

Úvod do GIS

Zpracování dat I

Pouze podkladová prezentace k přednáškám, nejedná se o studijní materiál pro samostatné studium.

Karel Jedlička



Uchovávání a zpracování dat

- Metadata
- Uchovávání dat
- Zpracování dat
 - Geometrické transformace
 - Zpracování obrazu
 - Převody mezi reprezentacemi
 - Vektorizace
 - Rasterizace
 - Interpolace
 - Tvorba TIN
 - ...



Metadata

- Co jsou metadata?
 - Obecně se dá říci, že jsou to **data o datech**, tzn. informace co popisovaná data obsahují a kde se nacházejí.
 - Tyto informace jsou zvláště důležité, pokud je zpracováváno několik druhů dat, či velký objem dat.
 - Pomáhají lépe organizovat a udržovat přehled na daty. Problematika tvorby metadat je úzce spjata s tvorbou digitálních dat a jejich převody.

Metadata

- V metadatech by měla být obsažena následující informace:
 - **Co je obsahem dat** (tématická složka).
 - **Rozlišení – prostorové** (jaké území zabírají), **popisné** (popis možných hodnot atributů a jejich význam) a **časové** (jakou dobu zahrnují – kdy byla data aktuální).
 - **Formát dat** (typ souboru, předávací médium).
 - **Datum pořízení dat** (případně aktualizace).
 - **Kontakt na pořizovatele a správce.**

Metadata

- Pro metadata je vhodné použít **standardní formát**.
- Dnes již existuje několik standardů:
 - ISO 19115 – metadata dat
 - ISO 19119 – metadata služeb
 - ISO 19139 – reprezentace metadat
 - **ISO 15836 – Dublin Core**
- INSPIRE v ČR:
 - <http://inspire.gov.cz/implementacni-pravidla/metadata>
 - <http://geoportal.gov.cz/web/guest/validate/metadata>

Uchovávání a zpracování dat

- Uchovávání dat

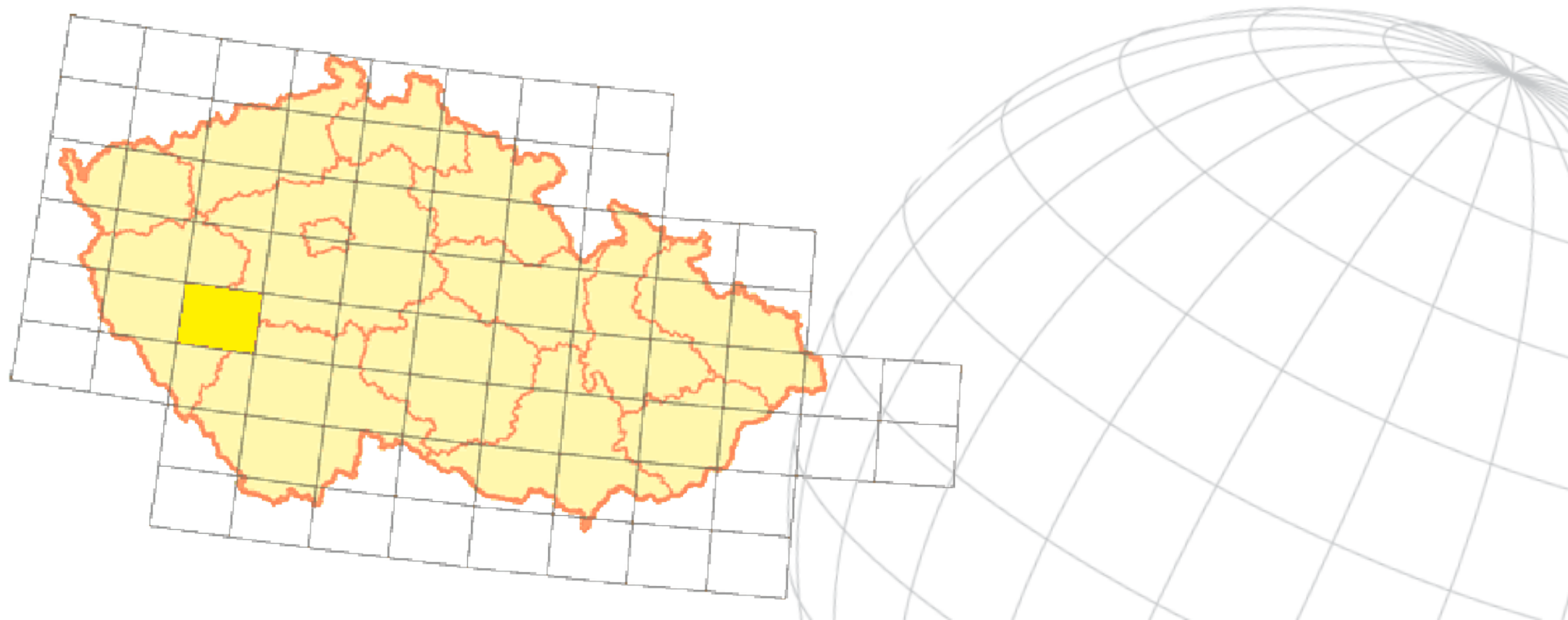
- Geografická data se obvykle před uchováním rozdělují na oblasti a podle těchto oblastí se ukládají na disk:

- Pravidelné
 - Nepravidelné
 - Bezešvé



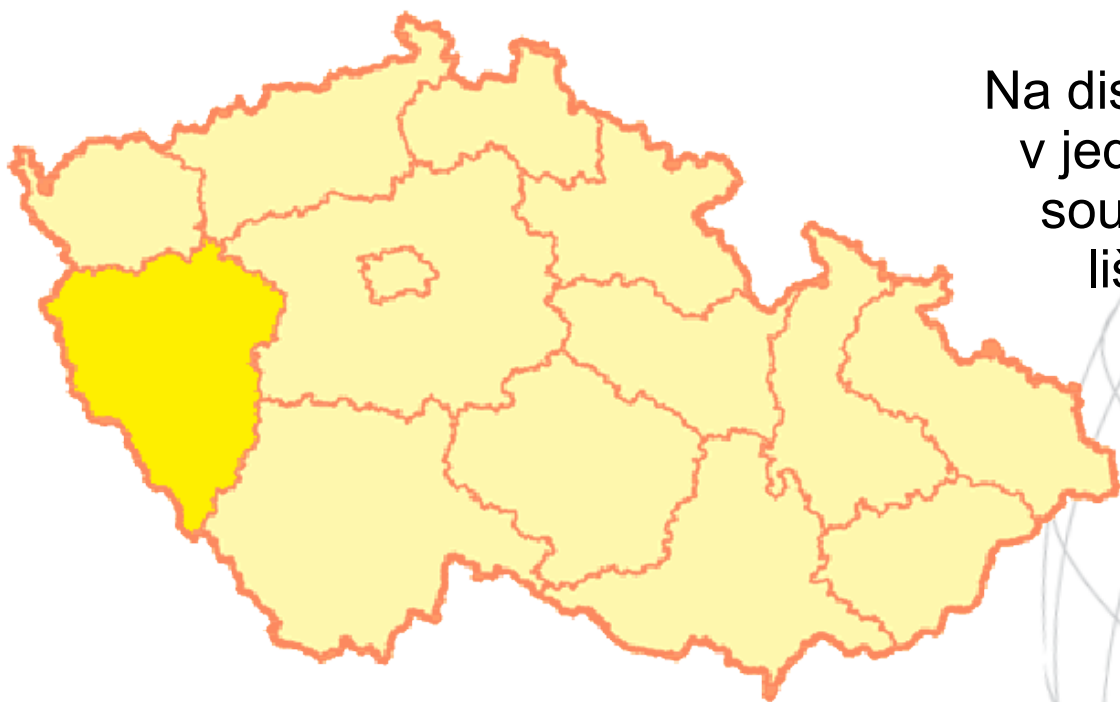
Uchovávání a zpracování dat

- Dělení na oblasti
 - **Pravidelné** (např. mapové listy).
Na disku je každý mapový list v jednom souboru (resp. ve více souborech se stejným jménem, lišících se pouze příponou) či adresáři.

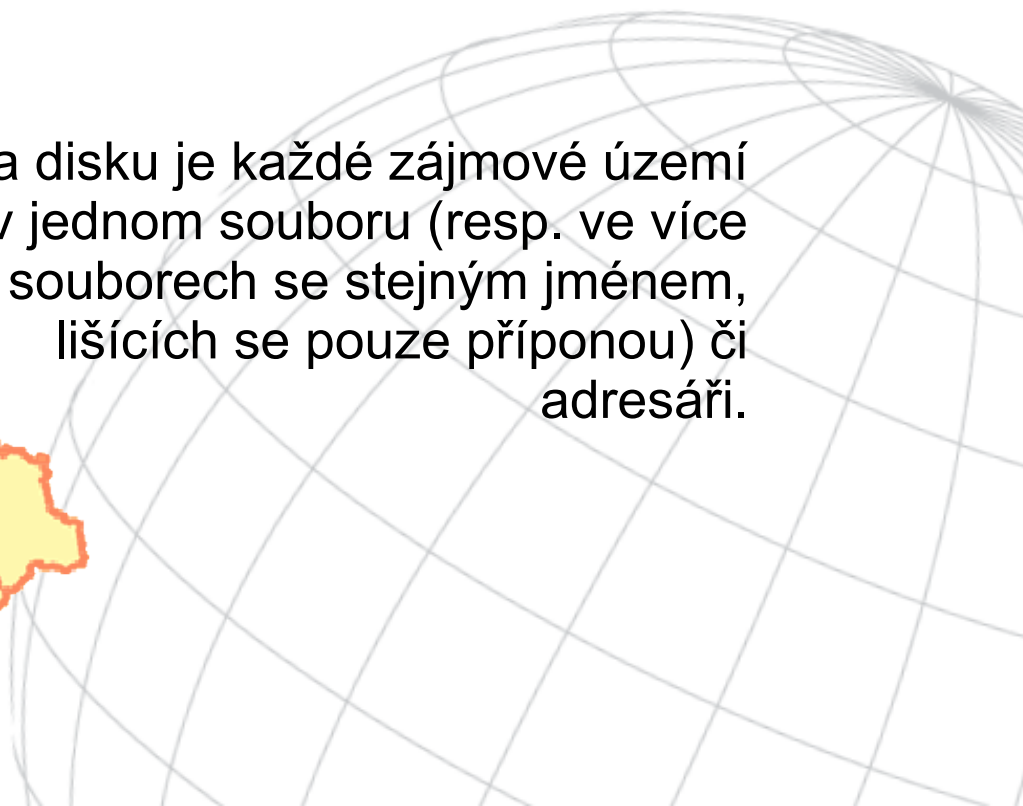


Uchovávání a zpracování dat

- Dělení na oblasti
 - **Nepravidelné** (mapové listy, zájmové území - katastrální území, území národního parku, okresu, kraje ...).



