

Elizabeth Simão Carvalho

Adérito Fernandes Marcos

Visualização da Informação

Relatório Técnico
Área de Computação Gráfica

Visualização da Informação

O tema deste relatório é a visualização da informação. Esta é uma área actualmente muito activa e vital no ensino, na pesquisa e no desenvolvimento tecnológico. A ideia básica é utilizar imagens geradas pelo computador como meio para se obter uma maior compreensão e apreensão da informação que está presente nos dados (geometria) e suas relações (topologia). É um conceito simples, porém poderoso que tem criado imenso impacto em diversas áreas da engenharia e ciência.

O relatório está dividido em duas partes. Na primeira parte, é abordado o problema da visualização exactamente no que diz respeito à subtil correlação existente entre as técnicas (e respectivas metáforas), o utilizador e os dados. Na segunda parte, são analisadas algumas aplicações, projectos, ferramentas e sistemas de Visualização de Informação. Para categorizá-los, serão considerados sete tipos de dados básicos subjacentes a eles: unidimensional, bidimensional, tridimensional, multi-dimensional, temporal, hierárquico, rede e *workspace*.

Information Visualization

The theme of this report is information visualization. Nowadays, this is a very active and vital area of research, teaching and development. The basic idea of using computer generated pictures to gain information and understanding from data and relationships is the key concept behind it. This is an extremely simple, but very important concept which is having a powerful impact on methodology of engineering and science.

This report is consisted of two parts. The first one, is an overview of the subtle correlation between the visual techniques, the user perception and the data. In the second part, several computer applications, tools, projects and information visualization systems are analyzed. In order to categorize them, seven basic types of data are considered: one-dimensional, two- dimensional, three-dimensional, multidimensional, temporal, hierarchic, network and workspace.

Índice

Modelação de Dados e Visualização

1.1. Introdução	1
1.2. Modelação dos Dados	2
1.3. O Utilizador e a Visualização	4
1.3.1. Principais Atributos	6
1.3.2. Análise das Tarefas	8
1.4. O Mapeamento	8
1.5. Representação Visual	10
1.6. Interacção	19
1.7. Conclusões	20

A área de Visualização da Informação

2.1. Introdução	22
2.2. Unidimensional	23
2.3. Bidimensional	26
2.4. Tridimensional	30
2.5. Multi-dimensional	32
2.6. Hierárquicos	35
2.7. Redes	37
2.8. Workspace	39
2.9. Temporal	42
2.10. Frameworks	45
2.11. Conclusões	46

Lista de Figuras

Figura 1.1. Exemplos de visualização	1
Figura 1.2. Mapeamento da realidade para o utilizador	2
Figura 1.3. Codificação gráfica segundo Cleveland e McGill's	5
Figura 1.4. Codificação gráfica considerando factores psico-físicos	5
Figura 1.5. Processo de mapeamento	8
Figura 1.6. Técnica <i>Basket Weave</i>	11
Figura 1.7. Glyphs 3D e 2D	12
Figura 1.8. <i>Focus</i> + contexto e <i>Stretching</i>	13
Figura 1.9. Coordenadas paralelas: 3D, 2D e com extrusão	14
Figura 1.10. <i>Linking Wings</i>	15
Figura 1.11. Cone Tree, Perspective Wall, Outliner, Árvore Hiperbólica e Treemap	16
Figura 1.12. Visualização de <i>clusters</i>	16
Figura 1.13. Textura, neblina e não linear e <i>Level-of-detail</i>	17
Figura 1.14. Starfield	18
Figura 1.15. Rank-by-Feature	19
Figura 1.16 Visualização de 1 milhão de itens	20
Figura 2.1. txtkit	24
Figura 2.2. Virtual Shakespeare	24
Figura 2.3. Sistema de Visualização por padrões sequenciais	25
Figura 2.4. TOPIC ISLANDS™	25
Figura 2.5. ExpO	26
Figura 2.6. ThemeRiver	26
Figura 2.7. PhotoFinder	28
Figura 2.8. GeoVista Studio™	28
Figura 2.9. CounterPoint	29
Figura 2.10. YMap	29
Figura 2.11. WebBook and Web Forager	31
Figura 2.12. VxInsight	32
Figura 2.13. WilmaScope	32
Figura 2.14. HCE	33
Figura 2.15. GlassEye	34
Figura 2.16. Snap Together Visualization	34
Figura 2.17. ADVIZOR/2000 - Perspectiva ancorada	35
Figura 2.18. Treemap 4.1.1	36
Figura 2.19. DBT	36
Figura 2.20. Space Tree	37
Figura 2.21. TaxonTree	37
Figura 2.22. Graphvis	38
Figura 2.23. ILOG Discovery	40
Figura 2.24. Improvise	40
Figura 2.25. Análise da Web log data	41
Figura 2.26. InfoVis Toolkit	41
Figura 2.27. SpotFire	42
Figura 2.28. Timesearcher	43
Figura 2.29. Discussão de e-mails	43
Figura 2.30. FishCal	44
Figura 2.31. Lifelines em casos legais	44
Figura 2.32. Vizeit, construído a partir da UVP	45
Figura 2.33. Piccolo	46
Figura 2.34. A periodic Table of Visualization Methods	47

Lista de Tabelas

Modelação de Dados e Visualização

Tabela 1.1. Prioridades dos atributos visuais versus tipo de dado.....	3
Tabela 1.2. Características principais de algumas técnicas de visualização.....	10
Tabela 1.3. Avaliação de técnicas segundo Schumman.....	11
Tabela 1.4. Avaliação de técnicas segundo Steiger	12
Tabela 1.5. Mapeamento de atributos segundo os modelos RGB e HLS de cor	12

Modelação de Dados e Visualização

1.1. Introdução

O objectivo básico e principal subjacente à qualquer visualização é sempre o mesmo: a partir de um grande volume de dados extrair o máximo de informação de uma forma rápida, clara e precisa. Para isso, são utilizadas metáforas visuais adequadas. Várias são as áreas do conhecimento que necessitam dela: medicina, física, geologia, contabilidade, economia, administração, etc. (figura 1.1). Por outro lado, a tecnologia subjacente é vital para o seu futuro (por exemplo, as próximas gerações de — PCs, sistemas de armazenamento, dispositivos de visualização, sistemas de comunicação, ferramentas analíticas) bem como um melhor conhecimento das características psicológicas, físicas e cognitivas humanas.

Muitas são as técnicas de visualização existentes [68]. Cada uma delas, permite analisar os dados com metáforas visuais diferentes. Porém, ao aplicar-se uma determinada técnica de visualização, diversos factores devem ser considerados. A classificação dos dados e das tarefas, a característica cognitiva do potencial utilizador ou o tipo de ênfase visual desejada ao resultado final (conteúdo versus quantidade), são alguns exemplos desses factores. A questão a responder é sempre a mesma: Como conseguir uma visualização efectiva?

Este capítulo aborda o problema da visualização exactamente no que diz respeito à subtil correlação existente entre as técnicas (e respectivas metáforas), o utilizador e os dados. Sem ser abrangente ou exaustivo, este capítulo fornece uma panorâmica do actual estado da modelação da visualização.

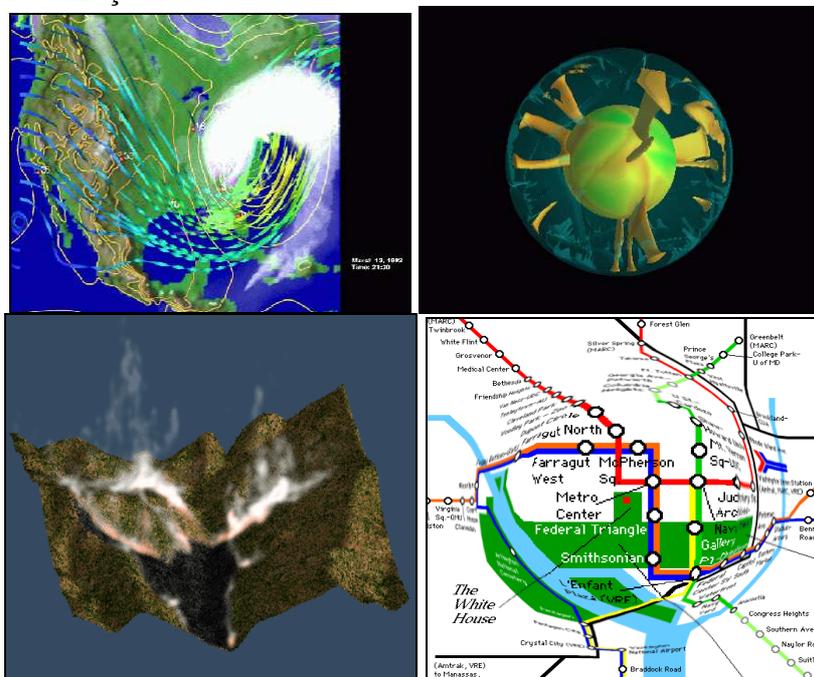


Figura 1.1. Exemplos de visualização: Tufão [66], sistema de convecção do magma da Terra, incêndio [67] e linhas de metro [69]

1.2. Modelação dos Dados

Os dados podem ser analisados de diversas formas. Podemos de uma forma geral os caracterizar segundo a sua: natureza (hidráulico, geológico, contabilístico, linguístico, etc.), dimensão de seu domínio, distribuição nesse domínio, característica temporal (estática ou dinâmica), forma como estão organizados, estado físico (gasoso, líquido ou sólido), localização geográfica, entre outros. A sua caracterização expressa uma interpretação conceptual dele segundo diferentes perspectivas.



Figura 1.2. Mapeamento da realidade para o utilizador

Existe actualmente uma grande variedade de técnicas de visualização utilizadas para mapear a informação (figura 1.2), que por sua vez é constituída por um conjunto de dados. Um factor prioritário utilizado para a escolha das técnicas, é a identificação do modelo de dados. O modelo deve “traduzir” de forma conveniente a relação existente (ou a que se pretende que exista) entre o mundo real e os respectivos dados.

A solução adoptada na modelação científica é [70] sempre a de derivar com rigor um modelo e de executar aproximações até este ser processável por computador. Um outro detalhe, é a importância de enfatizar a informação de interesse contida nos dados é eliminar os eventuais ruídos (filtrar, limpar, normalizar, etc.). Por último, o modelo de dados deve apresentar características que expressem o grau de confiança destes, a sua amplitude de valores (média, mínimo e máximo) e a descrição do momento temporal (vários são os significados que o tempo pode possuir – tempo de gravação, tempo de simulação, tempo de computação, etc.).

Nominal, ordinal, quantitativo, ponto, escalar e vector

Os dados podem conter informação numérica que discrimina quantidades ou uma ordem sequencial, ou ainda algum tipo de “rótulo” (*label*) que identifica uma classe ou conjunto. Por outro lado, cada dado pode ser considerado como sendo uma posição espacial num espaço de dimensão n (ponto), ou possuir um valor numérico (escalar). Os escalares também podem constituir conjuntos discretos de funções contínuas com 1 (linear), 2 (imagem), 3 (volumétrica) ou mais dimensões. Num vector, cada elemento é considerado uma recta com um determinado comprimento, num espaço de dimensão n .

De uma forma geral, [71] algumas prioridades visuais podem ser observadas consoante o tipo de dado. A tabela 1.1 apresenta algumas das situações existentes, estando a prioridade ordenada de forma crescente (1ª linha apresenta a maior, enquanto que a última a menor).

Quantitativo	Ordinal	Nominal
Posição	Posição	Posição
Comprimento	Densidade	Matiz
Ângulo	Saturação	Densidade
Inclinação	Matiz	Saturação
Área	Comprimento	Forma
Densidade	Ângulo	Comprimento
Saturação	Inclinação	Ângulo
Matiz	Área	Inclinação
Forma	Área	Área

Tabela 1.1. Prioridades dos atributos visuais versus tipo de dado (nominal, ordinal ou quantitativo)

Domínio e Distribuição

Considere-se que um modelo empírico tenha sido criado a partir de um conjunto de dados, de forma a definir uma “entidade”, a qual se deseja visualizar. Esta entidade terá um determinado intervalo de valores. Os dados contínuos são dados que são representados por uma entidade que é descrita como uma função ($F(x)$) ou conjunto de funções, compostas por variáveis independentes e dependentes (parâmetros). A função F pode ser uma dos muitos tipos diferentes existentes: escalar, vectorial ($F = (f_1, f_2, \dots, f_k)$), ou até mesmo tensorial de segunda ordem ($F = (f_{jk})$), ou mais [72]. Podem estar distribuídos de uma forma regular ou irregular. A visualização de dados contínuos é em geral feita segundo o proposto também na tabela 1.1.

Os dados descontínuos, ao contrário dos anteriores, não possuem uma função que possa representá-los. A sua topologia/estrutura é utilizada normalmente para analisá-los – sequencial (texto), hierárquico, relacional, etc.

Segundo Nielsen [75], a dimensão do domínio ao qual a entidade pertence, permite fazer ainda mais algumas distinções:

- I. A entidade é definida pontualmente ao longo de um domínio contínuo – por exemplo, num mapa a altitude é definida separadamente para cada ponto.
- II. A entidade é definida ao longo de regiões existentes no domínio contínuo – por exemplo, um mapa que mostre as densidades populacionais por Concelho.
- III. A entidade é definida como um grupo de dados enumerados – por exemplo, um gráfico que mostre o total de carros vendidos por ano.

Lattices e Malhas

Os *lattices* [67] são utilizados para descrever diferentes arranjos de dados contínuos. A dimensão de um *lattice* indica a ordem dos dados: zero é desordenado, uma dimensão indica um vector de dados, e duas, uma matriz de dados. Esta classificação é mais orientada para os dados e menos para a entidade matemática subjacente.

As malhas [79] descrevem a geometria de um conjunto de dados contínuos. Por exemplo, um conjunto de dados topológicos [84] dispostos numa grelha, consistem em pontos

distribuídos uniformemente (com distâncias uniformes). Uma malha pode ser estruturada e regular, ou não.

Metadados

Um outro aspecto importante que define [82] um conjunto de dados é a sua informação documentada, isto é, o seu metadado. O metadado é uma informação suplementar que descreve a estrutura, a escala, a orientação, o tipo, o momento temporal, ou qualquer outra coisa, relacionada com a história e propriedades de um conjunto particular de dados.

1.3. O Utilizador e a Visualização

O próprio utilizador é um elemento fundamental no processo de visualização [99]. Algumas questões devem estar sempre presentes no desenvolvimento de um sistema de visualização:

- ✓ Quais características permitem (ou não) que o utilizador interprete correctamente e da forma desejada uma visualização?
- ✓ Como é possível ter a certeza de que a visualização está sendo correctamente interpretada?

É necessário o conhecimento de como funciona a percepção humana a nível biológico, psicológico e cognitivo, e de teorias da percepção [95]. Por outro lado, é necessário distinguir entre características gerais e [87] individuais – cor, ilusão óptica, modelo cognitivo visual do utilizador, capacidade motora, etc. A validação da correcta visualização implica em ter informação acerca dos objectivos, hábitos e expectativas visuais, do nível cultural e habilidades do utilizador. A aplicação de testes para avaliação estatísticos podem auxiliar muito nesta tarefa.

Um outro detalhe, é que outros sentidos humanos também podem contribuir para a compreensão de uma visualização: som, olfacto, pressão, etc. Os atributos visuais, por sua vez, influenciam a compreensão consoante: cor, tamanho, orientação, brilho, quantidade, movimento e forma [98]. A interpretação destes pode ser resumida a:

- *Reacção inata aos atributos visuais* – natural e usualmente simples, inconsciente.
- *Reacção adquirida aos atributos visuais* – através de treino (educação por ex.), normalmente mais complexa.
- *Atributos visuais ilusórios* – ilusões visuais que se encontram documentadas, porém, pouco compreendidas.

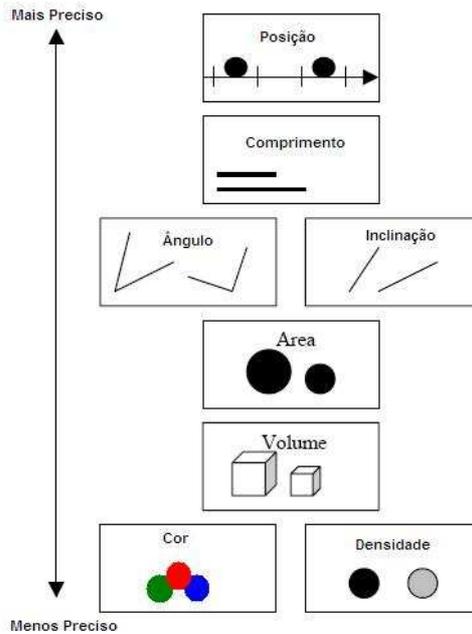


Figura 1.3. Codificação gráfica segundo Cleveland e McGill's

Embora a codificação gráfica [37] de valores segundo o esquema na figura 1.3, proposta por Cleveland e McGill's, auxilia a comparação quantitativa de informação, o mesmo não é possível no caso de informação não quantitativa, pois esta exige outro grau de percepção. Por exemplo, a textura não está referida no esquema acima. Como consequência disso, ela foi aprimorada, de forma a levar em conta resultados de análise aos factores psico-físicos e tarefas perceptivas. A figura 1.4 mostra o resultado obtido.

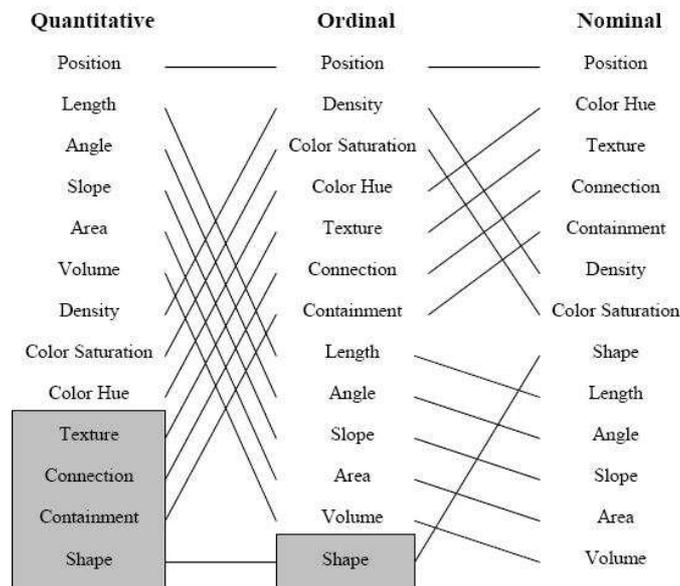


Figura 1.4. Codificação gráfica considerando factores psico-físicos

A maior parte das ferramentas falha em dar suporte na visualização dos dados que estejam em falta, sendo esta uma informação valiosa para o utilizador. Cyntrica Eaton e al. [25] identificaram causas de possíveis faltas, e categorizaram as técnicas de visualização tendo como base o impacto que os dados em faltam causam na região dependente na área de

visualização. Como conclusão final, é dada ênfase a necessidade dos sistemas e/ou aplicações de visualização indicarem com clareza a existência de falta de informação, além de sugerir alguma forma de resolve-lo.

