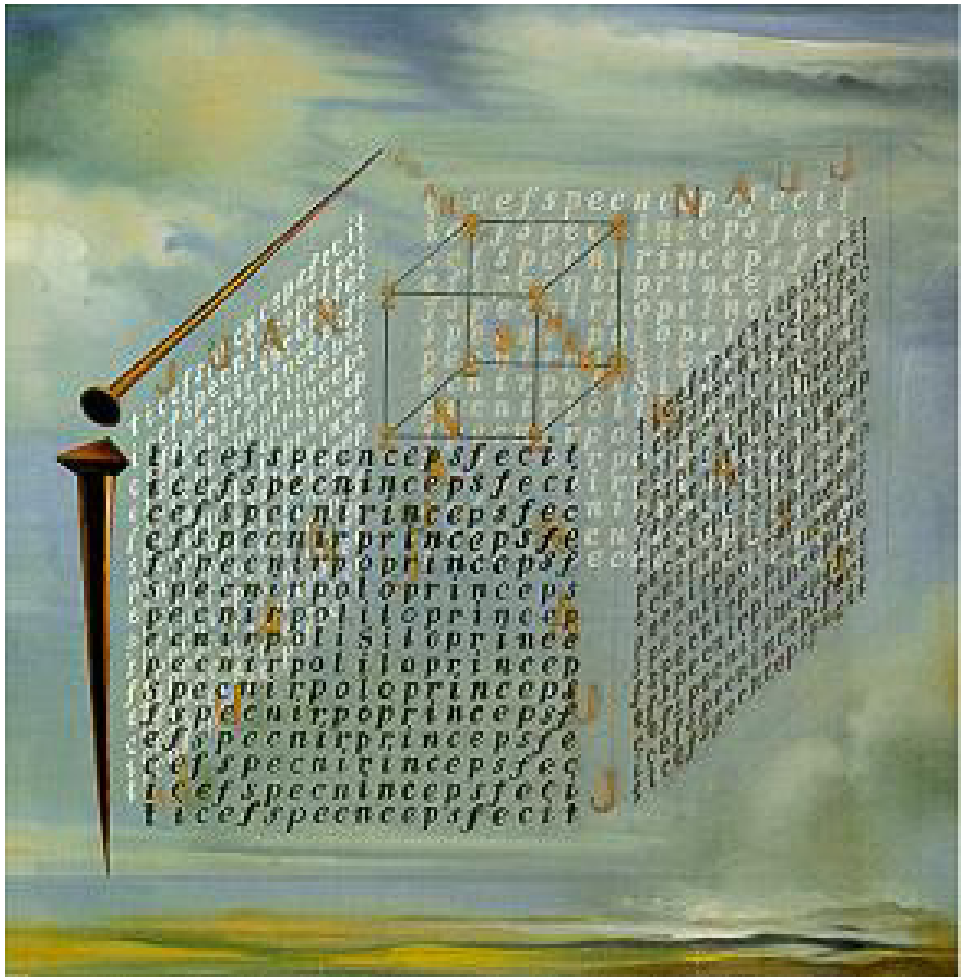
JACOBS UNIVERSITY

Raster Databases

- tutorial -

VLDB 2007
Vienna, 25-sep-2007



Peter Baumann

Jacobs University Bremen,
rasdaman GmbH

About the Presenter

www.faculty.jacobs-university.de/pbaumann

- Professor of Computer Science
  - research focus: large-scale multi-dimensional raster services
  - ...and application in geo, life science, Grid, and e-learning
  - geo raster service standardization: OGC
  - research spin-off: rasdaman GmbH
- Jacobs University Bremen
  - Private research university, est. 1998 by State of Bremen
  - >1100 Studenten, 91 nations, 25% German
  - ACQUIN accredited
  - Transdisciplinary, international, multi-cultural, all-english
- "Smart Systems" CS graduate program
  - MSc, PhD

Roadmap

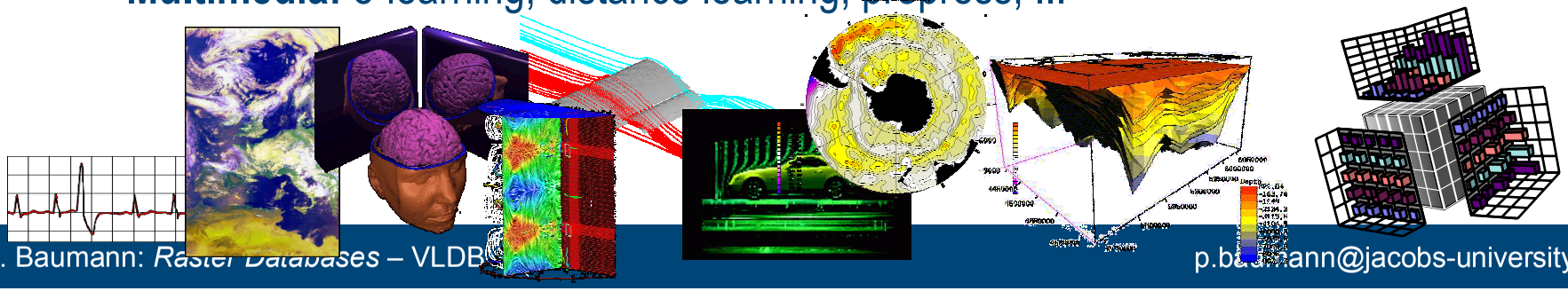
JACOBS UNIVERSITY

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

Why (Large) Arrays?

JACOBS UNIVERSITY

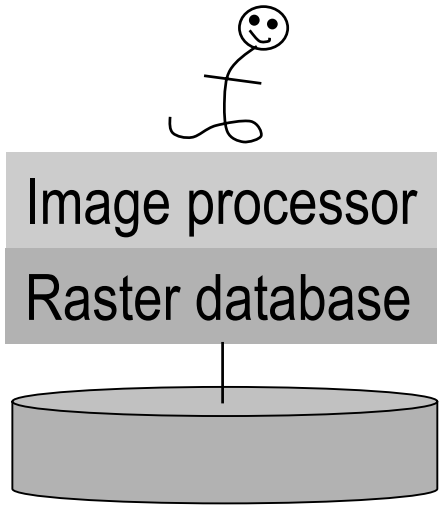
- Key characteristics: Dimensional, gridded (Euclidean space), large
  - raster = array = Multidimensional Discrete Data (MDD)
- Sensor, image, statistics data
  - **Life Science:** Pharma/chem, healthcare / bio research, bio statistics, genetics
  - **Geo:** Geodesy, geology, hydro/ocean, meteorology, earth system research, ...
  - **Management/Controlling:** statistics / Decision Support, OLAP, Warehousing, ...
  - **Engineering & research:** Simulation & experimental data in automotive/shipbuilding/ aerospace industry, turbines, process industry, astronomy, experimental physics, high energy physics, ...
  - **Multimedia:** e-learning, distance learning, prepress, ...



Raster Services: Differentiation

JACOBS UNIVERSITY

- multimedia databases
  - Analyse images, then drop them and work on auxiliary structure
- image processing
  - Advanced processing of rasters, but not on objects >>> main memory size
- image understanding, computer vision
  - General recognition probabilistic
  - databases to deliver exact results whenever possible
- Statistical DB / OLAP: dense vs sparse



high-level analysis

selection, data reduction

p.baumann@jacobs-university.de

Why Array *Databases*?

JACOBS UNIVERSITY

- Why should we bother?

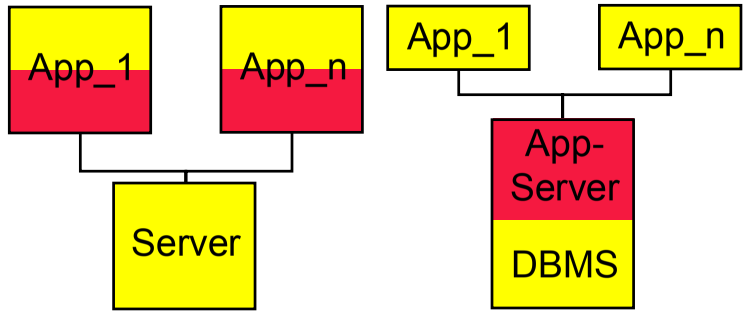
  ...because it's tons of data, that's us!

  - Multi-Terabyte objects, soon multi-Petabyte archives

- What can we offer?