Na disku je každé zájmové území v jednom souboru (resp. ve více souborech se stejným jménem, lišících se pouze příponou) či adresáři.

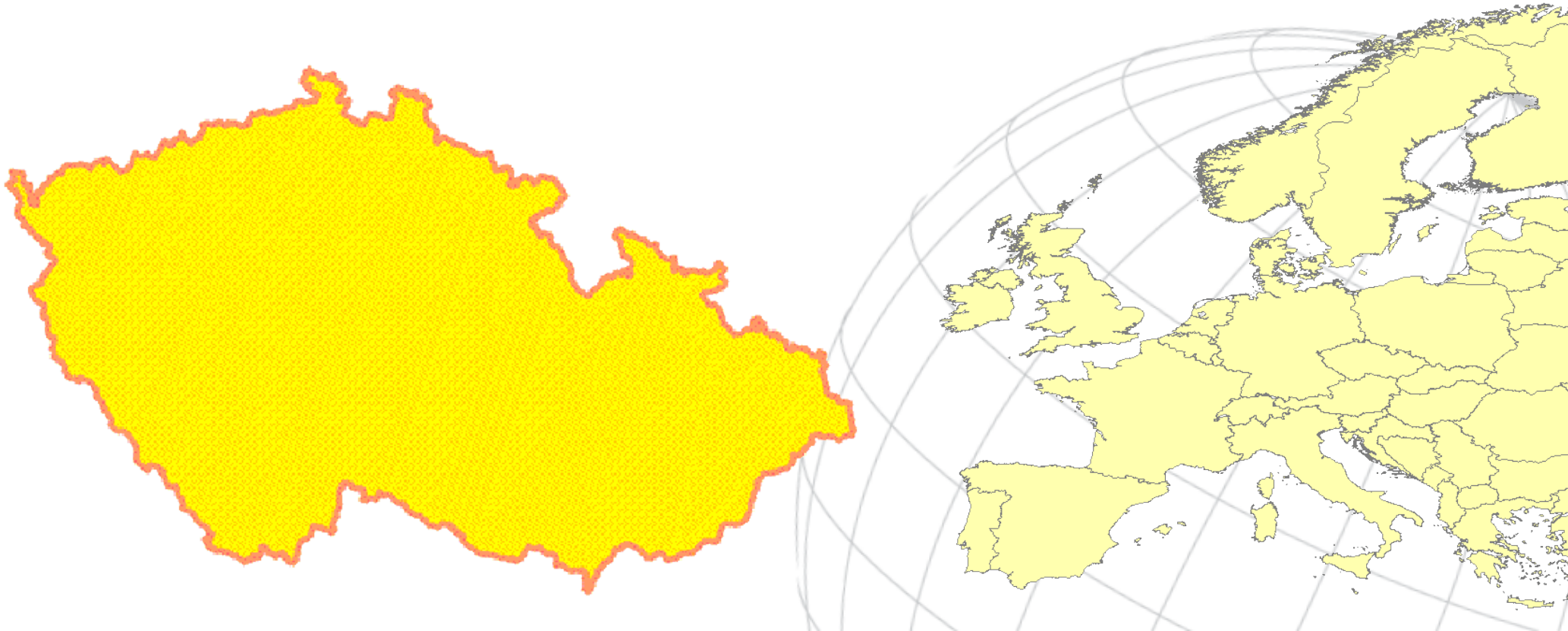


Uchovávání a zpracování dat

- Dělení na oblasti

- **Bezešvé** (Seamless)

- Celé zájmové území je uloženo v jednom souboru, adresáři či databázi).



Uchovávání a zpracování dat

- Metadata
- Uchovávání dat
- Zpracování dat
 - Geometrické transformace
 - Zpracování obrazu
 - Převody mezi reprezentacemi
 - Vektorizace
 - Rasterizace
 - Interpolace
 - Tvorba TIN
 - ...



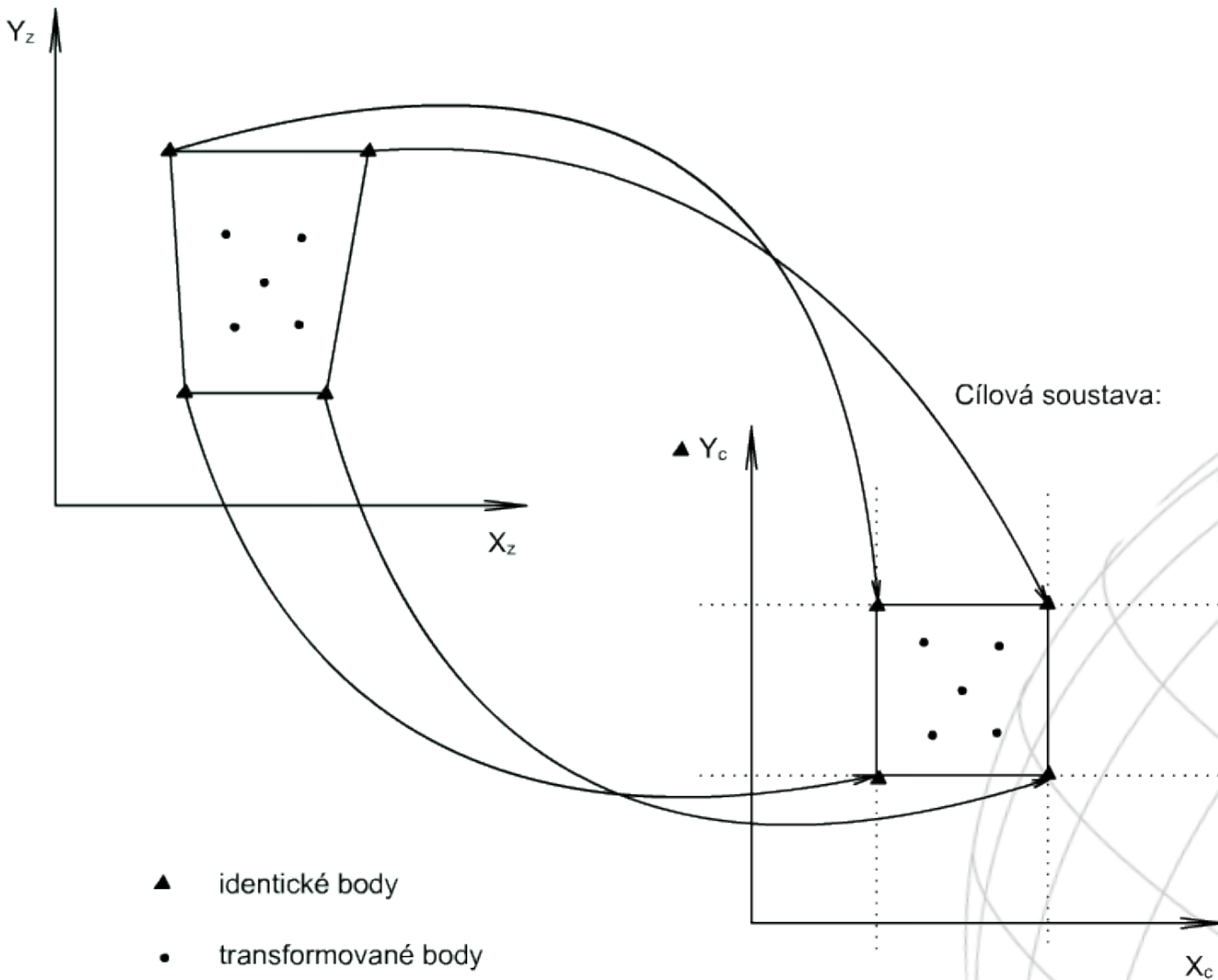
Geometrické transformace

- V GIS se používají 2D i 3D transformace.
- Problematice 3D transformací mezi souřadnicovými systémy se věnuje matematická kartografie.
- My se zde budeme zabývat pouze jednoduchými 2D transformacemi (matematické pozadí bude jen nastíněno).
- Jedná se v podstatě o stejnou problematiku, jaká byla již zmiňována u digitalizace a u scannování!
- Nyní se na ni podíváme podrobněji.

Geometrické transformace

- **Transformace** mezi rovinnými pravoúhlymi souřadnicemi jsou založeny na poznání přesné polohy vybraných **identických bodů**.

