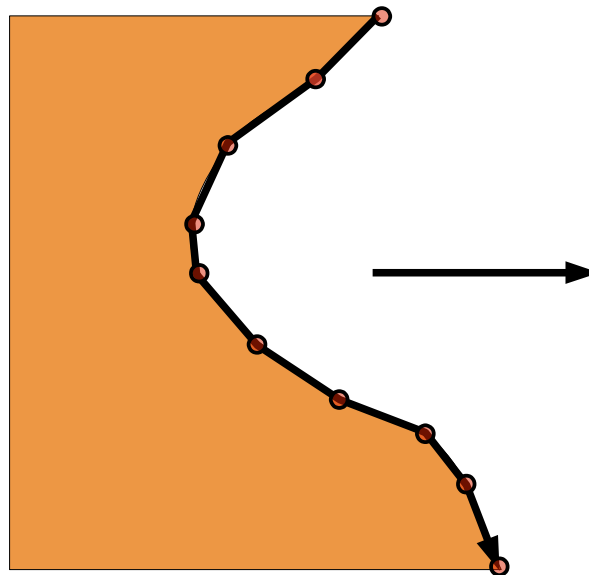
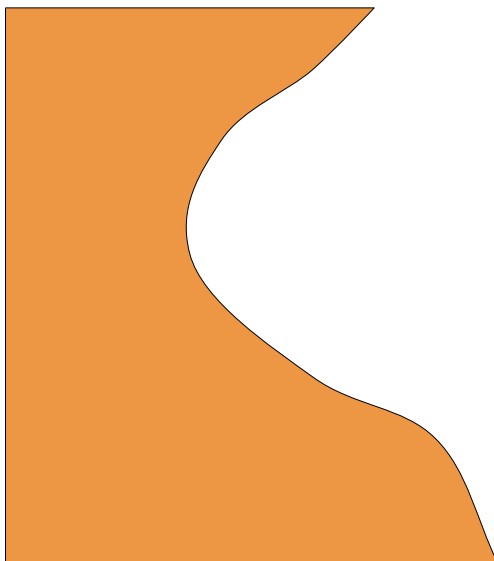
- ✦ Incisura ischiadica major
 - ✦ pro experta zdroj informací nejen o pohlaví jedince



Semilandmarky



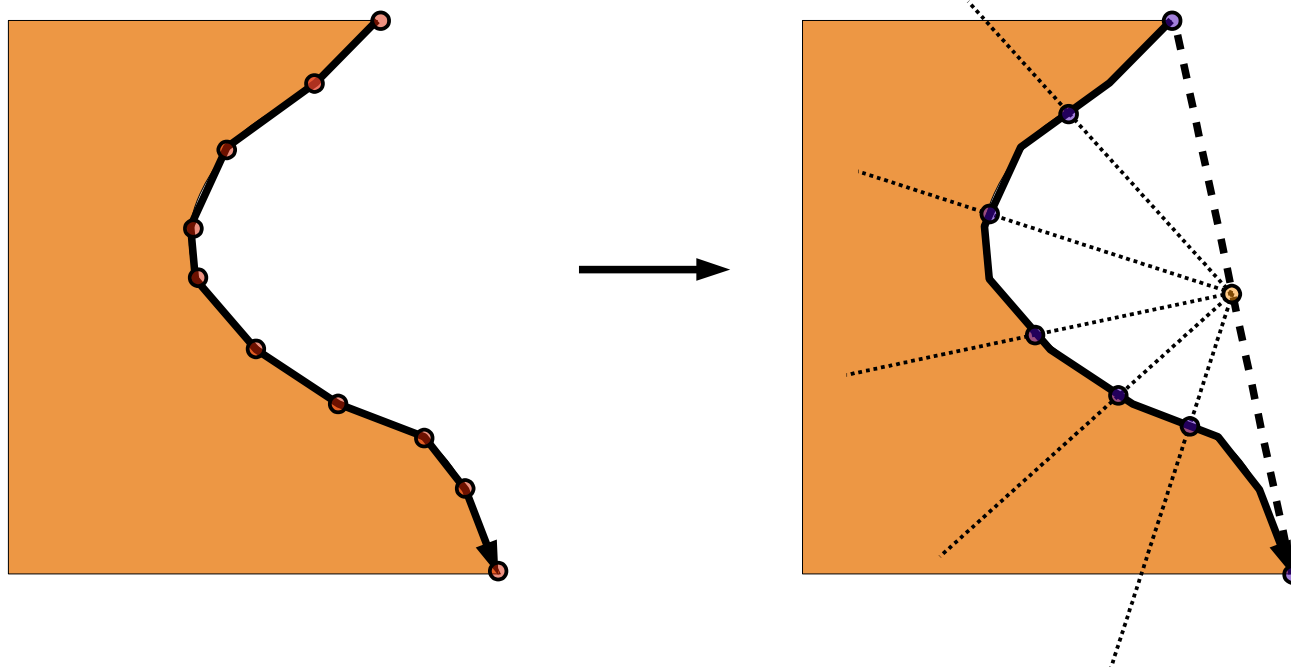
- dělení podle délky křivky (arc) na potřebný počet n úseků