  ...„Classical“ database benefits, for a new data type:

  - information integration
  - flexibility
  - scalability
  - ...plus all our further assets



Roadmap

JACOBS UNIVERSITY

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

History

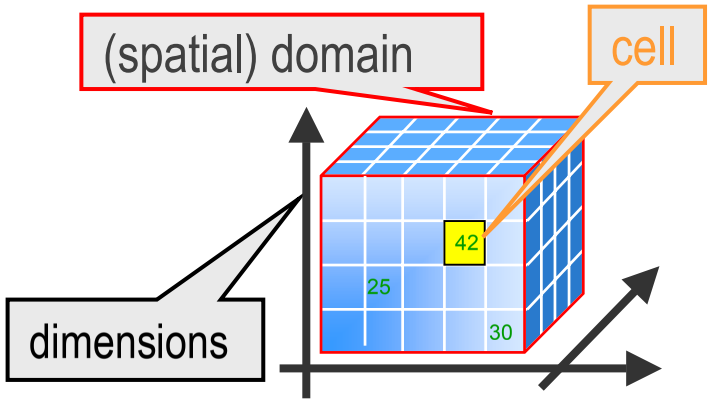
JACOBS UNIVERSITY

- Database view on raster images (eg, [XXX]):
  - „*image data...matrix of pixels*“, but: „*data appear just as a string of bits*“ → BLOBs
- Steps towards array support:
  - Image partitioning (tiling) in standardised files, API access library [Tamura 1980]
  - Fixed set of imaging operators (scaling, rotation, edge extraction, thresholding, ...) [Chang, Fu 1980; Stucky, Menzi 1989; Neumann et al 1992]
  - PICDMS [Chock, Cardenas 1984]: image stack (same res); no nesting; no architecture
  - rasdaman array algebra [Baumann 1991] & system [Baumann 1994+]
  - AQL [Libkin, Machlin, Wong 1996; Machlin 2007]
  - AML [Marathe & Salem 1997, 1999]; RAM [Ballegooij, de Vries, Kersten 2003]; [Ordinez, Garcia 2007]
  - ESRI ArcSDE, Oracle GeoRaster [200x]

Conceptual Modelling: Array Algebra

JACOBS UNIVERSITY

- Array = function:
  - a: $X \rightarrow F$, $a = \{ (x,f): x \in X, a(x)=f \in F \}$
    for finite multi-dimensional interval $X \subset \mathbf{Z}^d$, d>0, algebraic structure F
  - d: Dimensionality of a,
    X: spatial domain,
    F: Value set (*range*), Pixel, Voxel, ...



- 3 primitives:
  - Array constructor
  - Condenser
  - Sort
- Inspired by AFATL Image Algebra [Ritter et al 1990], basis for rasdaman system

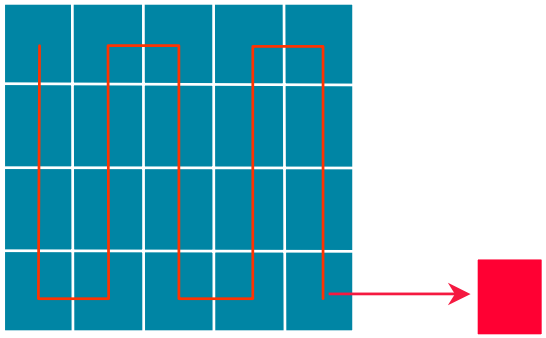
Array Operations: MARRAY

JACOBS UNIVERSITY

- Array constructor: $\text{MARRAY}_{X,p}(e(p))) := \{ (p,f): f = e(p), p \in X \}$
  - for n-D finite interval X, expression `e(p)` potentially containing occurrences of p, of result type F
  - Ex: $\text{MARRAY}_{X,p}(a[p] + b[p]) =: a + b$
    $\text{MARRAY}_{X,p}(p[0])$
- Shorthand: "induced operations"
  - (X = sdom(a) = sdom(b), a:X→F, b:X→G and f:F→F', g:F×G→G'):
  - $f_{ind} : X^F \to X^{F'}$, $f_{ind}(a) = \text{MARRAY}_{X,x}(f(a(x)))$ unary induced operation
  - $g_{ind}: X^F \times X^G \to X^{G'}$, $g_{ind}(a,b) = \text{MARRAY}_{X,x}(g(a(x), b(x)))$ binary induced operation

Array Operations: COND

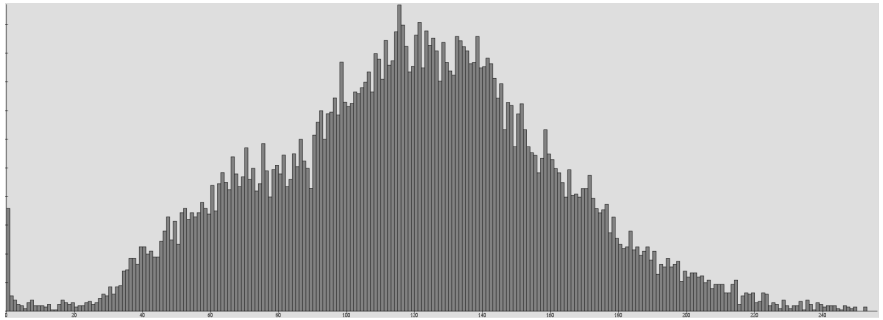
- Condenser: $\text{COND}_{o,X,x}(e(a,x)) := e(a,p_1)\ o\ e(a,p_2)\ o \ldots o\ e(a,p_n)$
  - n-D finite interval X, o commutative, associative, `e(a,p)` expression potentially containing a and p_i
  - Ex: add_cells(a) := $\text{COND}_{+,\text{sdom}(a),p}(a[p])$
- Shorthands:
  - count_cells(), avg_cells(), max_cells(), min_cells(), some_cells(), all_cells()
  - cf. Relational aggregates

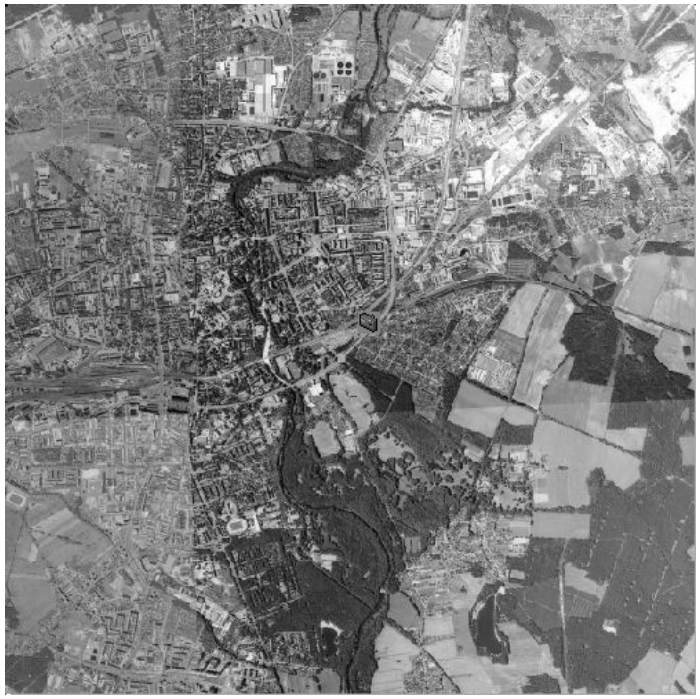


Example: Histogram

JACOBS UNIVERSITY

- Histogram of an n-D array over 8-bit unsigned integer:
  - $H(a) = \mathrm{MARRAY}_{a,[0:255]}(\mathrm{count\_cells}(a = n))$
- MARRAY can change cell type, dimension, domain!
  - sdom(H(image)) = [0:255]

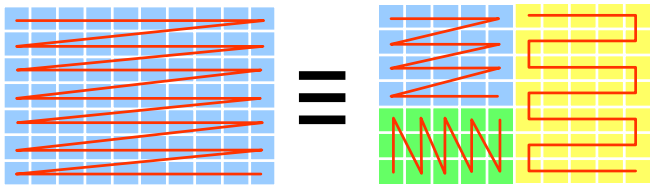




Properties

JACOBS UNIVERSITY

- Array Algebra declarative wrt array addressing
  - MARRAY: implicit iteration; COND: associative + commutative aggregator functions
  - tile-based processing:

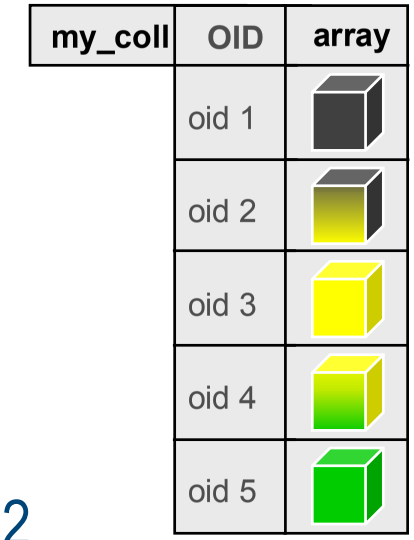


- Array algebra safe in evaluation
  - Array indexing without recursion
  - [Machlin 2007] goes beyond
- Expressive power: AML, Array Algebra equal to relational + ranking [Libkin, Machlin, Wong 1996]
  - In practice: filters, convolutions, statistics, ...

