

wat?

Wat? WAAROM?

Wanneer?

Waar?

**Waar? Waar?
WIE? Wanneer?**

Waar?

WIE?

SPATIO-TEMPORELE DATA

DATASTRUCTUREN

Dr. D.P. Huijsmans

College 9 30 oktober 2013

Universiteit Leiden LIACS

1

VOORBIJ TABELLEN

- Eerste generatie database systemen: tabellen
- Tweede generatie: GIS (geografische Informatie systemen)
- Derde generatie: spatio-temporeel DBS
- Vierde generatie: voorspellend STDBS

- Eisen die opnemen van plaats en tijd stellen aan database operaties
- Datastructuren voor tijd aspecten
- Datastructuren voor vorm aspecten
- Algebra's voor tijd, vorm en voorspel functies

GIS GEOGRAFISCH INFORMATIE SYSTEEM

- “Kaart in database” systeem voor inbreng, beheer, wijzigen, analyseren, modelleren en visualiseren van locatie gebonden gegevens
- Vanaf rond 1980: op basis van op bepaald tijdstip geldende situatie
- 2D vectordata
- Data elementen: punt, lijnstuk, polygoon (veelvlak)
- Toepassingen:
 - Architectuur: gebouw/wegen/leidingnet ontwerp
 - Kadaster: vastleggen en bijwerken huidige situatie grondgebruik in administratieve eenheid
 - Wegennet en Route planning
 - Uitvoer: geplotte kaart of vector displays

VOORBEELDEN VECTORPLOT EN RASTERDISPLAY BEGIN JAREN (19)80



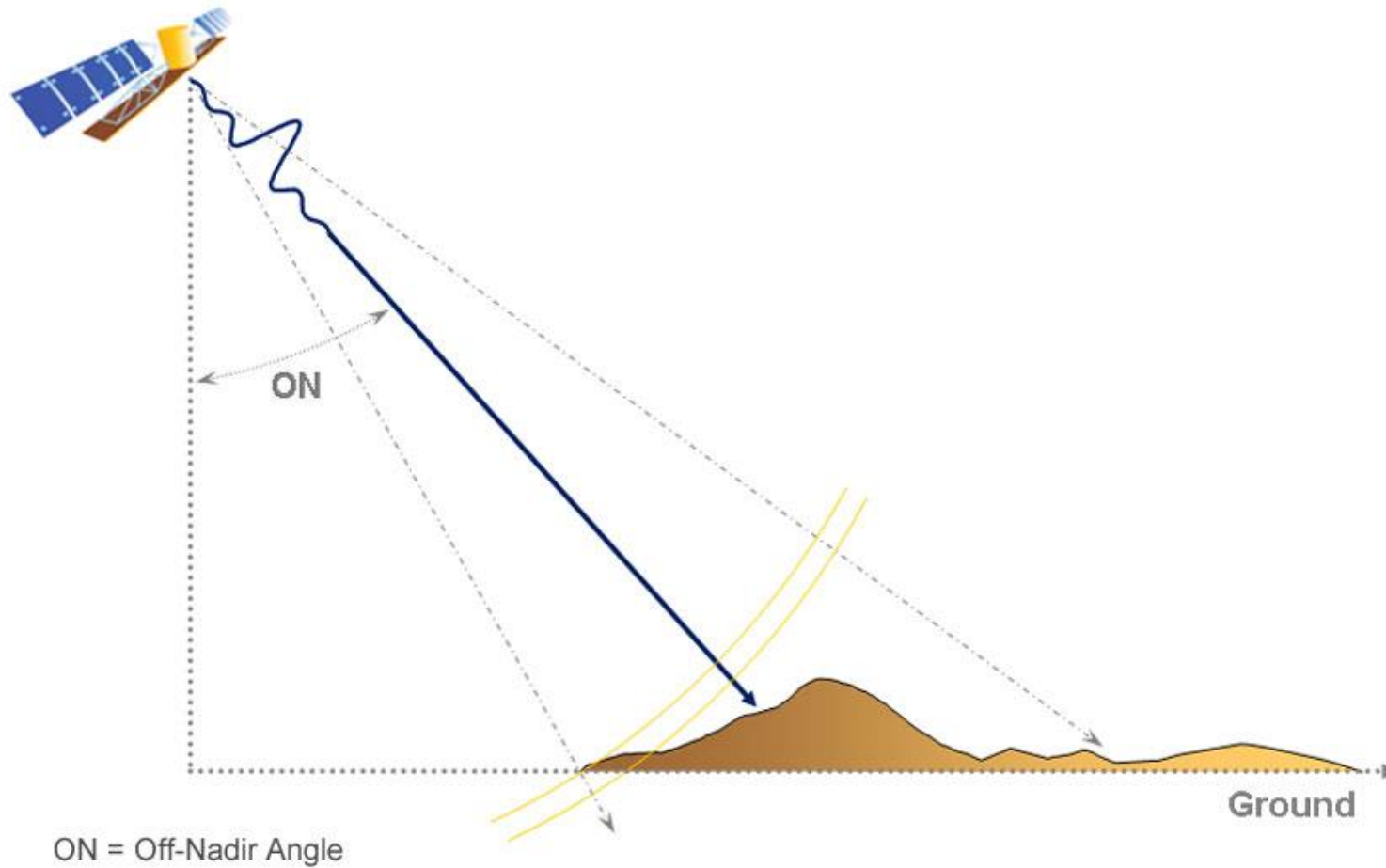
REMOTE SENSING GIS

- Vanaf ~ 1985 dankzij opkomst beeldgeheugenkaarten en raster displays
- Gelaagd raster model op basis van gelaagd beeldbewerkingmodel “thematic mapper”
- Lagen kunnen verschillende thema’s representeren of ontwikkeling in de tijd van thematisch gegeven
- Ontwikkeling gedreven door aardsatelliet stroom opnames in verschillende spectraalbanden

REMOTE SENSING



OPNAME GEOMETRIE



VOORBEELD STDB: GENEALOGISCHE DATABASE

- Oude vorm met tabel informatie:
- Gebaseerd op akten Burgerlijke Stand, kerkelijke registers, notariële archieven, kadastrale informatie
- Typisch records met veld informatie
- Opslag in tabel vorm
- SQL Query met traditionele tekst interface
- Naam: begin van naam gelijk
- Mogelijkheden plaats en/of tijd restricties:
- Plaats: naam uit lijst (gemeente, provincie)
- Tijd: datum range: >< mogelijk (alleen op jaar)

GENLIAS SEARCH INTERFACE

Zoeken in Genlias database



	Eerste persoon	Tweede persoon
Familienaam	<input type="text"/>	<input type="text"/>
Patroniem	<input type="text"/>	<input type="text"/>
Voornaam	<input type="text"/>	<input type="text"/>
Tussenvoegsel	<input type="text"/> <input type="checkbox"/> Zonder	<input type="text"/> <input type="checkbox"/> Zonder
Rol	- alle rollen - <input type="button" value="v"/>	- alle rollen - <input type="button" value="v"/>

Zoekmethode voor namen

Druk slechts één keer.
Het kan enkele minuten duren totdat het zoekresultaat verschijnt.