1.3.1. Principais Atributos

Cor

Não existe nenhuma teoria completa e é descrita geometricamente em função da gama de cores [111]. A sua dimensão em termos de percepção é caracterizada pelo seu matiz, saturação e intensidade (brilho) [100]. Estes três componentes podem variar de forma independente ou não. Em termos gerais se tem que:

- ❑ **Matiz** – a sua utilização é efectiva para dados do tipo ordinal e nominal (escala de cores) [105]. Deve-se ter os seguintes cuidados na sua utilização:
 - ☞ Azul (cores “frias”) – para valores afastados do ponto de observação, pequenos, negativos, ou associados com temperaturas baixas.
 - ☞ Vermelho (cores “quentes”) – para valores próximos, altos, positivos, associados a temperaturas altas ou algo de perigoso.
 - ☞ A forma do objecto representado segundo uma escala similar a do arco-íris pode não ser prontamente visível.
 - ☞ O matiz pode apresentar uma aparência diferente de acordo com a cor de fundo utilizada.
 - ☞ Descontinuidade nas escalas de cor.
 - ☞ Efeito de “*color-blindness*”.
 - ☞ Pequenos objectos azuis: desvantajoso aquando da utilização de fontes de luz pontuais com comprimento de onda curta.

- ❑ **Saturação** – a sua utilização é efectiva para dados do tipo ordinal. Deve-se ter os seguintes cuidados na sua utilização:
 - ☞ Quando se deseja interpretar o brilho e a saturação independentemente.
 - ☞ Deve se utilizar escala de cor 2D para uma maior efectividade.

- ❑ **Brilho** – A sua utilização é efectiva para dados do tipo ordinal e quantitativo [74]. Deve-se ter os seguintes cuidados na sua utilização:
 - ☞ Os objectos brilhantes num fundo escuro parecem ser maiores do que os objectos escuros num fundo claro.
 - ☞ A utilização de brilho gradiente fornece uma melhor interpretação da distância ou profundidade.
 - ☞ Alterações de brilho ou a utilização de um brilho total não são apreendidos linearmente.
 - ☞ O contraste do brilho influencia a percepção do brilho.
 - ☞ Devem se utilizar escala de cor 2D para uma maior efectividade.

Textura

A sua utilização é efectiva para dados do tipo nominal [91]. Deve-se ter os seguintes cuidados na sua utilização:

- ☞ Não sobrepor texturas.
- ☞ As texturas podem dar origem a outras impressões, por exemplo, densidade.
- ☞ Incluir legendas.

Orientação

Deve-se ter os seguintes cuidados na sua utilização [100]:

- ☞ A simetria deve preferivelmente ser mantida em relação ao eixo vertical.
- ☞ A familiaridade da forma está frequentemente associada a orientação.
- ☞ Devem ser utilizadas várias orientações de forma a garantir uma correcta visualização dos objectos.

Movimento

Devem ser mostrados numa velocidade mínima de 10 quadros/segundo.

- ☞ Piscar duas ou mais imagens para realçar diferenças ou similaridades.

Profundidade

Deve ser utilizada para enfatizar a existência e melhorar a percepção de estruturas 3D [114]. Características:

- ☞ A aplicação de brilho gradiente auxilia a percepção da profundidade.
- ☞ A utilização de projecção perspectiva.
- ☞ A oclusão distingue o que é frontal e o que é posterior.
- ☞ A transparência ou opacidade também distingue o que está a frente ou atrás.
- ☞ Utilização de sombra para simular superfícies.
- ☞ A rotação auxilia a percepção 3D.
- ☞ Efeito *stereo*: realidade virtual [125], *shutter glasses*, [115], etc.

Por fim, o contexto visual onde as técnicas são inseridas, também é importante [108]. As cores devem ser utilizadas segundo os seus significados mais padronizados (de acordo com os hábitos do utilizador), o mesmo ocorre com a utilização de símbolos (*glyphs*) [120]. A anotação é outro detalhe que também auxilia a interpretação de atributos visuais:

- ☐ Rótulos, títulos, legendas.
- ☐ Escala de cor e brilho.
- ☐ Escalas de distâncias que relacionam as coordenadas do mundo real com as existentes no ecrã.
- ☐ Símbolos de orientação (por ex. norte/sul, leste/oeste).
- ☐ Animações com indicação temporais: duração, momento, etc.

1.3.2. Análise das Tarefas

A aplicação pode ser dependente ou não da classificação dos objectivos da visualização [111]. A visualização tem como tarefa básica traduzir/representar correctamente o enfoque que o utilizador dá a um determinado domínio de interesse durante a interpretação de uma imagem. Para que a visualização de uma tarefa seja efectiva, é necessário que a visualização seja devidamente personalizada de acordo com a natureza desta última.

As tarefas são frequentemente determinadas de uma maneira informal ou implícita e são subdivididas em tarefas de alto, médio ou nível atómico [121]. Uma vantagem é que uma única representação pode servir para uma tarefa inteira, pertencente a um determinado nível. Um outro aspecto, é a avaliação do progresso do processo de mapeamento da tarefa *versus* visualização proposta. Este deve sempre ter como base a utilização de procedimentos específicos, que permitam a sua correcta avaliação [104].

Com base nas técnicas visuais que já foram desenvolvidas e propostas, pode-se afirmar que existe um conjunto de tarefas de alto nível que estão sempre presentes, independentemente do tipo de dado em análise, sendo elas [10]:

- Vista geral (*Overview task*) – vista geral de toda a informação disponível.
- Zoom – foco especial num determinado conjunto de informação.
- Filtragem – execução de pesquisas dinâmicas.
- Detalhes consoante solicitação (*Details-on-demand task*) - detalhes são fornecidos conforme solicitação do utilizador.
- Relação (*Relate task*) – identificação de relações entre itens
- Historial (*History task*) – memória das acções executadas.
- Extracção e Relatório (*Extract and Report task*) – gravação de subconjuntos dos dados ou de determinadas visualizações.

1.4. O Mapeamento

Para a exploração dos dados é necessária a existência e utilização sistemática de metodologias, sistemas inteligentes de visualização e conceitos [118]. A figura 1.5 mostra a ideia conceptual subjacente ao mapeamento da informação na visualização.



Figura 1.5. Processo de mapeamento

A eliminação de qualquer limitação é importante, isto é, não se deve padronizar as técnicas de visualização em função da origem dos dados (medicina, satélite, mecânica, etc.), dos formatos e estruturas requeridas por eles (por exemplo, seleccionar uma técnica porque esta requer os mesmos formatos e estruturas que os dados), ou da dimensão que os dados

possuam (por exemplo, é possível dar o mesmo tipo de tratamento e representação a dados de um campo escalar 2D e dados de uma superfície 3D [105]).

O método básico para conseguir escolher a representação ideal é abstrair as propriedades chaves do comportamento dos objectos físicos [111]. Edward Tufte [108] propõe alguns princípios para visualizar a informação:

- ❑ Mostrar os dados — induzir o utilizador a pensar em relação ao conteúdo mais do que na metodologia, o desenho gráfico, a tecnologia, ou qualquer outra coisa;
- ❑ Evitar distorcer o que os dados devem transmitir — não apresentar muitos números em espaços pequenos;
- ❑ Revelar os dados em vários níveis de detalhe, a partir de uma panorâmica geral para uma mais detalhada;
- ❑ Servir a um propósito claro e aceitável — descrição, exploração, tabelas, ou decoração;
- ❑ Estar intimamente integrado com as descrições estatísticas e verbais de um grupo de dados.

O mapeamento não é tarefa fácil, pois as representações visuais geradas pelo computador nem sempre são claras. As aproximações e artifícios utilizados, sempre geram efeitos indesejáveis. Vários são os modelos que têm sido propostos para definir as estratégias de mapeamento, todos apresentam entretanto, prós e contras. Eles podem ser divididos nas seguintes categorias:

- ❑ *Blind Matching Procedure* – os dados são separados em elementos. É executado o cruzamento de elementos de dados *versus* primitivas visuais, por ex. 4 variáveis e 4 atributos visuais disponíveis. O número total de combinações possíveis de se obter entre os atributos é definido por: $(n>m): n!/(n-m)!$, onde n é o nº. de características existentes nos dados e m o nº. de atributos visuais.
- ❑ *Renaissance teams* – propõe um grupo formado por especialistas em visualização e na área destino da aplicação. Apresenta um custo elevado.
- ❑ *Bottom-up* – as imagens são geradas a partir das primitivas visuais ou uma linguagem gráfica [77]. Depois as imagens são verificadas com um conjunto de condições (características dos dados, ambiente computacional, utilizador, realidade, efectividade e expressividade).
- ❑ *Generate-and-test* – Consiste em testar as representações visuais que tenham sido geradas a partir do modelo anterior.
- ❑ *Top-down* – Permite somente técnicas complexas e coerentes [78]. Escolhe a técnica que melhor se encaixe.
- ❑ *Code based reasoning* – é um procedimento que escolhe as representações visuais a partir de um conjunto de representações (consideradas) bem sucedidas. É utilizado normalmente, após a aplicação do modelo anterior (*Top-down*).

Muitas ferramentas computacionais têm sido desenvolvidas, com o objectivo de executar o mapeamento de forma “automática”. Elas apresentam um conjunto de estratégias sistemáticas que analisam os mais diversos factores que podem intervir na escolha da representação gráfica: domínio, tipo de dados, dimensão, semântica, etc. [76]. Um bom exemplo é dado pelo **VISTA** – *VISualization Tool Assistant* [77] ou o modelo proposto por Robertson [78].

1.5. Representação Visual

Uma técnica de visualização pode ser simples ou composta. No segundo caso, a técnica é, na realidade, composta por um conjunto de técnicas, que devidamente combinadas, resultam numa nova. Vários foram os estudos efectuados de forma a identificar melhor a relação técnica *versus* dados. Os dados considerados pontuais, são visualizados normalmente através de gráficos do tipo *scatter plot* de 1 até n dimensões.

Técnica	Dados	Efectividade
Gráficos de linha	Escalares 1D contínuos	☞ Não é necessária uma interpolação “mental”
<i>Scatter Plots</i>	Multi variado de dimensão n	☞ A posição é a primeira impressão ☞ Animação para um efeito 3D ☞ Dimensão >2 , utiliza projecções
Glifos e ícones	Multi variado	☞ Vários atributos visuais associados a um único glifo ☞ Distinção entre interpretação macroscópica/microscópica dos glifos
Curvas de nível	Escalares 2D	☞ Anotação de curvas seleccionadas
Superfície		☞ A cor ou nível de cinzento devem ser significativos para o utilizador
Imagem		☞ Necessidade de escalas de cor e brilho
Cor	Até 3 conjuntos de escalares definidos em relação ao mesmo plano	☞ Utilizar o modelo RGB para os conjuntos de dados de mesmo tipo ☞ Utilizar os modelos HSV e HLS quando forem tipos de dados diferentes ☞ Efectivo para correlações/associações entre os dados.
Volumes cortados	Escalares 3D	☞ Utilização de um sistema de coordenadas apropriado de forma a identificar as posições dos planos de corte ☞ Animação, superfícies escondidas e geometria perspectiva para um efeito 3D
<i>Basket Weave</i>		☞ Utilização de um sistema de coordenadas apropriado de forma a identificar as posições dos planos de corte ☞ Superfícies escondidas e geometria perspectiva para um efeito 3D
<i>Render de Superfícies</i>		☞ Expansão 3D das superfícies – iluminação, sombra, projecção
Visualização volumétrica		☞ Transparência/opacidade
<i>Metaballs</i>		☞ Função implícita e valor limite
<i>Cuberilles</i> ou esferas	Escalares 3D discreto	☞ O espaço aberto entre objectos permite uma visualização mais penetrante
Setas	Campos de vectores	☞ Problemas em 3D devido a ambiguidade em relação a direcção ☞ Evitar excesso de dados de 1 só vez
Fitas		☞ Necessidade de <i>steering</i> interactivo ☞ Necessidade de ferramentas e métodos de interacção especiais
<i>Cone Tree</i>	Conteúdo de uma base de dados hierárquica	☞ Os cones devem ser girados de forma a mostrar informações
Diagramas de Fluxo	Módulos de programas e suas relações	☞ Geração interactiva de uma rede de programas
<i>Ball-and-Stick</i>	Estruturas moleculares	☞ As sombras e reflexos melhoram o efeito 3D
Falta de Dados	Indica a diferença entre os dados reais e os auferidos	☞ Utilização de cor para os dados medidos e brilho para os dados em falta

Tabela 1.2. Características principais de algumas técnicas de visualização

Os dados escalares de 1 dimensão utilizam também gráficos, porém estes são do tipo barra, histograma, *pizza* ou diagramas xy . Os dados escalares 2D são visualizados com técnicas do tipo curva de nível ou superfícies 2D. No caso 3D, as técnicas englobam o desenho de superfícies isométricas [130], volumes [86], *cuberilles* [93], ou a aplicação de

algoritmos que desenham curvas de nível como (figura 1.6) fitas largas (técnica designada de *Basket Weave*). Os dados vectoriais 2D são normalmente representados através de vectores, linhas de fluxo, ou com a técnica designada de *bump shading*. No caso 3D são utilizadas representações do tipo vectores num plano ou vectores num volume [122], *hedgehogs* [88] ou ainda fitas [94]. Muitas delas são adaptações para 3D das técnicas utilizadas para 2D. Com a variação de atributos, vários resultados visuais são possíveis. A tabela 1.2 resume as principais técnicas e suas características.

Tratando-se de dados multi-dimensionais (quando a dimensão do dado é superior a 3) são utilizadas técnicas mais sofisticadas [80]. A secção seguinte discute algumas das técnicas mais utilizadas para solucionar a visualização deste tipo de dados, e não só [81].

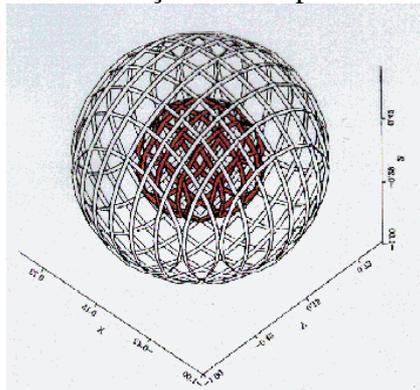


Figura 1.6. Técnica *Basket Weave*

Dados Multi-dimensionais

Os dados podem ser constituídos por colecções de amostras. Cada amostra pode ser um item de dimensão n [119]. O conjunto destas amostras constitui um conjunto de dados de elevada dimensão. A visualização deste tipo de dados pode tornar-se por vezes confusa e complexa. Cinco grupos de técnicas de visualização diferentes são utilizados basicamente para dados com esta característica:

- Mapeamento de atributos,
- Codificação geométrica,
- Som,
- Redução de dimensão e,
- *Clustering* [73].

	Mapeamento de Atributos	Glifos	Coordenadas Paralelas	Redução da dimensão
Dimensão	$n \leq 3$	$n > 3$	$n > 3$	$n \geq 3$
Identificação de parâmetros individuais	Só para 1 variável	Só para alguns glifos	Boa para pouca quantidade de dados	Depende do método de visualização
Descontinuidade	Boa		Boa	Boa
Correlação	Possível	Possível	Possível	Possível

Tabela 1.3. Avaliação de técnicas segundo Schumman [110]

Elas podem ser utilizadas em conjunto ou separadamente. (resultados de análises feitas por Schumman [110] e Steiger [126] a algumas técnicas podem ser vistas nas tabelas 1.3 e 1.4).

Técnica	O melhor	Tendências dos dados	Correlações entre dados	O pior
Scatter Plots	Grande nº de amostras	Mostra claramente	Entre planos 2D	Pouca qualidade com dimensões maiores que 8
Paralela	Foca bem correlações	Não é ideal	As gerais não é ideal	Reduz qualidade conforme nº de amostras aumenta
Glifos	É possível encontrar amostras com representação simbólica aplicável ao <i>glifo</i>	boa	boa	O tamanho do écran é problema e deixa de ser utilizável com elevado nº de amostras e dimensões

Tabela 1.4. Avaliação de técnicas segundo Steiger [126]

Mapeamento dos Atributos

Este é um dos métodos mais comuns. Ele utiliza *lattices* de uma ou duas dimensões para definir primitivas geométricas simples, como por exemplo, planos ou linhas de contorno. Os atributos destas primitivas geométricas podem ser utilizados para visualizar as variáveis restantes.

RGB	HLS
Variável 1 → Vermelho	Variável 1 → Matiz
Variável 2 → Verde	Variável 2 → Saturação
Variável 3 → Azul	Variável 3 → Brilho

Tabela 1.5. Mapeamento de atributos segundo os modelos RGB e HLS de cor

A cor é um [83] dos atributos mais utilizados, apesar de apresentar alguns inconvenientes: só pode codificar 3 variáveis, algumas pessoas têm problemas de percepção a cor, e exigência de legendas. Além disso, a cor implica considerações biológicas, psicológicas e técnicas [85]. Dependendo do modelo de cor utilizado, o efeito visual difere (tabela 1.5).

Codificação Geométrica

Consiste em mostrar dados de elevada dimensão num *lattice* de baixa dimensão. Diversos objectos distintos são mostrados dentro do *lattice* através do mapeamento dos dados de elevada dimensão para alguma característica geométrica ou atributo destes objectos. Os mais comuns são: *Glyphs* [112] (figura 1.7), ícones [116], Faces [110] *Chernoff*, *Data Jacks* e *m-Arm Glyph* [113].

