

Introdução à Visualização: um Panorama

Maria Cristina F. de Oliveira

Rosane Minghim

Depto. de Ciências de Computação
ICMC, USP

2010/2011/2018

Desafio

- É preciso tornar os dados acessíveis e fornecer às pessoas recursos adequados para interpretá-los
 - <http://www.gapminder.org/>
 - Ferramentas de visualização de dados
 - Organização com fins não lucrativos cujo objetivo é tornar acessíveis e transparentes dados sobre o desenvolvimento humano mundial
 - Fontes: UM, UNESCO, ONGs,
 - *Search statistics through Google and watch it move with Gapminder*

Panorama

- Motivação
- Visualização Científica
- Visualização de Informação
- Mineração Visual e Analítica Visual
- Desafios
- Bibliografia

3

Motivação

Metáforas visuais fazem parte do processo cognitivo humano...

Cognição: a aquisição ou o uso de conhecimento

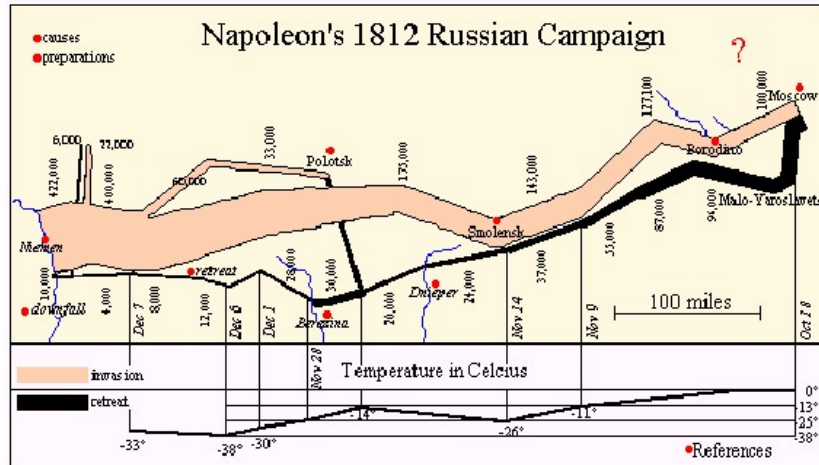
“Visualizar: construir na mente uma representação visual”

Shorter Oxford English Dictionary

Definição “tradicional”

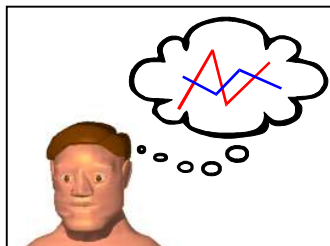
4

<http://www.edwardtufte.com/tufte/minard>

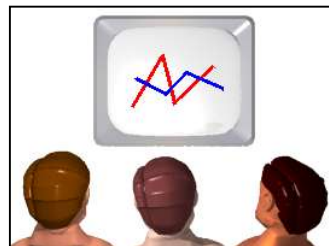


5

Visualização como Artefato



Construção Interna



Artefato Externo

Suporte à tomada de
Decisão

6

Visualização

- Área de pesquisa que estuda o uso de representações gráficas (metáforas visuais) para apoiar tarefas de exploração, análise e interpretação de dados
- Modelos gráficos acoplados a estratégias de interação: exploração dinâmica de dados
- A partir de 'dados', gera representações gráficas (imagens) interpretáveis pelas pessoas

7

Visualização

Hamming 1973: "*the purpose of computation is insight, not numbers*"

Card et al. 1999: "*the purpose of visualization is insight, not pictures*"

Principais objetivos desse "*insight*": descoberta, verificação de hipóteses, tomada de decisões, explicação

A Visualização é útil na medida em que amplia a nossa capacidade de executar essas e outras tarefas cognitivas

8

Visualização

- **Motivação**
 - Muito fácil coletar e armazenar dados
 - Muito difícil processar, analisar e interpretar todos os dados coletados, identificar o que é relevante
 - Volume dos dados muito grande
 - Dimensionalidade dos dados muito alta
 - Natureza dos dados muito diversa
 - Registros, textos, imagens, vídeos, voz, ...
- **Desafios para pesquisadores...**

9

Visualização

- **Contraposição com**
 - **Visão Computacional**
 - Ponto de partida são imagens adquiridas, que são processadas e analisadas para extrair informação útil
 - Análise das imagens pelo sistema computacional (processo automático)
 - Objetivo seria dotar um sistema de discernimento visual comparável ao do ser humano

10

Visualização

- Contraposição com
 - Computação Gráfica
 - Ponto de partida são modelos geométricos, a partir dos quais gera imagens 'realistas' – foco é 'ilusão', não interpretação
 - Entretanto, modelos geométricos precisos as vezes são gerados a partir de dados 'reais'... animações realistas podem ser geradas a partir de simulações de processos físicos...
 - CG ou Visualização??!