From Algebra To Query Language

JACOBS UNIVERSITY

- rasdaman ("raster data manager") middleware
  - in commercial use since 2001 (e.g. IGN-F: 13 TB ortho image, PostgreSQL)
- Data model: collections of typed arrays + OIDs
- Data definition language: rasdl [ODMG ODL]
  - Parametrised array constructor
  - Ex:
```
typedef marray
< unsigned char, [ 1:1024, 1:768 ]
> XgaGreyImage;
```
- Retrieval & manipulation language: rasql, based on SQL92
  - Select, insert, update, delete; speciality: partial update
  - Set oriented: all queries return sets, ...ahem: multi-sets, ...ahem: lists of arrays

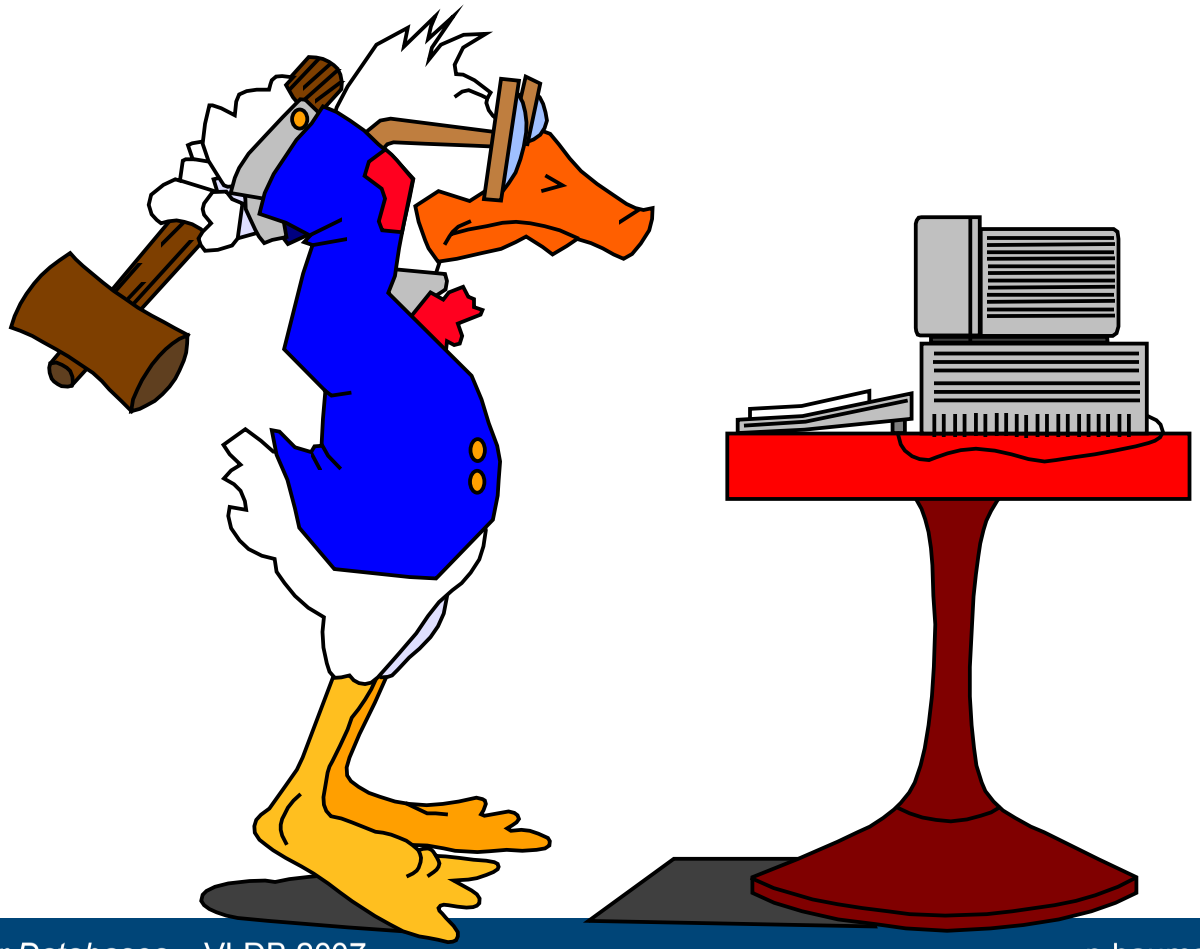


Inset: Types vs Type Constructors

- Remember: Marray is not a *type*, but a parametrized type *constructor*
  - Ex:
```
typedef marray
<    struct { double vx, vy; }, [ 0:*, 0:127, 0:63, 0:16 ]
> ECHAM_T42_Windspeed;
```
  - Cf. Stack: `Stack<>` is constructor, `Stack<int>` a concrete type
- Object-relational extensions allow user-defined data types, however not type constructors
  - Exception: Predator, U of Wisconsin-Madison

Demo

JACOBS UNIVERSITY



Oracle 10g/11g GeoRaster

JACOBS UNIVERSITY

- GeoRaster
  - Large 2-D geo raster images
  - Response to ESRI's ArcSDE 8
- Functionality:
  - (non-transparent) image pyramids
  - Subsetting, component extraction
  - reprojection?
- Observations
  - data independence?
    eg, pyramids visible
  - No SQL-integrated processing
  - No optimization found

```
declare
   g sdo_georaster;
   b blob;
begin
   select raster into g
   from uk_rasters
   where id = 4;
   dbms_lob.createTemporary(b,true);
   sdo_geor.getRasterSubset(
        georaster => g,
        pyramidlevel => 0,
        window =>
         sdo_number_array(0,0,699,899),
        bandnumbers => '0',
        rasterBlob => b );
end;
```

```
select g.green[0:699,0:899]
from   uk_rasters as g
where  oid(g) = 4
```

Roadmap

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

Storage Mapping

JACOBS UNIVERSITY

- Task: materialise finite interval $X \subset \mathbf{Z}^n$, find suitable (disk) access structure
  - Core structural property: Euclidean neighbourhood in $\mathbf{Z}^n$
  - Secondary, contents/app based: data density/ sparsity, data pattern, access pattern
- Excursion: difference to arrays in main memory
  - Ex: APL [Iverson 1968]
  - Assumption 1:
    access times independent from array position
    - *cost(„`a [x]`“) = const for all „`x`“*
  - Assumption 2:
    access times independent from access sequence
    - *cost(„`a [x] ; a [y]`“) = 2*cost(„`a [x]`“) = const for all „`x`“, „`y`“*

Storage Mapping: Variants

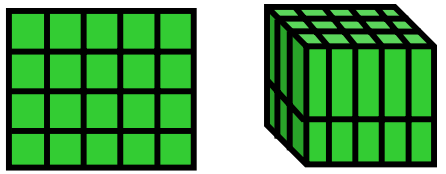
JACOBS UNIVERSITY

- BLOB (binary large object)
  - Coordinate free sequence
  - *Costs mainly position/dimension dependent*

- Sequence independent, coordinates explicit
  - Costs not position correlated, but high
  - *Sequence independent, coordinates explicit*

$$\{ (x_1,f_1), (x_2,f_2), ..., (x_n,f_n) \}$$

- Imaging, multidimensional OLAP
  - Partitioning, sequence within partition
  - *Costs low for bulk access, usually not location correlated*

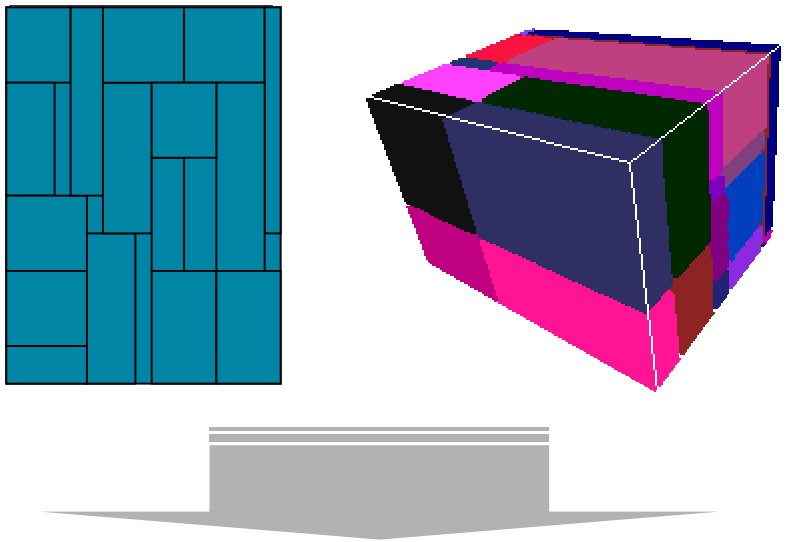


Partitioned Array Storage

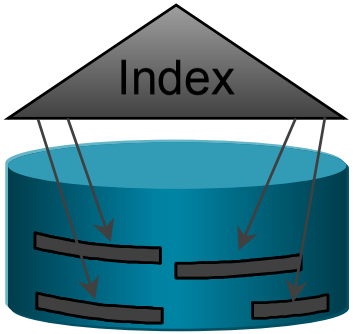
JACOBS UNIVERSITY

[Furtado 2000, Widmann 2001]

- multidimensional object
  ➠ multidimensional tiles
  - Tile = subarray
  - Also called "chunking" [Sarawagi, DeWitt]



- Tiles stored as BLOB in relational database
  - Compression
  - Geo index



Storage Layout: Tuning

- Parameters:
  - Tiling strategy
  - Geo index
  - Data format within tiles, incl compression
- Many dependencies
  - Access patterns, data contents
  - Buffer size, page size, CPU performance, bus bandwidth, ...