Beperk resultaten tot

Regio of land	<input type="button" value="- alle regio's en landen -"/> <input type="button" value="▲"/> Drenthe <input type="checkbox"/> Flevoland <input type="checkbox"/> Friesland <input type="button" value="▼"/>
Plaats	<input type="text"/> <input type="button" value="Plaatsnamenlijst"/>
Bron	<input type="button" value="- alle bronnen -"/> <input type="button" value="▲"/> Burgerlijke stand - Geboorte <input type="checkbox"/> Burgerlijke stand - Huwelijk <input type="checkbox"/> Burgerlijke stand - Overlijden <input type="button" value="▼"/>
Periode	<input type="text" value="1903"/> t/m <input type="text" value="1905"/>

GEWENSTE EXTRA TEKSTUELE SELECTIE MOGELIJKHEDEN

- Ongelijk aan
- Ongeveer gelijk aan (binnen opgegeven edit afstand)
- Alle aktes uit de database die dezelfde perso(o)n(en) in een gevonden akte bevatten

GEWENSTE GEOGRAFISCHE FUNCTIONALITEIT

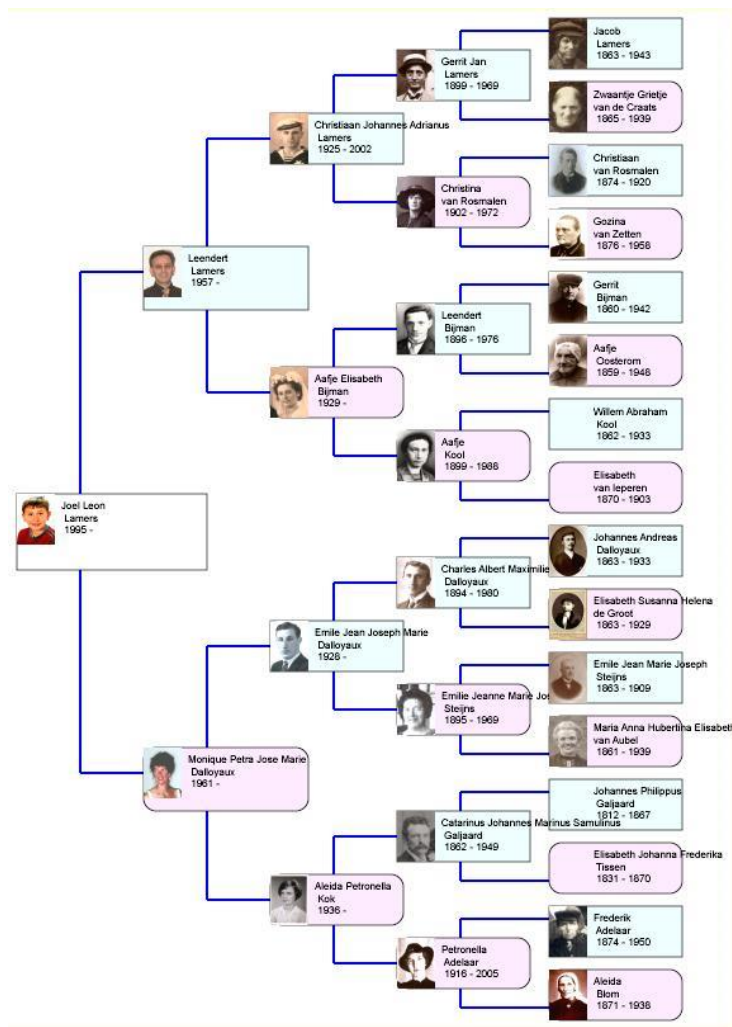
- (zwaarte)Punt selecties:
- Nearest Neighbour selectie: neem alles mee binnen bepaalde afstand van een opgegeven plaatsnaam of locatie (lengte en breedtegraad)
- Plaatsnamen rond locatie: geef een lijst van alle plaatsnamen rond opgegeven geografische positie (lengte en breedtegraad met afstand)
- (grens)lijn selecties:
- Neem alles mee dat rechts/links, N/O/Z/W van grenslijn ligt
- Gebied selecties:
- Neem alles mee dat binnen door polygoon omschreven gebied valt

GEWENSTE EXTRA FUNCTIONALITEIT VISUALISATIE

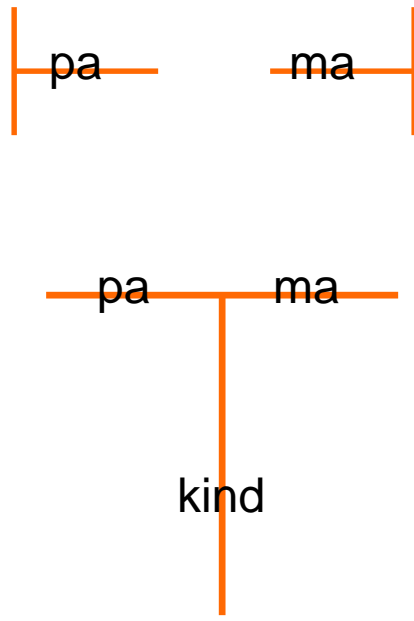
- Overzichtkaart:
- Verspreiding gekozen selectie over land, provincie, gemeente
- Tijdsbalk: visuele sortering akte info op datum
- Stamboom: toon ouder, kind relaties