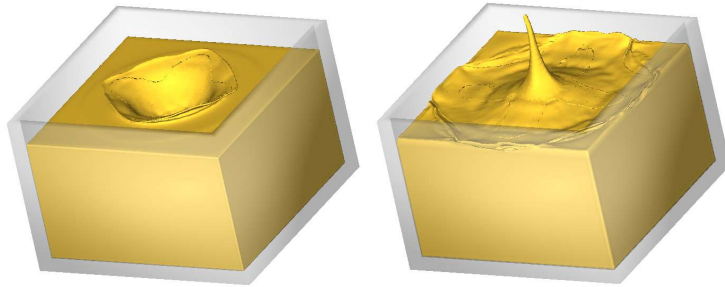
11

Reconstrução



Créditos: Helton Biscaro, LCAD, 2005

Escoamentos



Computing & Visualization in Science, 2000
Int. J. Numerical Methods in Fluids, 2002

13

Visualização

- Visualização no Processo de Análise de Dados
- Conceitos básicos
 - Comp. Gráfica, Dados, Malhas, Técnicas, Pipeline de Visualização Exemplos
- Visualização Planar e Volumétrica
 - Escalares, Vetoriais, Tensores
 - Imagens e Volumes
 - Visualização Médica
- Representação de Dados
- Visualização de Informação
 - Pontos, Atributos, Grafos, Projeções, Aplicações
- Mineração Visual de Dados

Visualização

- Científica

Dados Médicos, Meteorológicos, Fluidos

- de Informação

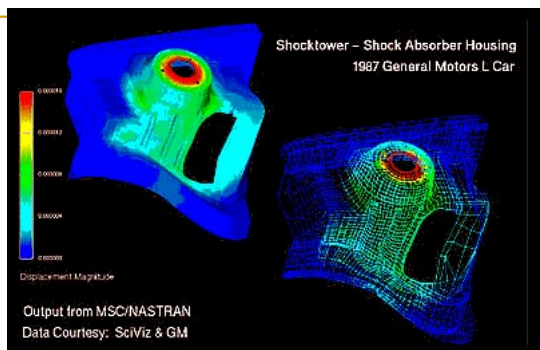
Bases de Dados, Web, Carga em Redes, Relações Humanas

- de Software

Projeto, Depuração, Teste

ViSC

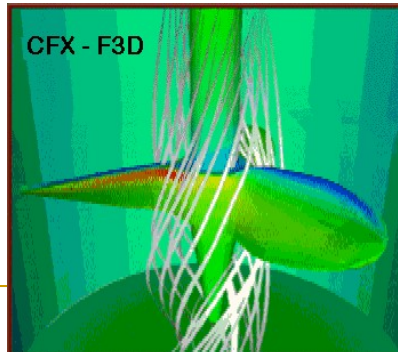
Exemplos



Engenharia

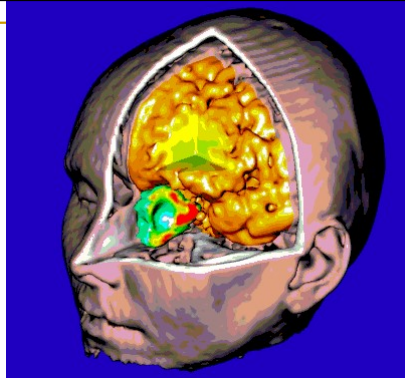
Simulação - CFD

CFX - F3D



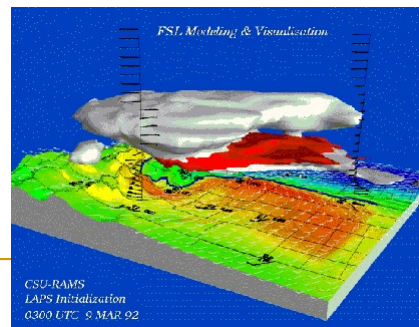
ViSC

Exemplos



Medicina

Meteorologia

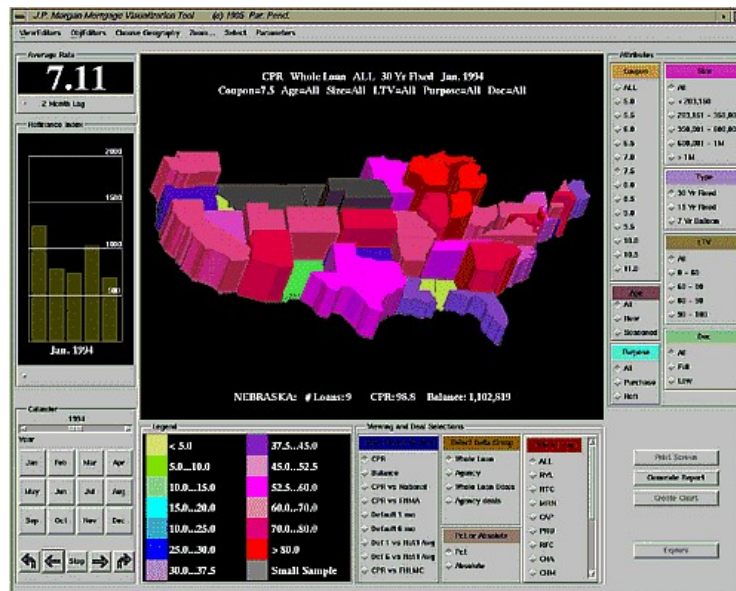


InfoVis

Exemplo

Dados

Financeiros



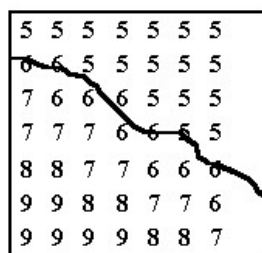
Visualização Científica

- Dados 'científicos'
 - Têm uma representação geométrica conhecida, real ou imaginária. P.ex., órgãos do corpo humano, moléculas, ...
 - Resultantes de processos físicos, medidos, coletados ou simulados: atributos espaciais (e temporais) são determinantes para a visualização
- Interação com CG, HCI, matemática...
 - 'Reconstrução' de modelos geométricos, simulação numérica, *rendering* adequado dos modelos