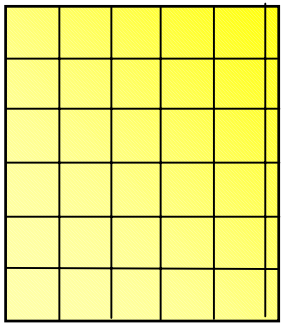
- In rasdaman:
  - Controlled via API, eg rasj class `RasStorageLayout`
  - Storage layout determined during insertion
  - Reorganisation = copying (beware!), possible via API

Tiling Strategies

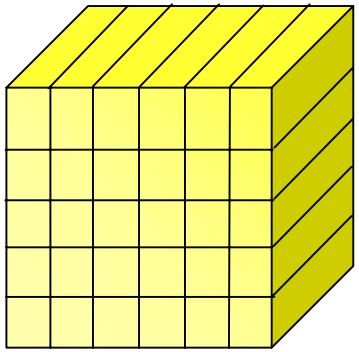
JACOBS UNIVERSITY

- Goal: faster tile loading by adapting storage units to access pattern
- Issues
  - When is tiling optimal? Tiling strategies?
- 3 sample tiling strategies [Furtado 1999]:

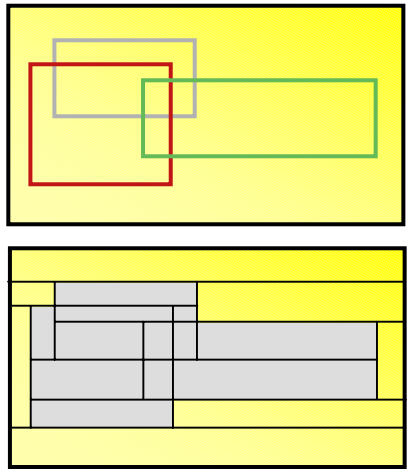
regular



directional



area of interest



Tile Based Compression

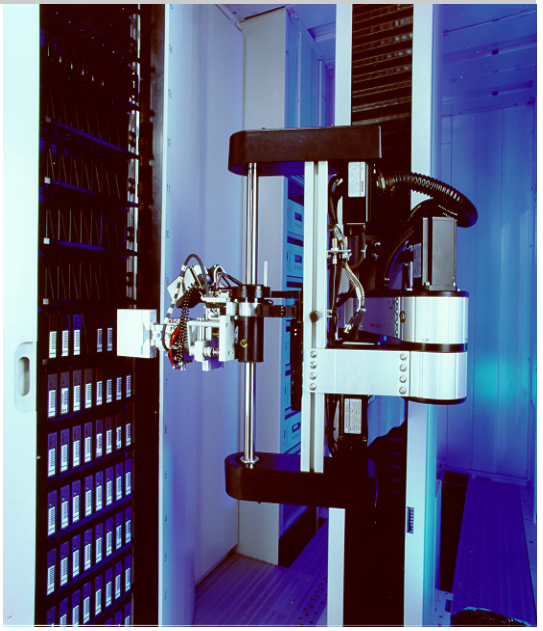
JACOBS UNIVERSITY

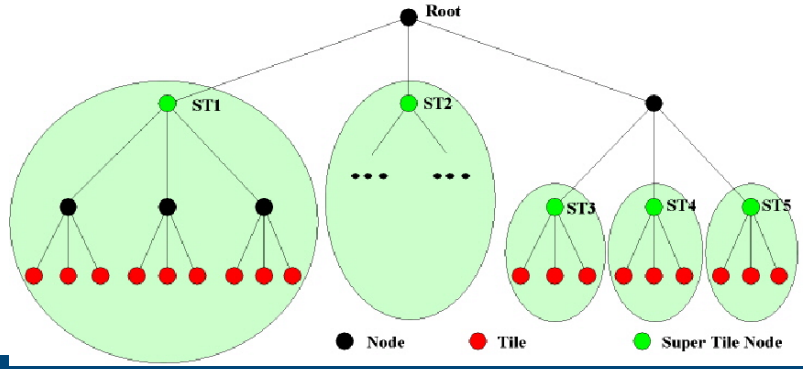
- Starting point: tiles are unit of access → compression units are tiles, too
- Degree of freedom (tuning parameter):
  mixing of compression methods according to access pattern
  - uncompressed ← hot spots
  - Fast & less storage gain ← high volume, frequent access
  - Slow & high storage gain ← infrequent access, high volume

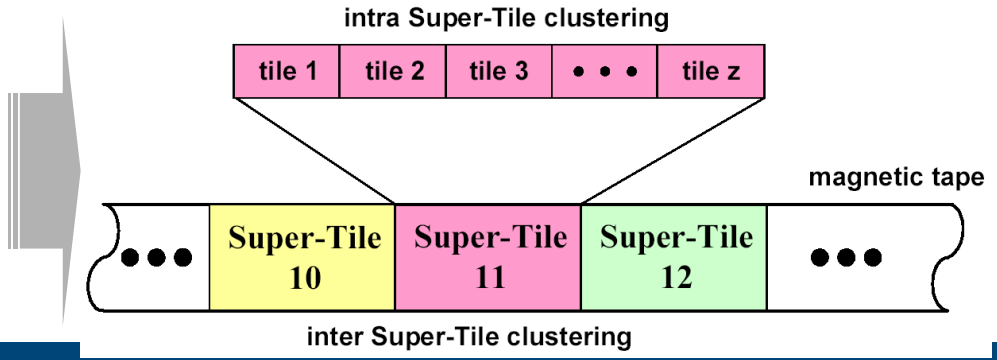
Hierarchical Storage Media Management

JACOBS UNIVERSITY

- near-line tape archives as storage extensions [Sarawagi, Stonebraker 1994]
- Issue: respect spatial clustering
  - access locality, long positioning times!
- super tile = tile set under some index node [Reiner 2001]
  - Natural unit, comfortable to handle (eviction information in index node!)







p.baumann@jacobs-university.de

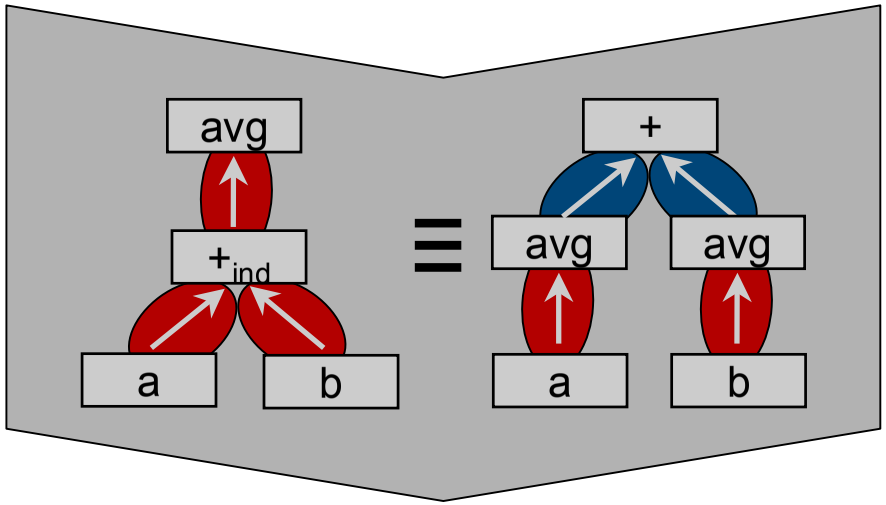
Roadmap

JACOBS UNIVERSITY

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

Query Optimization

```
select avg_cells( a + b )
from   a, b
```



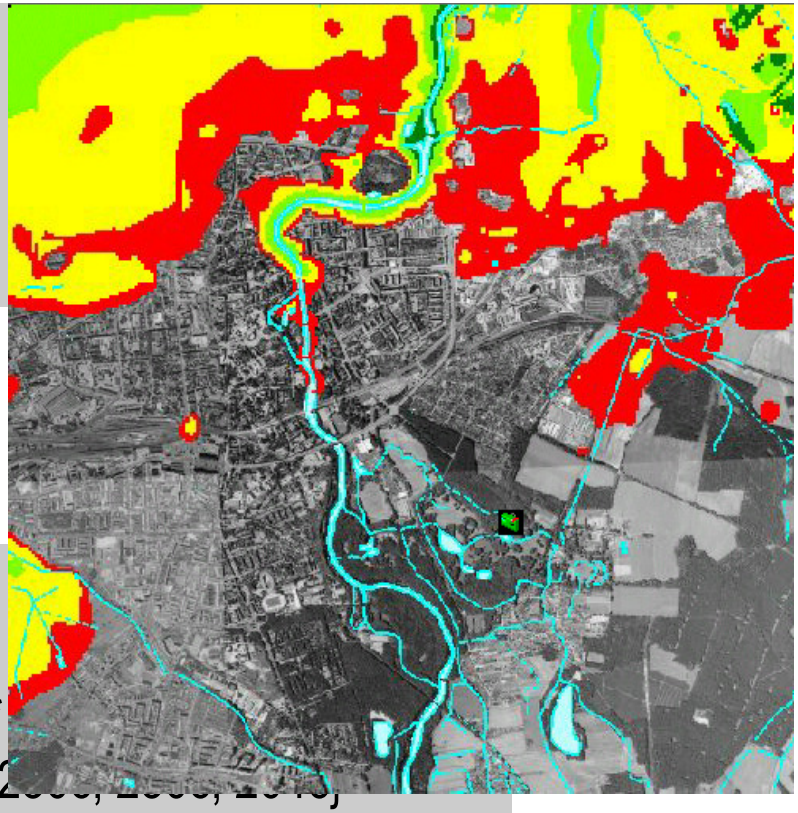
Tile stream
high traffic

Scalar stream
low traffic

```
select avg_cells( a )
     + avg_cells( b )
from   a, b
```

- *understood:*
  heuristic optimization
- *partially understood:*
  cost-based optimization

Optimisation Does Pay Off!