TERUG IN DE TIJD ALTIJD 2 OUDERS

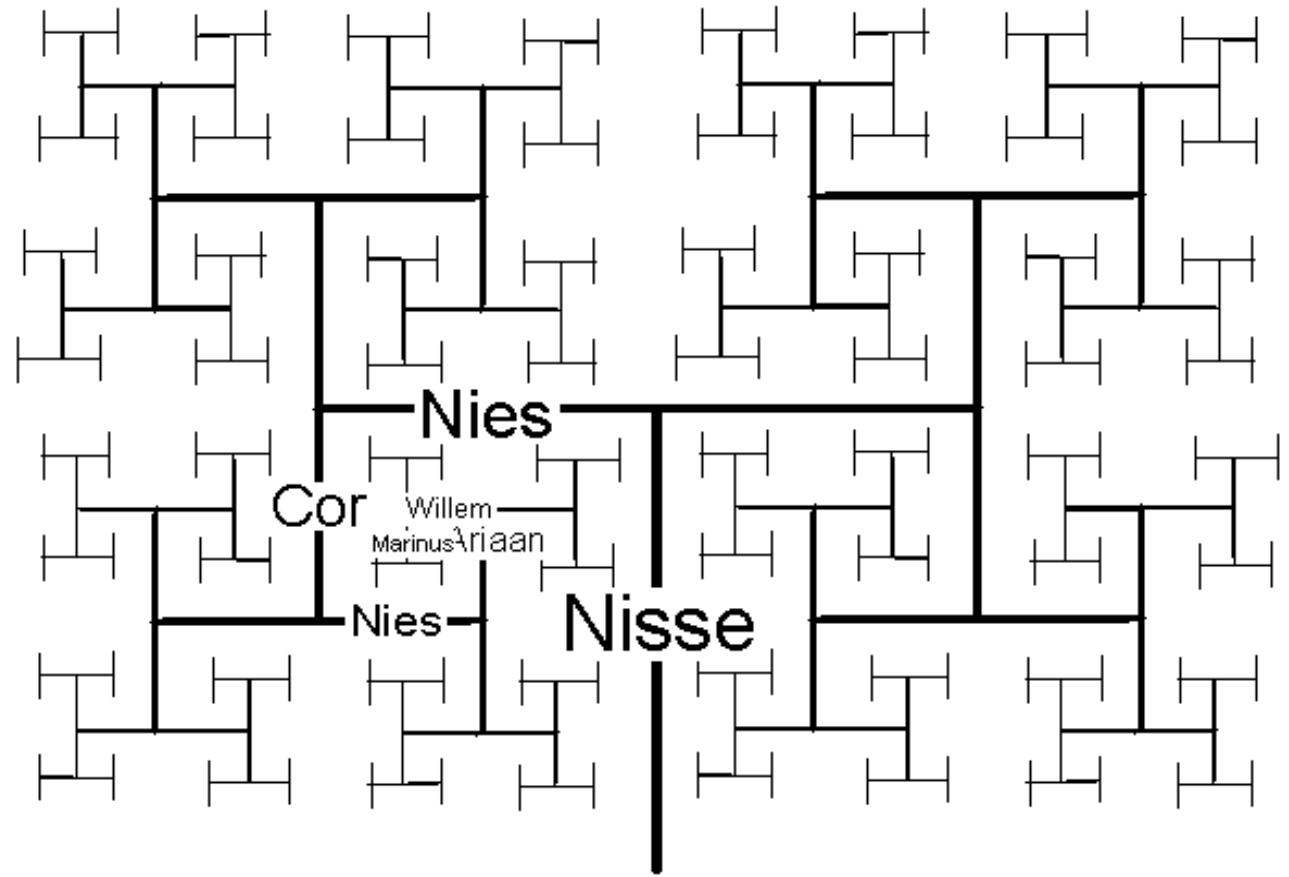
COMPLETE BINAIRE BOOM



OVERZICHT ALLE DIRECTE VOOROUDERS MET T-FRACTAL 7 GENERATIES-DIEP 255 PERSONEN ALS COMPLEET



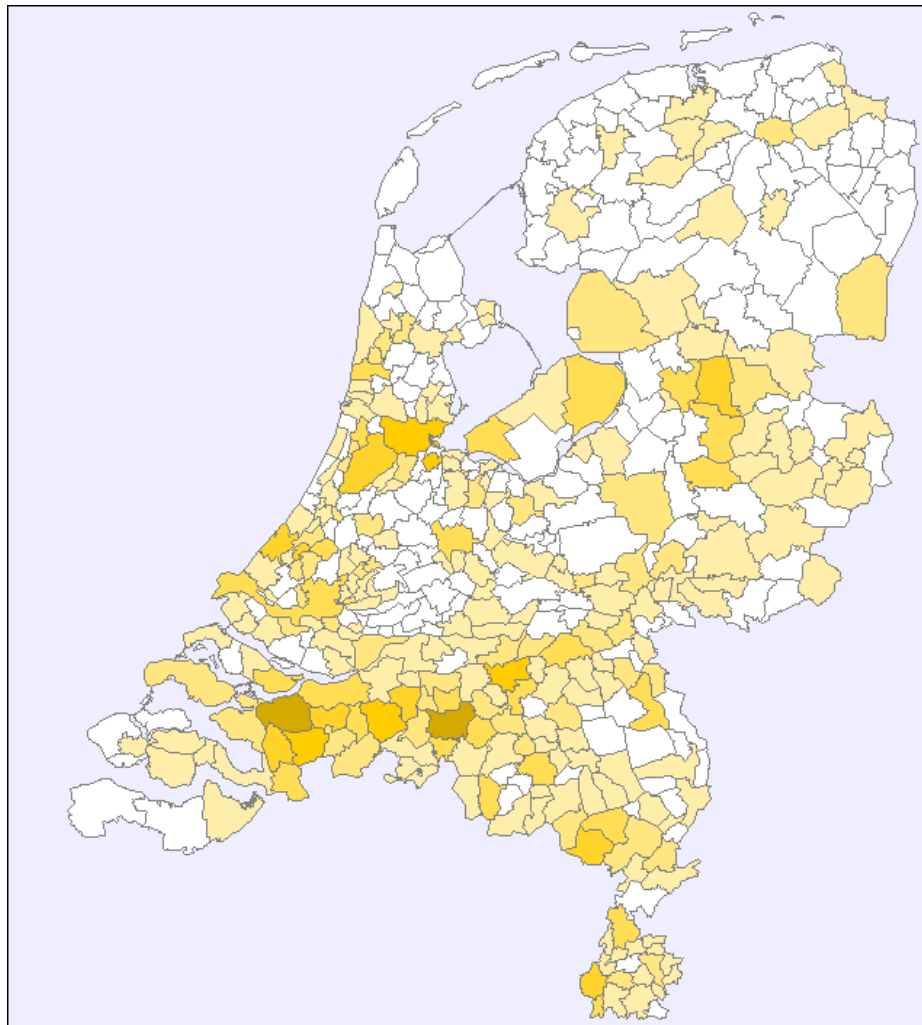
Kwartierstaat
Met Namen
stamreeks voor
7 generaties











Family Tree

Family Tree

VERSPREIDING FAMILIENAAM HUIJSMANS (AANTALLEN PER GEMEENTE IN 2010)



IPHONE TIJDLIJN

<p>Dec 15 1999 Apple registers iPhone.org and make it redirect to Apple.com</p>	<p>Sept 11 2002 Q: Will there be an iPhone? Jobs: "One never knows. We don't usually discuss products we haven't announced." <i>International Herald Tribune</i></p>	<p>Oct 14 2004 iPhone TM Apple files for iPhone Trademark in Canada</p>  <p>Dec 16 2004 Apple and Motorola admit they work on a phone that will use a mobile version of iTunes</p>	<p>Oct 18 2005 Steve Jobs: "There are more phones yet to roll out." <i>Fortune</i></p> <p>Sept 16 2006 iPhone TM Apple files for iPhone Trademark in New Zealand</p>	<p>Nov 30 2006 Patent is granted for Apple iPhone</p>	<p>Dec 18 2006 Linksys, owned by Cisco, launched a VoIP phone named iPhone. Cisco owned the iPhone name since 2000.</p> <p>Jan 10 2007 Cisco sues Apple over use of iPhone trademark</p>	<p>Jun 29 2007 The D Day! iPhone finally launched.</p> <p>Jul 04 2007 Since there is no SDK Apple publishes Web Development Page for the iPhone.</p>	<p>Sept 19 2007 iPhone confirmed for T-Mobile Germany, also on Nov 9.</p>	<p>Oct 31 2007 iPhone is named invention of the year by Time Magazine.</p> <p>Mar 06 2008 Apple releases The SDK for the iPhone</p>	<p>Jul 11 2008 iPhone 3G hits the streets. Starting in 22 countries it will be available in 70 countries at the end of 2008</p>
<p>Dec 18 2002 iPhone TM Apple files for iPhone Trademark in Singapore</p>	<p>Dec 3 2002 iPhone TM Apple registers iPhone Trademark in Australia</p> 	<p>Sept 7 2005 Apple & Motorola's joint venture, iTunes enabled phone named ROKR is released. Steve calls it: "an iPod Shuffle on your phone"</p> 	<p>Sept 25 2006 iPhone TM Ocean Telecom Services (thought to act in the name of Apple) files for iPhone Trademark in the US</p>	<p>Jan 09 2007 Official! Steve Jobs announced at MacWorld that they will release the iPhone.</p> 	<p>May 16 2007 Engadget publishes an unconfirmed rumor, that iPhone will be delayed. The news makes Apple stocks drop almost 3% (that's around \$4 billion).</p>  <p>Feb 22 2007 Apple and Cisco settle their dispute, deciding to share the iPhone name.</p>	<p>Sept 18 2007 iPhone comes to UK via O2. Will be available on Nov 9 2007.</p>  <p>Sept 05 2007 iPod Touch is released. Also there is a \$200 price cut for the iPhone. The 4Gb model is discontinued.</p>	<p>Oct 16 2007 Orange will bring the iPhone to France starting Nov 29 2007</p> 	<p>Feb 05 2008 Apple releases a 16 GB iPhone</p> <p>Jun 09 2008 iPhone 3G is announced at WWDC 2008. Features 3G, GPS and third party apps, all at a lower price.</p> 	