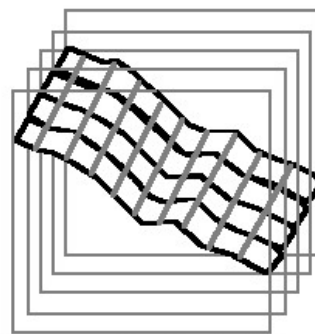
19

Visualização Volumétrica Reconstrução de Superfícies e Volumes

Contouring (object order)

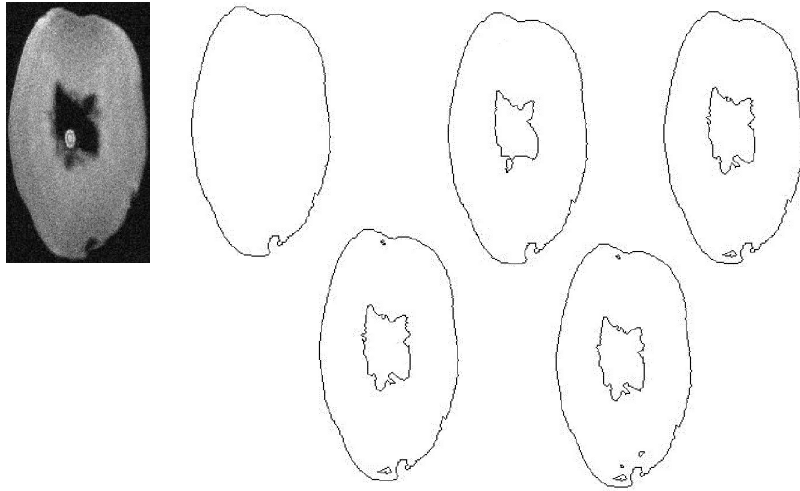


Contouring one slice
T=6

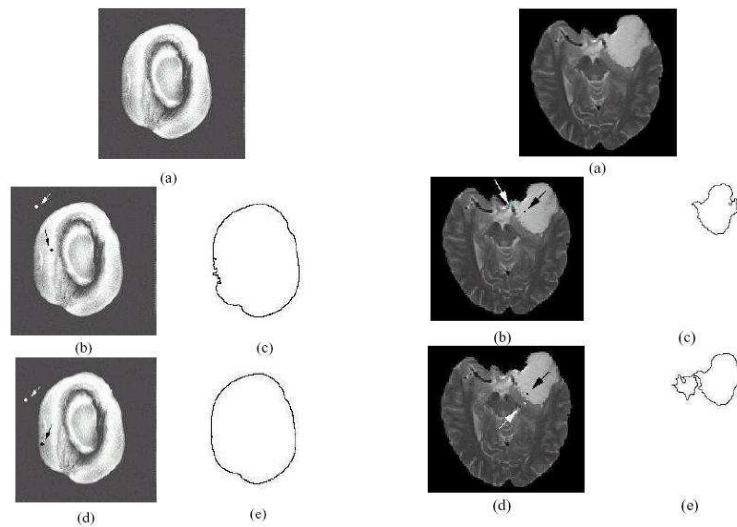


Contouring five identical slices

Exemplo: contornos

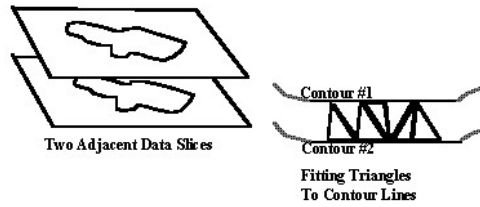


ESCOLHA DAS SEMENTES



Visualização Volumétrica Ajuste de Modelos

Connecting Slices

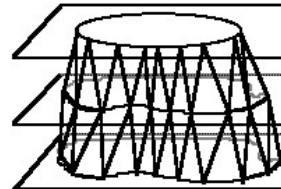


Connecting Slices (cont.)

Find one Closed
Curve Contour

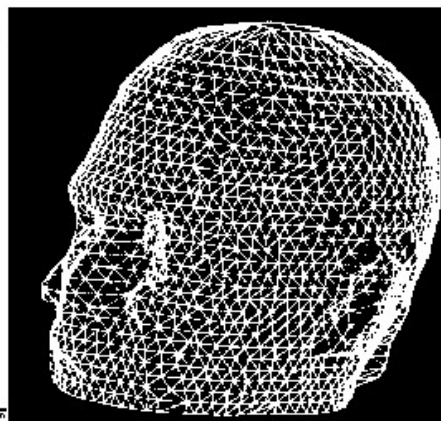
Connect Curves
with Triangles

Render the
triangles

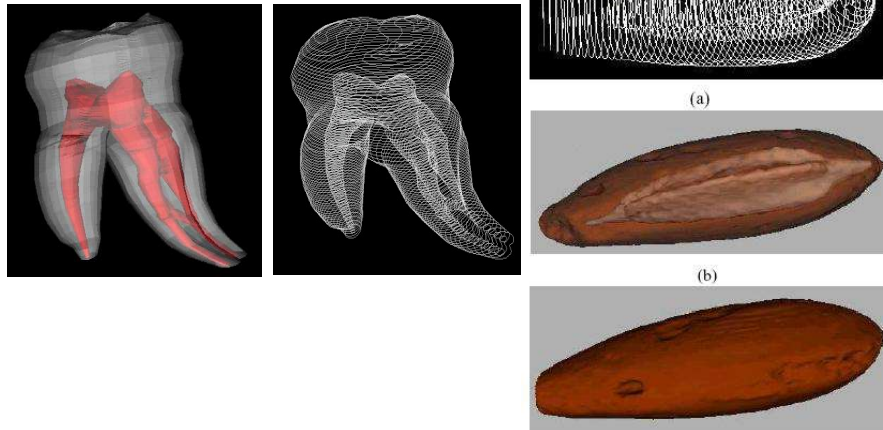


Visualização Volumétrica Ajuste de Modelos (contornos)