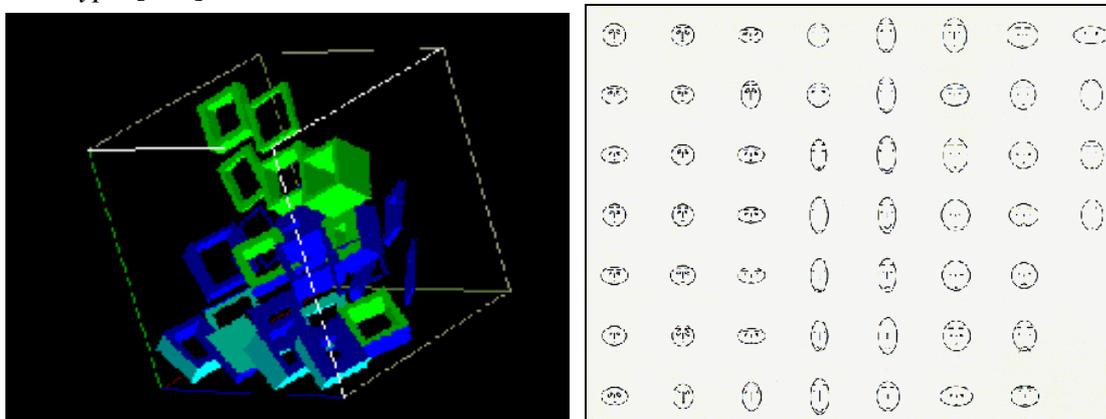


Figura 1.7. Glyphs 3D e 2D

Som

A utilização de som permite não sobrecarregar a visão. Por outro lado, ele pode ser produzido consoante os parâmetros que são mapeados para um determinado dado [107]. O tom, a intensidade sonora, ou o timbre são características sonoras que podem facilmente variar e expressar o comportamento de um grupo de dados.

Redução de Dimensão

Uma outra forma de visualizar dados com elevada dimensão é através da redução desta. Com técnicas que permitem variar o foco de atenção na informação visualizada, se consegue “reduzir” a dimensão real existente.

Para alterar o foco de atenção, técnicas [73] que permitem esticar ou distorcer o espaço visual, são trivialmente utilizadas. Alguns exemplos (figura 1.8) são:

- Visualização olho de peixe (*fisheye*) [17],
- Foco + contexto [69],
- Esticar uma folha elástica (*stretching rubber sheet*) e,
- Distorção orientada de visores (*oriented distorted displays*).

O termo ampliação não linear foi introduzido para descrever o efeito em comum que estas aproximações causam. A propriedade básica da ampliação não linear é a não oclusão da área local a qual pretende focar e a de preservar uma visualização geral do contexto global presente na cena [73]. Tradicionalmente os sistemas não lineares de ampliação são definidos com a identificação de um centro de foco. Este centro de foco pode ser um ponto, uma linha, ou uma região. A área próxima ao centro é ampliada, enquanto que as demais áreas (designadas de contexto) são reduzidas ou desfocadas.

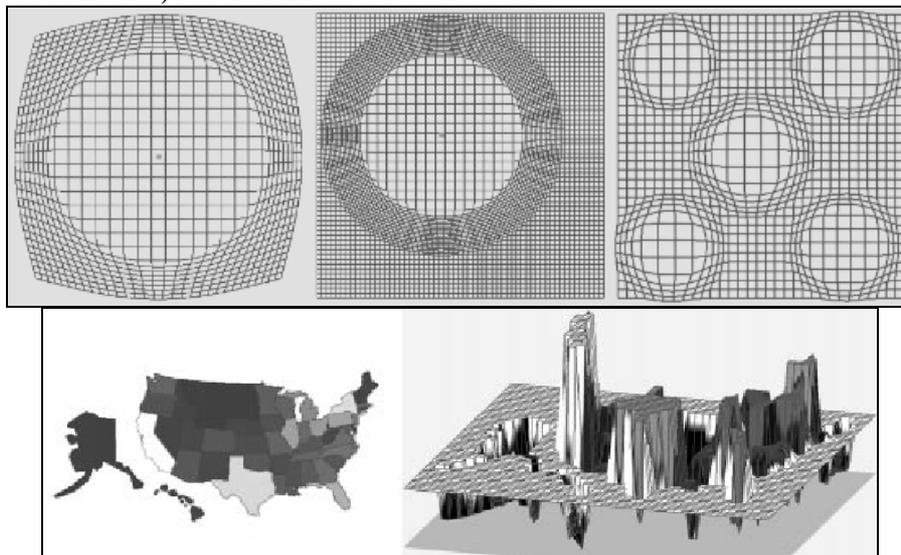


Figura 1.8. *Focus + contexto e Stretching*

A literatura sobre visualização da informação, de uma forma geral, estabelece que as interfaces baseadas em *overview + detail* apresentam um alto grau de usabilidade, porém, no caso de interfaces baseadas em zoom, os resultados ainda estão divididos. Kasper Hornbaeck e al. [26] compararam estes dois paradigmas utilizados em interfaces e concluíram que o grau de satisfação por parte do utilizador é maior no caso das interfaces *overview + detail*,

entretanto, no caso de tarefas de navegação ou visualização com mais de um nível, as interfaces zoom permitem interações mais rápidas.

Uma consequência da ampliação é que cada visualização só contém informação parcial a respeito dos dados. Isto pode ser compensado através da criação de ligações entre diversas visualizações focadas. A visualização paralela (também designada de geométrica) pode ser vista em janelas separadas ou com o auxílio de técnicas de projecção geométrica ou hierárquica.

Projecção Geométrica

As técnicas de projecção geométrica visam obter interessantes projecções de conjuntos de dados multi-dimensionais. Esta classe de técnicas inclui técnicas de exploração estatística (análise de factor, alteração de escala multi-dimensional, etc.).

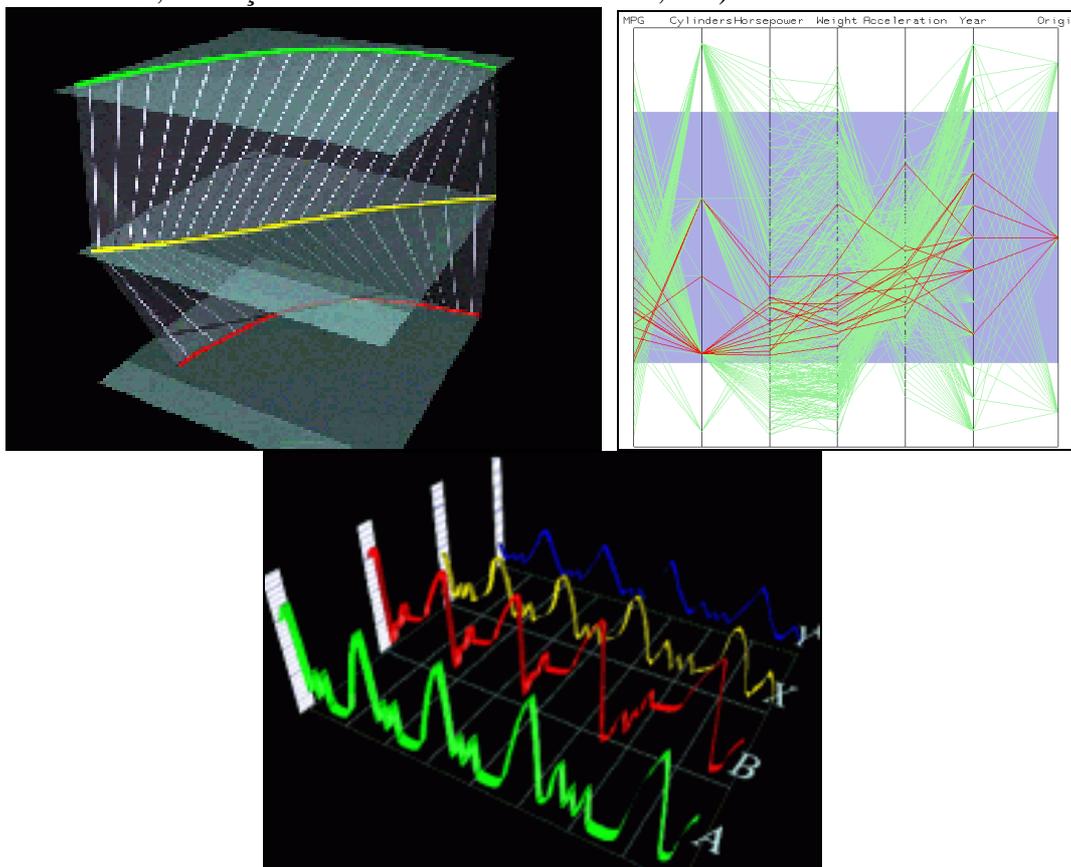


Figura 1.9. Coordenadas paralelas: 3D, 2D e com extrusão

As técnicas de projecção também incluem a técnica de visualização (figura 1.9) de coordenadas paralelas [97]. A ideia básica é a de mapear um espaço de dimensão k para o visor 2D, utilizando k eixos equidistantes entre si e paralelos a um dos eixos do visor. Os eixos correspondem as dimensões e estão linearmente ordenados do menor até o maior valor existente naquela dimensão. Cada item é um dado representado como uma linha poligonal, a qual intercepta o eixo no ponto com o valor considerado para aquela dimensão.

De uma forma geral, os métodos existentes focam mais a representação estatística da quantidade e não fornecem nenhum suporte a topologia que o fluxo apresenta. Alguns métodos [110] têm sido desenvolvidos (principalmente com base nas técnicas de projecção

geométrica) com o objectivo de contornar essa “falha” (coordenadas paralelas por extrusão - figura 1.9, ligação com asas - Figura 1.10 ou ainda, coordenadas paralelas 3D – figura 1.9).

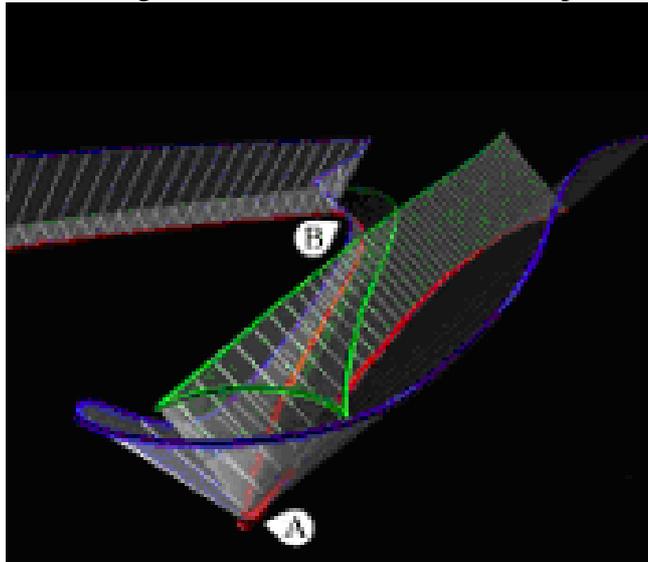


Figura 1.10. *Linking Wings*

Hierárquicas

As técnicas hierárquicas (Figura 1.11) subdividem um espaço de dimensão k e apresentam os sub-espacos obtidos segundo um critério hierárquico. A técnica de sobreposição dimensional é um exemplo [110]. Ela subdivide espaços de dimensão k em sub-espacos 2D. A *Perspective Wall* mostra a informação segundo uma metáfora de montra de informação. A *Treemaps* [44] foi desenvolvida para visualizar a informação tanto de forma quantitativa como hierárquica, e em especial nas árvores com distribuição equilibrada de ramos. Ela se baseia na construção de vários rectângulos que vão se subdividindo segundo algoritmos variados. Ben Shneiderman e Martin Wattenberg propõem [27] em um algoritmo que desenha *treemaps* de uma forma ordenada, garantindo que os itens próximos entre si numa dada ordenação, sejam representados visualmente próximos. A *Cone Trees* representa em 3D uma informação hierárquica e permite a visualização da estrutura na íntegra, segundo uma disposição espacial cónica. As *Bubbles Trees* [6] introduzem um mecanismo de visualização que é baseado na propriedade natural da árvore – de uma forma recursiva, criam-se sub-categorias na árvore, tendo como base as sub-árvores que a compõem. Cheops é um método baseado na compressão da visualização de dados hierárquicos, mantendo entretanto, o contexto pertinente a hierarquia complexa, e dando acesso fácil a detalhes. A árvore hiperbólica [42] dispõe os ramos da árvore circularmente (ver imagem abaixo), sendo mais apropriada para os casos de árvores com graus variados de profundidade nas hierarquias.

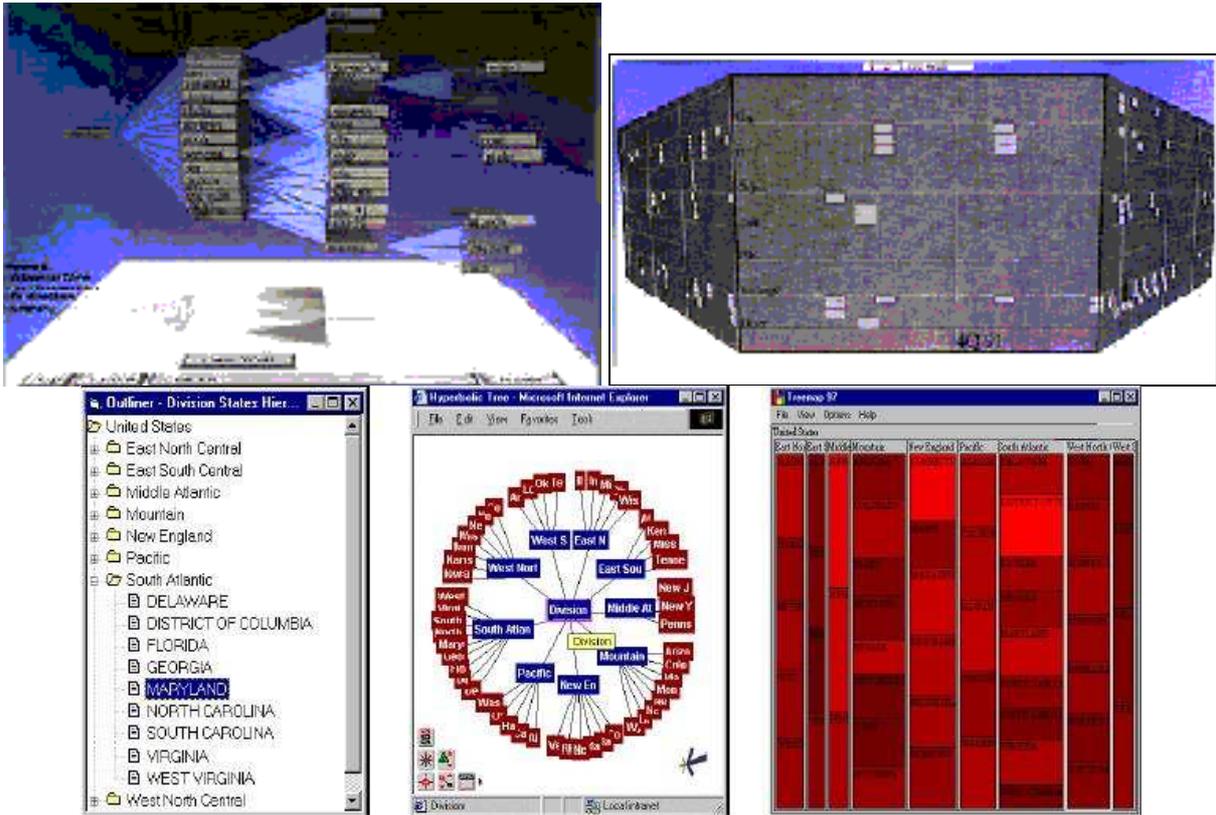


Figura 1.11. Cone Tree, Perspective Wall, Outliner, Árvore Hiperbólica e Treemap

Data Mining e Clustering

O *Clustering* consiste em dividir um conjunto de n pontos em m subconjuntos, de forma a minimizar a soma das distâncias (ou qualquer outra métrica similar) existentes entre cada ponto e o centro do *cluster* (conjunto) [117]. Ela substitui todos os pontos presentes num determinado *cluster*, por um único ponto, localizado no centro deste, que equivale a média dos demais. É utilizada normalmente em conjunto com o diagrama de Voroni. A escolha dos pontos iniciais para cálculo do centro do *cluster*, implica numa melhor ou pior performance da aplicação desta técnica. Uma outra variação, é a utilização hierárquica de *clusters* [73].

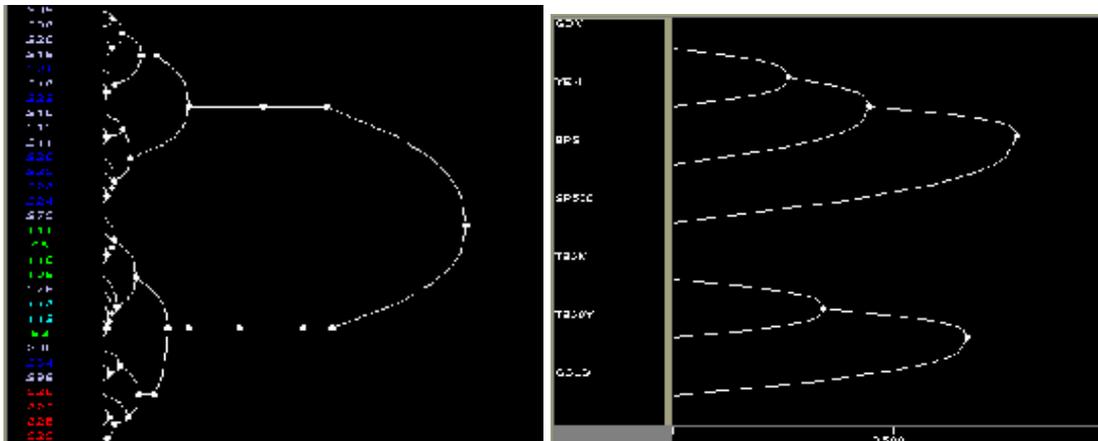


Figura 1.12. Visualização de *clusters*

O *clustering* não é uma técnica de visualização em si, porém de *data mining* (extração de dados) [101]. *Data mining* é um conjunto de metodologias para a exploração automática de complexas relações existentes em grandes volumes de dados. Muitas novas técnicas de visualização têm sido criadas para o *clustering* e o *data mining*. Ben Shneiderman em [12] propõe algumas recomendações básicas e importantes a serem seguidas aquando da criação de visualizações para o *data mining*. A figura 1.12 ilustra algumas das técnicas utilizadas.

Outras Técnicas Auxiliares

A transparência pode ser utilizada para realçar determinadas zonas da visualização [73]. A utilização de técnicas de luz e sombra, e até mesmo neblina (figura 1.13), são normalmente aplicadas com esse intuito na visualização resultante.

Os recortes de volumes [79] permitem extrair superfícies e visualizar zonas de especial interesse existentes nesses [86] volumes, enquanto que a variação do nível de detalhe [69] (figura 1.13), permite focar a atenção consoante a posição e distância de observação da cena. Quanto mais perto o observador está, mais detalhada é a visualização. Neste caso, a utilização de técnicas fractais, por exemplo, permitem criar representações interessantes [93].

A utilização de animação permite tratar a informação sob uma forma dinâmica e temporal, e avaliar a sua evolução [109]. A aplicação de texturas permite criar efeitos diferenciados que facilitam a percepção da forma, e realce de zonas de interesse [91].