Zdrojová soustava:

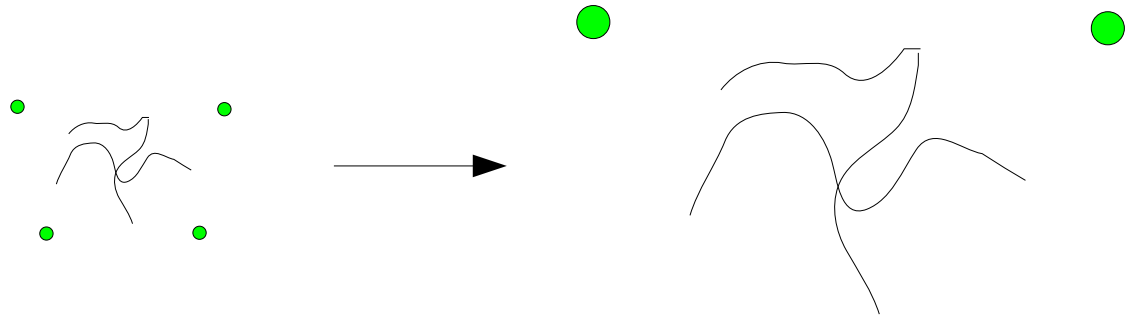


- ▲ identické body
- transformované body
- transformace

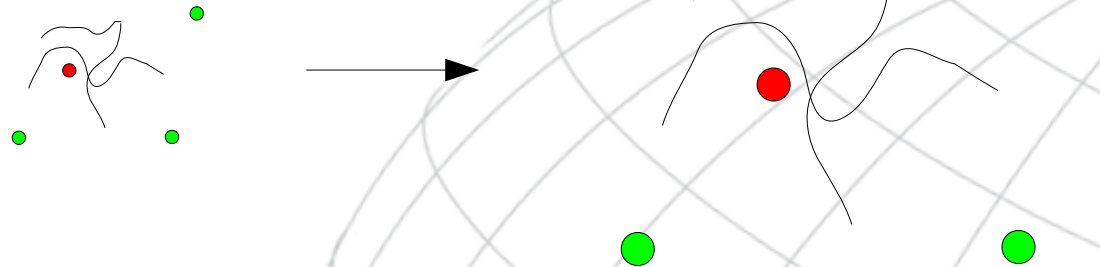
Geometrické transformace

- Volba identických bodů – **připomenutí**

» dobře



» špatně



- U výběru dvojic identických bodů je vhodné mít na paměti, že je nutné je vybírat **co nejbližše okrajům** transformovaného území, aby nebyly způsobeny nežádoucí deformace na okrajích.

Geometrické transformace

- Identické body a transformační koeficienty
 - **Transformační koeficienty** jsou hodnoty, vypočtené z dvojic identických bodů, kterými se vyjadřuje přechod od zdrojové souřadnicové soustavy do cílové.
 - U transformace se ale obvykle **používá více identických bodů, než je nutné** pro výpočet transformačních koeficientů.
 - Hodnoty transformačních koeficientů se pak vypočtou např. **metodou nejmenších čtverců**, kde se minimalizuje suma rozdílů v poloze mezi souřadnicemi transformovaných bodů.
 - Příklad: posun a změna měřítka – **transformace**.

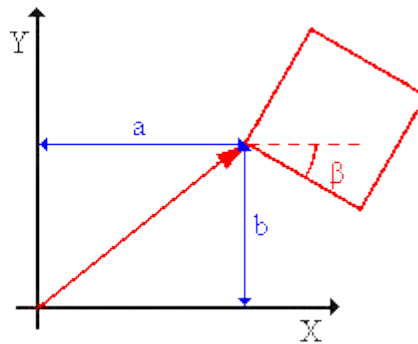
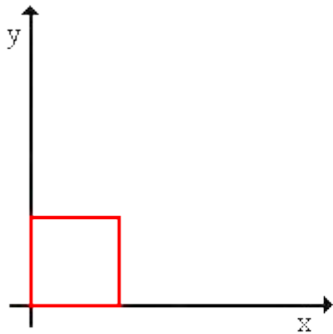
Geometrické transformace

- Transformace souřadnicového systému mezi rovinnými pravoúhlými souřadnicemi:
 - **Lineární konformní** transformace (LKT).
 - **Afinní** transformace (polynomická prvního řádu)
... a polynomické transformace vyšších řádů.
 - **Projektivní** transformace.
 - ... další specializované transformace.



Geometrické transformace

- Lineární konformní transformace



$X(X,Y)$ - nové souřadnice
 $x(x,y)$ - staré souřadnice
 B - úhel otočení
 m - změna měřítka
 $p(a,b)$ - posun

Zápis rovnicí

$$X = m \cdot \cos(B) \cdot x - m \cdot \sin(B) \cdot y + a$$

$$Y = m \cdot \sin(B) \cdot x + m \cdot \cos(B) \cdot y + b$$

Zápis maticí

$$X = m \cdot R \cdot x + p = m \cdot \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

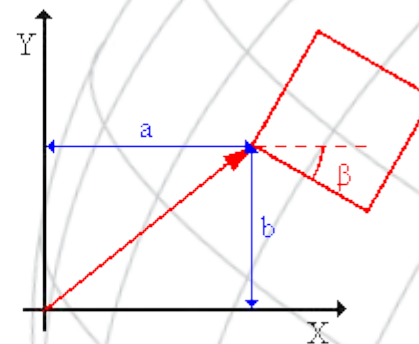
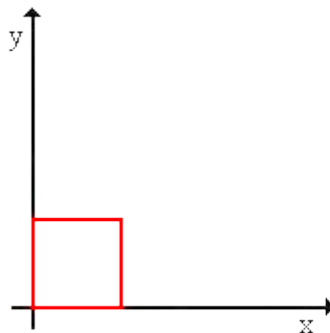
Transformační koeficienty
 (m, B, a, b) lze vypočítat již ze dvou dvojic identických bodů $(X_1, Y_1), (X_2, Y_2)$ a původní $(x_1, y_1), (x_2, y_2)$.

Helmertova transformace – speciální případ LKT; $m = 1$

Geometrické transformace

- Lineární konformní transformace:
 - posun,
 - rotace,
 - uniformní změna měřítka (v obou osách stejná),
 - zachovává tvar objektu!

 - Je potřeba dvou dvojic identických bodů.



Geometrické transformace

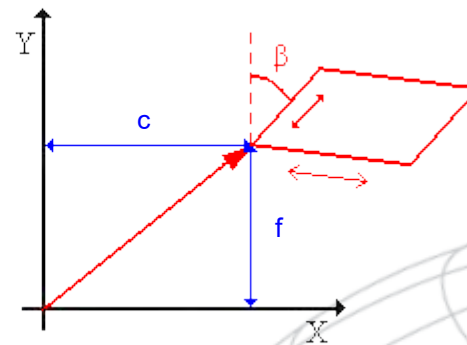
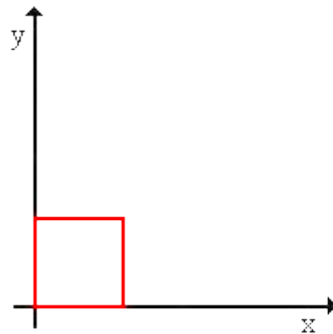
- Afinní transformace

- $X(X, Y)$ - nové souřadnice
- $x(x, y)$ - staré souřadnice
- A - regulární matice
- $p(c, f)$ - posun

Zápis rovnicí

$$X = a \cdot x + b \cdot y + c$$

$$Y = d \cdot x + e \cdot y + f$$



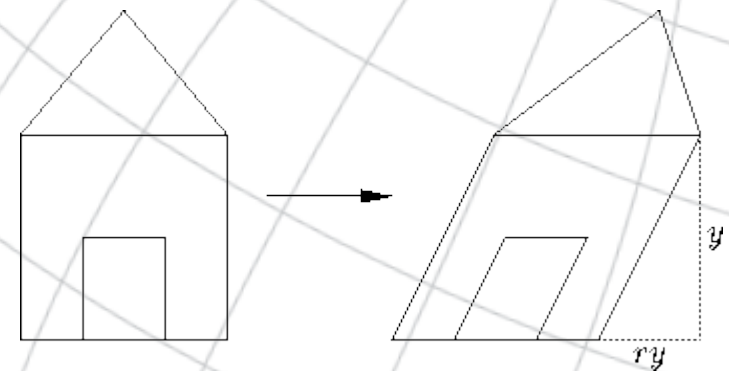
Zápis maticí

$$X = [X, Y, W] = [x, y, w] \begin{pmatrix} a & d & 0 \\ b & e & 0 \\ c & f & 1 \end{pmatrix}$$

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

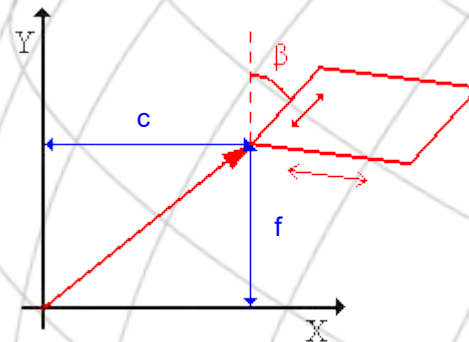
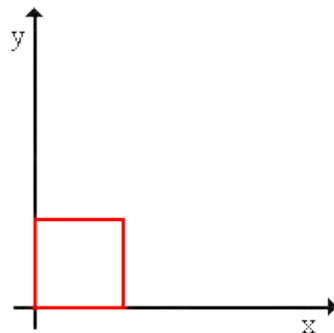
$$p = (c, f)$$

Transformační koeficienty (a, b, c, d, e, f) lze spočítat ze tří dvojic identických bodů.



Geometrické transformace

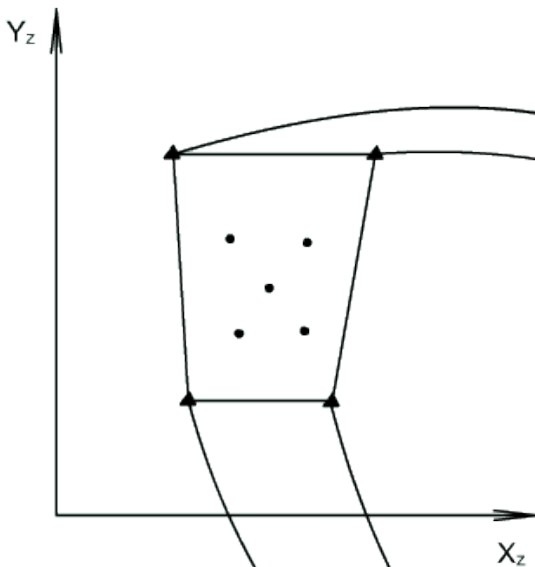
- Afinní transformace:
 - posun,
 - rotace,
 - neuniformní změna měřítka
(v každé ose jinak – **zkosení**),
 - „z obdélníka kosodélník“.
 - Je potřeba tři dvojic identických bodů!