Copyright (c) 2008 iPhoneGold.org

GEWENSTE TEMPORELE FUNCTIONALITEIT

- Nauwkeurigheid tijdstip:
- Gelijkheid tot op jaar en/of maand en/of dag etc
- Tijdstip 1D relaties: voor, na, range, overlap binnen gekozen nauwkeurigheid etc
- Verschil in tijdstip: voor, na, rond, buiten interval

GEWENSTE EXTRA SPATIO-TEMPORELE FUNCTIONALITEIT

- Grafische weergave (mogelijk animatie) van selectie middels locatie, tijd attributen:
- binnen totale gebied; binnen tijdspanne:
- Op niveau provincie, gemeente, woonadres

DATA ELEMENTEN EN OPERATIES VOOR TEMPORELE GEGEVENS

- Genormeerde lineaire representatie voor tijdstip(pen):
- Juliaanse Dag i.p.v. JJJJMMDD
- Temporele operaties:
- Puntoperaties: $<, \leq, =, \neq, \geq, >$ (numeriek)
- Tijdrange operaties: $<, \leq, =, \neq, \geq, >$ (interval aritmetiek rond overlap)

DATA ELEMENTEN EN FUNCTIONALITEIT VOOR LOCATIE INFORMATIE

- Data elementen: punt, lijnstuk, gebied
- Functionaliteit: 2D relaties, richting, vorm algebra
- 2D relaties: afstand tot, binnen/buiten afstand, afstandrange
- Richting: L/R, W/N/O/Z
- Vorm algebra: vereniging, doorsnede, complement etc.
- Topologische eigenschappen: overdekt, genest, raakt....

DATASTRUCTUUR OPGEBOUWD ROND SPATIO-TEMPORELE ELEMENTEN

- Bij een grafische interface met de database is een opzet te prefereren die:
- Data elementen ordent op locatie en/of tijd middels datastructuren die een verzameling objecten middels punten, lijnstukken, polygonen en 2D,t kleinstomsluitende MER's opslaan en toegankelijk maken (2D en 3D R-tree achtige aanpak)

TWEE HOOFDSOORTEN STDBS'EN: GRID OF VECTOR BASED

- Voor de beschrijving en opslag van locatie data zijn 2 hoofdrichtingen in gebruik:
- **Space driven** representation:
- Een globaal sampling grid is/wordt over de data gelegd en op elke mogelijke positie wordt vastgelegd wat de (eventueel gediscrètiseerde) waarde(s) zijn van het gegeven; snelle toegang via herhaalde hiërarchische opdeling v/d ruimte
- **Data driven** representation:
- De data zelf wordt gebruikt om een opdeling van ruimte en/of tijd te maken die snelle toegang tot elk data element geeft

KD TREE

EEN DATA DRIVEN HIERARCHISCHE TWEEDELING

- Een kd-tree is een Binary Space Partitioning (BSP) tree voor k dimensies
- (b.v. 2d geografisch plus 1d tijd)
- Het maakt snel zoeken in een meer-dimensionale ruimte mogelijk, door deze ruimte herhaald op te delen op grond van de data waarden zelf en is uitermate geschikt om snel de naaste burenen te vinden (NN search)
- Via een vaste volgorde wordt elk van de betrokken dimensies om beurten gebruikt om de k -dimensionale ruimte recursief te halveren
- Een goed gebalanceerde kd-tree bouwt men op door van de complete verzameling datapunten uit te gaan en steeds m.b.v. de mediaan punten $-/+$ over de 2 halfspaces te (her)verdelen
- Vraag: waarin verschilt nu een 2d-tree van een BST?

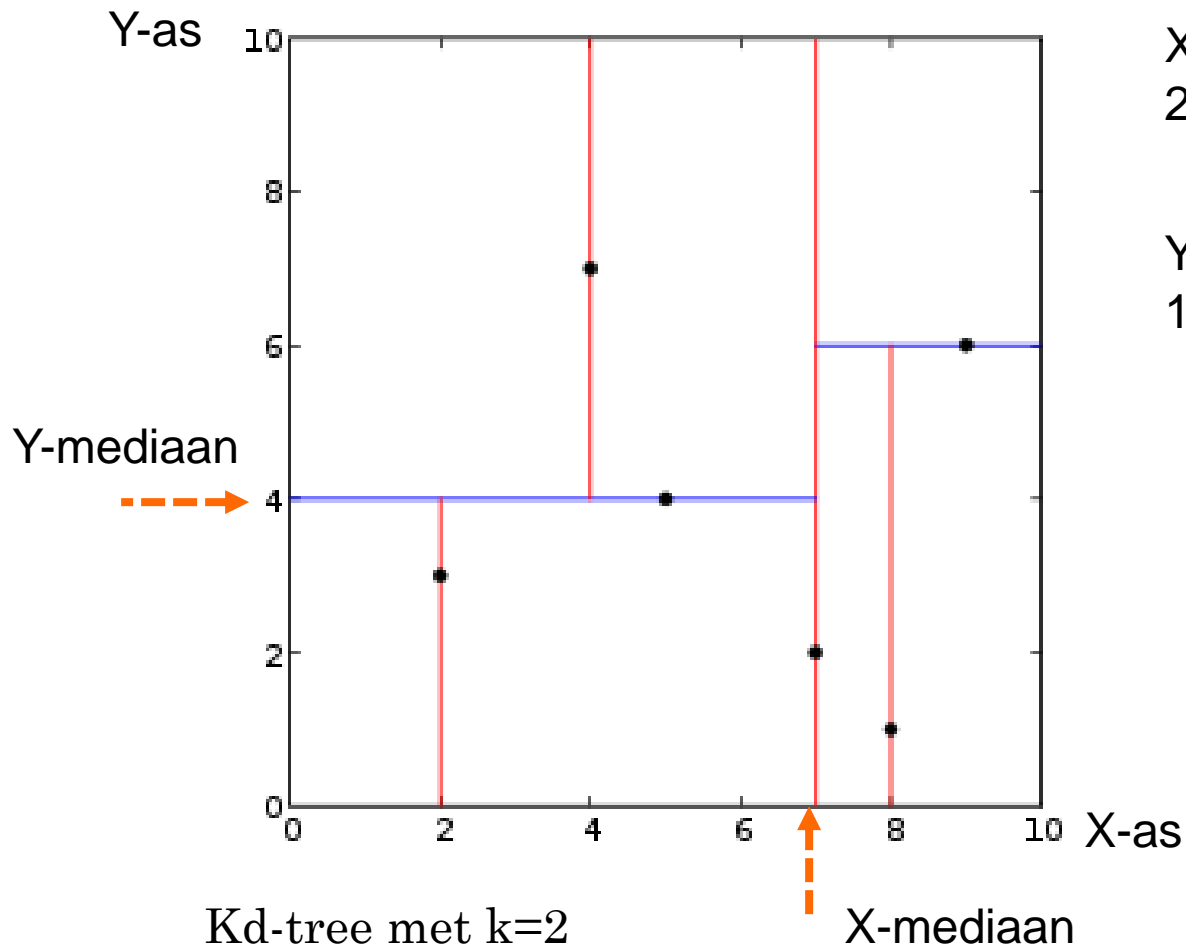
KD-ALGORITME

- **function** *kdtree* (*list of points* pointList, *int* depth)
- {
 - **if** pointList **is empty** **return nil**;
 - **else**
 - {
 - *// Select axis based on depth so that axis cycles through all valid values*
 - **var int** axis := depth **mod** k;
 - *// Sort point list and choose median as pivot element*
 - **select** median **by** axis **from** pointList;
 - *// Create node and construct subtrees*
 - **var tree_node** node;
 - node.location := median;
 - node.leftChild := *kdtree*(points **in** pointList **before** median, depth+1);
node.rightChild := *kdtree*(points **in** pointList **after** median, depth+1);
 - **return** node;
 - }
- }