Tessellated Head



Modelos reconstruídos dente e mamão



Visualização Volumétrica Ajuste de Superfícies - *Marching Cubes*

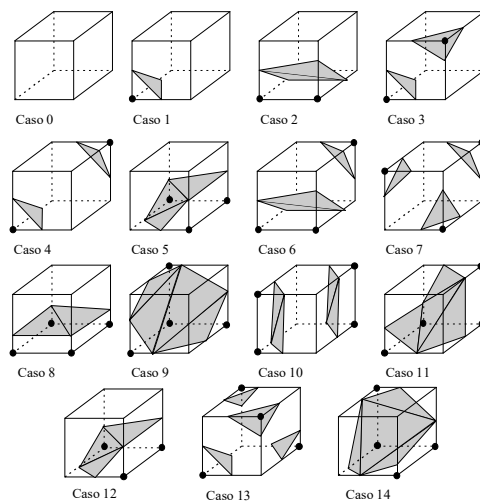
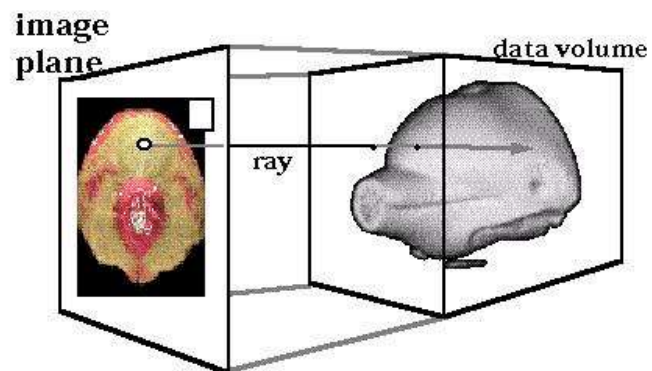


Figura 3.3 - Tabela de casos de triangulações para o *marching cubes*.

Visualização Volumétrica
Rendering Volumétrico Direto (DVR)

Métodos na ordem da imagem (Ray-casting)

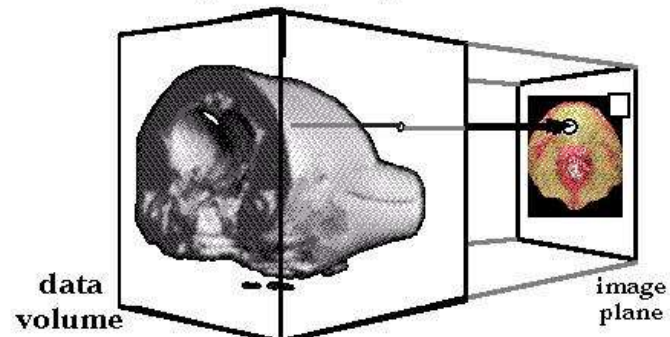
Ray-casting Algorithms



Visualização Volumétrica
Rendering Volumétrico Direto (DVR)

Métodos na ordem do Objeto (projeção)

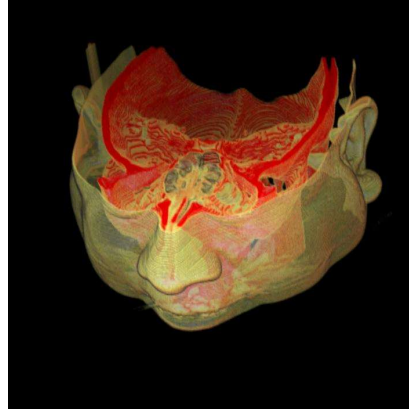
Projection Algorithms



Visualização Volumétrica

DVR

Exemplo



Visualização de Informação

- Dados 'abstratos'
 - Não têm uma representação geométrica (espacial) inerente; qualquer representação adotada é arbitrária (atribuída)
 - Ex.: transações de clientes em uma base de dados, acessos a um *site* na Web, movimentação financeira na bolsa de valores, hierarquia de diretórios, coleções de textos, ...
 - Em geral, coletados, medidos ou 'criados'
- Interação com HCI, 'design' gráfico, semiótica, estatística, mineração...

Visualização de Informação

- Visualização Multidimensional
- Visualização Exploratória
- Metáforas visuais para ajudar pessoas a explorar/analisar dados
- Modelos gráficos acoplados a estratégias de interação ⇒ processos dinâmicos de exploração

31

Visualização de Informação

- Dados estruturados
 - Tabelas de registros com múltiplos atributos (numéricos ou categóricos)
- Dados não estruturados
 - Textos, imagens, ...

32

Visualização de Informação

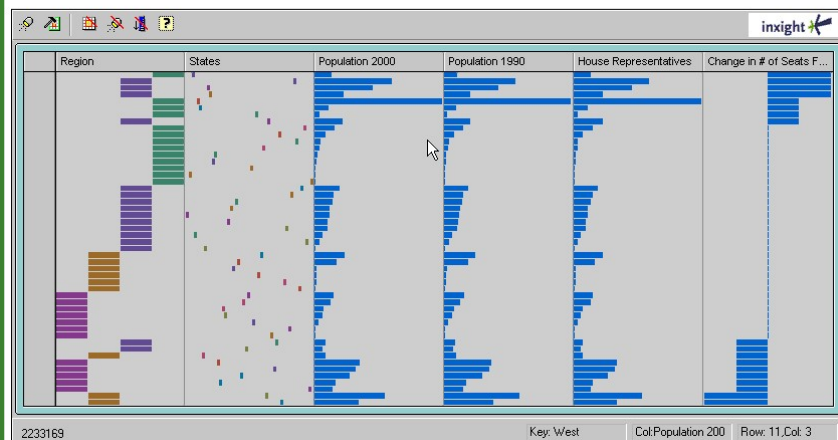
Exemplos de Produtos Comerciais e Recursos