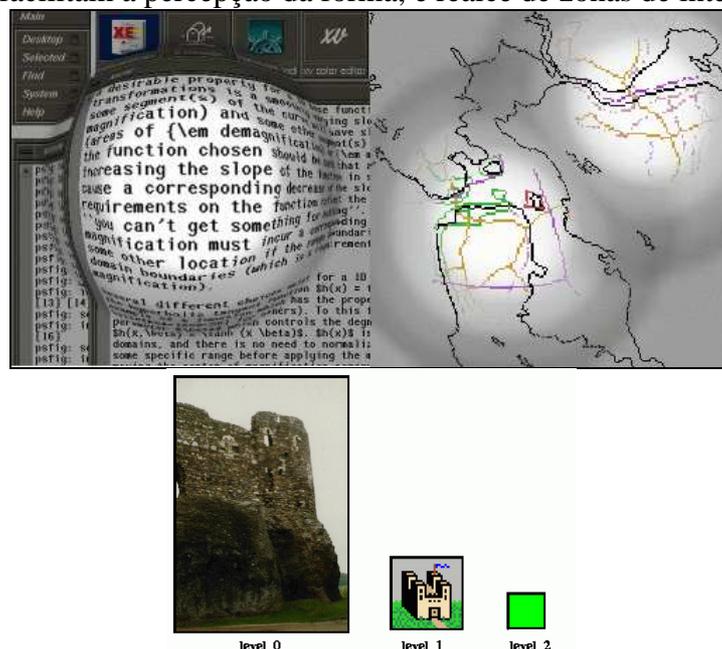


Figura 1.13. Textura, neblina e não linear e *Level-of-detail*

A interactividade permite que a visualização seja analisada de uma forma mais controlada [124]. *Computational steering* [123] é o nome da técnica que permite que a simulação e a visualização estejam integradas num único processo, e que permite um controlo total por parte do utilizador deste processo.

A segmentação permite dividir a visualização em partes menores, enquanto que a extração de formas facilita a identificação de determinadas formas (e não conteúdos)

presentes na visualização [102], enquanto que a utilização de ambientes de realidade virtual permite uma maior “imersão” na visualização e respectiva compreensão [125].

A visualização hierárquica [28] emparelhada (*PairTrees*) integra *treemaps*, diagramas de nós-ligações, mapas de cores e outras técnicas de visualização para dar suporte a exploração de conjuntos de dados em múltiplos níveis de abstracção. Ela permite que o utilizador explore dados heterogéneos utilizando hierarquias baseadas em relações de agregação, subclasses e outras.

A visualização de dados multi-dimensionais provenientes de bases de dados, comumente chamados de cubos de dados, está a se tornar cada vez mais trivial. Stephen Eick [96] propõe visualizações baseadas em projecções perspectivas.

A visualização em colaboração [128] permite que múltiplas pessoas possam executar anotações e interagir na criação de uma visualização em tempo real, estando fisicamente distantes umas das outras. A evolução deste tipo de visualização resultou nos ambientes tele-imersivos. Eles integram ambientes em colaboração [LeJJ-99] criados em realidade virtual, juntamente com capacidades de áudio e vídeo-conferência. Apesar de não serem directamente técnicas de visualização, auxiliam na eficácia destas.

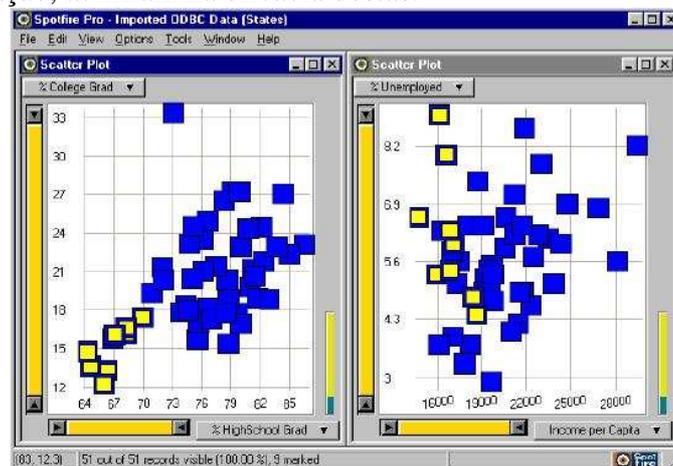


Figura 1.14. Starfield

A *Starfield* [37] é similar a um gráfico de pontos, porém possui capacidades de zoom, *panning*, fornecimento de detalhes (fig. 1.14) conforme solicitado, rotações, etc. Esta visualização combina simultaneamente a visualização de grande número de pontos de dados individuais com uma interface que permite facilmente executar o zoom, filtrar e pesquisar dinamicamente.

A visualização *Snap-Together* (STV) [8, 42] é, na realidade, uma arquitectura que permite aos utilizadores conectarem ferramentas de visualização de tal forma, que seleccionar, navegar e pesquisar são actividades que decorrem de uma maneira coordenada. Em adição a STV suporta várias ferramentas de visualização, o que permite combinar visualizações *Starfield* com tabelas, sumários e *web* browsers, por exemplo. A STV utiliza as ferramentas ODBC da Microsoft para ligação a base de dados, sendo a importação de dados directa.

Brushing and Linking [9] é uma técnica de análise exploratória de dados utilizada na visualização de conjunto de dados em múltiplas janelas. Quando o utilizador selecciona itens em uma das janelas de visualização, os itens são automaticamente colocados em destaque (*highlight*) em todas as demais.

Metaballs [132] são objectos de dimensão n com aspecto orgânico que são criados a partir de funções implícitas. Esta técnica foi inventada nos anos 80 por Jim Blinn. Cada *metaball* é definida como uma função de dimensão n . Um valor limite é definido para controle do volume da *metaball*.

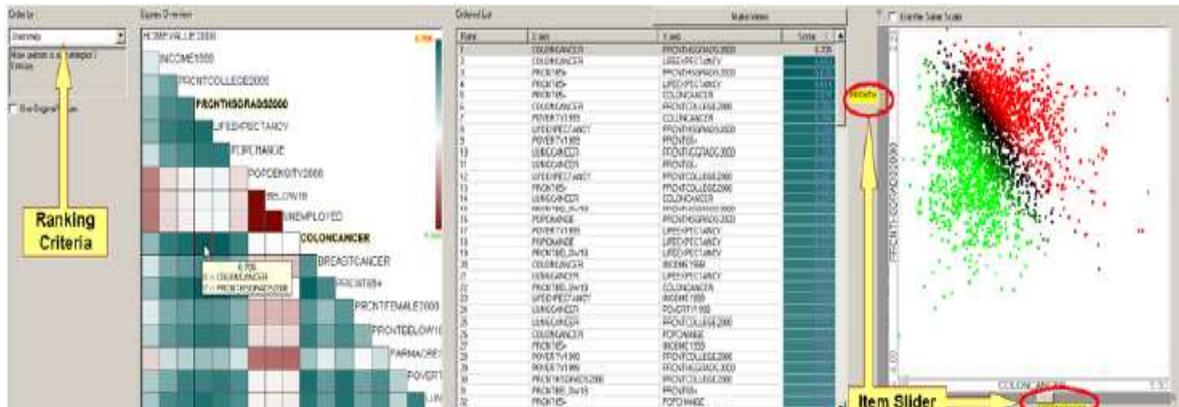


Figura 1.15. Rank-by-Feature

Na visualização *Rank-by-Feature* [33, 34] o utilizador pode seleccionar (fig. 1.15) um critério interessante de pontuação, sendo em seguida todas as projecções paralelas possíveis dos dados de multi-dimensão avaliadas segundo o critério de pontuação definido. O resultado da pontuação é apresentado segundo uma grelha codificada por cores (designado de “*Score Overview*”), bem como numa tabela (designada de “*Ordered List*”), cujas linhas representam uma projecção e está colorida segundo o critério de pontuação. O utilizador pode manualmente percorrer as projecções através da alteração da dimensão de um eixo, com o auxílio dos manípulos gráficos associados ao eixo.

1.6. Interação

A interacção é identificada pelas acções que o utilizador executa durante o processo de visualização [88]. Existem várias formas de interacção. Algumas formas podem implicar na alteração da cor, do ponto de observação da visualização ou de técnicas de sombreamento. Outras, podem incluir alterações nas próprias técnicas [106] de visualização ou dos parâmetros utilizados para a geração de dados (*steering* interactivo). Outras formas de interacção estão relacionadas com a exploração dos dados: exploração de seus valores e eliminação de dados oclusivos.

Na interacção do utilizador com a visualização, as características humanas de resposta são importantíssimas [74]. De uma forma geral, os utilizadores preferem respostas rápidas. O *feedback* deve decorrer em menos de 1 décimo do segundo, para que as tarefas de entrada (fornecimento de informações) executadas pelo utilizador sejam contínuas. Respostas com mais de 15 segundos causam quebras no processo de interacção. Por outro lado, respostas muito rápidas não permitem que o utilizador pense. Um outro factor importante é o equilíbrio entre a velocidade e a produtividade desejada, pois as duas não variam proporcionalmente - a rapidez pode implicar taxas de erro maiores. De uma forma geral, o tempo de resposta deve ser personalizado consoante a tarefa a executar. Testes empíricos podem ser úteis para avaliar o tempo de resposta.

As técnicas de interacção são várias: selecção [103], movimentação, teste, com auxílio de *widgets* [127], etc. A interface desempenha um papel fundamental, pois pode facilitar ou

inibir a interacção. Um outro detalhe é a consistência, pois nem sempre o que é considerado “padrão” implica esta.

Por exemplo, Jean Fekete e Catherine Pleasant propõem uma técnica interactiva em [24] para visualizar 1 milhão de itens num ecrã de resolução 1500 x 1200 *pixels*, sem efeito de agregação. A técnica depende intensamente da capacidade que a placa gráfica do computador possui. Ela inclui efeitos animados, para auxiliar a percepção de valores fronteiriços e tendências existentes nos dados, e pesquisas dinâmicas utilizando o *Z-buffer* da placa gráfica. A utilização de atributos visuais não padronizados, como a transparência e a visão estéreo, permitem inspecções (fig. 1.16) temporárias efectivas.



Figura 1.16 Visualização de 1 milhão de itens

Outro bom exemplo de interacção é a pesquisa dinâmica. Elas são um método para inspecionar de forma interactiva dados na visualização da informação. Por outro lado, é importante a criação de interfaces que permitam que estas decorram de uma forma eficiente. Evan Golub e Shneiderman [35] propõem a utilização de métodos que incluem controle de manípulos, representação visual e visualização/actualização de resultados, tendo como base o Dynamic HTML para potenciar mais a pesquisa dinâmica. Hochheiser e Shneiderman [43] sugerem o desenvolvimento de widgets específicos para análise de séries temporais – *timeboxes*, para garantir a pesquisa dinâmica interactiva de dados com forte componente temporal.

1.7. Conclusões

A visualização é um processo bastante complexo. Factores cognitivos [83], físicos, psicológicos, culturais, entre muitos outros, podem influenciar, de forma decisiva, o sucesso de uma visualização. O conhecimento do modelo visual do utilizador, a caracterização dos dados a serem visualizados e das tarefas a serem executadas, são alguns exemplos da informação básica necessária para poder se executar um mapeamento eficiente da informação. Muitos modelos já foram propostos, porém, a visualização [92] ainda não possui uma “fórmula” absoluta [104]. Por outro lado, existem poucos testes para verificar e garantir a fiabilidade dos resultados fornecidos pelos sistemas de visualização [129], ou pelas técnicas a nível algorítmico [75].

Visualizar informações disponíveis na *Internet* ou [90] nas *intranets*, tornou-se uma aplicação trivial na visualização da informação, e [89] no próximo milénio será necessário o desenvolvimento de novos modelos para dados e de técnicas inteligentes de exploração. Estas técnicas deverão permitir a extracção de informação a partir dos vastos repositórios de dados existentes, muitos destes, oriundos de complexas simulações computacionais. Isto auxiliará também a investigação de tecnologias de visualização mais efectivas (*hardware* e *software*).

No novo milênio, será trivial a execução de pesquisas tendo como base na representação pictural da informação (QBPE – *Query By Pictorial Example*), em lugar da textual. Este é o futuro da visualização.

A área de Visualização da Informação

2.1. Introdução

A quantidade de informação actualmente disponível é imensa, pois diversas são as fontes existentes. A Visualização da Informação viabiliza que as pessoas possam processar toda essa informação, tirando partido de suas capacidades inatas de percepção visual e utilizando menos o sistema cognitivo cerebral. Por exemplo, se considerarmos uma tabela de dados relativos ao stock de uma empresa, para determinar-se os valores máximos e mínimos, será muito mais difícil fazê-lo directamente a olhar para a tabela, em lugar de utilizar-se um gráfico com essa informação. Excelentes exemplos podem ser encontrados no website “VisualComplexity” [133] que é na actualidade um dos melhores repositórios de apontadores para trabalhos nesta área ou ainda no “A Periodic Table of Visual Methods” [134].

A cor, o tamanho, a forma, a noção de distância e movimento são características básicas de apreensão do sistema perceptivo humano. Todas as aplicações orientadas a Visualização da Informação têm como base estas características, que combinadas e utilizadas de forma adequada, podem maximizar em muito a percepção da informação.

Uma boa visualização deve potenciar não só uma clara compreensão da informação, mas também aumentar a quantidade de informação que é “compreendida” e “apreendida” através da mesma. A identificação de valores máximos e mínimos, de desvios, erros, tendências ou intervalos e até mesmo de “conjuntos” (*clusters*) de informação, são alguns dos objectivos que devem estar sempre presentes na Visualização da Informação, pois só assim, é possível uma melhor compreensão de sistemas complexos, descobrir informação “escondida” ou ainda dar apoio a decisões. Vários estudos têm sido feitos para avaliar a validade dos sistemas de visualização, segundo critérios bem definidos [16].

É importante ser feita a distinção entre as áreas de Visualização da Informação e Científica. Ambas são áreas da computação gráfica orientadas para a visualização de dados com o auxílio das mais diferentes técnicas, porém, a Visualização Científica foca dados que descrevem objectos físicos, fenómenos e medições de cariz científico, tendo surgido no início dos anos 80, enquanto que a Visualização da Informação está mais orientada para o desenvolvimento de metáforas para dados que não são inerentemente espaciais, e surgiu a partir dos anos 90. Portanto a diferença entre elas reside basicamente na natureza dos dados que são utilizados para análise – científicos ou não científicos, já que partilham objectivos e desafios bastante semelhantes. Muitas são as técnicas de visualização comuns a estas duas áreas, sendo difícil determinar também neste sector uma clara fronteira. Outro detalhe é que esta distinção tem sido tema de intenso debate na comunidade científica, como pode ser visto em [2], não havendo até a presente data um claro consenso.

Este capítulo analisa algumas aplicações, projectos, ferramentas e sistemas de Visualização de Informação. Para categoriza-los, serão considerados sete tipos de dados básicos subjacentes a eles: unidimensional, bidimensional, tridimensional, multi-dimensional, temporal, hierárquico, rede e *workspace*. Esta sistemática de categorização, embora válida para o objectivo deste capítulo, é restritiva sob o ponto de vista analítico da questão, já que muitos deles estão orientados para mais de um tipo de dado, sendo portanto mistos em relação

aos dados que lhe estão subjacentes. Por exemplo, muitos sistemas serão categorizados como *workspaces*, embora eles englobem muito mais do que essa categoria.

Por fim, serão também analisadas algumas das *frameworks* que estão a ser propostas, com o objectivo de criar uma estrutura comum e padronizada para todos os desenvolvimentos de software que são feitos para esta área.

2.2. Unidimensional

As listas sequenciais, baseadas normalmente em texto, são um bom exemplo deste tipo de dado. Exemplos típicos incluem os códigos fonte de programas de computadores, textos com muitas linhas e resultados de pesquisas em documentos. Via de regra, os potenciais utilizadores de aplicações desta natureza desejam transitar entre visualizações focadas e visualizações globais da lista de caracteres que eles estão a analisar, efectuar comparações entre os elementos presentes na lista, ou ainda, simplesmente aceder a um determinado elemento da lista.

A solução, mais adoptada neste caso, utiliza métodos que permitem rolar o texto, ordena-lo ou ainda executar alguma busca. Uma busca típica na WWW (*World Wide Web*), com o auxílio de um browser como o Explorer™, é um bom exemplo disso. Por outro lado, as limitações impostas nesse tipo de buscas, ilustram bem a inadequação destas técnicas a longo prazo.

Uma aplicação, baseada na Visualização da Informação, pode oferecer, nesses casos, uma resposta sempre mais efectiva às necessidades do utilizador, além de vistas globais e compactas das listas em si. Com a alteração da cor, tipo de fonte, tamanho, orientação, posição e escala, a visualização pode tornar-se muito mais rica, permitindo de uma forma fácil comparar, filtrar ou pesquisar dados na lista.

Como objectivos básicos na visualização de dados desta natureza na Visualização da Informação, temos:

- Visualização de sumários dos itens ordenados, de forma a encontrar os elementos mais importantes.
- Percorrer as listas segundo uma ordenação variável.
- Filtragem de itens indesejáveis.

Exemplos de Projectos e Ferramentas

O ttxtkit [29] é uma ferramenta de código (Fig. 2.1) aberto para a exploração de grandes quantidades de texto escrito em vários idiomas. É uma aplicação multi-utilizador que foca em especial o processo de leitura e propõe soluções como uma série de decisões e eventos. Para expandir uma determinada vista de actividade o ttxtkit agrupa todos os dados oriundos de todas as pesquisas dos utilizadores e os utiliza para criar um mecanismo de sugestões automáticas de conteúdo, através da filtragem dessa informação. O software corre em MAC OS X 10.3 e requer acesso a Internet.

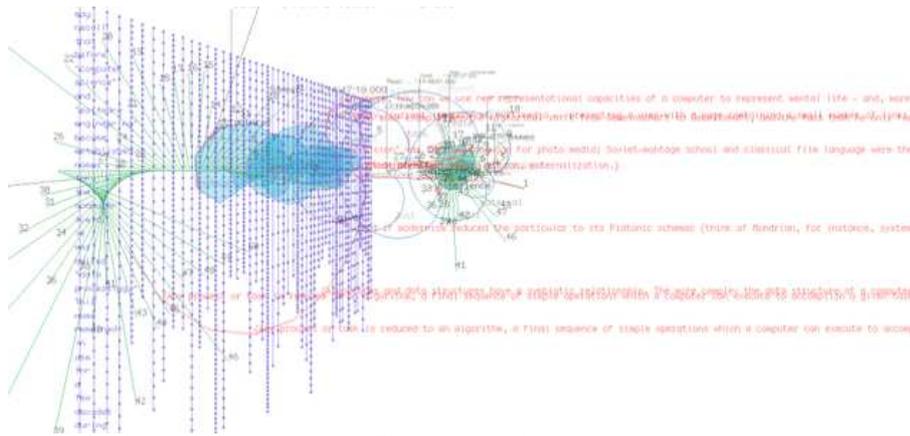


Figura 2.1. txtkit

O Virtual Shakespeare [5] permite que o utilizador navegue por vários textos de peças escritas por Shakespeare, visualizando-os de forma bastante variada e interessante. Uma única visualização pode conter mais de 100 mil palavras. A este nível, as palavras (figura 2.2.) são ilegíveis, mas os padrões e a estrutura geral de cada peça é perceptível. Através da alteração escala, detalhes, como o a frequência de caracteres e tamanho dos diálogos, o tamanho relativo de cada acto ou o diálogo legível da peça, podem ser visualizados em diferentes momentos.