De K2-tree is net als de BST een data-driven opdeling, maar statistisch beter verdeeld (mediaan i.p.v. willekeurig opdeelpunt)

KD-TREE VOORBEELD

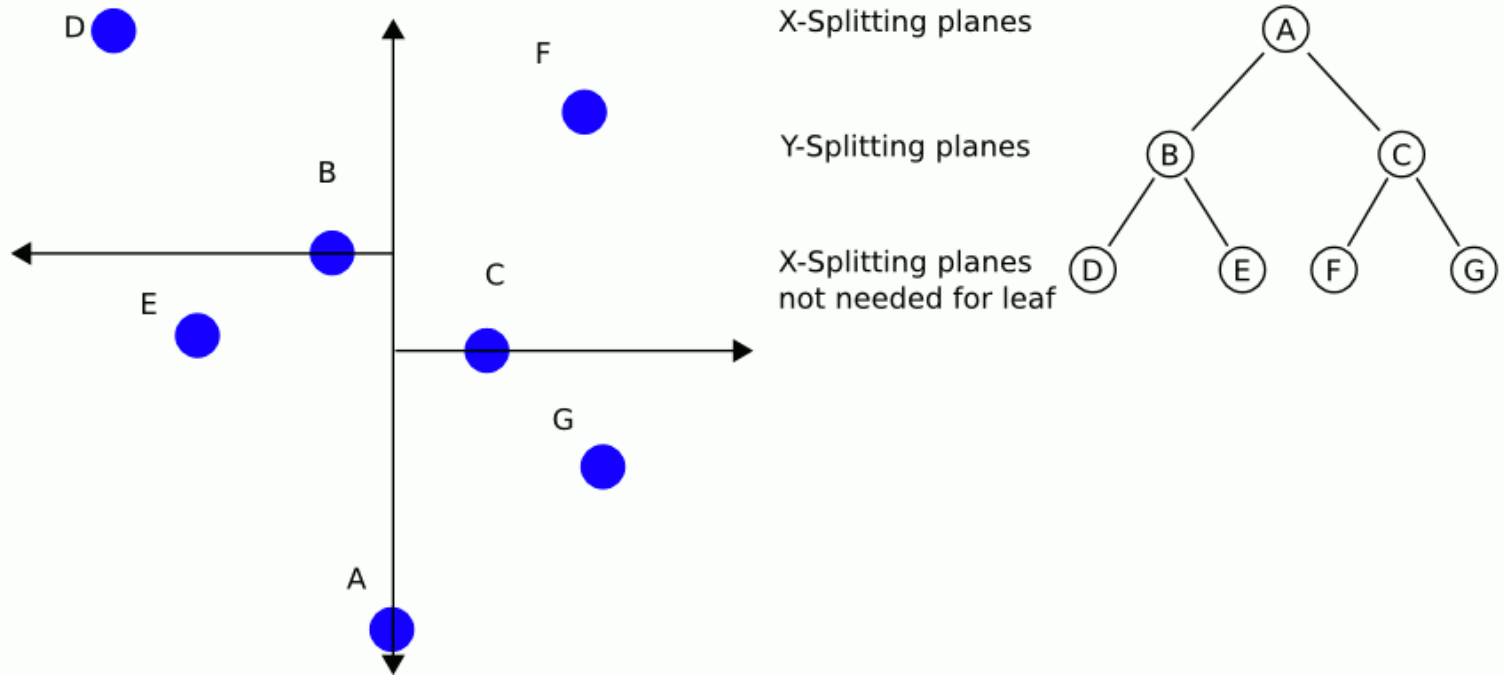
- pointList = [(2,3), (5,4), (9,6), (4,7), (8,1), (7,2)]



X-gesorteerd:
2,4,5,7,8,9

Y-gesorteerd:
1,2,3,4,6,7

NN SEARCH IN KD-TREE



OPDELING KD-TREE

- Stel $d=4$; 3 ruimtelijke en 1 tijd dimensie
- Opdelingsvolgorde: x,y,z,t
- Betekenis 2 kinderen bij tweedeling:
 - X: linker en rechter
 - Y: boven en onder
 - Z: voor en achter
 - T: eerder en later

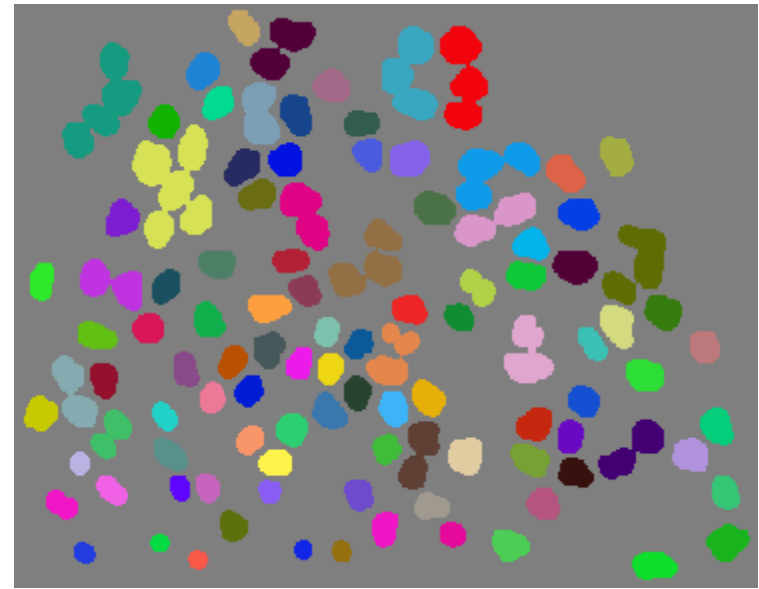
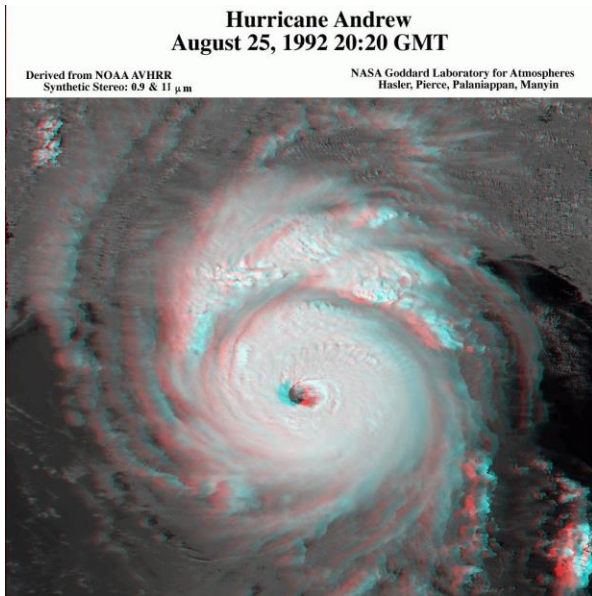
SPACE DRIVEN ALTERNATIEF

- Als je er voor kiest om niet de data de opdeling te laten sturen, maar een hierarchisch sampling grid dan kunnen de volgende grid datastructuren gebruikt worden:
- 2D: quadtree en gelaagd raster model $\Sigma 2D$
- 2D,t: $\Sigma 2D,t$
- 3D: octtree