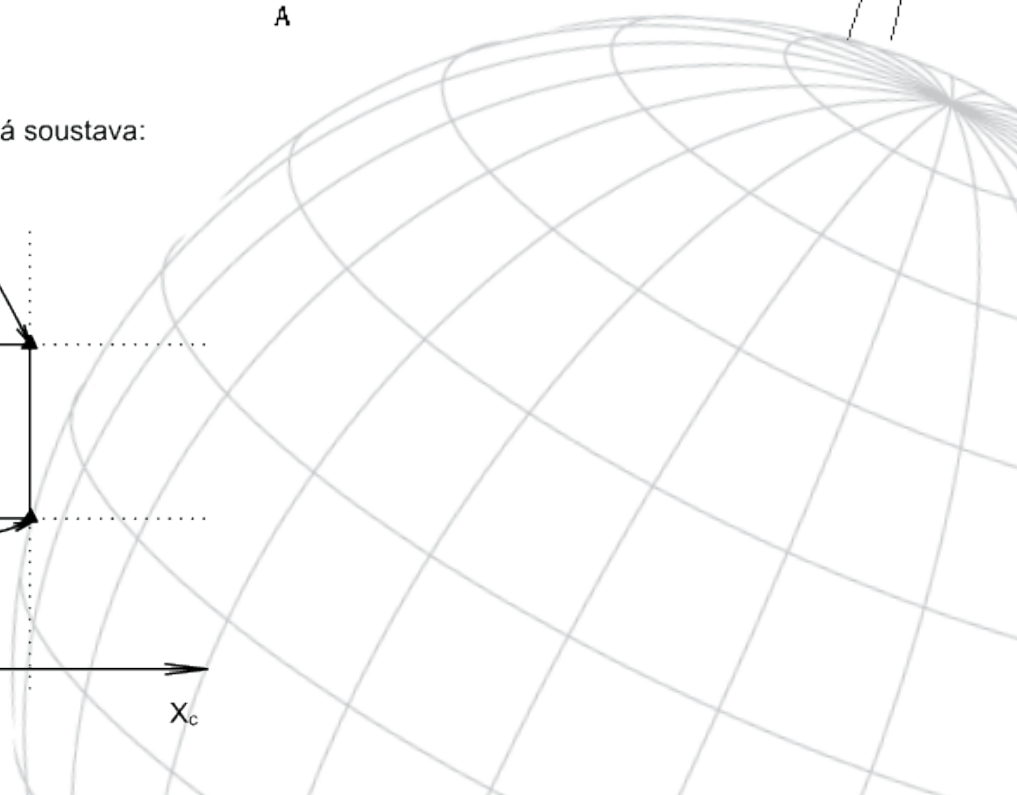
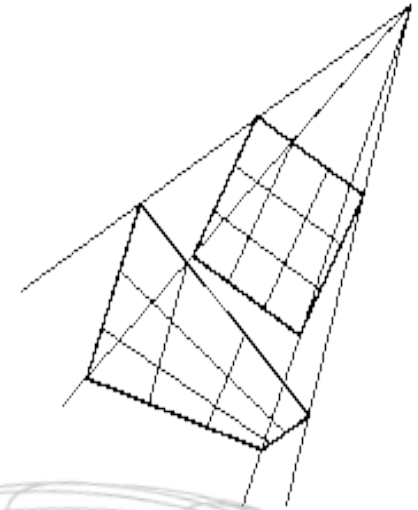
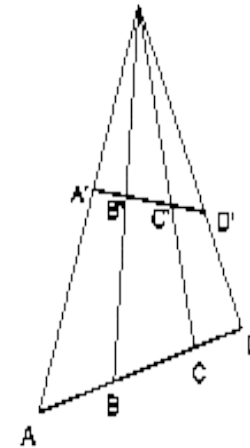
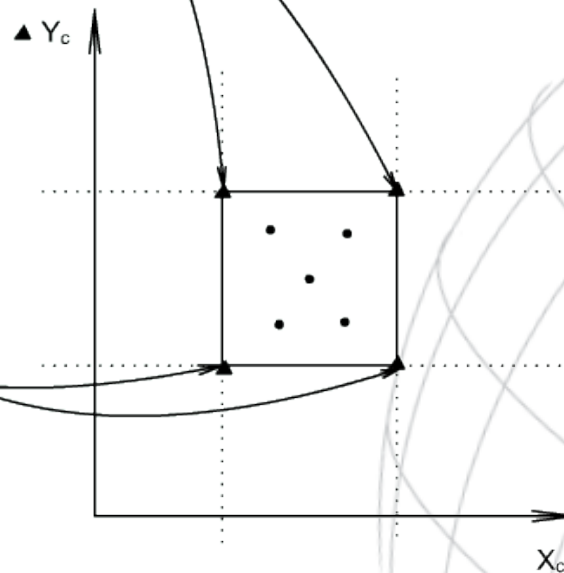
Geometrické transformace

• Projektivní transformace

Zdrojová soustava:



Cílová soustava:

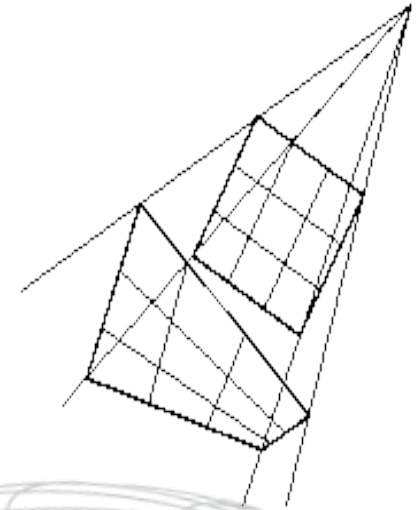
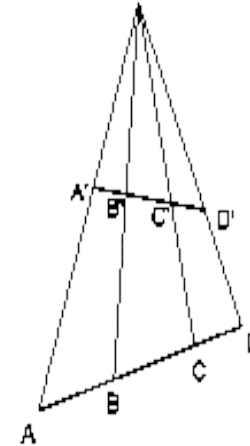


- ▲ identické body
- transformované body
- transformace

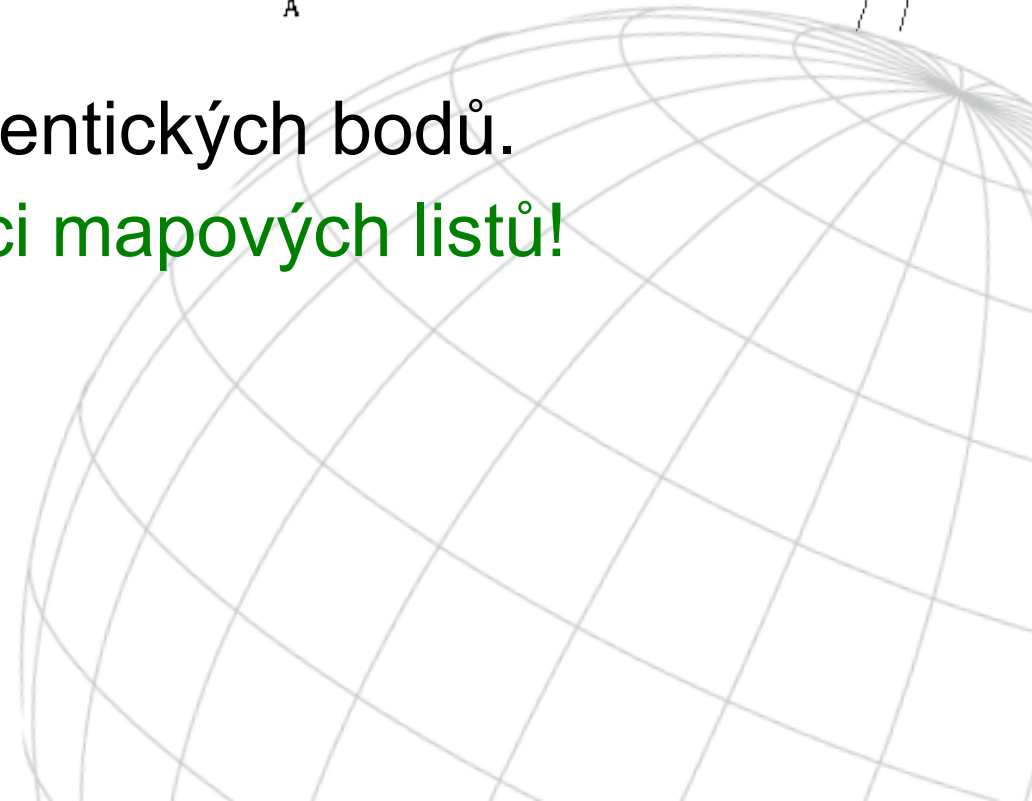
Geometrické transformace

- Projektivní transformace:

- posun,
- rotace,
- „z obdélníka lichoběžník“.