Semilandmarky II



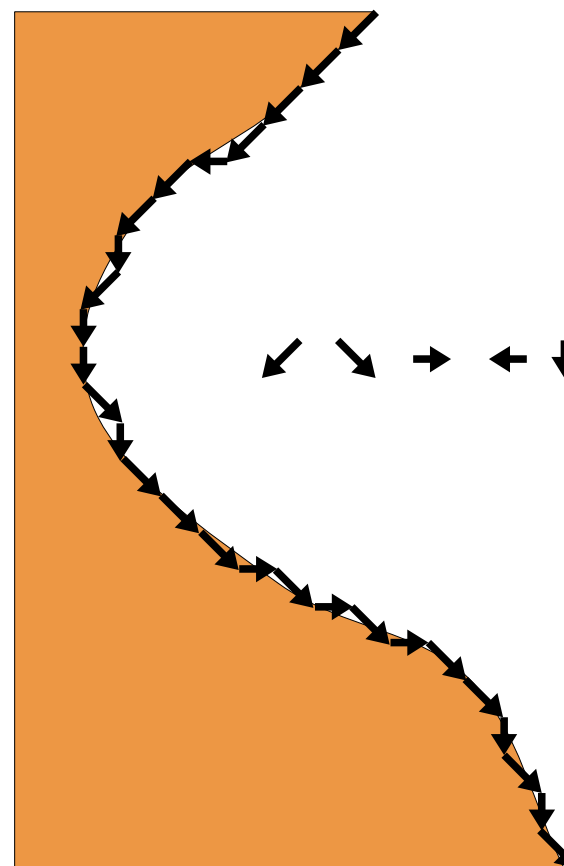
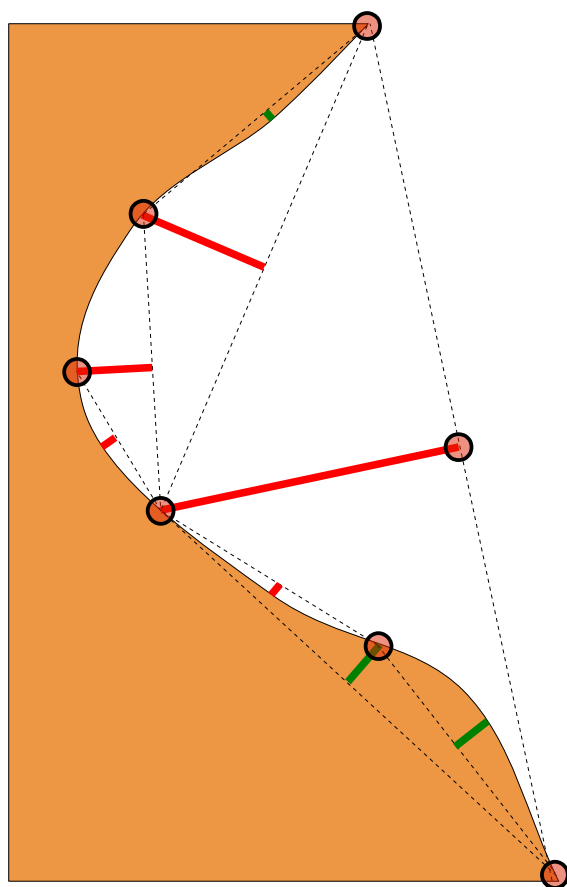
♦ dělení podle úhlu