RASTER DATA TYPEN

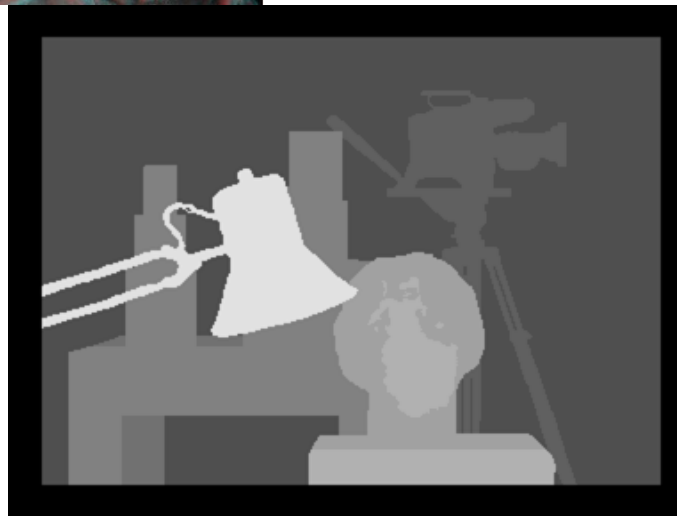
- We kunnen een beeld met n rijen en m kolommen als voorbeeld nemen van een rasterdatastructuur
- In feite geeft het raster middels een volledige graafrepresentatie ($n \times m$ pixels op kruispunt van kolom n en rij m) een opslag van pixelwaarden als:
 - 1 byte (helderheid)
 - 3 bytes (RGB waarden)
 - m bytes (remote sensing spectraal band waarden)
 - 1 byte (patroonherkenningslabel)
 - 1 integer (diepte kaart)
 - 1 bit (voor- of achtergrondpixel)
 - Etc.

VOORBEELD RASTER DATA



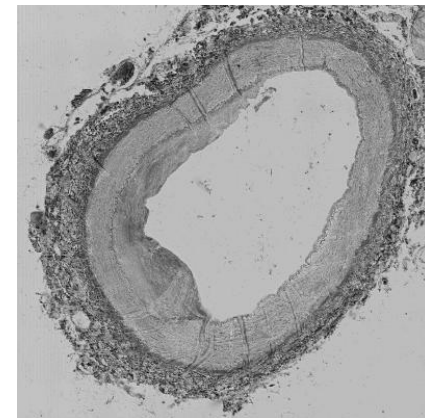
Remote
sensing

RGB beeld



Dieptekaart

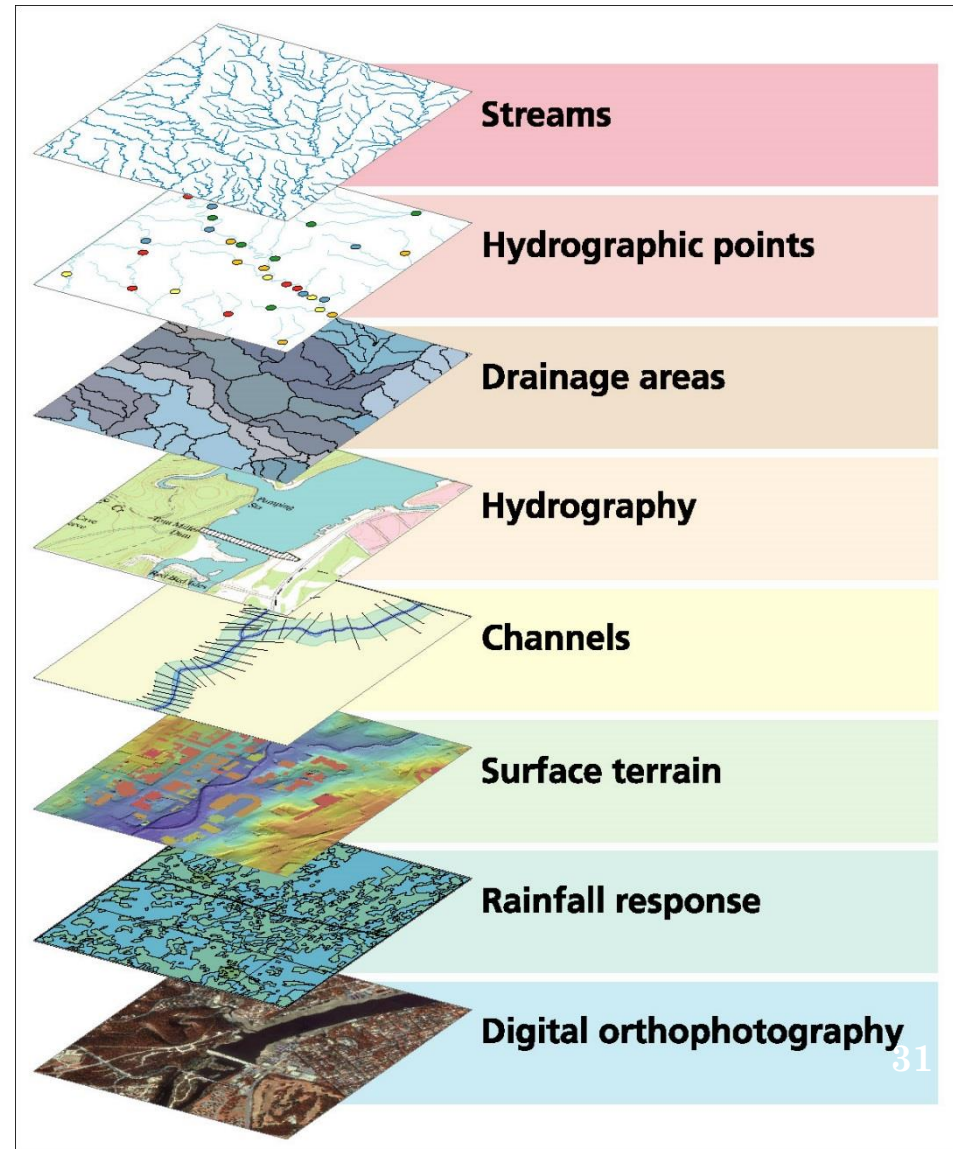
Labeled pixels



Cross section

GELAAGD RASTER MODEL: REMOTE SENSING

- Elke laag bevat 2D informatie m.b.t. een bepaald gegeven op een bepaald tijdstip of binnen een bepaald tijdsinterval
- Hoe visualiseer je het totaal aan punt/grens/gebieds informatie dat af te leiden is uit de som van al die gegevens?



Bron: Bart Kuijpers Spatial and Spatio-Temporal Data Models for GIS

VORM ALGEBRA FUNCTIONALITEIT

- Bottleneck STDBS'en: moeilijk geschikt te krijgen voor alle mogelijke combinaties van elementaire data eenheden
- Manipulatie raster data is eenvoudiger, weinig primitieven, beperkt aantal algebraïsche bewerkingen
- Manipulatie vector data is moeilijker; lastige combinaties verschillende primitieven en veel bijzondere randgevallen

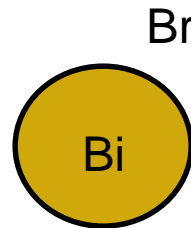
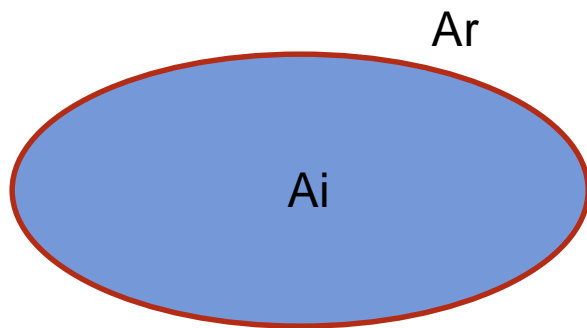
“RASTER IS FASTER BUT VECTOR IS CORRECTOR” (JOSEPH BERRY)

VECTOR	Points	Lines	Areas
Feature data			
Areal units			
Networks			
Sampling records			
Surface data			
Label/text			
Symbols			
Relations			