- Je potřeba čtyř dvojic identických bodů.
- **Vhodná pro transformaci mapových listů!**

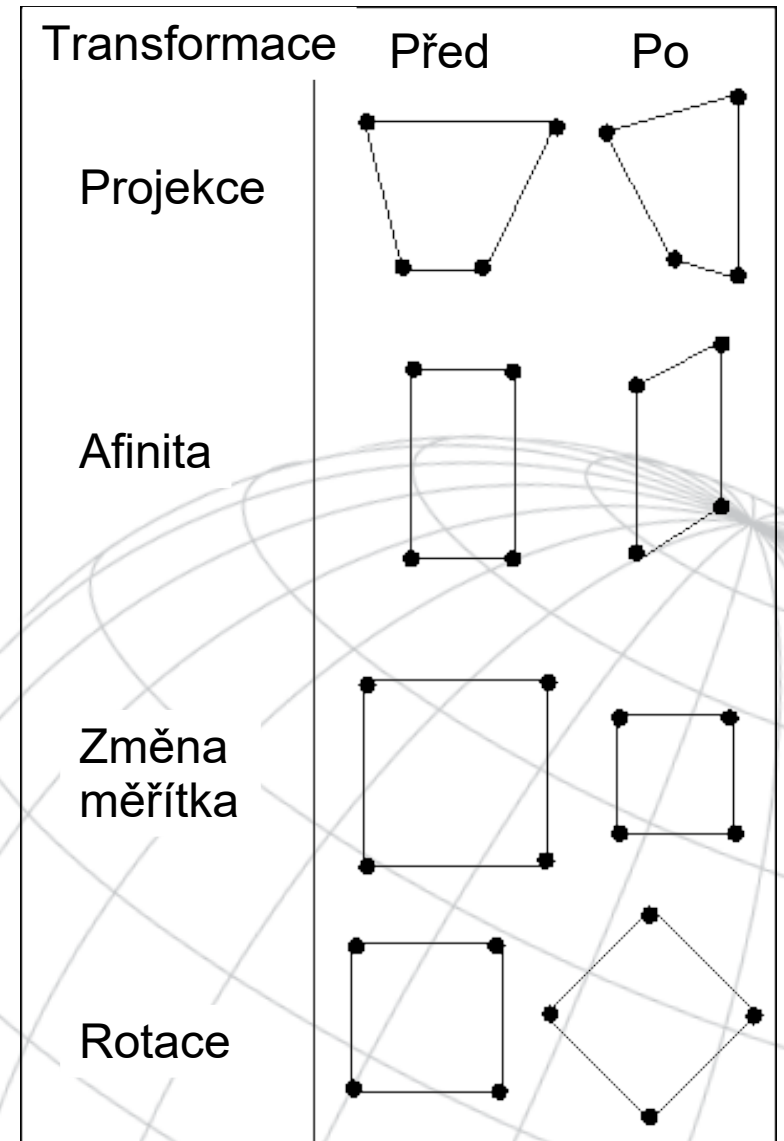


Geometrické transformace

- **Polynomické transformace** druhého a **vyšších řádů**:
 - Používají se pro deformace mapového listu, které mají lokální charakter, případně při komplikovanějším průběhu těchto deformací. Prakticky se však používají pouze řády 2 a 3, jelikož vyšší řády nepřinášejí podstatnější zvýšení přesnosti, spíše naopak.
 - Obdélník deformují polynomické transformace vyšších řádů obdobným způsobem, jako afinní transformace. Jeho hranice v cílové soustavě pak tvoří křivky. Zatímco u prvního řádu to byly úsečky, u druhého řádu se jedná o části parabol, ...
- Další specializované transformace ... KMA/TGI

Geometrické transformace

- Projevy transformací



Geometrické transformace

- Geometrické operace (nad vektory)
 - **Interaktivní editace prvků** – obvykle standardní nástroje CAD jako kopírování, posuny, rotace, mazání, spojování a rozpojování segmentů.
 - **Snižování počtu vrcholů**/ředění (weeding/coordinate thinning)
 - nástroj vycházející z generalizace a používaný hlavně po digitalizaci a vektorizaci. Spočívá v odstranění nadbytečného počtu vrcholů z linie.
 - **Zvyšování počtu vrcholů**/zhušťování (densification) – opak ředění
 - umělé vkládání dalších bodů na linii. Pouze pro vektorová data.
 - **Proložení bodů křivkou** – použití po digitalizaci a vektorizaci vrstevnic. Proložení křivkou dodá vrstevnicím přirozený vzhled. Na rozdíl od zhušťování nezachovává 100% tvar křivky.
 - Další geometrické operace probereme u generalizací.

Zpracování obrazu

- U rastrových systémů a u systémů DPZ se pro úpravu rastrů používají nástroje nazývané souhrnně "**Image Processing**". Je to poměrně rozsáhlá problematika (na KIV a KKY se image processing učí celý semestr, částečně je probírán také v předmětu KKY/DPZ).
- Metody zpracování obrazu mají v GIS relativně omezené využití, resp. **používají se ve Fotogrammetrii a DPZ**.
- A GIS využívá jejich výsledků!

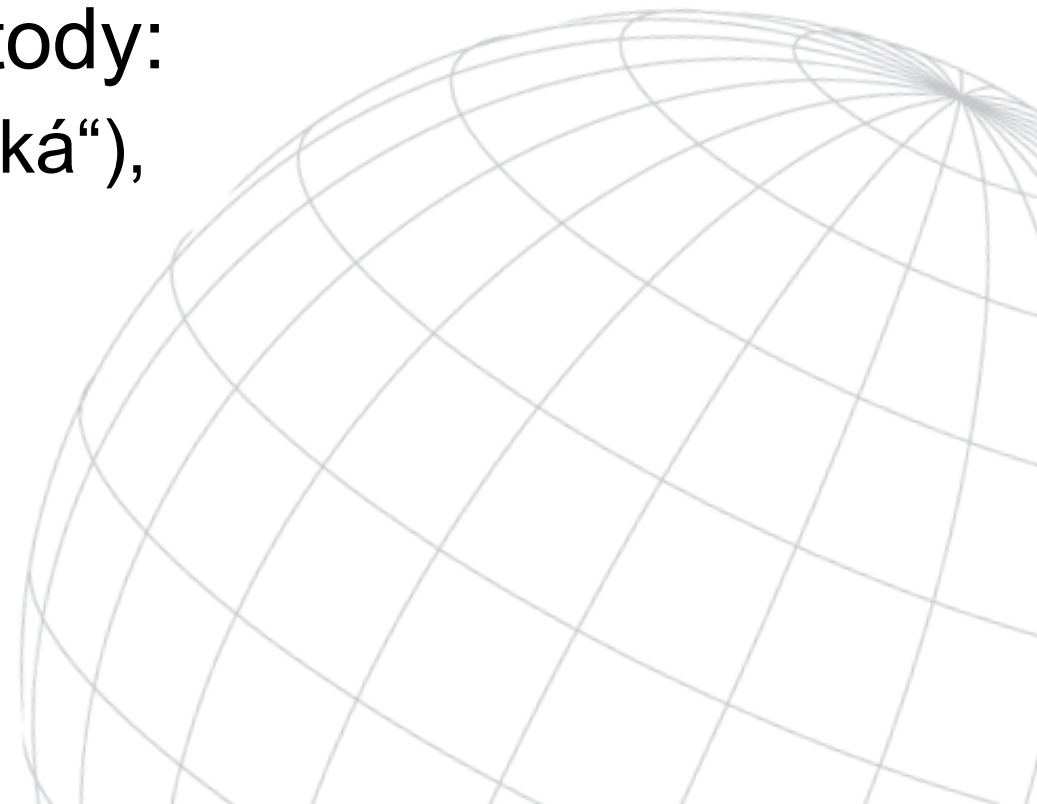
Zpracování obrazu

- Některé z funkcí zpracování obrazu:
 - filtrace (hledání hran, vyhlazování, ostření),
 - úpravy histogramu (roztahování), prahování,
 - změna jasu/kontrastu,
 - odstraňování nečistot,
 - mozaikování,
 - klasifikace,
 - převzorkování (nejbližší sused, bilineární, bikubické).



Vektorizace

- Vektorizace je **převod rastru na vektor**. Jedná se o složitější proces než rasterizace (je nutné rekonstruovat jednotlivé vektorové objekty z jejich spojitě rastrové podoby).
- Existují tři základní metody:
 - ruční (a „čtvrt automatická“),
 - poloautomatická,
 - automatická.



Vektorizace

- **Ruční**
 - Vše dělá operátor (případně za asistence počítače při přichytávání vektorových prvků na existující rastrovou kresbu – tzv. „**čtvrtautomatická**“). Jedná se o nejméně náročný způsob na hardware a software, ale nejdéle trvající (záleží na podkladech). Vhodný pro staré podklady nebo „řidké“ podklady, kde operátor musí velice často rozhodovat, co k čemu patří. Příklad systému – Kokeš, MicroStation I/RAS.

Vektorizace

- **Poloautomatická**
 - Operátor zvolí počátek rastrové linie, systém se pokusí identifikovat rastrový objekt, ukáže operátorovi směr, kterým se vektorizace bude ubírat, a při potvrzení ze strany operátora, se vydá vektorizovat, dokud nenarazí na nějakou překážku (mezera, křižovatka) či sporný bod, kde se zastaví a čeká na operátorovu odezvu (jestli má pokračovat, v jakém směru má pokračovat, ...).
 - Existují dva módy poloautomatické vektorizace, podle způsobu přichytávání:
 - na střed rastru (používaný pro vektorizaci linií),
 - na okraj rastru (používaný pro vektorizaci polygonů).

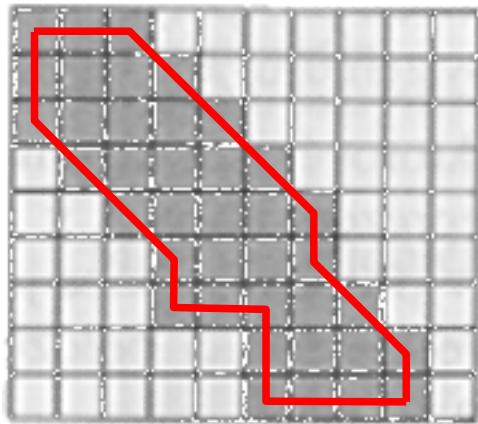
Vektorizace

- Poloautomatická
 - **Přichytávání na okraj** je pro počítač výrazně jednodušší, jelikož vektorizační software pouze hledá hranu v rastrovém obrazu, které se drží.
 - **Přichytávání na střed** je složitější a pro identifikaci středu vektorizovaného objektu se využívá principu nazývaného „**skeletizace**“, který vychází z principů používaných v automatické vektorizaci.