Jiné reprezentace křivek

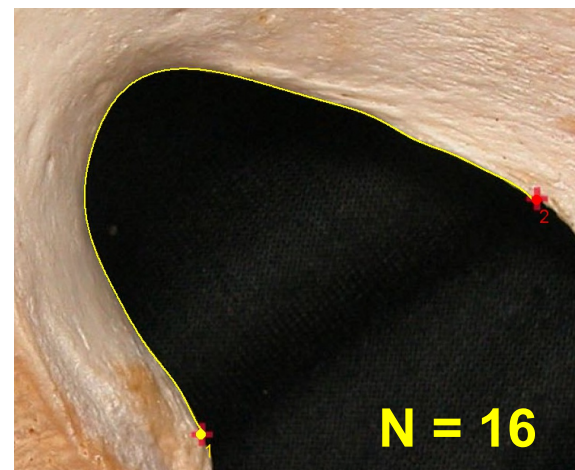
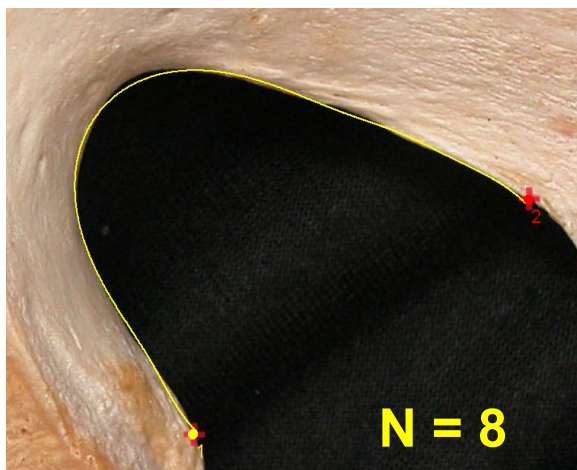
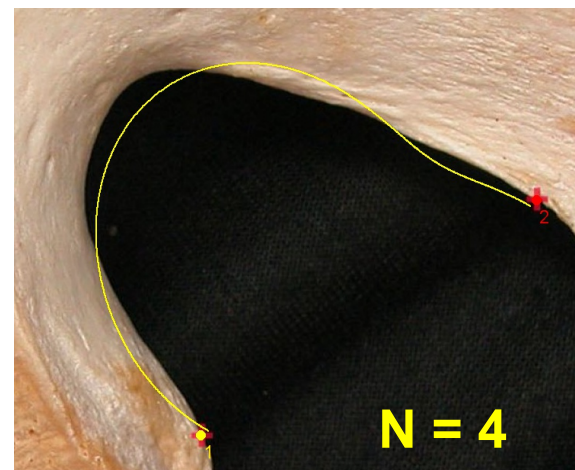
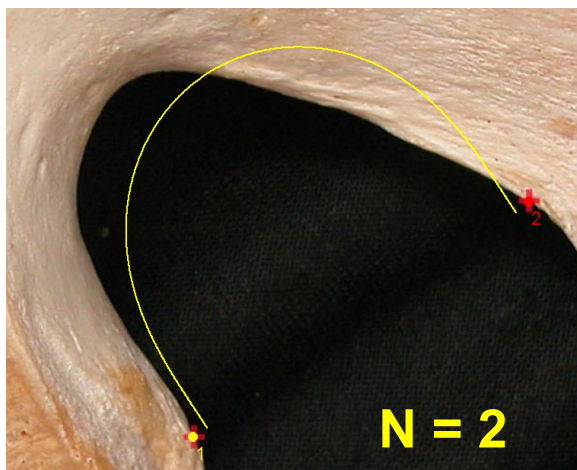
- mnoho různých způsobu jak zachytit tvar kontury