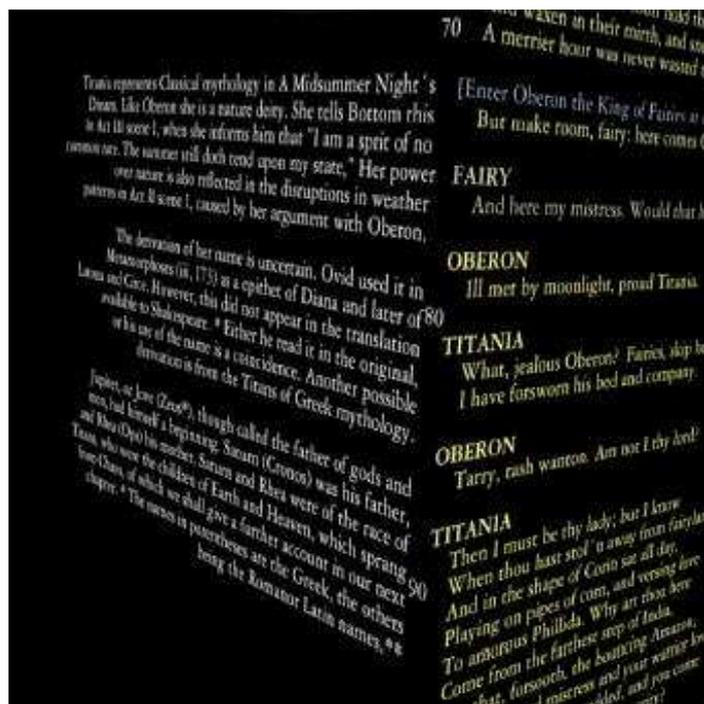


Figura 2.2. Virtual Shakespeare

Pak Wong e al. [4], do Pacific Northwest National Laboratory, propõem um sistema (Figura 2.3) de visualização que se baseia na utilização de técnicas de prospecção (*data mining*) para a descoberta de padrões sequenciais em grandes conjuntos de dados. Um padrão sequencial, na prospecção de dados, é uma série finita de elementos como $A \rightarrow B \rightarrow C \rightarrow D$, onde A, B, C e D são elementos do mesmo domínio. A prospecção de padrões sequenciais é executada para encontrar eventos discretos, que ocorrem frequentemente segundo um determinado padrão ao longo do tempo. Da mesma forma que a associação e o agrupamento (*clustering*), a prospecção (*mining*) de padrões sequenciais, é uma das técnicas mais populares de busca de informação, que utiliza avaliações estatísticas para extrair informação útil à partir de grandes conjuntos de dados.

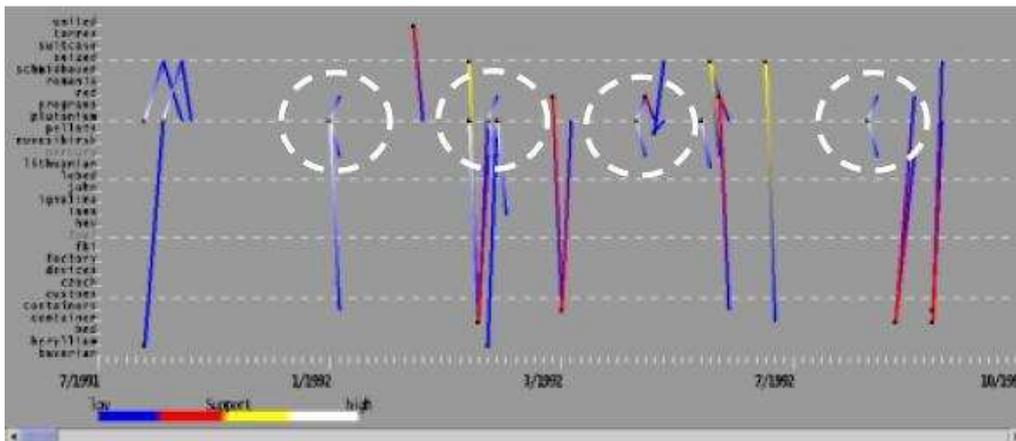


Figura 2.3. Sistema de Visualização por padrões sequenciais

TOPIC-OF-GRAPHY™ [7] aplica *wavelets* ao sinal digital construído a partir das palavras existentes num determinado documento. O resultado da transformação obtida pela aplicação das *wavelets*, é utilizado para analisar as características do fluxo da narrativa no domínio da frequência, tal como as alterações de tema, que podem ser comparadas ao restante tema com o auxílio de métodos estatísticos. As características temáticas de um documento podem ser analisadas com diferentes grau de detalhe, desde partições de texto do tamanho de secções, até partições com apenas poucas palavras. Com base nesta tecnologia está a ser desenvolvido um protótipo designado de TOPIC ISLANDS™ (Figura 2.4) para analisar texto. O protótipo permite a criação de sumários com níveis de detalhe diferenciados (de acordo com os interesses do utilizador), pesquisa do conteúdo do texto, entre outras coisas.

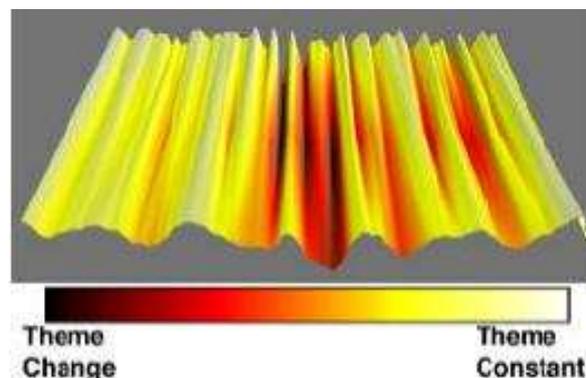


Figura 2.4. TOPIC ISLANDS™

ExpO é uma interface flexível, baseada [21, 32] num browser hierárquico, que permite pré-visualizar resultados de pesquisas em grandes bases de dados *on-line*. Este browser permite, ao contrário dos normalmente existentes, que o utilizador possa percorrer todos os atributos e relações existentes na base de dados. Todos os atributos da base de dados (que sejam apropriados para tal) podem ser utilizados para mostrar a distribuição da informação (fig. 2.5). As visualizações incluem gráficos de barra e há possibilidade de serem expandidas. Para depurar e percorrer os dados antes de efectivamente lê-los, são utilizados os meta-dados disponíveis. Com isto, há uma redução drástica do tempo gasto em acessos remotos.

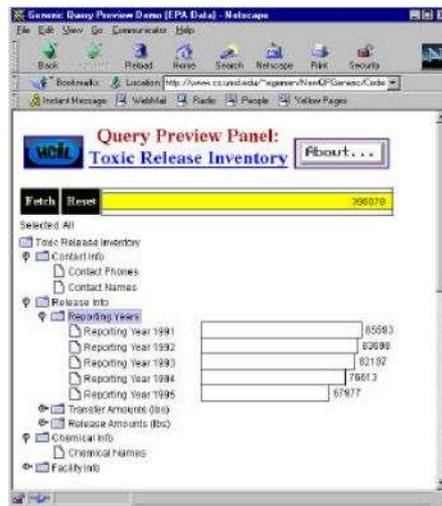


Figura 2.5. ExpO

A visualização no ThemeRiver [41], tem como objectivo principal, evidenciar as variações temáticas ao longo do tempo, em grandes colecções de documentos. As alterações são mostradas dentro de um contexto temporal e relacionadas com os eventos externos correspondentes. O foco na alteração temática dentro de um contexto temporal permite que o utilizador consiga detectar padrões e tendências. Por exemplo, uma alteração abrupta na intensidade temática seguida de um evento externo pode ser um indício de uma relação causal. O fluxo (fig. 2.6) da esquerda para a direita é interpretado como um movimento ao longo do tempo e a distância horizontal entre dois pontos no “rio” define um intervalo de tempo. Em qualquer ponto temporal, a distância vertical, ou a largura do “rio” indica a expressividade colectiva dos temas seleccionados. As “correntes” coloridas que flúem pelo “rio” representam temas individuais. A largura vertical de uma “corrente” alarga ou estreita de forma a indicar a expressividade individual do tema.

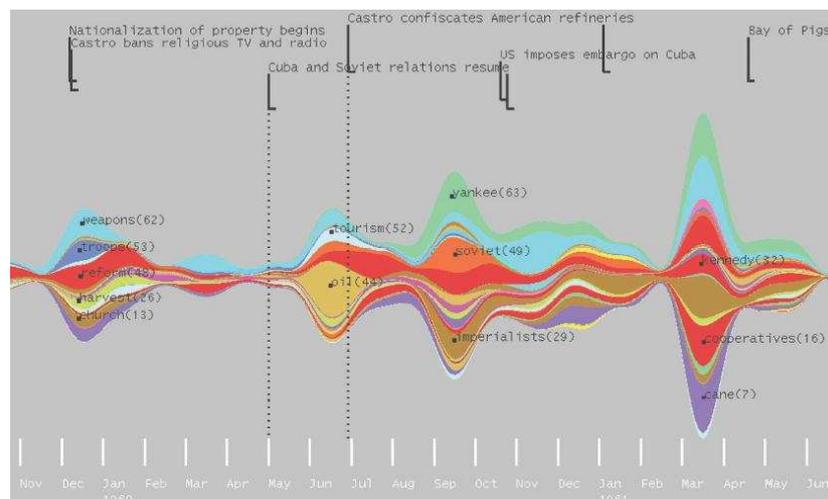


Figura 2.6. ThemeRiver

2.3. Bidimensional

Este ambiente de informação combina a representação e tipo de dados bidimensional. Exemplos de aplicações em ambientes 2D incluem: Sistemas de informação geográfica (SIG), desenho de hardware, *design* de jornais e fotografias. Os SIGs incluem visualizações

cartográficas da informação, sendo bons exemplos deste tipo de aplicação o Arcview™ e o GeoMedia™. Os programas que permitem criação de desenhos estilizados e artísticos, como o Pad++, incluem o zoom semântico. O Photoshop™ é um software que permite os desenhistas e fotógrafos criarem trabalho artístico de uma forma original (corrigindo cores, retocando e compondo imagens digitalizadas). No *design* de edições de jornais, os ambientes bidimensionais provam também ser de grande utilidade, pois facilitam a localização de artigos ou áreas em branco.

A melhor maneira de definir uma ambiente bidimensional de dados, é inicialmente determinar quais tipos de respostas as representações visuais devem dar resposta. A direcção, a localização e a distância são características, que podem efectivamente ilustrar o ambiente de visualização 2D. Por exemplo, um objecto está a direita ou esquerda? Quão grande é Lisboa em comparação com o Porto? Quais são os sítios localizados ao norte de Lisboa? Guimarães está mais perto de Braga ou de Viseu?

Um ambiente visual bidimensional se sobrepõe a outros ambientes de visualização, quando tenta classificar tipos de dados. Na realidade, todas as visualizações de dados são mostradas numa superfície 2D (ecrã do computador). Daí, muitas aplicações podem ser falsamente classificadas como bidimensionais, quando são temporal ou rede, por exemplo. Há uma certa confusão entre o conceito de representação 2D dos dados e o tipo de dados 2D.

Os dados do tipo bidimensional apresentam um número de atributos, que é normalmente utilizado directamente para sua representação visual em 2D. Exemplos são a longitude e latitude, largura e altura, etc. Por vezes, os dados não são do tipo bidimensional, e o ambiente, mesmo assim, é considerado bidimensional. Os dados com mais de duas dimensões também podem, em alguns casos, serem representados segundo um modelo 2D. Por exemplo, se considerarmos uma aplicação que desenhe o mapa de Portugal e que mostre toda a informação associada a cada cidade seleccionada, a visualização pode ser multi-dimensional, pois mais de dois atributos podem estar associados a cada cidade. Embora isto seja possível, neste caso, ainda será considerada como sendo uma aplicação 2D, pois o objectivo principal dela é dar a direcção e posição. Esta subtil distinção é a base que deve ser utilizada para identificar um dado de tipo 2D.

Como objectivos básicos na visualização de dados desta natureza temos:

- Mapas geográficos;
- Desenho de *hardware*;
- Zoom semântico;
- Sistemas de rastreio;
- *Design* de jornais;
- *Browser* de fotos e imagens.

Exemplos de Projectos e Ferramentas

O PhotoFinder [55] é uma aplicação orientada a anotação de fotos digitais. A anotação é uma tarefa que consome bastante tempo, tediosa e com grande probabilidade de erro. Em função disso, boa parte das pessoas que possuem bibliotecas digitais de fotos, acabam por não o fazer. Permitir que o utilizador possa simplesmente arrastar (figura 2.7) nomes a partir de uma lista e os largar directamente sobre a foto a ser etiquetada, facilita em muito esse tipo de

actividade. Desde que os nomes sejam entrados numa base de dados, a procura de todas as fotos de uma determinada pessoa é dramaticamente simplificada.

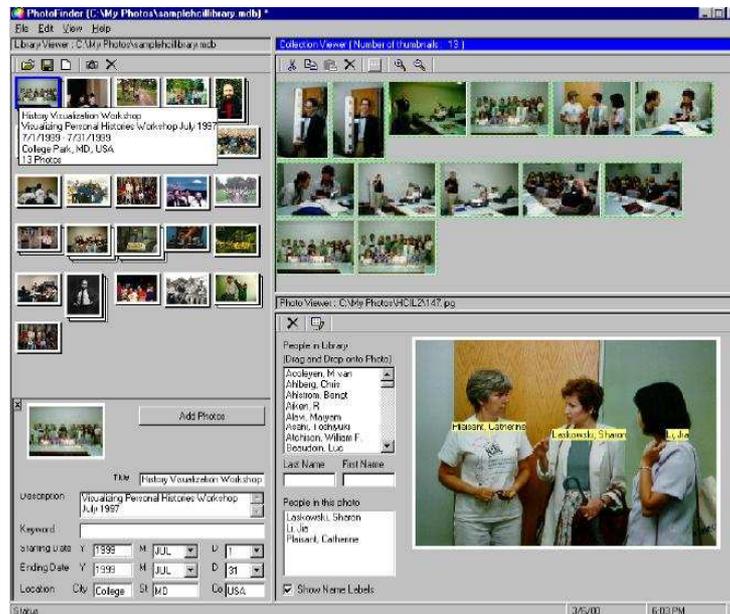


Figura 2.7. PhotoFinder

O GeoVista Studio [1] é um ambiente de desenvolvimento de código aberto, orientado para dados de natureza geo-referenciada. É possível construir com rapidez aplicações para a visualização e processamento de dados geográficos, bastando desenvolver o código apropriado. O objectivo fundamental deste projecto é melhorar a análise geo-científica, criando um ambiente totalmente operacional, que inclui uma vasta gama de ferramentas de análise, em especial, visuais e (figura 2.8) estatísticas.

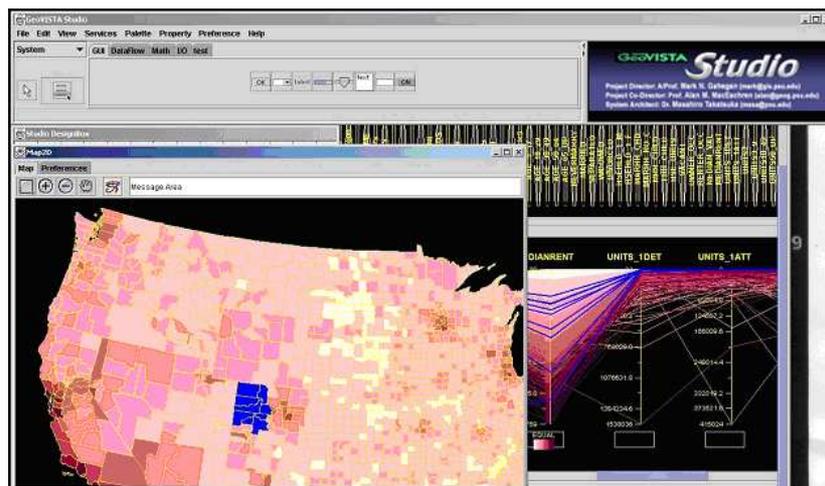


Figura 2.8. GeoVista Studio™

Counterpoint [20] é uma ferramenta de apresentação, que oferece graus diferenciados de zoom. Ela suporta a construção de apresentações no estilo *slide show*, com o conteúdo disposto num espaço de dimensão 2.5. Como no caso de outras interfaces que permitem a variação do zoom, CounterPoint possibilita navegações na apresentação, com a inclusão de efeitos animados, conforme ocorrem transições entre as visualizações. A ferramenta suporta

ainda, a execução de caminhos sequenciais automatizados de visualizações durante uma apresentação, bem como interactivas. Também podem ser definidos múltiplos (fig. 2.9) caminhos dentro de um ambiente único de apresentação.

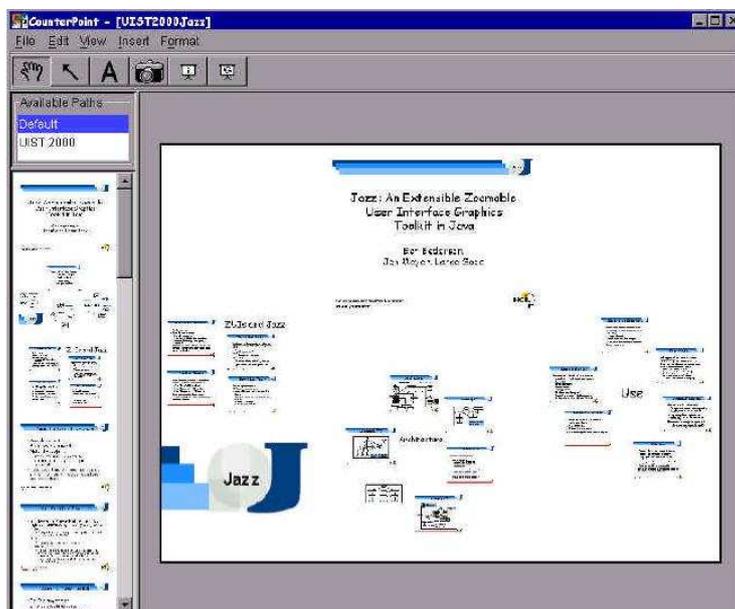


Figura 2.9. CounterPoint

Dynamaps [9, 18] é uma ferramenta generalista de visualização de mapas, que permite pesquisar e analisar dinamicamente dados geográficos, representados segundo mapas de cores (*choropleth maps*). O utilizador pode utilizar as cores para codificar as variações em cada região geográfica e filtrar as áreas, que não atendem a um determinado critério desejado. Dynamaps permite visualizar as distribuições dos dados, além dos seus detalhes e, ao mesmo tempo, obter vistas genéricas. É possível relacionar dados estatísticos e geográficos, descobrindo as tendências e os seus limites. Outro projecto muito similar a este é o YMap [14]. Este, porém, permite a visualização simultânea de várias alternativas (fig. 2.10).