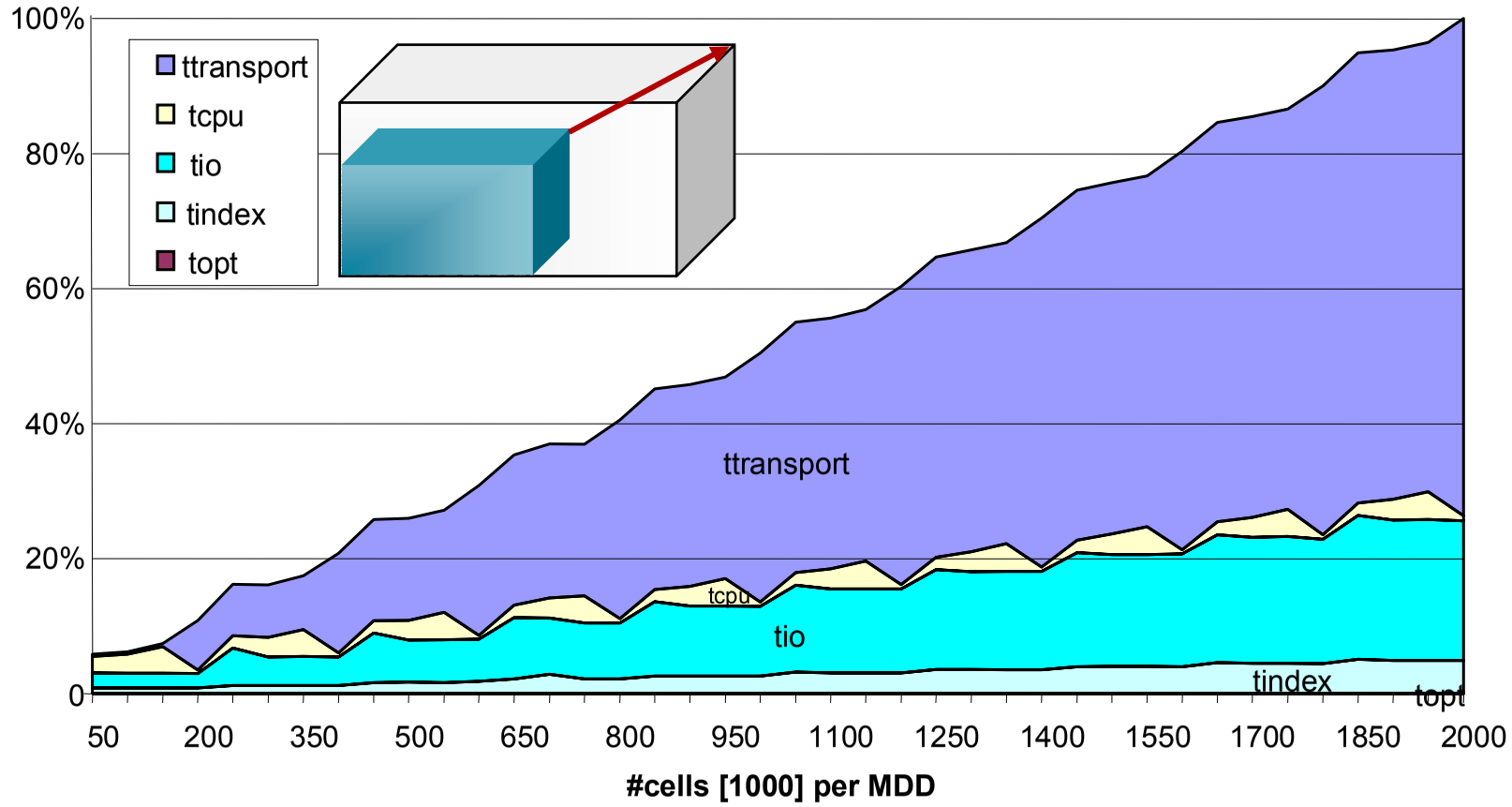
- Complex queries give more space to optimizer
- Typical OGC *Web Map Service* query:

```
select jpeg(
                scale(bild0[...],[1:300,1:300])          * {
     overlay ((scale(bild1[...],[1:300,1:300])<71.0)) *
     overlay bit(scale(bild2[...],[1:300,1:300]), 2)  * {
     overlay bit(scale(bild2[...],[1:300,1:300]), 5)  * {1c, 1c, 1c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 7)  * {102c, 102c, 102c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 6)  * {255c, 255c, 0c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 3)  * {191c, 242c, 128c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 4)  * {191c, 255c, 255c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 1)  * {0c, 255c, 255c}
     overlay bit(scale(bild2[...],[1:300,1:300]), 0)  * {102c, 102c, 102c}       )
from ...
```

Benchmarks: Data Access

JACOBS UNIVERSITY

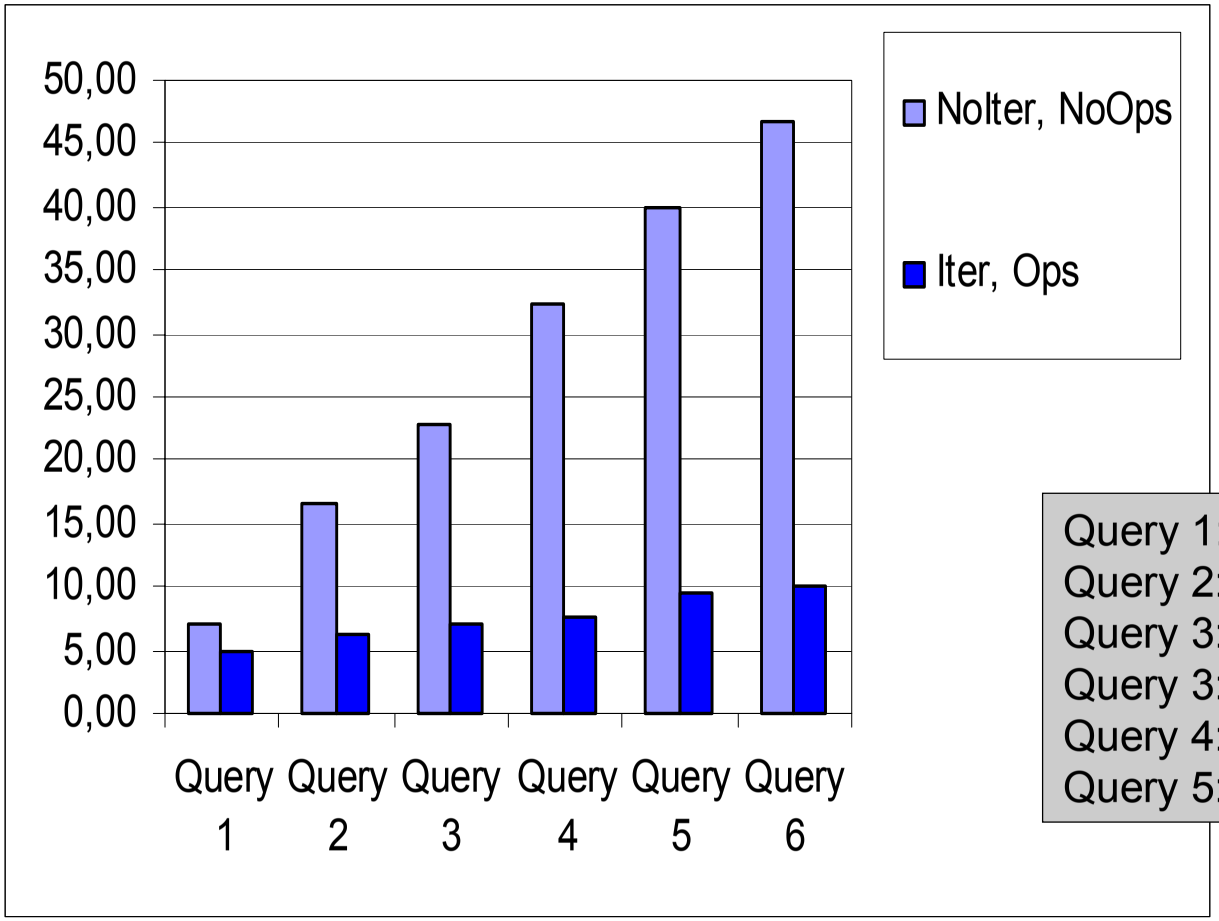
[Ritsch 2000, Widmann 2001]



Benchmarks: Data Processing

JACOBS UNIVERSITY

[Ritsch 2000, Widmann 2001]

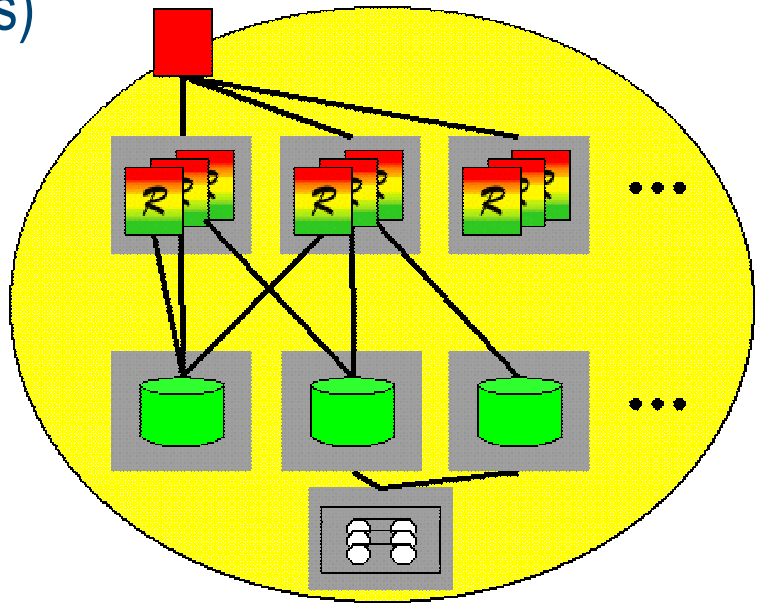


Query 1: access to 2-D object
Query 2: + 1 induced operation
Query 3: + 2 induced operations
Query 3: + 3 induced operations
Query 4: + 4 induced operations
Query 5: + 5 induced operations