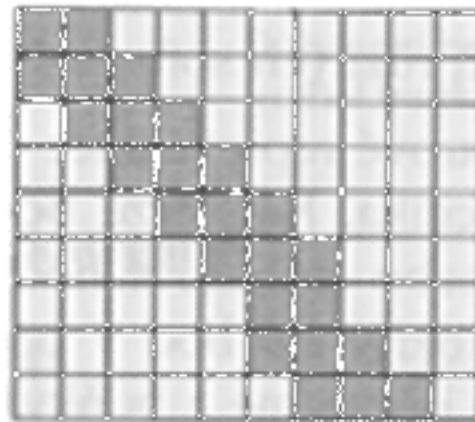
Vektorizace

- Vektorizace linií – skeletizace

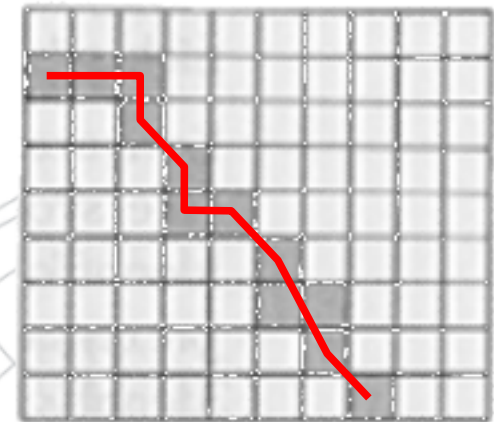
1)



2)



3)



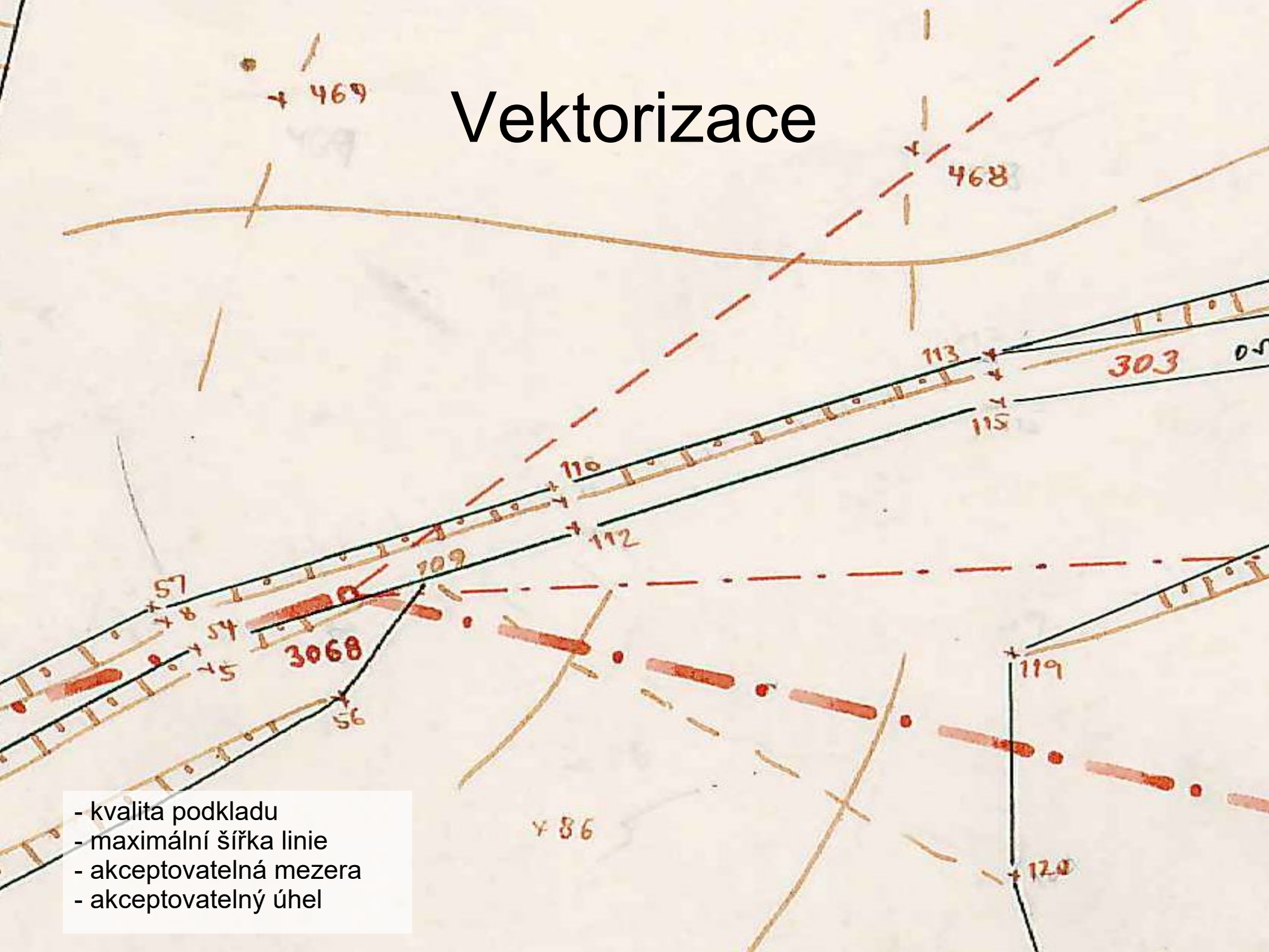
- Vektorizace polygonů



Vektorizace

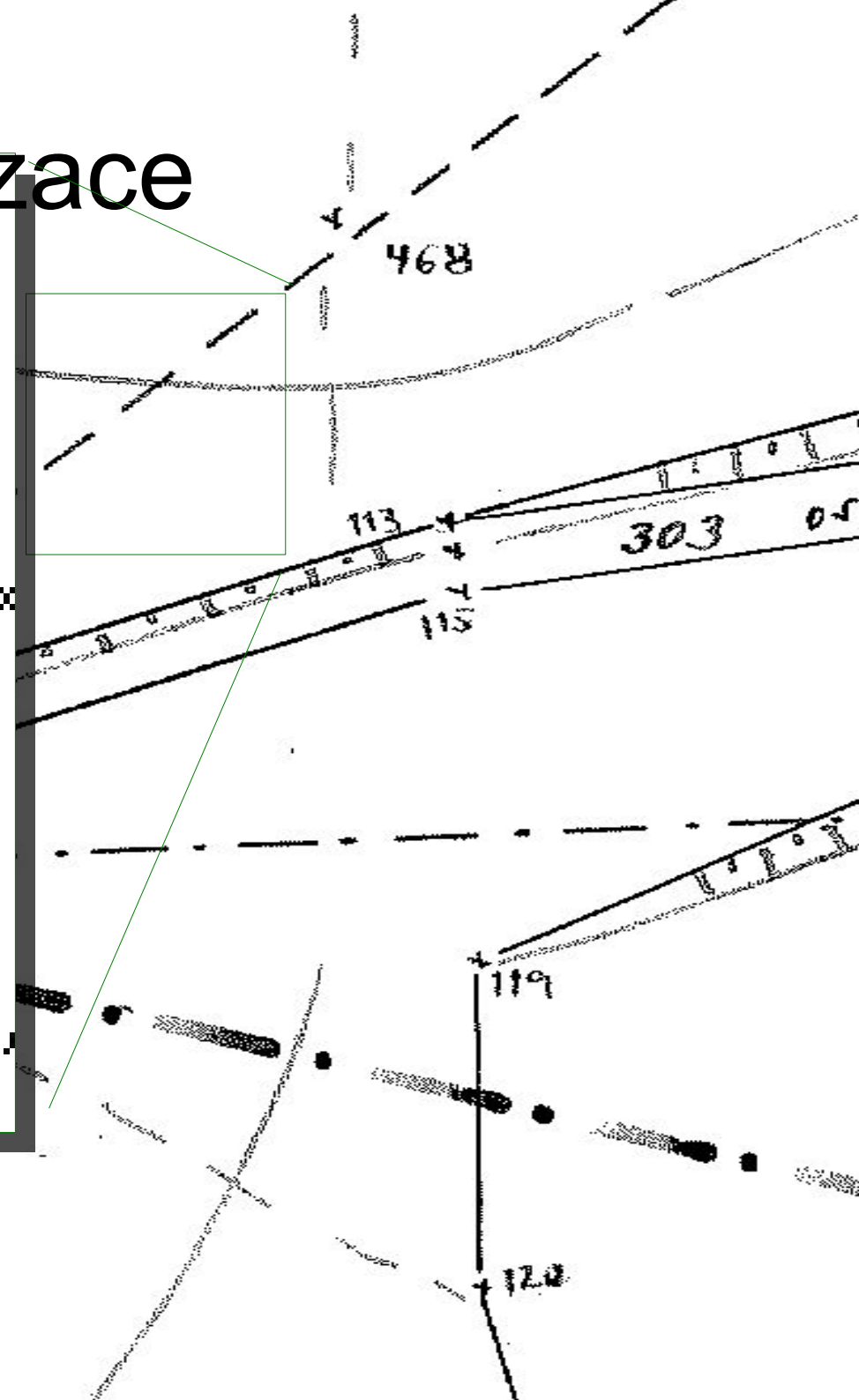
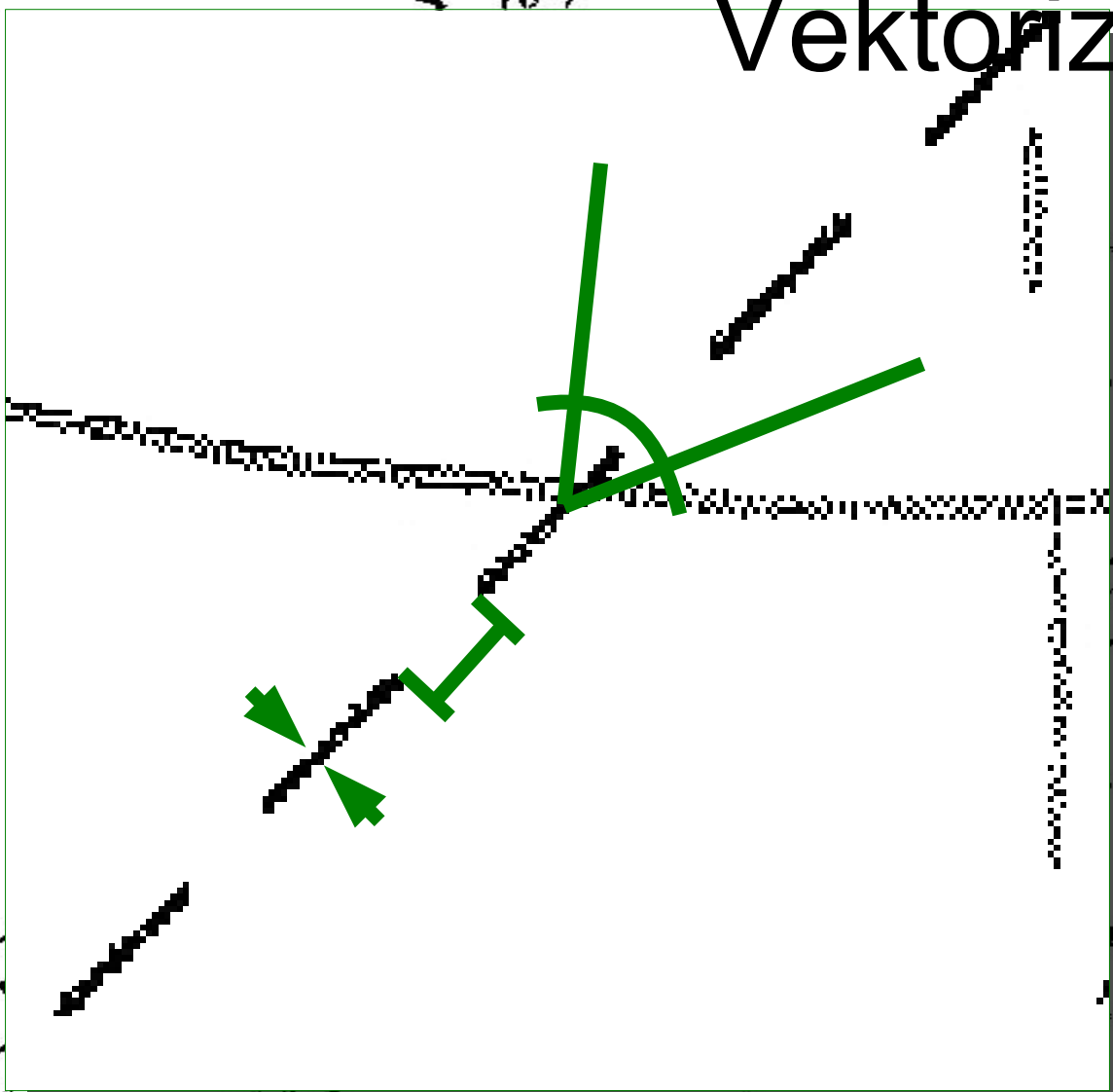
- Poloautomatická
 - Jelikož nascanované mapy jsou různé kvality (spíše horší než lepší), systémy pro poloautomatickou vektorizaci umožňují obvykle **nastavit několik důležitých parametrů** pro zautomatizování činnosti.
 - Mezi tyto parametry patří:
 - kvalita rastrového podkladu (jestli jsou objekty homogenní oblasti či ne),
 - maximální přípustná šířka linie,
 - akceptovatelná mezera v rastrové linii (při digitalizaci čerchovaných a jiných čar),
 - akceptovatelný úhel mezi částmi linie a variabilita (jak reaguje systém na změny šířky pouze v jednom směru).