- plotly
- D3
- R, Python
- Tableau – www.tableau.com

33

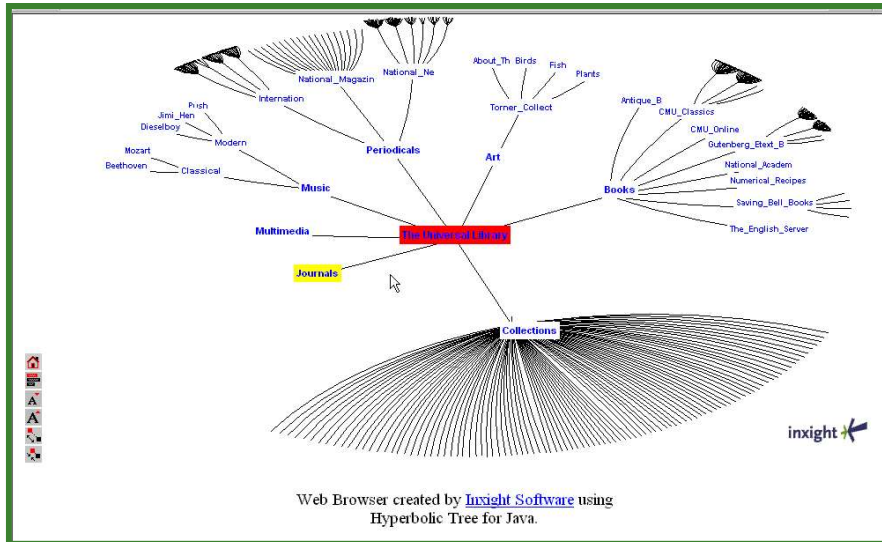
This data includes the resident population of the 50 States from Census 2000 and from Census 1990. It also includes the changes in the number of House Representatives in each state.

- 1) Which states gained the most seats due to the 2000 Census?
- 2) Which states lost the most seats?
- 3) Which 5 states are the most populated? How about 10 years ago?
- 4) Which region of the country gained seats and which region lost seats?



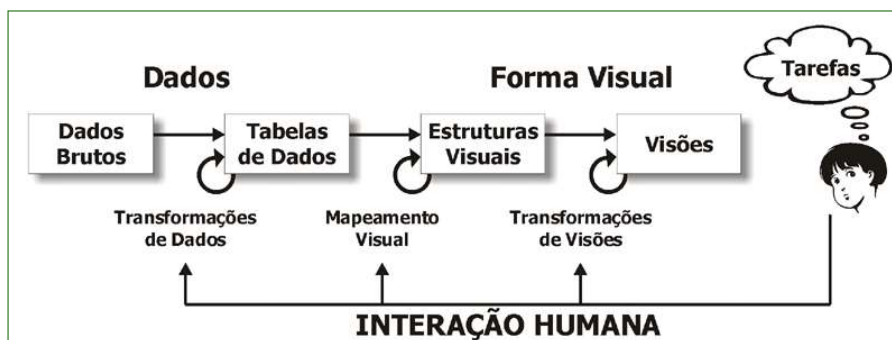
34

Exemplo: *Hyperbolic Tree*



35

Processo de Visualização: Modelo de Referência (Card et al. 1999)



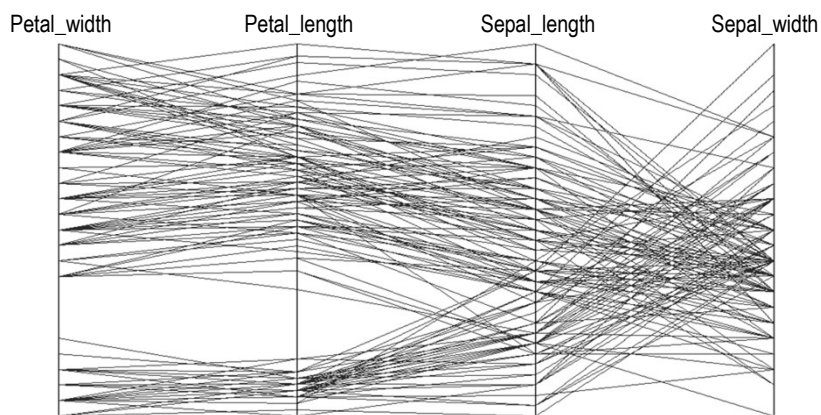
36

Comparação ViSc x InfoVis

- Em ambos os casos, ocorre um processo de 'especialização' da informação: dados são mapeados em um espaço 2D ou 3D
- Especialização determina a representação geométrica visível com a qual o usuário interage
- *ViSc*: geometria do modelo (explícita ou não) **determinada** pelo domínio
- *InfoVis*: geometria do modelo **atribuída** pelo 'designer' da representação visual