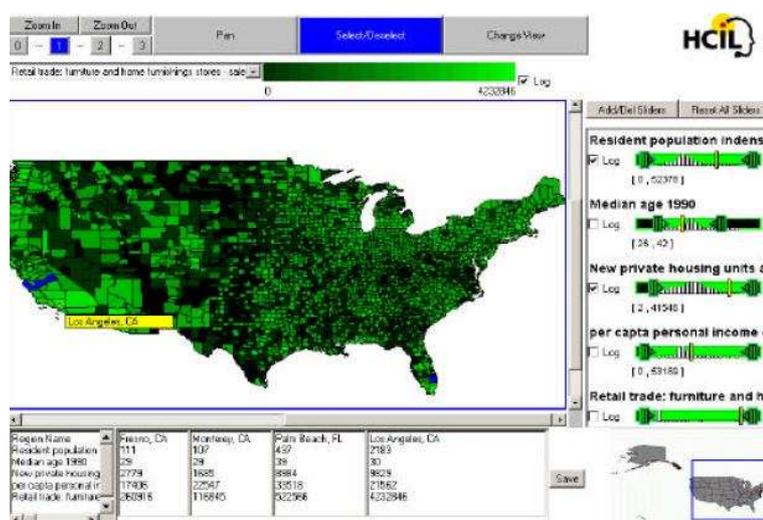


Figura 2.10. YMap

2.4. Tridimensional

Este tipo de visualização deve ser aplicado, quando ver o objecto em 3D ou o que está em seu interior, é fundamental para a execução de uma dada tarefa por parte do utilizador. Embora a criação de ferramentas para navegar e visualizar em 3D sejam complexas, há uma grande solicitação para a criação destas. Por exemplo, o *National Library of Medicine* desenvolveu o projecto Visible Human [60]. Ele é um repositório de imagens sobre o corpo humano, que podem ser utilizadas para criar visualizações 3D. Muitas aplicações têm sido criadas com base nessas imagens, permitindo uma interacção total com as visualizações em 3D do corpo humano.

A visualização volumétrica é a forma mais utilizada em 3D. Ela implica em criar visualizações muito próximas do mundo real, num ecrã 2D de um computador, além de eventual interacção em ambiente imersivo virtual. As aplicações actuais desse tipo, incluem áreas como a medicina, ensino, meteorologia e arquitectura, entre outras.

Tem vindo a crescer cada vez mais a utilização de realidade virtual nessas aplicações. O VRML (*Virtual Reality Modeling Language*), por exemplo, é uma ferramenta bastante utilizada na criação de universos virtuais, que permite modelar desde lugares históricos até planetas, disponibilizando as visualizações interactivas na Internet.

O conceito de universo 3D também é por si confuso. Existem basicamente quatro interpretações diferentes deste:

- A primeira considera que o universo 3D é composto por objectos 3D reais, como os presentes em visualizações nas áreas da medicina, arquitectura ou química. Nesses casos, há sempre o interesse por parte do utilizador em visualizar os objectos no seu interior, pois isto é fundamental para a compreensão da visualização. Navegar para cima, baixo, frente e trás dentro do ambiente imersivo apresentado, fornece ao utilizador a percepção de contenção espacial;
- A segunda é a de que o universo 3D é artificial e sintético. As visualizações criadas são universos fictícios, que entretanto, se apresentam como bastante reais. Esse tipo de visualização implica em passeios virtuais por ambientes fictícios, e não possui como ideia chave explorar o interior de objectos, como é no caso anterior. As aplicações que simulam visitas virtuais a museus ou lugares históricos, são bons exemplos deste caso;
- A terceira categoria de visualização em 3D, inclui objectos em que é possível aplicar o nosso sentido real de 3D para os fazer parecer tridimensionais. Por exemplo, árvores hierárquicas, redes de comunicação, sistemas SIG, objectos multi-dimensionais e temporais, podem ser modelados em 3D, mas isto não significa que estes objectos fazem realmente parte de um universo 3D. Existem muitos produtos que utilizam este conceito 3D, como por exemplo o [61] WebBook and Web Forager (figura 2.11);



Figura 2.11. WebBook and Web Forager

- A quarta categoria engloba objectos em que a criação da 3ª dimensão é forjada. Bons exemplos são os gráficos em barras ou queijo construídos a partir de dados 2D, mas que são visualizados em 3D (muito utilizado no Excel™, por exemplo). Este tipo de visualização, apesar de apelativo, é bastante questionável em termos de qualidade da compreensão que se pode obter.

Como objectivos básicos na visualização de dados desta natureza temos:

- Percepção da contenção especial existente num ambiente;
- Pesquisar e explorar o ambiente segundo posições espaciais diferentes e intuitivas (acima e abaixo, dentro e fora, para frente e trás).

Exemplos de Projectos e Ferramentas

A VxInsight [52] é uma aplicação que combina técnicas de visualização científica com as de informação. É passível de uma fácil padronização e extensão. É uma ferramenta para análise das relações existentes em grandes bases de dados (figura 2.12). O objectivo principal é revelar a estrutura implícita dos dados, facilitando descobrir padrões e ligações e/ou dependências. Na próxima versão, irá incorporar várias outras técnicas como o mapa de contornos e a *treemap*.

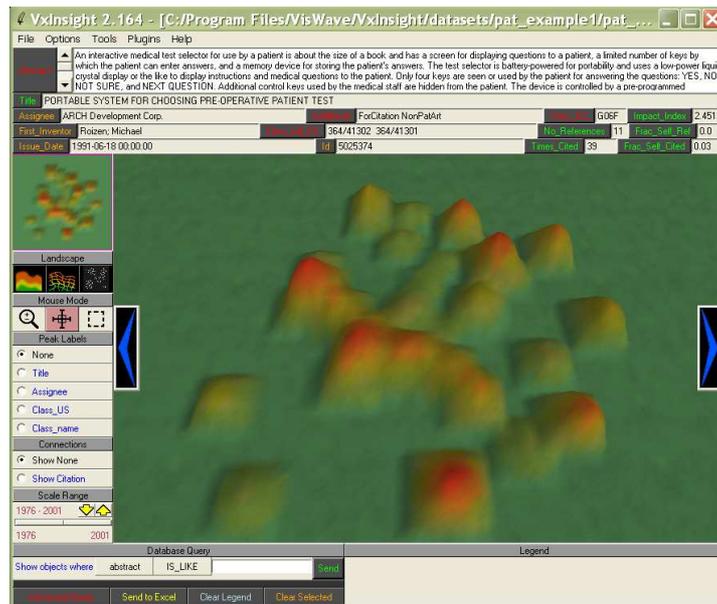


Figura 2.12. VxInsight

WilmaScope [51] é uma aplicação em JAVA3D capaz de criar animações tridimensionais de grafos em tempo real. Ela permite (fig. 2.13) a visualização 3D de grafos, a criação rápida de protótipos a partir de algoritmos, a produção de imagens estáticas 3D com alta resolução, a modelação do tipo *cluster*-grafo, o controlo de luz ambiente e movimentos, a criação de extensões personalizadas com interface através da sua API, a navegação interactiva, entre outras coisas.

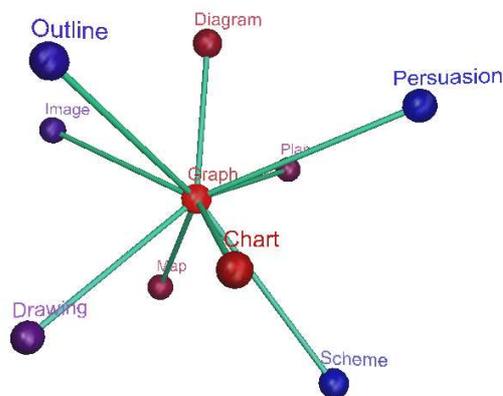


Figura 2.13. WilmaScope

2.5. Multi-dimensional

Esta visualização, no caso da visualização de informação, mostra dados que primariamente não são espaciais. O número de atributos associado a cada item é sempre mais que três. Os dados utilizados podem ser abstractos e de natureza estatística, como por exemplo: fenómenos, produção fabril, base de dados cinematográfica, distribuição de livros numa biblioteca, etc.

A visualização unidimensional, bidimensional, tridimensional e temporal podem ser vistas como subconjuntos da visualização multi-dimensional. Existe sempre um determinado atributo predominante em cada uma destas. Por exemplo, no caso de ser temporal, o tempo é o

atributo de maior peso. No caso de ser 2D e 3D a posição é o atributo mais expressivo. No caso multi-dimensional, todos os atributos apresentam o mesmo peso.

A visualização científica também utiliza muito dados com esta dimensão, mas em geral, os atributos espaciais associados aos dados são utilizados na visualização. Bons exemplos são os sistemas de tomografia computacional e sistemas CAD (*Computer Aided Design*). Em ambos os casos, as coordenadas espaciais dos dados são fundamentais para as respectivas visualizações. Esta é a destriça fundamental entre a visualização científica de dados multi-dimensional e a de informação.

Como objectivos básicos na visualização de dados desta natureza temos:

- Perceber o todo ou parte de um conjunto de dados de n dimensões. São bons exemplos: encontrar padrões, relações, limites, falhas e *clusters*.
- Encontrar um item específico nos dados. O zoom, a filtragem e a selecção de um único item ou grupo de dados são bons exemplos deste caso.

Exemplos de Projectos e Ferramentas

O Hierarchical Clustering Explorer (HCE) [11, 19, 33, 34] utiliza o algoritmo hierárquico para criação de *clusters* sem um número pré-definido. Os utilizadores podem definir os grupos, com total feedback visual (dendograma e mosaico colorido), além de controlarem dinamicamente as pesquisas. O HCE (figura 2.14) implementa quatro técnicas genéricas, que podem ser utilizadas interactivamente na exploração dos *clusters* resultantes: visão global, combinada ou em detalhe do conjunto de dados, controles para pesquisas dinâmicas (com controlo de número de *clusters* visíveis por vez), visualizações coordenadas entre si e comparações entre os *clusters*. O HCE também utiliza projecções de reduzida dimensão para facilitar a visualização e utiliza como base a técnica *rank-by-feature*.

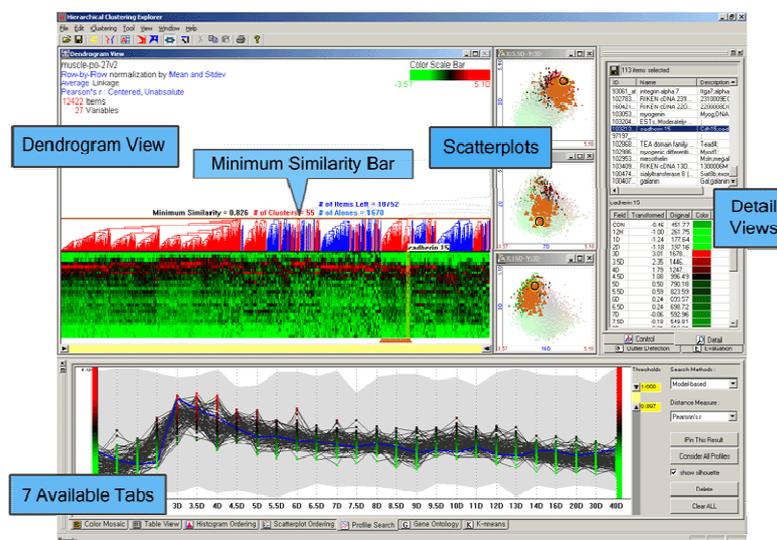


Figura 2.14. HCE

O GlassEye [38] é um browser criado para explorar o site que contém todo o catálogo musical das composições de Philip Glass (é um conteúdo tipicamente multimédia - música,

vídeos, letras, etc.). O protótipo (figura 2.15) foi construído tendo como base a combinação de técnicas como a *starfield*, a possibilidade de execução de pesquisas dinâmicas, e interfaces ZUI (*Zoomable Interfaces*), além de outras técnicas de interação e complementares de visualização (som e imagem).

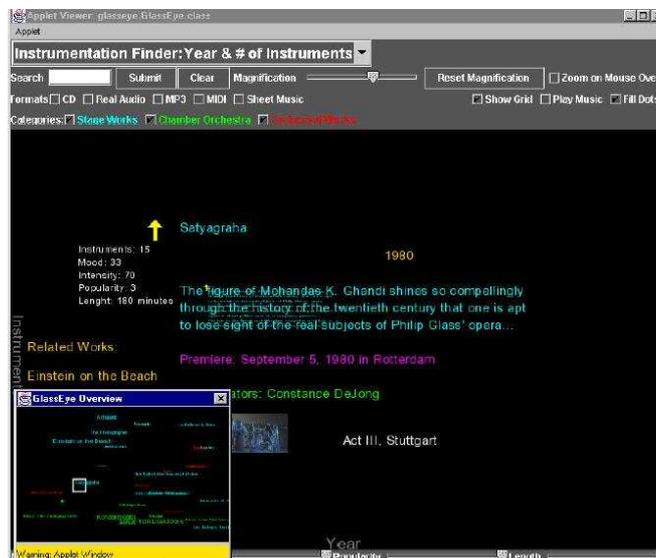


Figura 2.15. GlassEye

A *Snap Together Visualization* [50] é uma ferramenta que permite a criação e publicação na Internet de múltiplas vistas padronizadas de bases de dados relacionais e das suas relações. Existe sempre a coordenação entre as várias visualizações, pois tem como base conceptual a utilização da técnica que lhe dá o nome (STV): vistas dos esquemas relacionais das bases de dados, diagramas dos esquemas das interfaces com o utilizador, etc. O *Snap Together Visualization* (figura 2.16) permite a combinação de múltiplas base de dados em simultâneo, disponibiliza uma API simplificada para integração de componentes, corre como uma *applet* em browsers, permite a filtragem e pesquisa de dados, entre outras coisas.

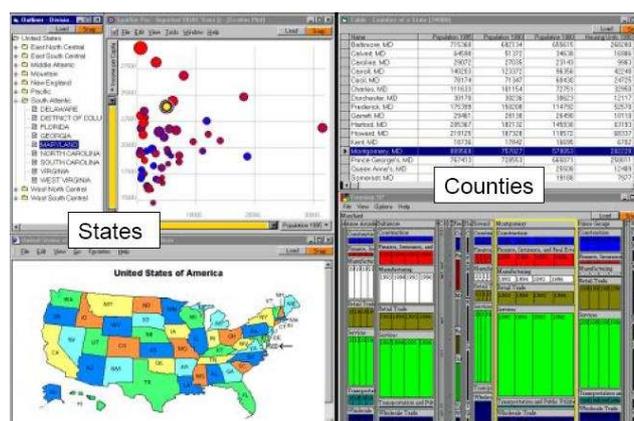


Figura 2.16. Snap Together Visualization

O ADVIZOR/2000 [54] é uma ferramenta que explora as técnicas de visualização para analisar base de dados de multi-dimensão. A metáfora visual e conceptual utilizada é a de cubos de dados. Os cubos podem ser visualizados segundo três perspectivas distintas: única, múltipla e ancorada (fig. 2.17). Cada uma delas combina tipos de gráficos e grau de dimensão diferentes. Para correr o ADVIZOR/2000, basta o utilizador criar uma tabela Pivot em Excel.

Os dados podem ter origem numa folha de dados do Excel, numa base de dados relacional ou num cubo OLAP.

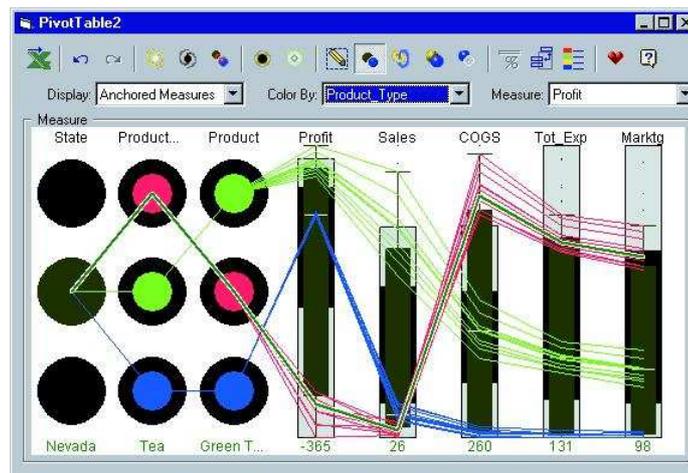


Figura 2.17. ADVIZOR/2000 - Perspectiva ancorada

2.6. Hierárquicos

Este tipo de dados também é conhecido como estrutura em árvore. Eles são conjuntos de dados dispostos em nós, onde cada nó possui apenas um nó pai (situado hierarquicamente acima deste), e pode ter vários nós filhos (situados hierarquicamente abaixo deste). Os nós e ligações podem ter múltiplos atributos. A própria estrutura hierárquica traduz informação relevante. A análise de hierarquias pode focar um único nó, ligação, colecção de nós ou até mesmo toda a estrutura.

As taxinomias, as estruturas organizacionais, a gestão de espaço de disco e as árvores genealógicas são bons exemplos de situações, onde a hierarquia está implícita na informação. A análise de hierarquias pode implicar na pesquisa de um determinado nó ou das anomalias e/ou duplicações na estrutura, até ao exame da estrutura em árvore como um todo.

Muitas das representações tradicionais de hierarquias são feitas em 2D, porém, outras técnicas têm sido desenvolvidas. O objectivo é permitir uma navegação e pesquisa da informação mais intuitiva e simples, garantindo que o utilizador perceba o contexto geral da árvore como um todo. A visualização de hierarquias está relacionada com a visualização de dados 3D, pois em muitos casos, as técnicas utilizadas são tipicamente 3D. Por outro lado, as hierarquias são similares aos dados multi-dimensionais no que toca o número de atributos, embora nestes, questões de hierarquia não existam. Em última análise, as hierarquias podem ser vistas como um caso especial das redes. A diferença reside em não existir o conceito de hierarquia nas redes, podendo haver mais de uma ligação e caminho entre nós, e com sentido bi-direccional na relação.