RASTER	Points	Lines	Areas
Feature data			
Areal units		-	
Networks	-	-	-
Sampling records		-	
Surface data		-	
Label/text	-	-	-
Symbols			
Relations			

TOPOLOGISCHE RELATIES IN EEN VECTORMODEL VOOR GEBIEDEN

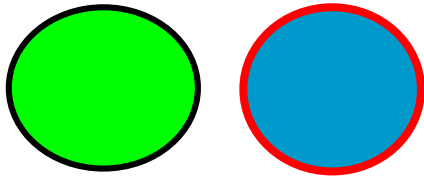
- Data elementen: punt, lijnstuk (nodes and arcs)
- Gebied: gesloten polygon is rand van omsloten binnengebied



	B_i	B_r
A_i	leeg	leeg
A_r	leeg	leeg

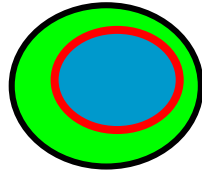
- Gebied kan als geheel $A=A_r + A_i$ of A_r dan wel A_i apart worden bedoeld.
- Afhankelijk van de posities van A en B kan de doorsnede van A en B delen al dan niet leeg zijn

SEMANTIEK VAN TOPOLOGISCHE RELATIES



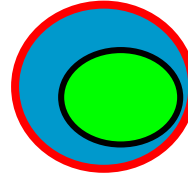
\emptyset \emptyset
 \emptyset \emptyset

apart



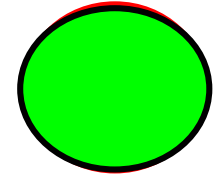
$\neg\emptyset$ $\neg\emptyset$
 \emptyset \emptyset

omvat



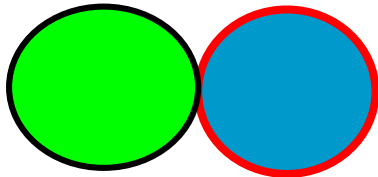
$\neg\emptyset$ \emptyset
 $\neg\emptyset$ \emptyset

binnen



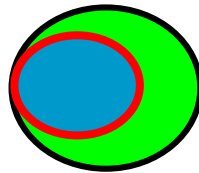
$\neg\emptyset$ \emptyset
 \emptyset $\neg\emptyset$

identiek



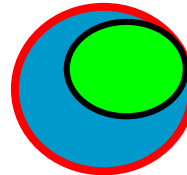
\emptyset \emptyset
 \emptyset $\neg\emptyset$

raakt



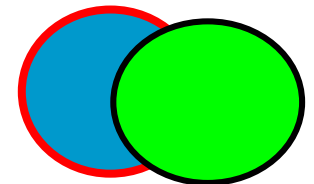
$\neg\emptyset$ $\neg\emptyset$
 \emptyset $\neg\emptyset$

bedekt



$\neg\emptyset$ \emptyset
 $\neg\emptyset$ $\neg\emptyset$

bedekt door

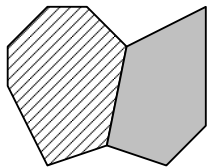


$\neg\emptyset$ $\neg\emptyset$
 $\neg\emptyset$ $\neg\emptyset$

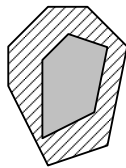
overlapt

SEMANTISCH VERSCHILLENDE TOPOLOGISCHE RELATIES

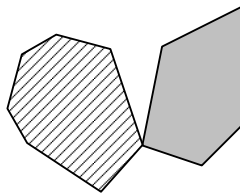
area-area relationships



adjacency

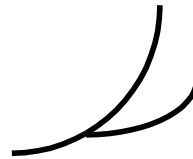


island

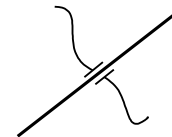


touch

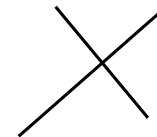
line-line relationships



branching off

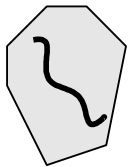


cross

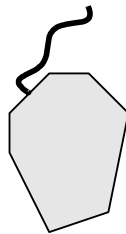


intersect

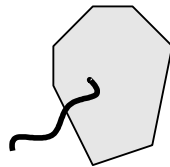
area-line relationships



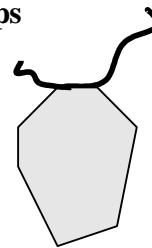
line in an area



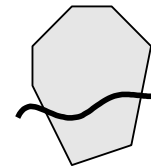
line ends at
an area



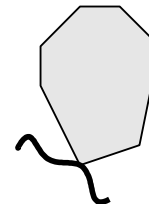
line ends in
an area



line is border
of an area



line intersects
area



line touches area

point-line relationships

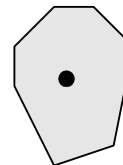


point on line

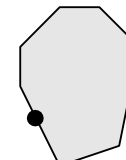


point beside
a line

point-area relationships

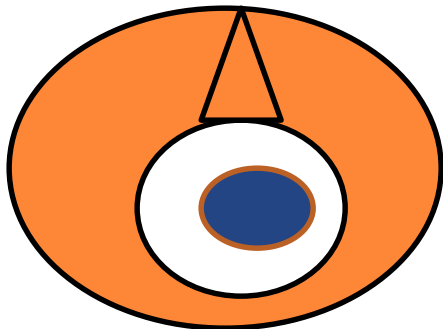
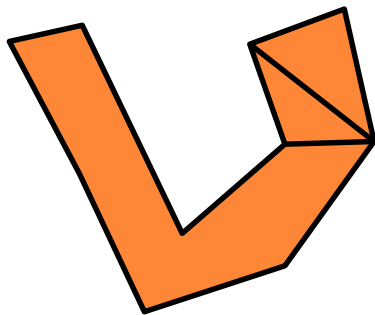


point in area



point on border
of an area

CONCAAF/CONVEX RESTRICTIES



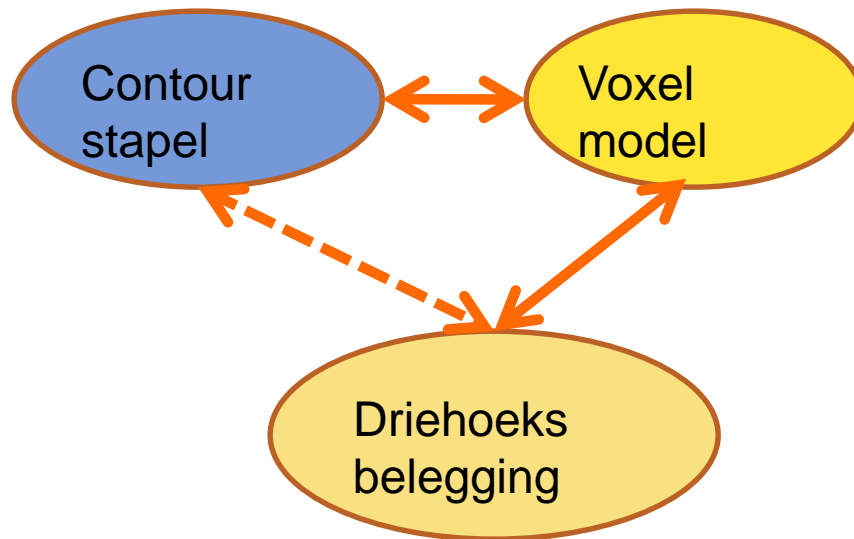
- Een polygon met meer dan 3 hoekpunten hoeft niet meer convex te zijn hierdoor kan het middelpunt buiten de polygon liggen!
- Gebied kan open binnengebied hebben
- Remedie: polygon is $\sum \Delta$