Query Parallelisation

JACOBS UNIVERSITY

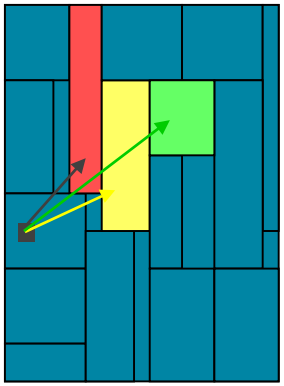
- easy: inter-query parallelization
  (one client – one dedicated server process)
  - Long-runners don't block service
  - higher throughput
- Non-trivial: intra-query parallelization
  (one client – several server processes)
  [Hahn 2003]
  - Idea: tiles dynamically assigned to processors
  - *Non-trivial array index patterns?*



Non-Local Access Patterns

JACOBS UNIVERSITY

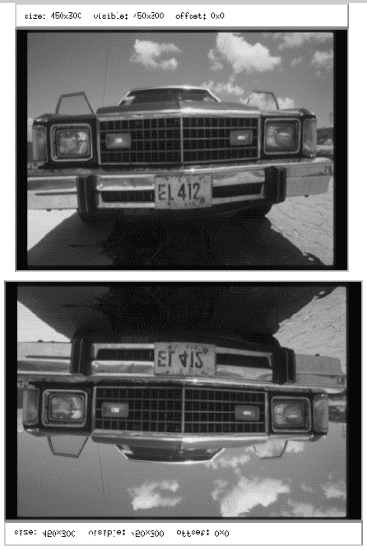
- **Problem**: how to efficiently evaluate tiles in face of non-trivial access patterns
  - `marray x in X`
    `values img[ f(x) ]`

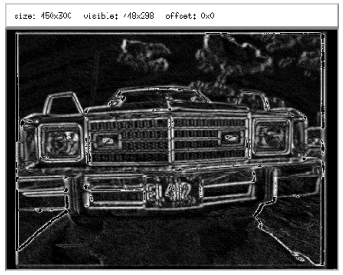


- **Use cases**:
  - mirroring: f(x) = hi - x
  - scaling: f(x) = x / s
  - Filtering:

```
marray x in sdom(img)
        values condense +
                over y in sdom(kernel)
                using a[x+y] * kernel[y]
```

1	3	1
0	0	0
-1	-3	-1





- **Approach**: address important cases first: const, linear expressions

Roadmap

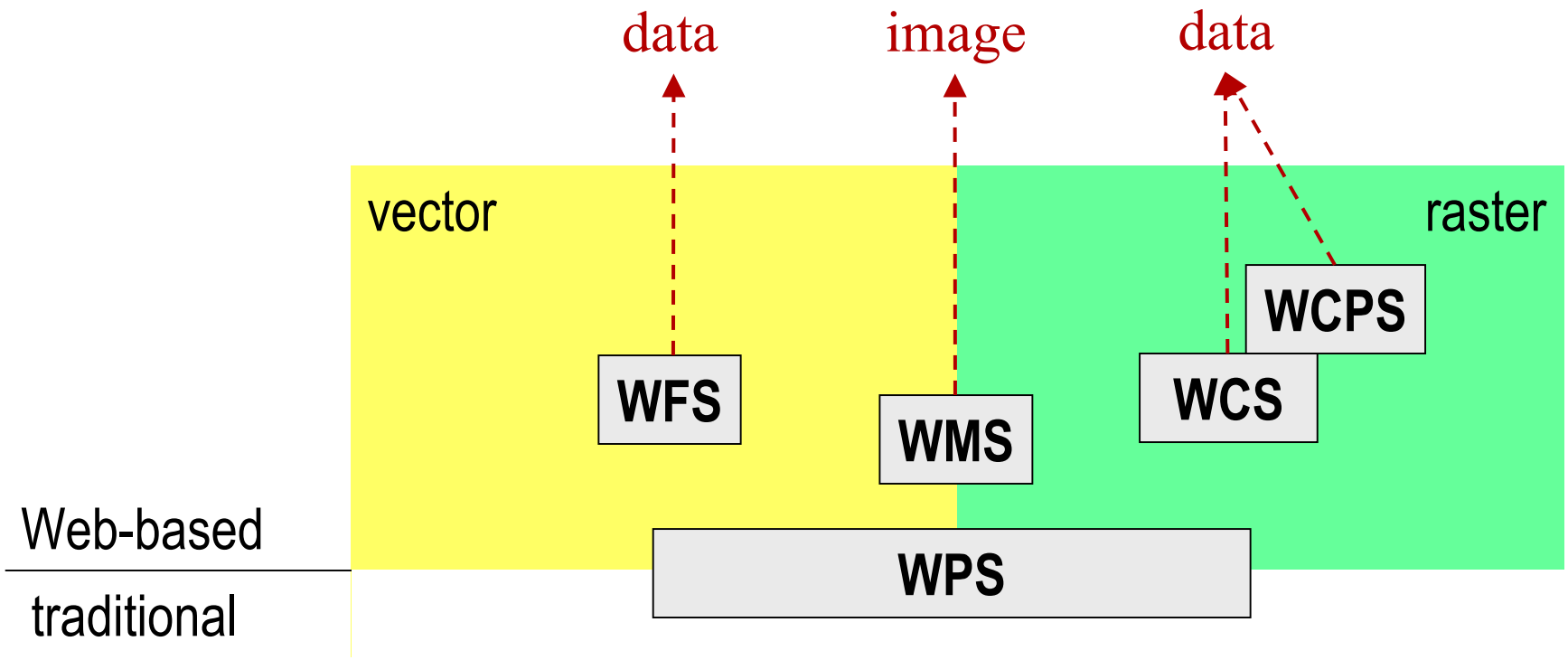
JACOBS UNIVERSITY

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

Geo Service Standardization

- OGC (Open GeoSpatial Consortium) driving geo service standards
  - Web-based modular, open, interoperable geo services
  - Liaisons with ISO TC 211, OASIS, CGI/IUGS; ...
  - www.opengeospatial.org
- Raster = coverage in OGC / GIS speak
  - Web Coverage Service Revision Working Group (WCS.RWG)
  - Web Coverage Processing Service Group (WCPS)
  - Coverages WG
  - GALEON OGCnetwork (Geo-interface to Atmosphere, Land, Earth, Ocean, NetCDF)

(Part of) The OGC Quilt



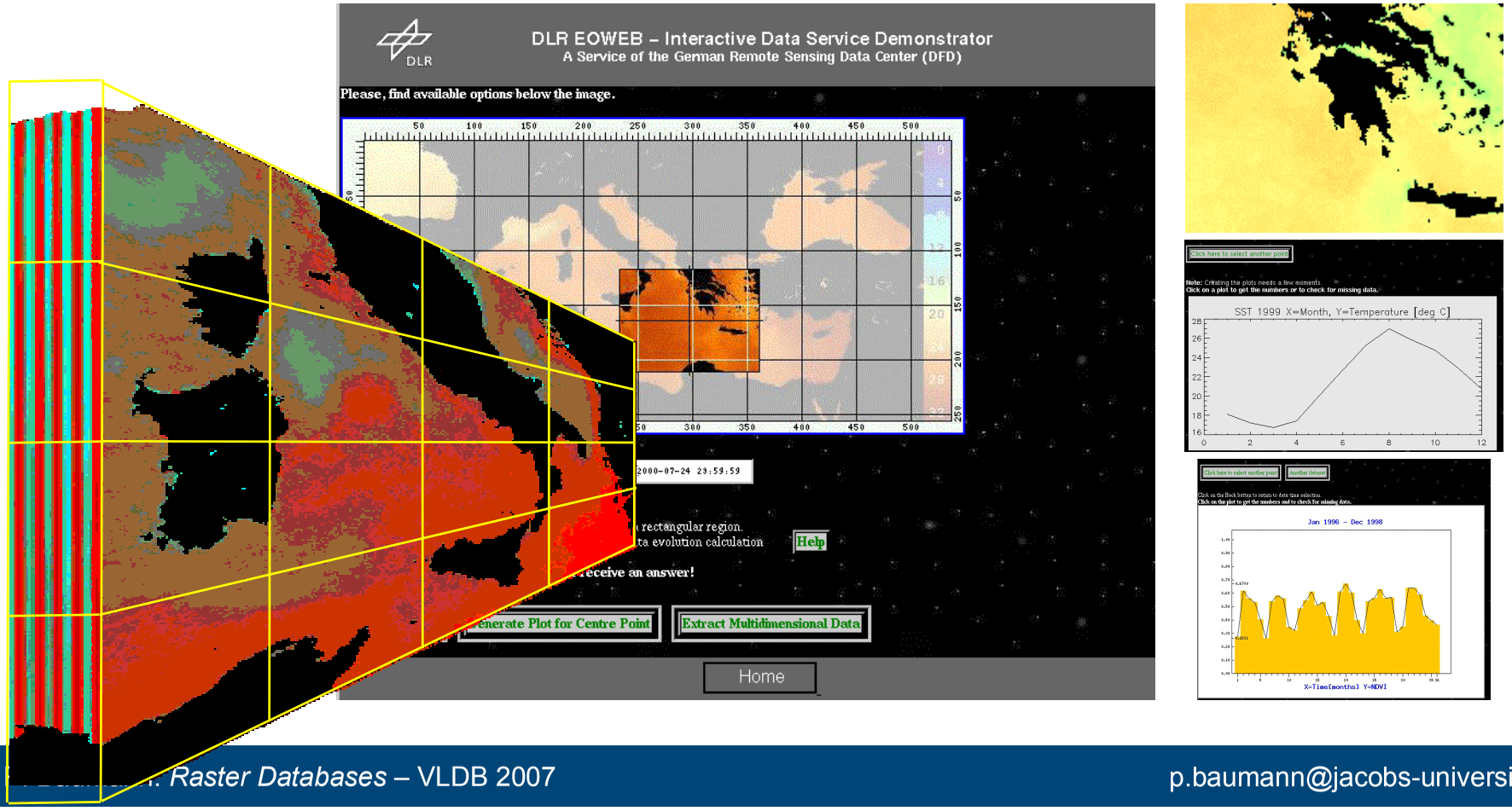
- *WMS "portrays spatial data → pictures"*
- *WCS: "provides data + descriptions; data with original semantics, may be interpreted, extrapolated, etc."*