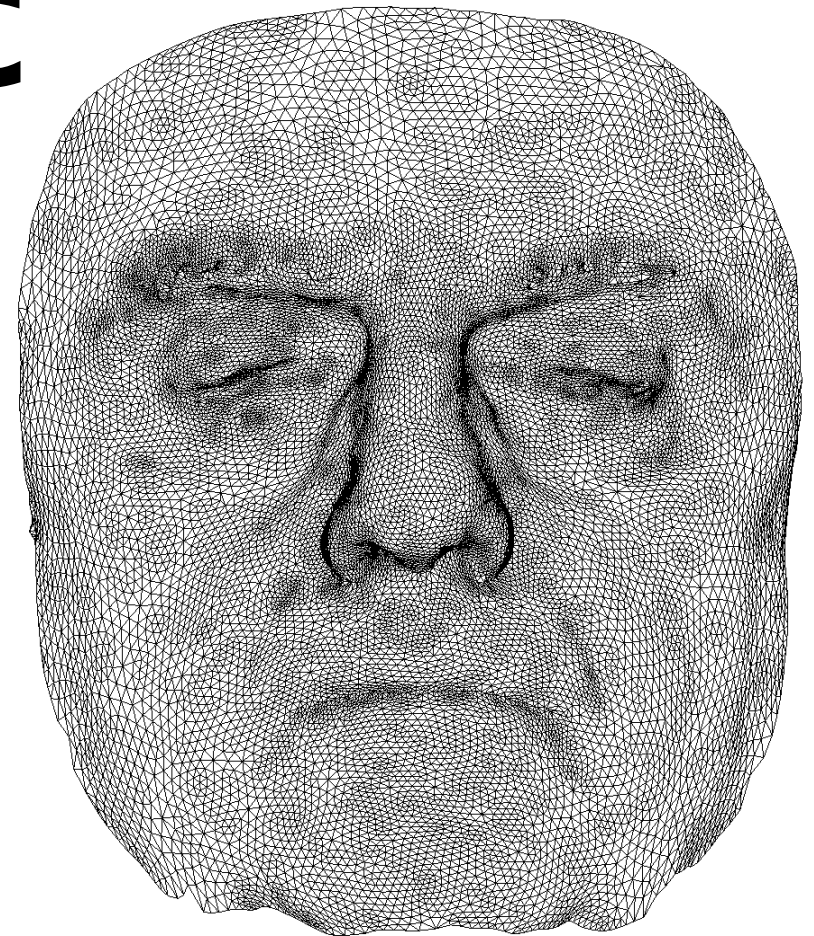
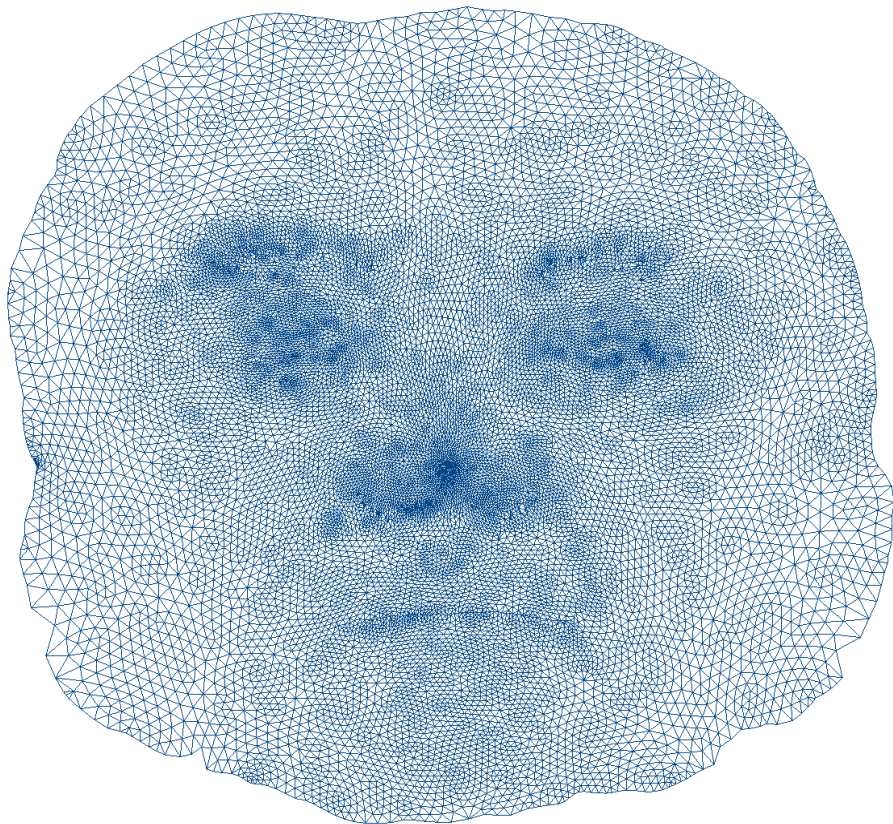
Cirkulární harmoniky



- ◆ Legendrové polynomy nebo Fourierovy koeficienty



Konec



Skupina počítačové grafiky a IP



- ◆ 3 full-time, několik externistů, ~10 PhD studentů
 - ◆ <http://cgg.mff.cuni.cz/>
- ◆ **široké spektrum přednášek (viz WWW)**
 - ◆ klasická i výpočetní geometrie, křivky a plochy, geometrická morfometrie
 - ◆ digitální analýza obrazu, rozpoznávání vzorů, komprese obrazu
 - ◆ 2D a 3D grafika, fotorealistické zobrazování
 - ◆ vědecké vizualizace, zpracování a zobrazování objemových dat
 - ◆ realtime grafika, virtuální realita
- ◆ specializace u Bc. státnic, speciální Mgr. obor
 - ◆ každoročně 5-10 absolventů
- ◆ PhD obor „Počítačová grafika a analýza obrazu“