DATASTRUCTUREN VOOR 3D VORMEN

- Hoofdkeuzes met elementaire data elementen:
- Oppervlakken (buiten- en binnen-oppervlak)
 - Driehoeksbelegging (vector)
 - Buiten- en binnencontouren (vector/sampled)
 - Splines etc. (parametrisch)
- Volumes
- Vector:
 - 3D:Binnen/tussen oppervlakken
 - Σ 2D:Binnen/tussen stapel gesloten contouren in parallelle plakken
- Sampled:
 - Voxels (3D pixel)
 - Octtree (Level of Detail hierarchie)
- CSG (Constructive Solid Geometry)

TRANSFORMATIES VAN 3D DATASTRUCTUREN

- Omdat voor bepaalde bewerkingen specifieke datastructuren het handigst zijn, zijn er ook functies ontwikkeld die 3D datastructuren in elkaar om kunnen zetten



Marching cubes:

voxel \rightarrow Δ belegging

Contourfilling:

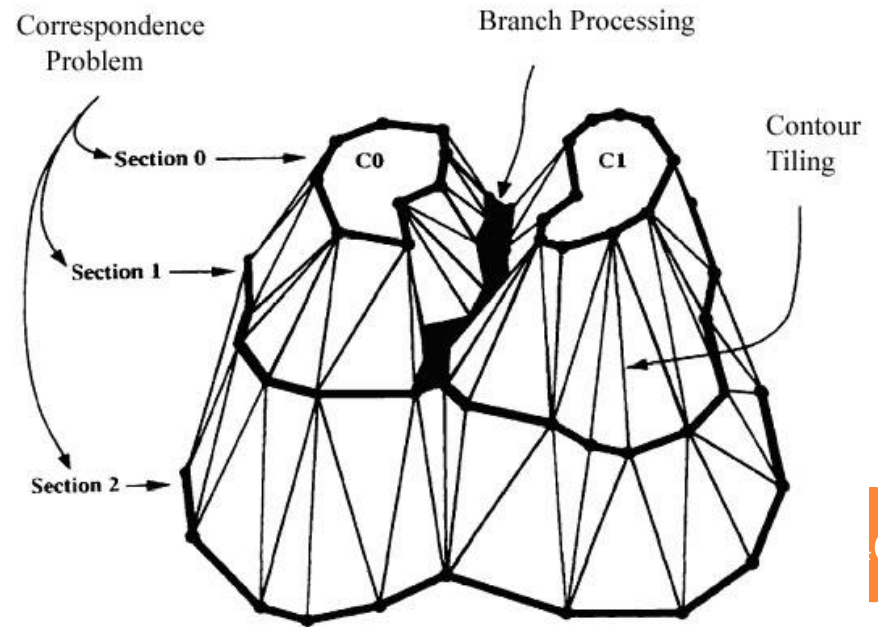
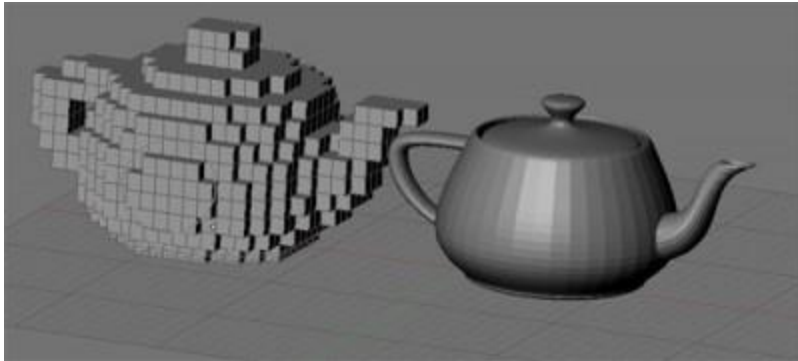
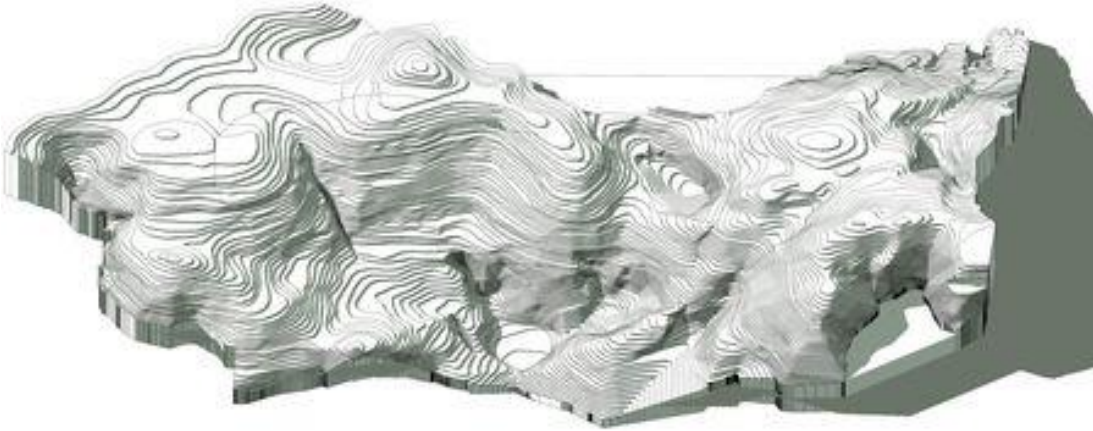
Contour \rightarrow voxel

Contouren voor invoer

Δ belegging voor visualisatie met shading

Gewenste datastructuur vaak afhankelijk soort taak

VOXELS, CONTOUREN, DRIEHOEKEN EN SPLINES



VISUALISATIE OP EEN ST DATABASE

SQL UITBREIDING

- Klassieke SQL queries kunnen veel spatio-temporele selecties/manipulaties niet uitvoeren
- Uitbreiding met spatio-temporele query formuleringen
- Soorten queries:
- Verzamelingenleer: vereniging, doorsnede, compl
- statistisch: min, max, gemiddeld, inhoud, ...
- Metrisch: afstand (in ruimte/tijd)
- Topologische: omvat, besloten in ...

EEN STDB VOOR DE TOEKOMST

- Het aantal spatio-temporele databases neemt exponentieel toe
- Modelleren van gedrag uit discrete waarnemingen (interpolatie)
- Extrapoleren naar de toekomst
- Genereren van waarschuwingen op basis van voorspeld (ongewenst) gedrag

STDB MET VECTOR EN RASTER DATA

- De meeste GIS en STDB opzetten beperken zich tot ofwel vector- ofwel raster gebaseerde systemen
- Uit raster gebaseerde data kan een benaderend beschijvingsmodel bepaald worden
- Weergave van model of fitten van rasterdata aan vectormodel vraagt om integratie van vector- en rasterdata in 1 applicatie

DE TOEKOMST IS AAN STDBS

- Om niet te verzuipen in de exponentieel groeiende hoeveelheid data
- Data heeft bijna altijd een plaats en/of tijd aspect
- Organiseer standaard op plaats en tijd
- Implementatie in STDBS: lastig
- Database gedeelte doenlijk
- Elementaire data-elementen: hou het eenvoudig
- Functionaliteit: mogelijk complexe bewerking van elementaire data-elementen
- Queries: lastig aan te leren/formuleren ST queries
- Visualisaties: vaak probleemspecifiek

SPATIO-TEMPORELE DATA

- Data mining
- Vector \leftrightarrow Raster
- Data-driven \leftrightarrow Space-driven
 - Interval aritmetiek
 - Vorm algebra
- Ingewikkelde queries
- Visualisatie lagen

DROZDEK

- Niets in Drozdek