Sample WCS Based 3-D Service

JACOBS UNIVERSITY

DLR-DFD: eoweb.dlr.de [Diedrich et al 2001]

based on rasdaman



p.baumann@jacobs-university.de

WCPS

JACOBS UNIVERSITY

- Request yields one or more n-D coverages
- Abstract syntax (requests shipped as XML):

```
for var in ( coverageList )
[ where condition(var) ]
return processingExpr(var)
```

- Example:

```
for m in ( ModisA, ModisB, ModisC )
where
    max( m.red > 127 )
return
    encode( m.red + m.nir,
            "tiff"
          )
```

→

```
( tiff_A,
  tiff_C )
```

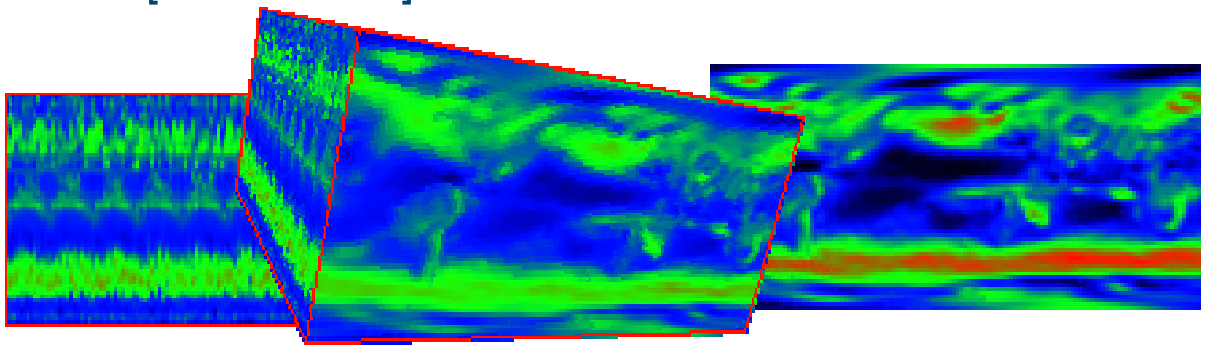
Climate Modelling

JACOBS UNIVERSITY

- Example: ECHAM T42 (cf. video)
  - 50+ physical parameters („variables“): temperature, wind speed x/y, humidity, pressure, CO2, ...
  - 2.5 TB per variable

dimension	extent
Longitude	128
Latitude	64
Elevation	17
time (24 min per time slice)	2,190,000 (200 years)

- observation:
  Huge volumes moved, only part needed (10:1)
  - [Kleese 2000]



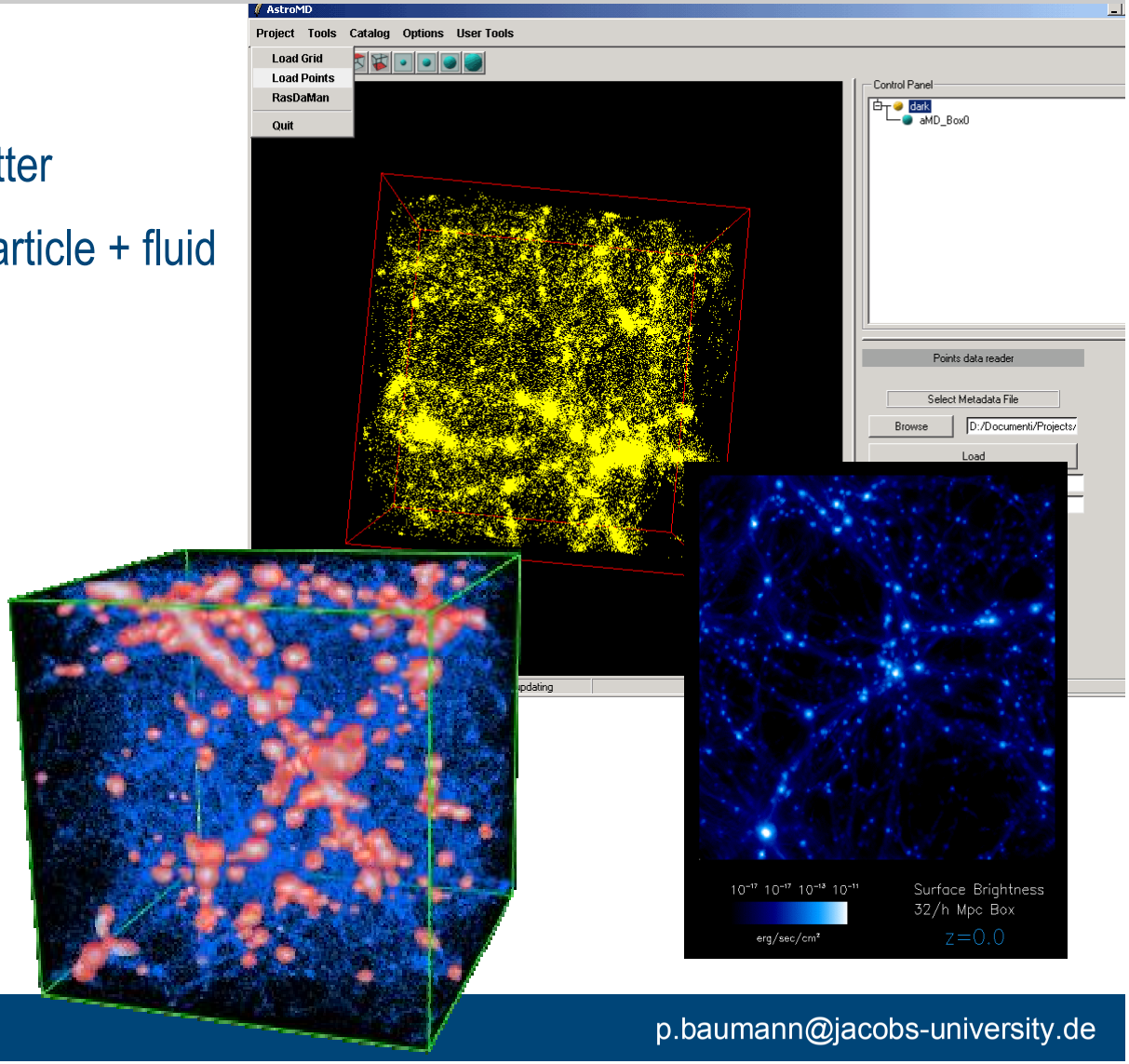
DKRZ: 24-node NEC SX-6

p.baumann@jacobs-university.de

Cosmological Simulation

JACOBS UNIVERSITY

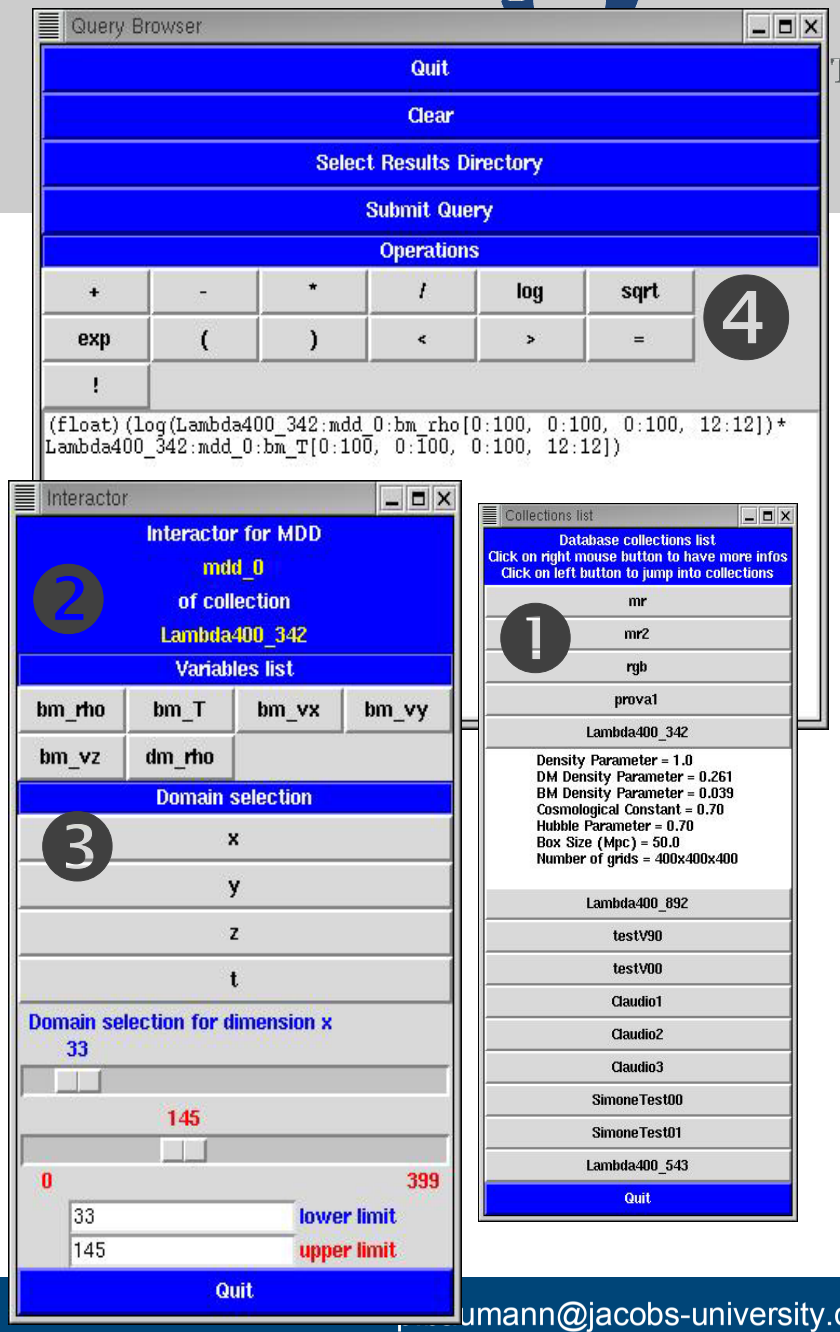
- Modelling domain: 4-D
  - Dark matter, baryonic matter
  - → Coupled simulation: particle + fluid
- Results are 3-D/4-D cutouts from universe
  - Eg, 64 Mpc^3 (Mega Parsec; 1 pc = 3.27 light years)
- Screenshots: AstroMD [Gheller, Rossi 2001]