37

Exemplo: Coordenadas Paralelas



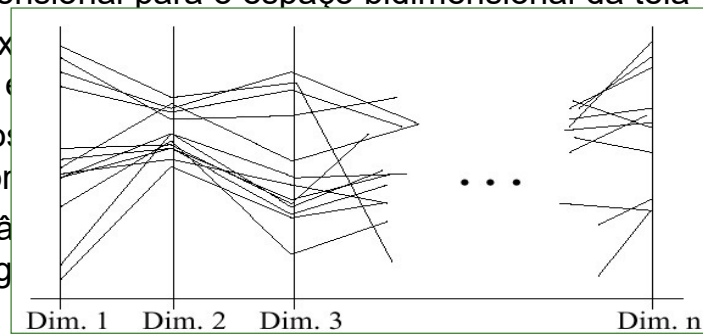
Iris flower data: 3 classes de flores

38

Exemplo: Coordenadas Paralelas

- Inselberg (1985) – geometria computacional
- Projeta instâncias definidas em um espaço n-dimensional para o espaço bidimensional da tela

- n-eixos dos eixos
- Eixos com valores
- Instâncias poligonais

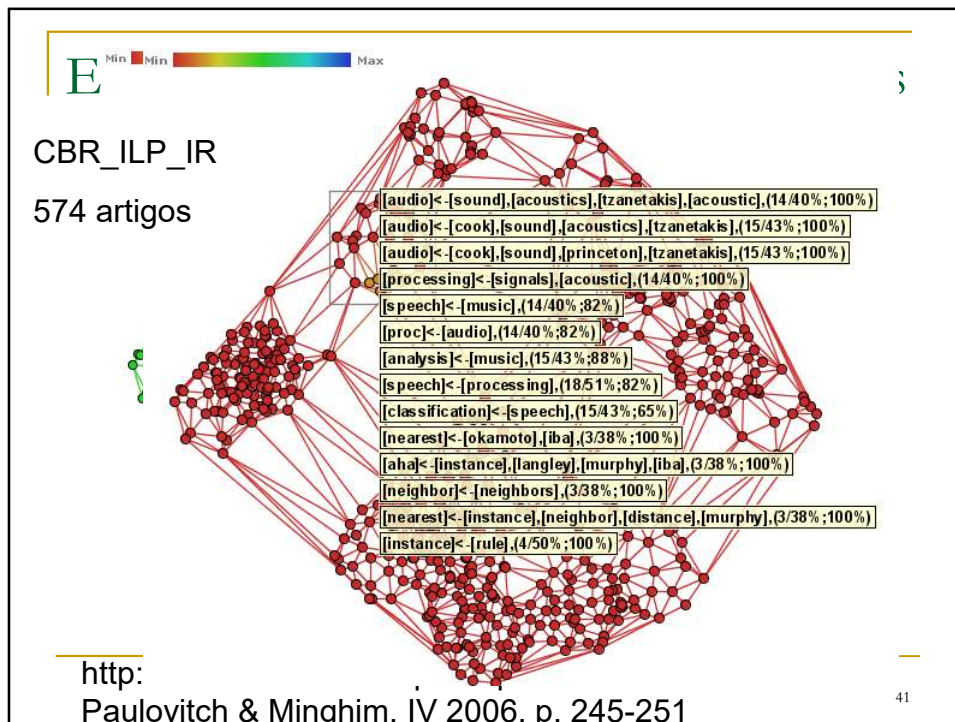


39

Exemplo: *Hyperbolic Tree*

- Demo
- <http://www.inxight.com/products/sdks/st/>

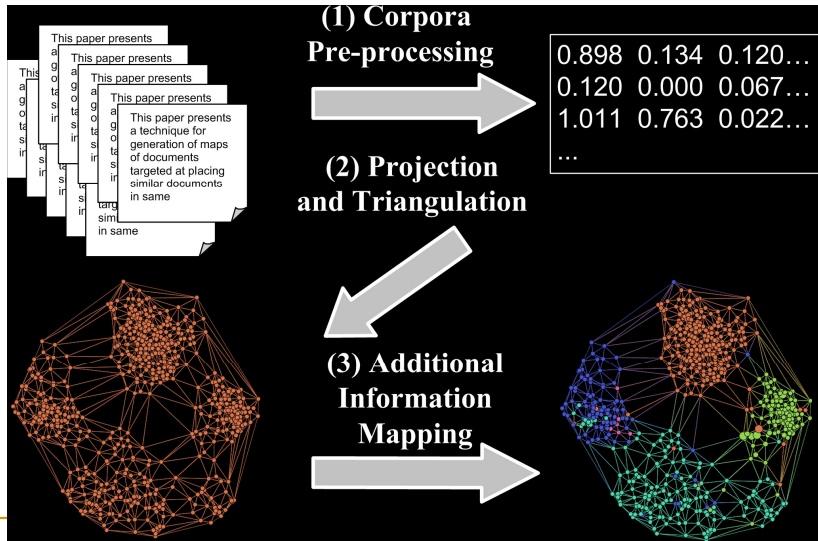
40



Mapas de documentos

- projeção 2D de instâncias de dados definidas em um espaço multidimensional (nD)
- Instâncias são corpus de documentos
 - Cada texto representado por um vetor nD
 - n determinado pelo número de termos na coleção
 - Em geral, n é grande, i.e., dimensionalidade alta
- Pontos próximos são 'similares'
 - Definição de 'similaridade' entre documentos
 - P.ex., medidas de distância entre os vetores nD

Mapas de documentos



Paulovitch & Minghim, IV 2006, p. 245-251

43

The screenshot shows the Projection Explorer Web (PEx-Web) interface. The main window displays a network graph with red nodes and edges, representing the relationships between documents. A search bar at the top contains the word "creative", and the "Number documents" field shows 987. The interface includes a "Retrieved Documents" list on the left, a "Commands" panel at the top right, and a "Nearest Neighbors" section at the bottom left. The URL at the bottom is <http://infoserver.lcad.icmc.usp.br/infovis2/PExWeb>.