Vektorizace



- kvalita podkladu
- maximální šířka linie
- akceptovatelná mezera
- akceptovatelný úhel

Vektorizace



- kvalita podkladu
- maximální šířka linie
- akceptovatelná mezera
- akceptovatelný úhel

486

Vektorizace

- Poznámky

- Schopnosti systémů se liší, ne všechny systémy pro poloautomatickou vektorizaci mají výše uvedené možnosti. Navíc **nastavení parametrů** je vždy třeba provést **empiricky**.
- Písma na mapách jsou často velmi nestandardními fonty => možnosti **využití OCR jsou omezené** převážně na analogové podklady vytvořené z digitálních dat (tištěné digitální mapy, výkresy).
- Příklad nástrojů pro poloautomatickou vektorizaci: MicroStation Descartes, I/Geovec.

Vektorizace

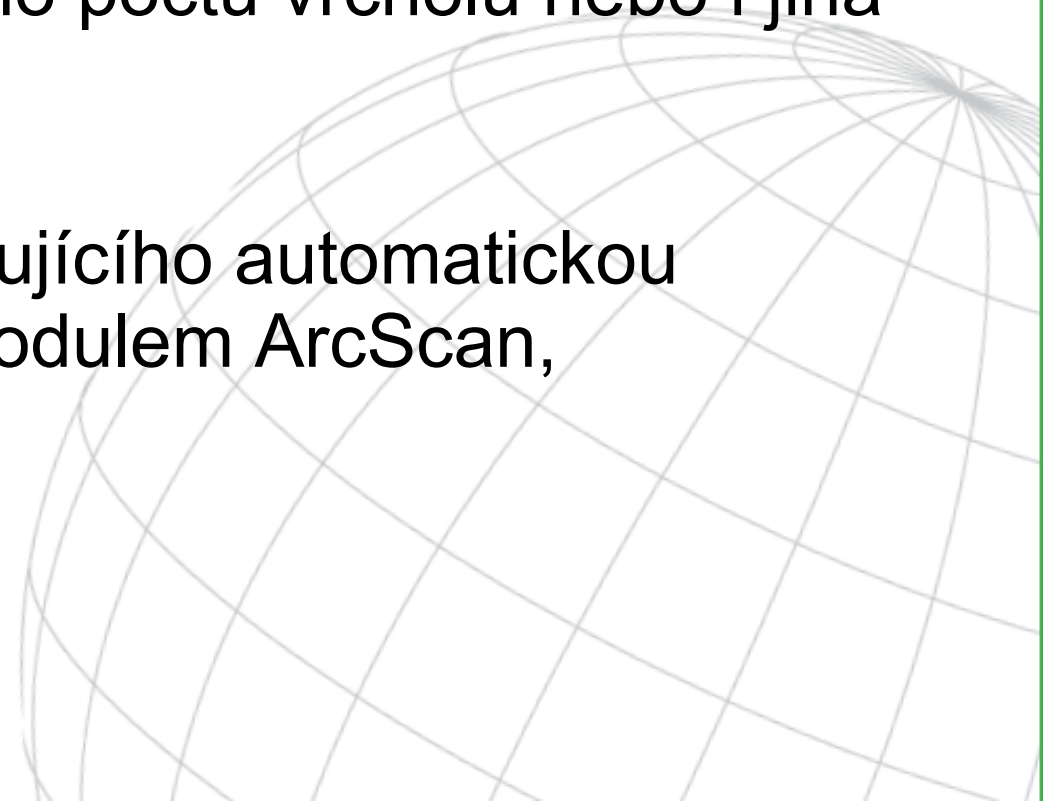
- **Automatická**
 - Při automatické vektorizaci probíhá převod rastr->vektor automatizovaně, **bez aktivní účasti operátora**.
 - Algoritmy automatické vektorizace vycházejí z algoritmů zpracování digitalizovaného obrazu a umělé inteligence.
 - Tuto metodu však většinou **nelze použít pro převod běžných analogových podkladů**, ale pouze pro již tištěné mapy z digitálních podkladů (podobně jako OCR).

Vektorizace

- Princip automatické vektorizace pro jednotlivé typy základních objektů:
 - **Body** – zpracovávající program vyhledá střed buňky reprezentující bod a zjistí jeho souřadnice a zaznamená je spolu s identifikátorem bodu v rastru (obvykle barva, či nějaká skalární hodnota).
 - **Linie** – automatická vektorizace linií funguje na principu hledání kostry (skeletu, odtud skeletizace) objektů, což je metoda velice často používaná pro ztenčování objektů. Po nalezení skeletu jsou pak pouze vyhledány na sebe napojené pixely (v rámci 4 nebo 8 okolí) a ty tvoří požadovanou linii.
 - **Polygony** – podobně jako u poloautomatické vektorizace jsou hledány hrany objektů a ty pak převáděny do linií. Poté se ze všech uzavřených liniiových objektů vytvoří polygony.