Como objectivos básicos na visualização de dados desta natureza temos:

- Obter uma noção global e hierárquica da estrutura;
- Encontrar *clusters*, duplicados, relações, caminhos e heranças existentes na estrutura;
- Descobrir atributos dos nós ou de ramos inteiros.

Exemplos de Projectos e Ferramentas

O Treemap 4.1.1 [6, 44] é uma aplicação orientada para a visualização de estruturas hierárquicas, e que tem a técnica que lhe dá o nome. Ela é muito efectiva na visualização dos atributos dos nós, utilizando a cor e o tamanho para expressar valores. Esta aplicação permite que o utilizador compare os nós e ramos percorrendo em profundidade a árvore, de forma a evidenciar padrões e excepções. O Treemap (figura 2.18) foi inicialmente proposto por Ben Shneiderman durante os anos 90, e desde então, foram desenvolvidas várias versões que têm sido utilizadas em diferentes áreas do conhecimento.

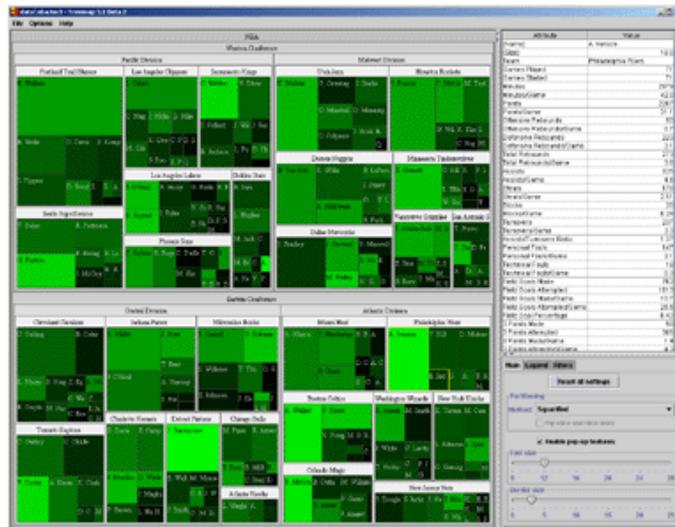


Figura 2.18. Treemap 4.1.1

O Domain Tree Browser (DTB) [13] é uma ferramenta de visualização do histórico de páginas *web* visitadas pelo utilizador. O objectivo é ser utilizado como complementar ao browser. Ele recebe os eventos do browser, toda a vez que uma hiper ligação é picada com o rato na página, e utiliza esses eventos para manter um historial das visitas as páginas. Ela constrói uma hierarquia conforme o utilizador atravessa as ligações. O DTB (fig. 2.19) mantém automaticamente o histórico das páginas visitadas, com o menor esforço por parte do utilizador. A ferramenta organiza as URLs visitadas com base nos domínios dos *websites*. É uma interface passível de zoom, sendo automaticamente redimensionável, de forma a garantir que todos os *thumbnails* caibam na janela.



Figura 2.19. DTB

Space Tree [36] é um browser de árvores que cria visualizações baseadas nos diagramas convencionais para árvores (treemaps, grafos, hiperbólica, etc.). Ele permite a alteração de escala dinamicamente, garantindo sempre um bom enquadramento e aproveitamento do espaço disponível no ecrã. Para esse propósito, também é utilizada a pré-visualização com ícones como sumário da topologia, pesquisa e filtragem integradas e alteração do ponto de (fig. 2.20) vista.

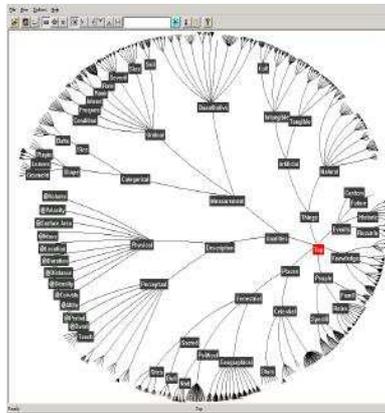


Figura 2.20. Space Tree

TaxonTree [39, 40] é uma aplicação para visualizar a hierarquia existente na taxinomia dos nomes de animais. Ela inclui facilidades para pesquisa e filtragem, que não exigem que o utilizador tenha um profundo conhecimento da área ou das palavras-chave para poder organizar os dados. O objectivo básico desta aplicação é o de determinar padrões de busca de informação em assuntos relacionados com o domínio da biodiversidade (Fig. 2.21).

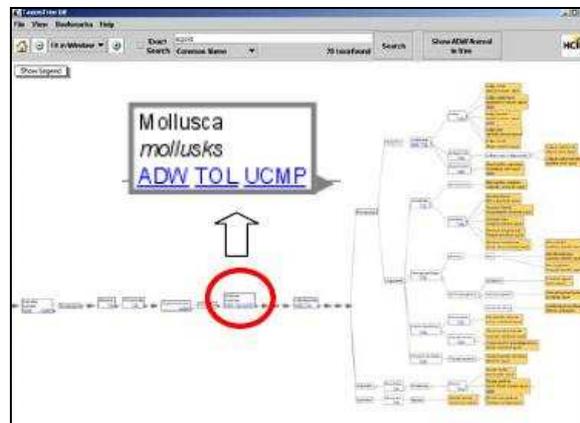


Figura 2.21. TaxonTree

2.7. Redes

A visualização em rede está normalmente associada a visualização de uma estrutura de dados com muitos itens. As redes são constituídas por nós e ligações, onde os nós representam um dado pontual e a ligação, a relação entre dois nós. Muito do trabalho realizado nesta área foi directamente importado da teoria dos grafos. Um grafo simples, com poucos dados, é fácil de entender, porém o mesmo não ocorre, quando a quantidade de dados cresce para os milhares. Outro detalhe, é que muita informação presente em grandes conjuntos de dados está “escondida”. Encontrar uma estrutura ou hierarquia, dentro de uma rede, pode

ser muito difícil, daí, as redes também terem sido categorizadas segundo as suas particularidades estruturais – cíclicas, *lattices*, directa versus indirecta, com raiz ou sem raiz.

As visualizações normalmente envolvem algum tipo de diagrama, e até mesmo efeitos como a distorção, animação ou vistas das ligações mais dependentes ou fortes existentes na rede. As áreas de aplicação da visualização em rede incluem base de dados (tanto o modelo da base como os seus itens), software (conexões estáticas entre módulos, classes, ligações dinâmicas entre processos, etc.), redes computacionais, WWW, redes de gestão social, SIG (relação geográfica entre localizações), tráfego aéreo, linhas de comunicação, etc.

Como objectivos básicos na visualização de dados desta natureza temos:

- Encontrar o caminho menor ou mais barato – por exemplo, instalações eléctricas num condomínio;
- Performance de programas – grafos auxiliam a encontrar possíveis estrangulamentos ou fluxos errados de dados;
- Gestão de tráfego – relacionar o tráfego a partir de diferentes nós, de forma a encontrar um determinado padrão;
- Edição de grafos,
- Relacionar facilmente os nós entre si.

Exemplos de Projectos e Ferramentas

BOOST Graph Library (BGL) [62] é uma biblioteca aberta, de alta performance, escrita em C++, especialmente concebida para o processamento de grafos. Parte dela é uma interface genérica que permite acesso à estrutura de grafos, escondendo entretanto, os detalhes da implementação. Os algoritmos permitem operar facilmente qualquer estrutura de dados do tipo hierárquico. O BGL tem sido utilizado como base de muitas aplicações orientadas para a visualização de redes, e não possui nenhuma interface para interacção.

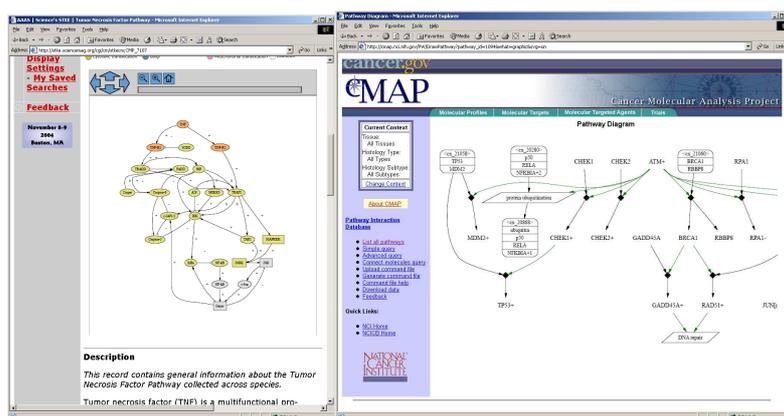


Figura 2.22. Graphvis

O Graphviz [63] é um software de código aberto para a visualização de grafos. Ele inclui programas para criar vários *layouts* de redes/grafos. Ele também possui uma interface gráfica (fig. 2.22) interactiva, ferramentas auxiliares e bibliotecas. Os programas para desenho de *layouts* do Graphviz utilizam descrições textuais dos grafos para criar vários diagramas em diferentes formatos, além disso, é possível alterar atributos visuais como a cor, forma dos nós, estilos de linhas, entre outros.

A JUNG [46] é uma biblioteca de código aberto (escrita em JAVA) que funciona como uma *framework* comum para análise e visualização de grafos e/ou redes. Para a visualização, a *framework* permite uma fácil construção de ferramentas interactivas para a exploração dos dados. Vários algoritmos estão disponíveis para desenho dos grafos segundo diferentes *layouts*, além de haver a possibilidade que o utilizador crie o seu próprio. A JUNG inclui também mecanismos para a filtragem, permitindo que o utilizador ou os algoritmos utilizados, foquem apenas uma determinada parte do grafo. Não possui interface gráfica, como no caso do BOOST.

2.8. Workspace

A visualização *workspace* é constituída por dois tipos de classes: (1) a tentativa de obter mais área de trabalho no ecrã do computador de uma forma organizada e (2) a tentativa de modificar o ambiente de trabalho com nova informação. Na primeira, não há inclusão de nova informação na visualização da *workspace*, porém, uma tentativa de organizar melhor esta de maneira a facilitar a visualização imediata de dados de interesse. Na segunda classe, o objectivo é adicionar mais e novos tipos de informação, como ocorre no caso da realidade virtual e/ou conferência virtual.

Como objectivos básicos na visualização de dados desta natureza temos:

- Permitir que o utilizador interaja com o ecrã do computador de forma mais eficiente;
- Permitir que pessoas fisicamente remotas entre si possam colaborar e interagir de forma concorrente;
- Sintetizar informação, conhecimento e resultados para a criação de soluções de alta qualidade;
- Organizar, interagir e executar pesquisas de informação relacionadas com tarefas de uma forma eficaz;
- Permitir um rápido acesso e reestruturação de informação pertinente a tarefas.

Exemplos de Projectos e Ferramentas

O ILOG Discovery [64] é uma ferramenta potente de visualização, que além de permitir uma manipulação fácil de janelas – não permite a sobreposição e minimiza (ícone) automaticamente, é possível ter acesso a um número ilimitado de vistas, que inclui uma série de potencialidades que o classificariam também em outras categorias de visualização (fig. 2.23). Ela permite, entre outras coisas: suporte a vários formatos de dados, visualizações de histogramas, distribuições, grafos, mapas de dados e hierarquias, análise paralela, navegação e execução de macros. Utiliza como base conceptual o modelo de visualização de dados linear.

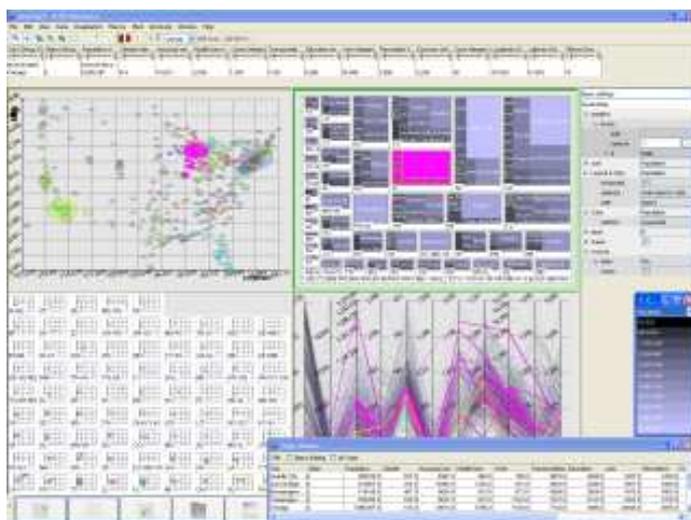


Figura 2.23. ILOG Discovery

O *Improvise* [65] possui três níveis distintos de funcionalidade: *browser* (navegação e selecção), *construção* (edição de pesquisas coordenadas e de *layouts*) e *meta visualização* (visualização *in situ* de coordenadas e outras estruturas). Ele suporta vistas hierárquicas utilizando uma variedade de gestores de *layout* (equivalente ao do JAVA AWT), utiliza editores gráficos para vistas, pesquisas e *layouts*, cria, elimina, modifica e interliga dinamicamente todos os objectos e salva/lê uma visualização armazenada em documento XML. É uma aplicação auto-suficiente que mostra uma ou mais visualizações, cada uma, numa janela própria. Cada janela contém um painel *desktop* com multi-páginas (figura 2.24). É também um caso que poderia ser classificado em outra categoria, além desta.

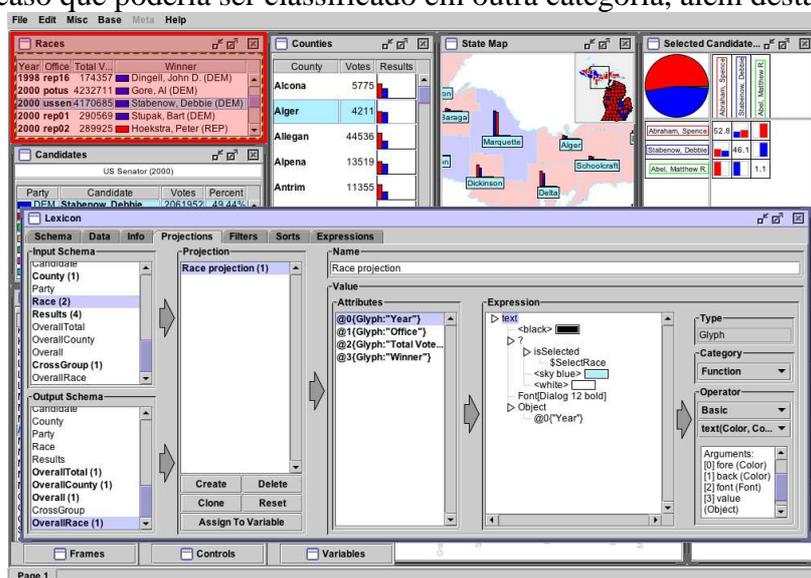


Figura 2.24. Improvise

A análise do *web server log* permite que os gestores de sites obtenham informação sobre os utilizadores de um site. Com a aplicação de multi-visualizações [56] coordenadas dos dados da *web log*, com granularidade diferente e complementadas com informação contextual, é possível melhorar bastante esta análise. Hochheiser [8] propôs uma aplicação (figura 2.25) onde a utilização da técnica STV (*Snap-Together-Visualization*) desempenha papel fundamental para a análise desse tipo de informação.

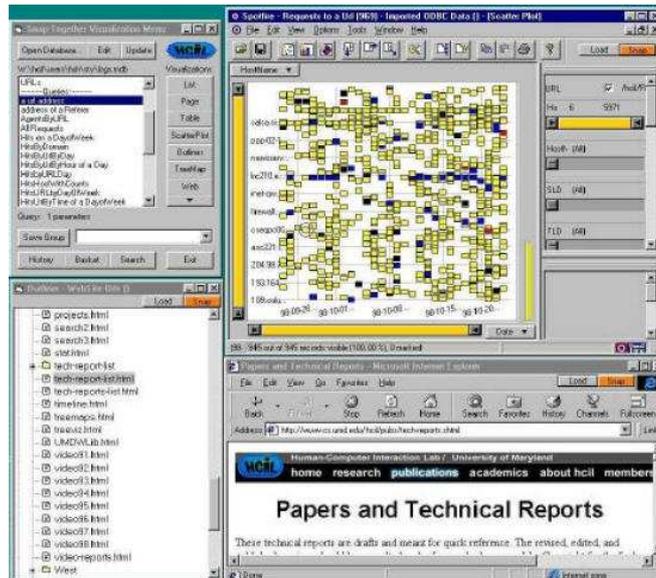


Figura 2.25. Análise da Web log data

O InfoVis Toolkit [49] é uma ferramenta gráfica interactiva escrita em Java que permite o desenvolvimento de aplicações e componentes para a visualização da informação. A estrutura de dados básica é uma tabela com colunas, sendo que cada coluna contém objectos homogéneos. As árvores e grafos são derivados a partir das tabelas. Permite a filtragem interactiva (pesquisas dinâmicas) de dados, pois se baseia num conjunto unificado de componentes. Além disso, oferece uma grande variedade de técnicas de visualização de (fig. 2.26) informação.

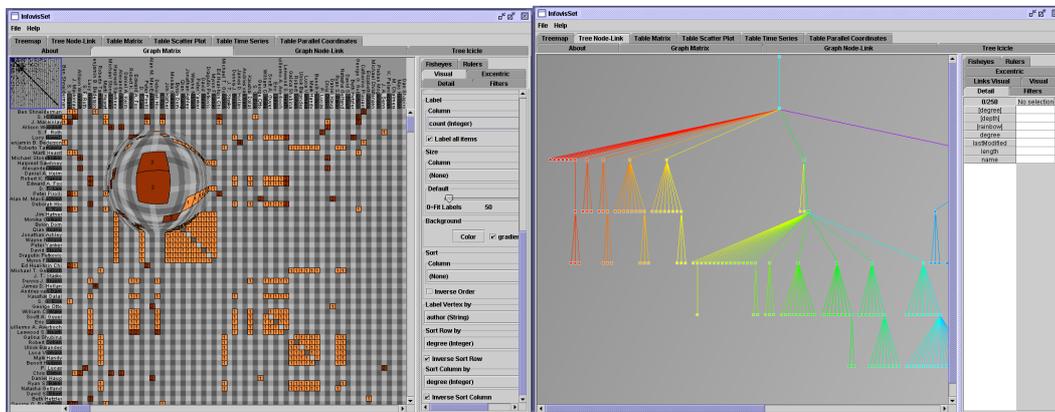


Figura 2.26. InfoVis Toolkit

O Spotfire [57, 37] é um sistema de visualização que inclui uma vasta gama de ferramentas para visualizar e analisar diferentes tipos de informação. É um ambiente multi-janelas, que utiliza várias técnicas de visualização, gráficos, metáforas, etc. Pode ser utilizado para vários propósitos e graças a sua API (*Application Programming Interface*), permite facilmente a sua extensão (fig. 2.27). É mais um caso que pode pertencer a mais de uma categoria.