44

Mapas de documentos

- **Projeção: redução de dimensionalidade**
 - Pontos próximos no espaço nD devem ser mantidos próximos no espaço 2D
 - Várias maneiras de projetar...
- **Mapas: exploração interativa de coleções de documentos texto**
 - IDMAP, ProjClus, LSP (*Least Squares Projection*)
 - PEX: Projection Explorer, e PEX-Web
 - R. Minghim, F. Paulovitch e R. Pinho

45

Visualização

- **Ferramentas de Análise de Dados**
 - Estatística, Mineração, Visualização
- **Por que Visualização**
 - Habilidade de expressar muita informação
 - Percepção de propriedades não antecipadas
 - Facilita a percepção simultânea de características dos dados em grande e pequena escala
 - Apoio a processos de formação de hipóteses
 - Apoio a tarefas de pré-processamento dos dados
 - Detecção de problemas, limpeza, seleção, ...

46

Visualização - Desafios

- Processos cognitivos difíceis de serem modelados
 - Ainda falta muito para que técnicas sejam usadas de forma efetiva por uma gama ampla de usuários
- Criar "boas" representações visuais não é trivial
- Disponibilizar ferramentas efetivas não é simples (funcionalidades e interação)
- Escalabilidade

47

Visualização - Desafios

- Excesso de dados, de natureza complexa
 - Enormes volumes, dimensionalidade alta (muitos atributos), diferentes tipos de dados, diferentes organizações, diferentes mídias, *streaming data*...
- Como tratar tanto volume e variedade?
- Como saber o que é realmente relevante?

48

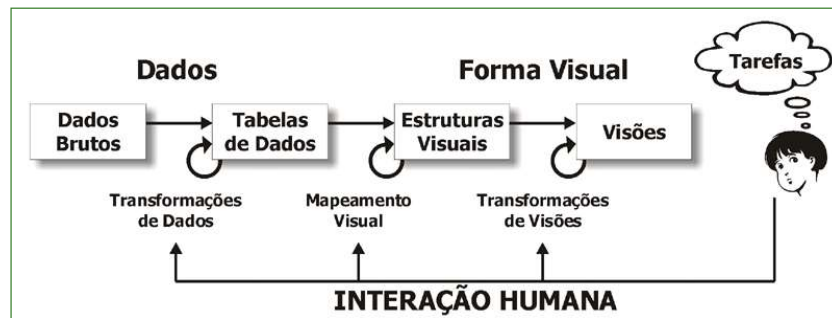
Visualização - Desafios

- Mineração
- Interação

Papel da interação

- Aumentar a escalabilidade visual, dada a limitação física dos dispositivos
 - Limitar a quantidade de informação
 - Manter o contexto global
 - Visão geral e detalhe
 - Filtragem
- Viabilizar a visualização simultânea de diferentes representações
 - Coordenação entre múltiplas visualizações
 - *Link-and-brush*

Papel da interação



51

Papel da interação

- *Visual Information Seeking mantra*
- Overview first, zoom and filter, then details-on-demand [Shneiderman, 1996]

52

Papel da Mineração

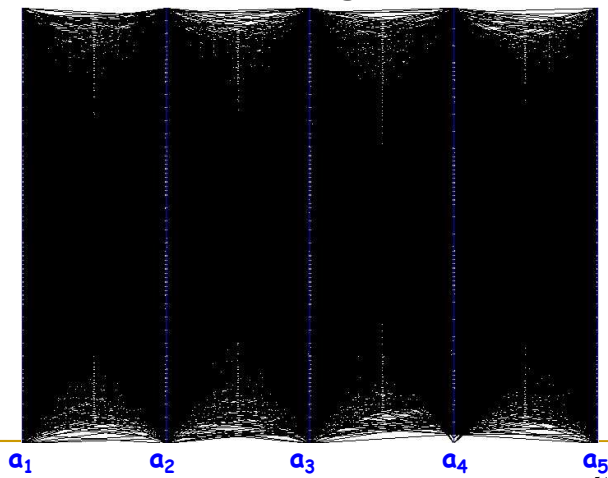
Visualização direta dos dados
vs
Visualização de conteúdo/informação
'relevante' embutido nos dados

53

Problema

Dados sintéticos, 7,500 registros 5-d

Como
analisar???



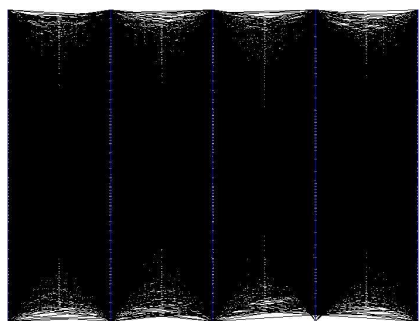
54

Problema

- Instância do problema (mais genérico) de mapear individualmente itens de informação
 - Sobreposição e sobrecarga visual
 - Várias abordagens na literatura buscam identificar e realçar informação relevante em visualizações 'sobrecarregadas'

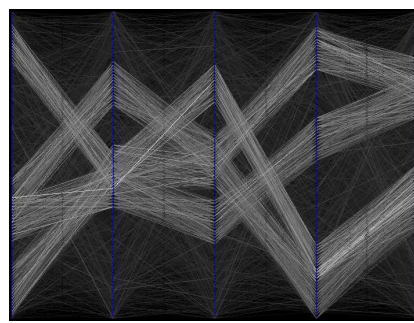
55

IPC Plots



a_1 a_2 a_3 a_4 a_5

O que???



a_1 a_2 a_3 a_4 a_5

Agrupamentos...