Vektorizace

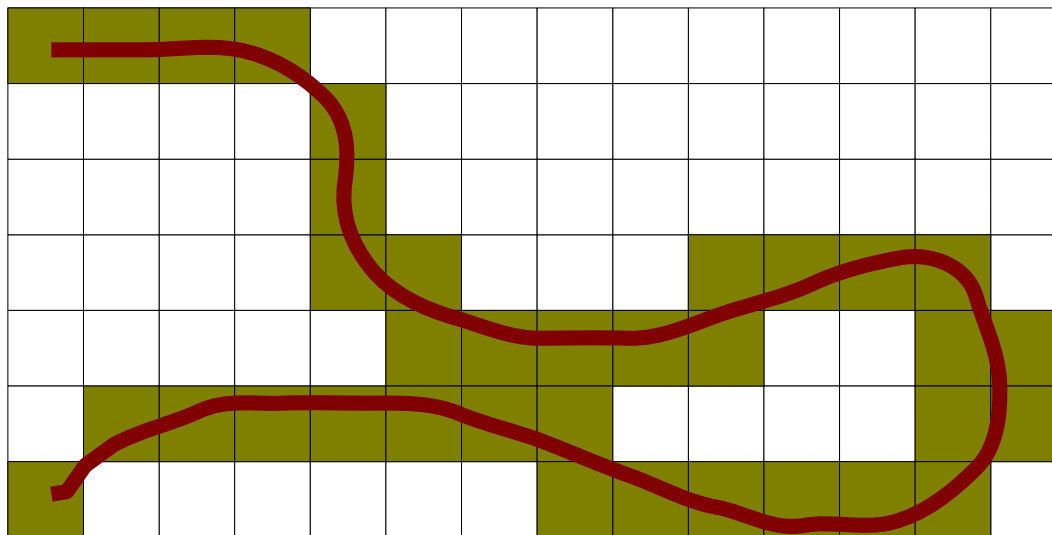
- Následné úpravy
 - Po vlastní vektorizaci často následuje proces, který odstraní nadbytečné informace získané při vlastním procesu vektorizace (nadbytečné body). Může to být odstranění nadbytečného počtu vrcholů nebo i jiná generalizační technika.
 - Příklad software umožňujícího automatickou vektorizaci: ArcGIS s modulem ArcScan, MicroStation Descartes.



Rasterizace

- Princip

- Provádí se jako překryt vektorové vrstvy na rastrovou mřížku (o určené velikosti buňky) a přiřazení hodnoty této buňky z vybraného atributu.



- Při rasterizaci je nejdůležitější určit správnou velikost buňky výsledného rastru (která bude dostatečně velká pro požadované účely, ale přitom nebude příliš velká pro možnosti hardware, které zpracovává rastr).

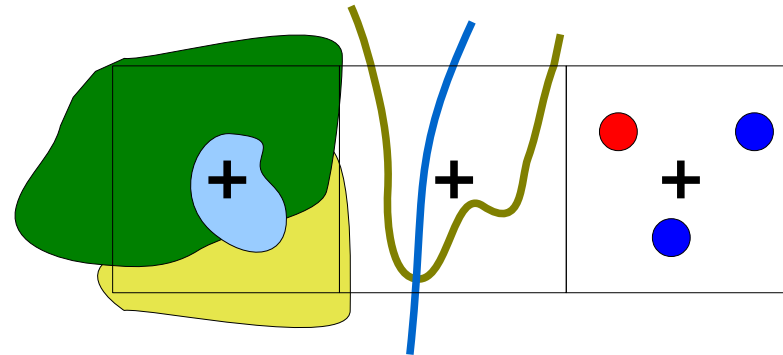
Rasterizace

- Metody řešení konfliktu

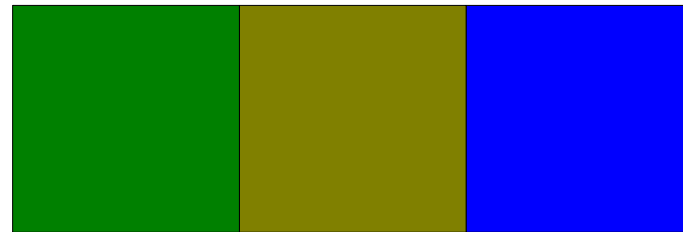
- Konflikty vznikají v případech, kdy jedna výsledná buňka obsahuje více různých objektů. Pro řešení konfliktu se používají 3 základní metody, z čehož první dvě lze použít používají pro body, linie i polygony a třetí lze použít jen pro převod polygonů:

- **Metoda dominantního typu** vychází z principu, že u buňky, do které zasahuje více objektů, se vyjádří podíl její plochy, zabíraný každým z objektů a hodnota objektu s největším podílem je pak buňce přiřazena (u bodů a linií se podíl plochy nahrazuje počtem a příp. délkou objektů, které buňka obsahuje).
- **Metoda nejdůležitějšího typu** buňce přiřadí hodnotu, která je považovaná za nejdůležitější z hlediska aplikace.
- **Centroidová metoda**, buňka má přiřazenou hodnotu definovanou polohou jejího středu při průmětu do vektorové reprezentace.

Rasterizace



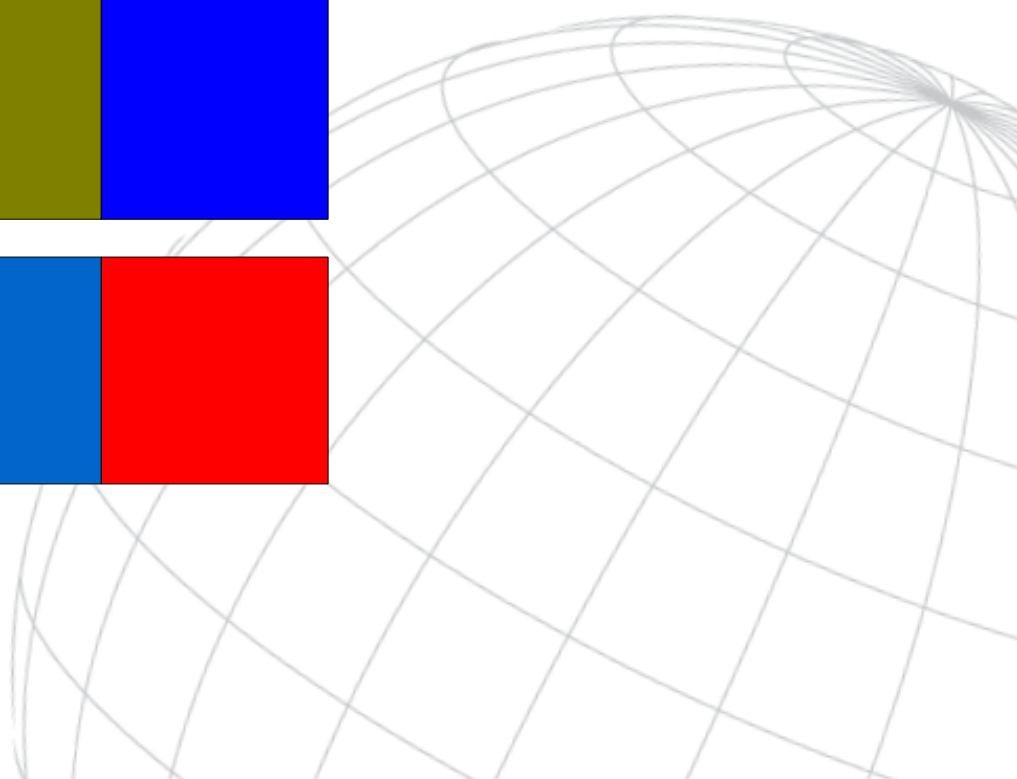
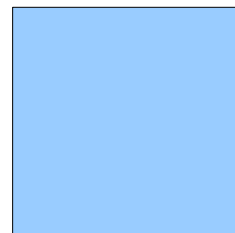
Dominantní typ



Nejdůležitější typ



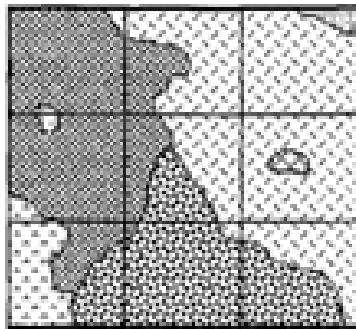
Centroidy



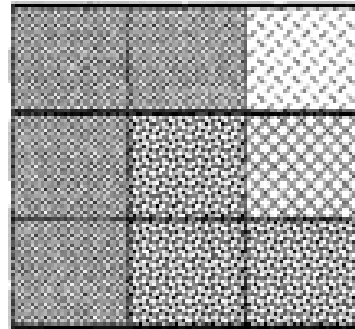
Rasterizace

- Řešení konfliktů pro polygony

Metoda centroidů



Vstupní vrstva
vegetační pokrývky

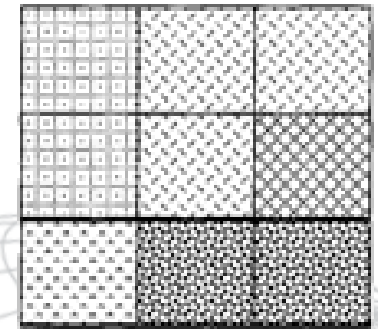


Výstupní rastr
(střed buňky)




Metoda nejdůležitější kategorie



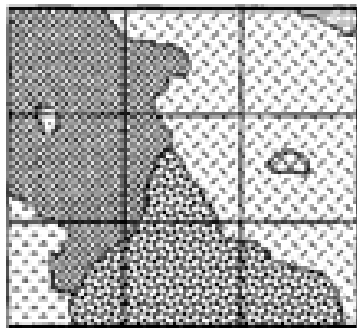
Vstupní vrstva
vegetační pokrývky



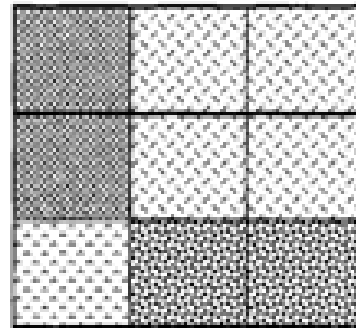
Výstupní rastr
(vážené třídy)

-  **Váha nejdůležitější kategorie**
-  **Váha pro druhou nejdůležitější kategorii**
-  **Kategorie bez přiřazení vah**

Metoda dominantního typu



Vstupní vrstva
vegetační pokrývky



Výstupní rastr
(střed buňky)

Vektor nebo rastr?

- Hybridní datový model – rozumný kompromis?
 - Některé vrstvy vektorové (silnice), některé rastrové (např. DMR)
 - Koncept duálních dat – uchovávat jak vektor, tak rastr (vrstevnice, rastr DMR) – jedna informace je uložena 2x. Vzniká riziko nekonzistence dat

Vytvoříme-li z vrstevnic rastrový DMR a z něj pak znovu vrstevnice, nebudou totožné!