p.baumann@jacobs-university.de

Cosmology (contd.)

- Guided retrieval:
  - Selection of objects ❶ and their cell components ❷
  - interactive setting of trim operations per dimension ❸
  - Augmented with induced operations ❹
- Suitable for expert users
- Details: cosmolab.cineca.it



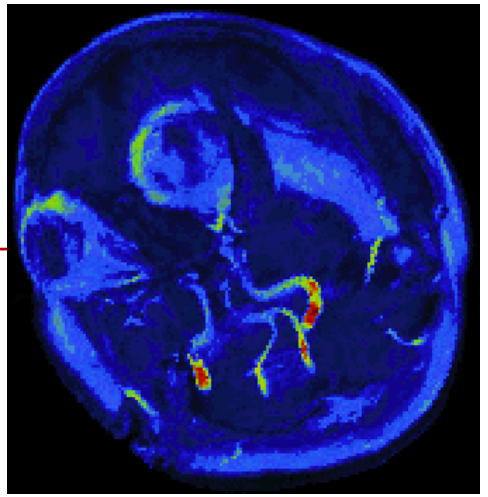
umann@jacobs-university.de

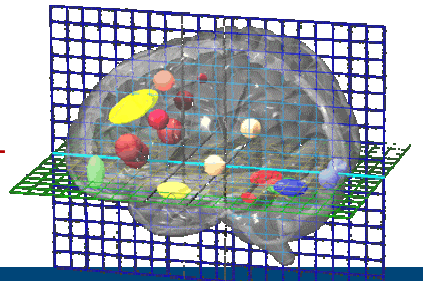
Human Brain Imaging

JACOBS UNIVERSITY

- Research goal: to understand structural-functional relations in human brain
- Experiments capture activity patterns (PET, fMRI)
  - Temperature, electrical, oxygen consumption, ...
  - → lots of computations → „activation maps“
- Example: “*a parasagittal view of all scans containing critical Hippocampus activations, TIFF-coded*.“

```
select tiff( ht[ $1, *:*, *:* ] )
from   HeadTomograms as ht,
       Hippocampus as mask
where  count_cells( ht > $2 and mask )
       / count_cells( mask )
       > $3
```





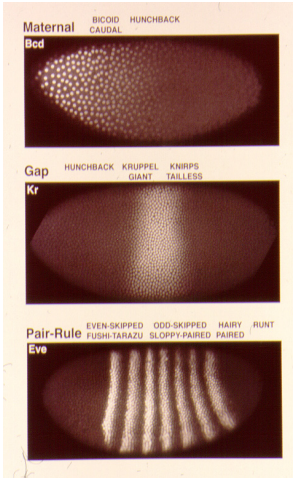
`$1` = slicing position, `$2` = intensity threshold value, `$3` = confidence

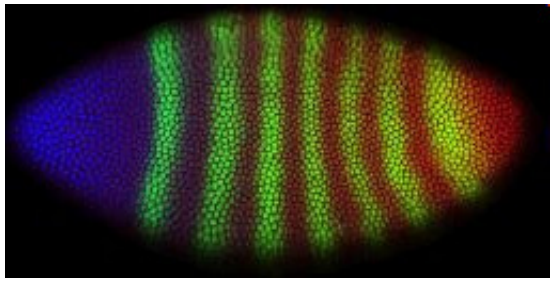
Gene Expression Analysis

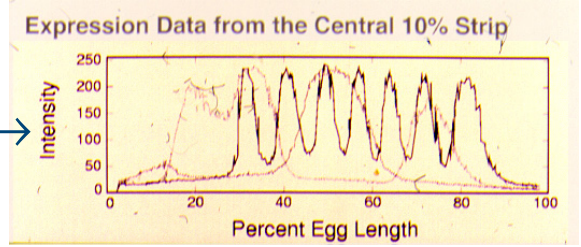
JACOBS UNIVERSITY

http://urchin.spbcas.ru/Mooshka/ [Samsonova et al]

- Gene expression = reading out genes for reproduction
- Research goal: capture spatio-temporal expression patterns in Drosophila







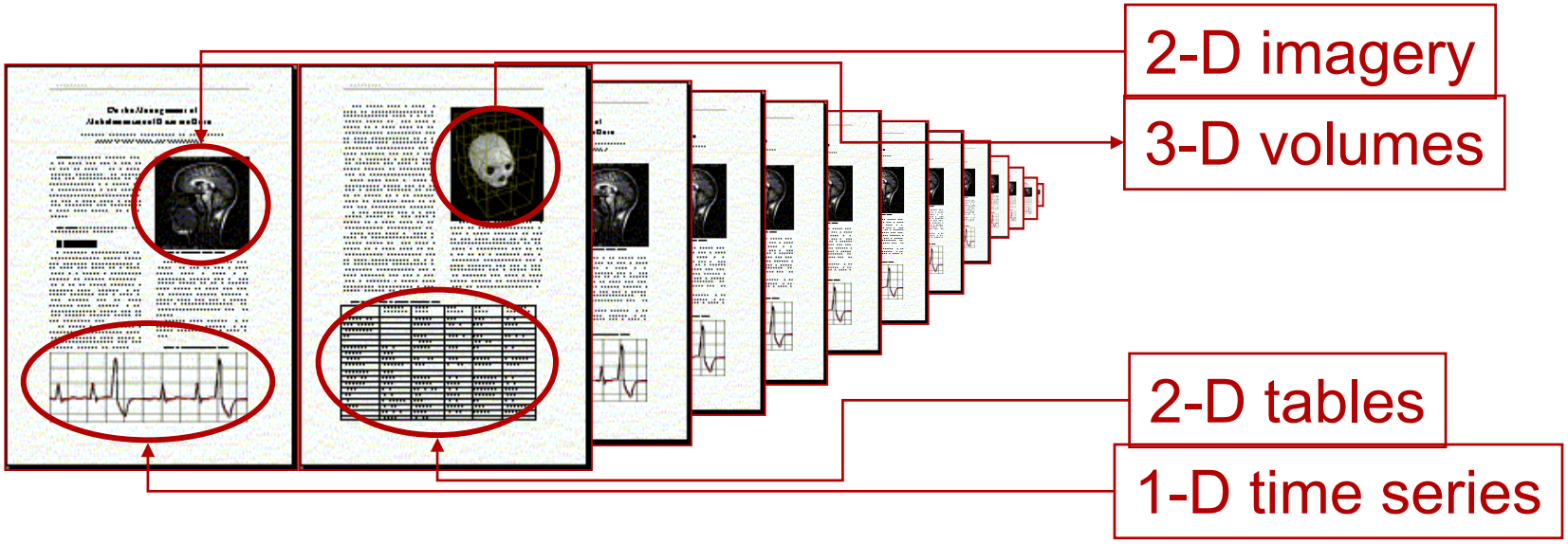

```
select jpeg( scale( {1c,0c,0c}*e[0,*:*,*:*]
                  +{0c,1c,0c}*e[1,*:*,*:*]
                  +{0c,0c,1c}*e[2,*:*,*:*], 0.2 ) )
from EmbryoImages as e
```

Roadmap

JACOBS UNIVERSITY

- Introduction
- Conceptual modelling
- Architecture
  - Arch I: Storage Management
  - Arch II: Query Processing
- Applications
- Wrap-up

Vision: Document Integrated Retrieval



„all clinical trials of drug X
where patient temperature > 40º C within the first 48 hours.“

Finally...

JACOBS UNIVERSITY

- value-added raster data services important + growing field
  - Service providers & users demand it
  - Currently driven by geo apps
  - "2D, 3D imagery next great challenge in geo databases" [Xavier Lopez, Oracle]
  - Many research issues in all facets
  - rasdaman system: commercialized + research vehicle
- contact: p.baumann@jacobs-university.de