56

Mineração Visual

- **Contraposição com Mineração**
 - Etapa do problema (mais geral) de extrair conhecimento de dados (Aprendizado de Máquina)
 - Ponto de partida são arquivos de dados, processados *automaticamente* com o objetivo de extrair modelos que 'explicam' os dados
 - Modelos têm por objetivo descrever ou prever o comportamento dos dados
 - Exemplos: classificar, agrupar, identificar associações, ...

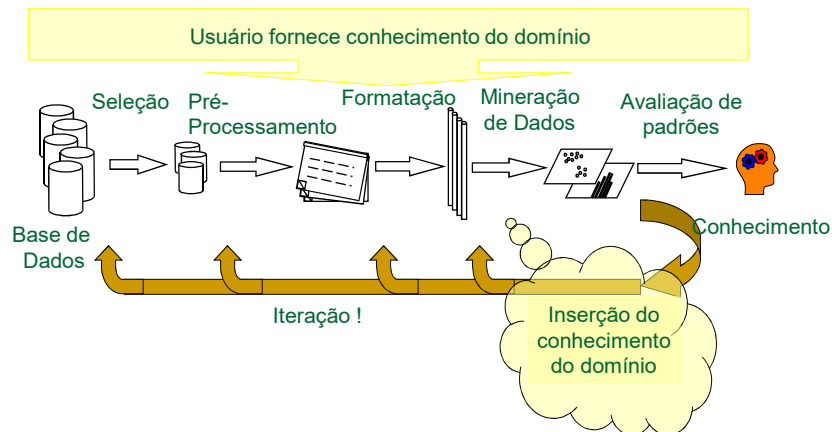
57

Mineração Visual

- **Convergência de Mineração e Visualização**
- **Metáforas visuais para apoiar usuários de tarefas de mineração**
- **Motivação**
 - 'Minerar' dados automaticamente é difícil, e nem sempre produz resultados compreensíveis/úteis
 - Visualizar dados brutos (sem minerar, ou extrair modelos) nem sempre é possível/interessante/útil

58

Processo de Descoberta "Centrado no Usuário"



Ankerst 2000

59

Mineração Visual de Dados – Definição (Ankerst 2000)

- “VDM é um passo no processo de extração de conhecimento (KDD) que utiliza a visualização como um canal de comunicação entre computador e usuário para apoiar a identificação de padrões novos e interpretáveis”
 - Posicionamento nas duas últimas fases do processo: mineração de dados e avaliação
 - Identifica três categorias de VDM

60

Categorias de VDM (Ankerst 2000)

- Visualização dos dados
 - Usuário tem total controle sobre a busca por padrões
 - Focalizar/delimitar espaço de busca
- Visualização dos resultados de uma mineração
 - Apoiar a interpretação dos modelos extraídos
- Visualização acoplada ao processo de mineração
 - Direcionar a busca
 - Fornecer conhecimento sobre o domínio, por exemplo, para adaptar um núcleo genérico (para diferentes aplicações) com a intervenção do usuário
- Mineração acoplada ao processo de visualização?

61

Categorias de VDM (Wong 1999)

- Fracamente acoplada
 - Visualização "intercalada" com estratégias analíticas de mineração
 - Apoiar pré-processamento, interpretação de resultados,...
 - Abordagem limitada: reforça limitações de uma e de outra...
- Fortemente acoplada
 - Visualização "integrada" em estratégias analíticas de mineração
 - Dar ao usuário maior controle e entendimento sobre o processo analítico, apoiando a tomada de decisões
 - Criação de representações visuais do espaço de busca

62

Desafios: *Visual Analytics*

- Science of Analytical Reasoning
- Ampliar a capacidade humana de análise de informação é problema estratégico
 - 2004: National Visualization and Analytics Center, EUA (<http://nvac.pnl.gov/>)
 - *Department of homeland security*
 - Foco: contra-ataque ao terrorismo

63

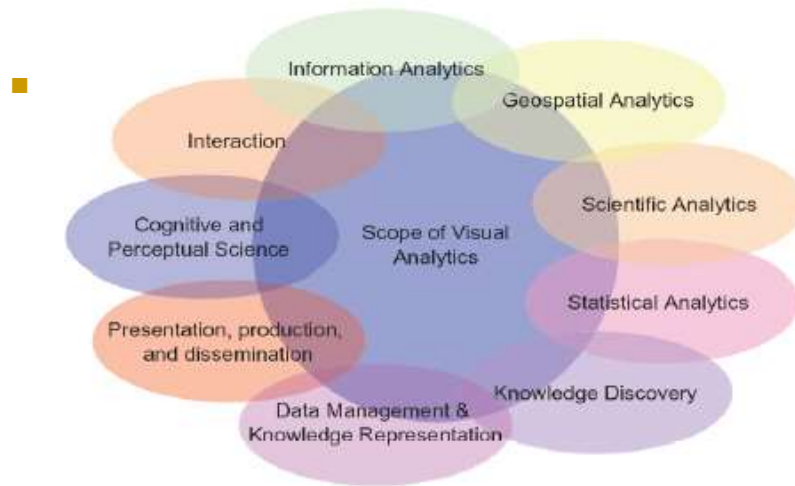
“Visual Analytics is the science of analytical reasoning facilitated by interactive user interfaces”

“Visual analytics requires interdisciplinary science beyond traditional scientific and information visualization to include statistics, mathematics, knowledge representation, management and discovery technologies, cognitive and perceptual sciences, decision sciences”

Illuminating the Path: The Research and Development Agenda for Visual Analytics, IEEE Press, 2005

64

Desafios: *Visual Analytics*



<http://nvac.pnl.gov/>