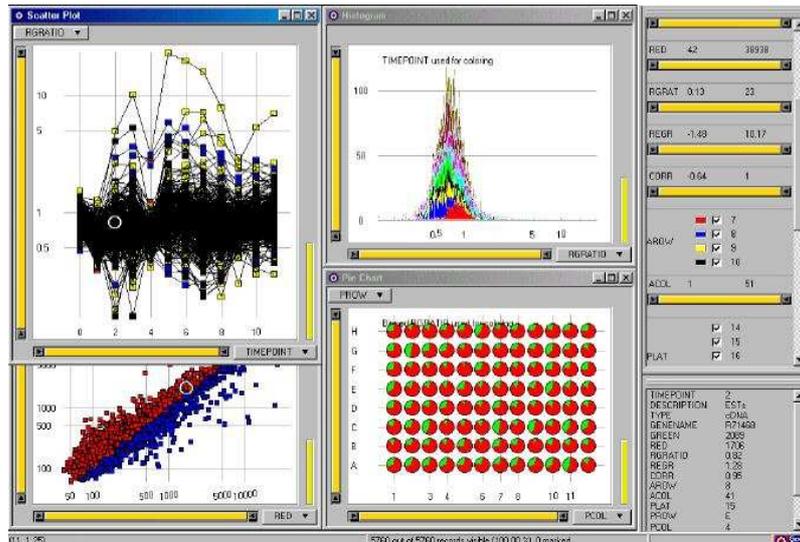


Figura 2.27. SpotFire

2.9. Temporal

A visualização de informação temporal possui duas qualidades fundamentais que a distinguem de simples listas padrão: (1) o facto de que os eventos temporais poderem ser simultâneos ou se sobreporem e (2) as múltiplas escalas subjacentes a dados temporais poderem exigir medições com alta precisão como mais grosseiras (desde segundos até milhares de séculos, por exemplo).

Alguns exemplos de visualizações temporais são as representações em vídeo e as animações. A ideia básica é mostrar os dados dinamicamente ao longo do tempo, permitindo ao utilizador analisar as variações que ocorrem, alterar a ordem temporal ou focar um determinado momento em especial, por exemplo. Neste tipo de visualização, é facilitada a observação de padrões comportamentais nos dados.

Como objectivos básicos na visualização de dados desta natureza temos:

- Visualizar eventos ou dados em sequência e/ou variando a ordem sequencial;
- Visualizar e analisar historicamente a evolução de dados;
- Encontrar inconsistências e/ou relações indesejáveis nos eventos ou dados.

Exemplos de Projectos e Ferramentas

Timesearcher [30] é um protótipo desenvolvido com o objectivo de permitir a pesquisa interactiva e a exploração em séries temporais. As *timeboxes* (Fig. 2.28) são utilizadas para criar pesquisas – é uma metáfora gráfica de manipulação directa que permite definir pesquisas nos conjuntos de séries temporais, além de auxiliar na triagem da informação. As *timeboxes* suportam formulação e modificação interactiva de pesquisas. O protótipo inclui janelas para a pesquisa e visualização de resultados. Este protótipo foi desenvolvido com o Piccolo Toolkit [45]. Alek Aris et al. desenvolveram quatro métodos para avaliar séries temporais com distribuições não uniformes no Timesearcher [31].

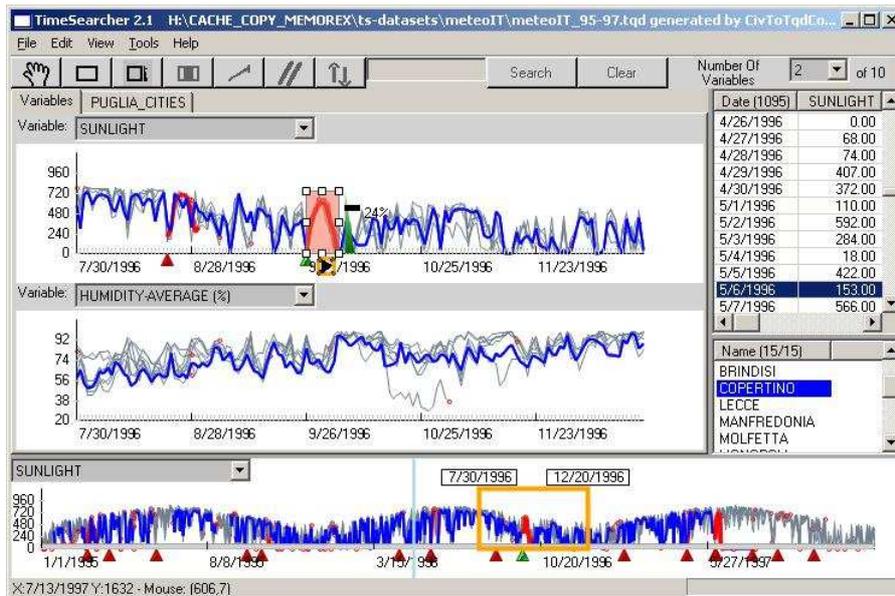


Figura 2.28. Timesearcher

Adam Perer [15] desenvolveu uma visualização interactiva que ajuda a percepção do contexto de discussões por e-mail. A visualização permite aos utilizadores acompanharem uma discussão ao longo do tempo, visualizando a lista de intervenientes em cada momento. A extensão (fig. 2.29) da linha de tempo no eixo horizontal fornece a duração de tempo decorrido desde a entrada da mensagem de e-mail na linha de discussão. Os pontos mais a esquerda, são as mensagens mais antigas, ocorrendo o oposto nas que estão situadas a direita.

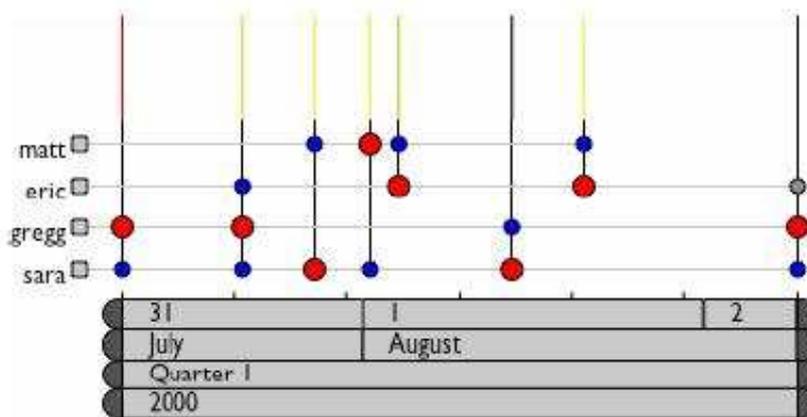


Figura 2.29. Discussão de e-mails

FishCal [17] é uma interface em forma de calendário para PDA. Ela permite o planeamento e análise de tarefas utilizando representações visuais do tipo *fisheye* das datas e da informação associada. Desta forma, é possível se obter desde visualizações genéricas como compactas de toda a informação presente no calendário. Da mesma forma, o utilizador consegue facilmente navegar e procurar informação, além de descobrir padrões e valores (fig. 2.30) limites.

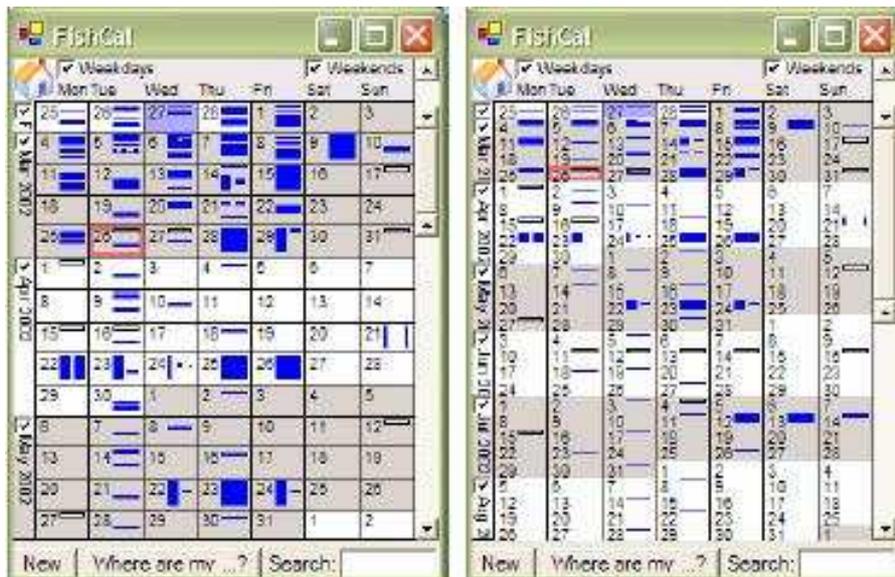


Figura 2.30. FishCal

Lifelines [23] é uma ferramenta para visualização temporal de informação, tendo sido já utilizada em diversas áreas. Chanda Harris e al.[22] utilizaram esta ferramenta para visualizar casos legais na justiça. O objectivo foi o de explorar visualmente a ligação existentes entre o caso original e as histórias directas e indirectas afins. Por exemplo, se o utilizador desejar saber como as regras e estatutos foram modificados ao longo de um determinado caso, ele pode analisa-los numa única visualização. Com a utilização da linha do tempo, o utilizador também pode seleccionar em qual momento temporal deseja começar a analisar o caso (figura 2.31).

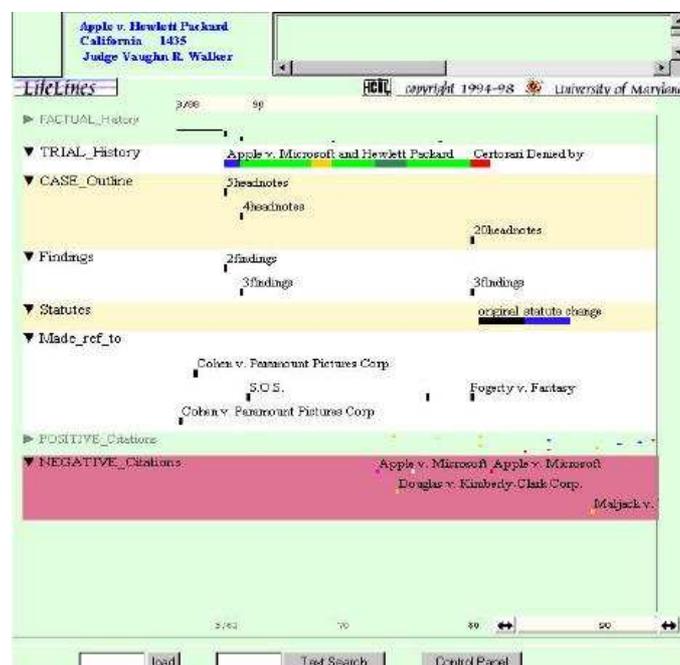


Figura 2.31. Lifelines em casos legais

2.10. Frameworks

O IVC [48] (*Information Visualization Cyberinfrastructure*) é uma *framework* para visualização da informação do tipo *plug-in*. Cada componente de software que faz parte do IVC pode ser “ligado” ou “desligado”. Desta forma, novos tipos de algoritmos, além de estruturas de dados, métodos, interfaces e até mesmo ferramentas inteiras podem ser facilmente integrados. A *framework* do IVC pode ser subdividida nos seguintes componentes: IVC *core*, modelo de dados, *plug-ins*, interface gráfica, armazenamento de dados, *scheduler* e *logger*. Todos os componentes, com excepção da *core*, podem ser “ligadas” ou “desligadas” consoante a análise de dados que se deseje fazer. Inclui actualmente vários algoritmos de análise, interacção, pesquisa, modelação, etc. O IVC foi implementado em J2SDK1.4.

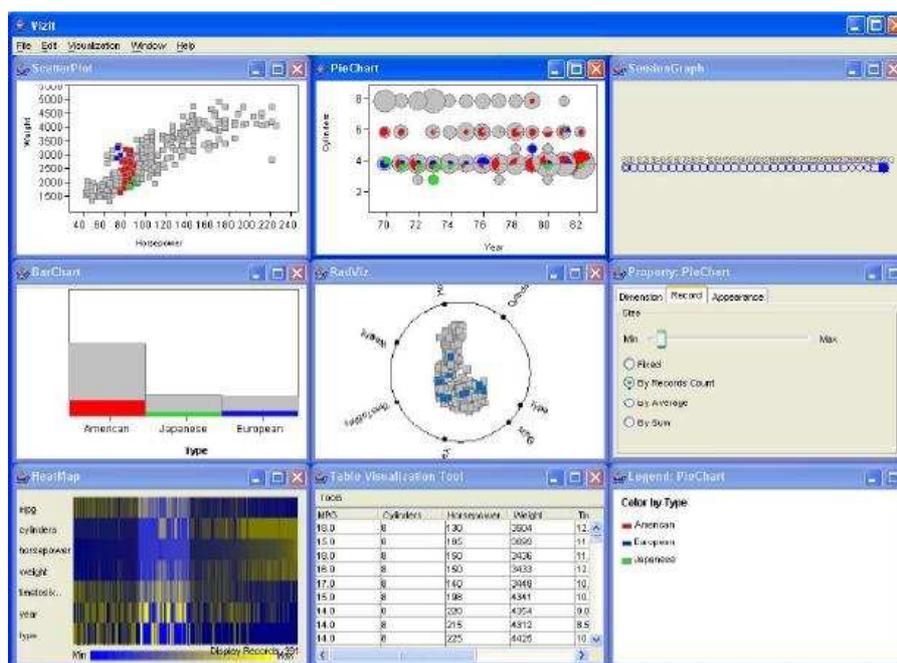


Figura 2.32. Vizit, construído a partir da UVP

A UVP [53] (*Universal Visualization Platform*) é a implementação de uma *framework* comum, a partir da qual várias visualizações de dados e aplicações para análise podem ser construídas. Esta *framework* tem como objectivo principal o desenvolvimento e partilha das ferramentas de visualização. Ela permite alcançar uma grande funcionalidade através de um modelo de dados comum. Outras ferramentas podem ser “ligadas” a *framework*, sendo o código automaticamente identificado e carregado dinamicamente. Ela é composta de quatro componentes principais agrupadas dentro de uma única API. Na figura abaixo (fig. 2.32), um exemplo de aplicação feita com base na UVP, a Vizit.

A Piccolo [45] suporta o desenvolvimento de programas estruturados para 2D, em geral, e de interfaces para o utilizador passíveis de zoom (ZUIs), em particular. A ZUI é um novo tipo de interface que permite que o utilizador execute o zoom da visualização resultante suavemente e progressivamente (para aproximar e afastar). A interface com os objectos que compõem esta ferramenta é feita com base num grafo. O utilizador pode (fig. 2.33) combinar, orientar e manipular os objectos que compõem a ferramenta, mantendo sempre o senso hierárquico existente entre eles. Como nos demais casos, a ocupação da área do ecrã é maximizada.

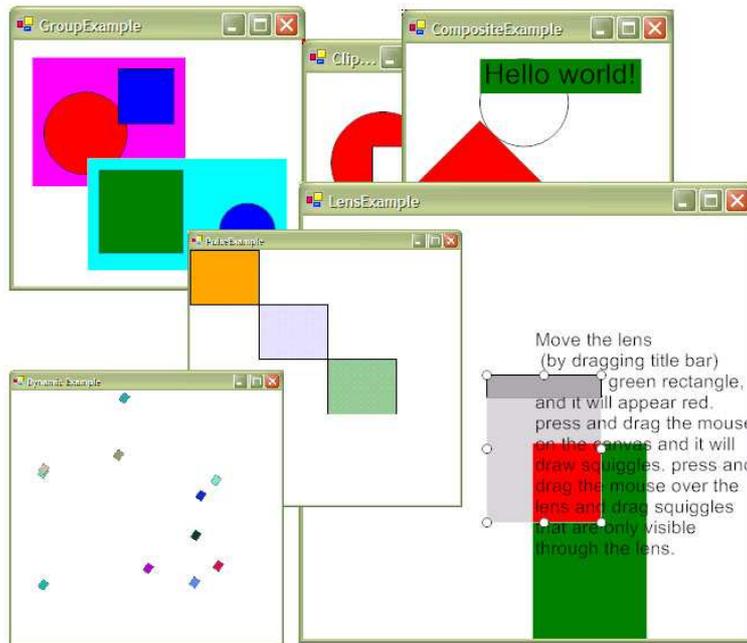


Figura 2.33. Piccolo

A Prefuse [47] é uma plataforma para construção interactiva de visualizações, que podem ser aplicadas a várias áreas. A arquitectura suporta vários tipos de interacção (selecção, focagem, reposição, zoom, *panning* e *brushing*), técnicas de visualização (animação, algoritmos de distorção e para *layout*), indexação e pesquisa de dados. A Prefuse foi construída totalmente em JAVA, utilizando a biblioteca gráfica 2D e o IDE Eclipse.

VTK 5.0 (*Visualization Toolkit*) [131] é uma ferramenta de código aberto gratuita para a computação gráfica tridimensional, o processamento de imagem e a visualização. Ela consiste numa biblioteca de classes em C++ e várias camadas de interpretadores incluindo Tcl/Tk, Java e Python. A VTK suporta uma grande variedade de algoritmos de visualização incluindo escalar, vector, tensor, textura, métodos volumétricos, além de métodos avançados de técnicas de modelação. O desenho e implementação desta biblioteca são orientados a objectos.

2.11. Conclusões

Uma infra-estrutura orientada para a Visualização de Informação deve ser composta por vários níveis, existindo um bom nível de comunicação entre eles. No nível mais próximo do utilizador, a visualização da informação requer um ambiente interactivo, onde várias visualizações podem ser criadas, combinadas, sincronizadas e configuradas rapidamente, de forma a potenciar a rica exploração de dados.

Para aceder a conjuntos de dados é exigido mais outro nível de infra-estrutura. Este deve garantir a leitura de ficheiros em diferentes formatos, provenientes de variadas fontes. Construir o algoritmo de visualização com técnicas de visualização apropriadas é o terceiro nível da infra-estrutura, sendo aconselhável a utilização de sistemas de visualização sempre que possível (evitar tempo/esforço em programação/desenvolvimento). Também é importante ter em mente a possibilidade de adicionar novas técnicas sempre que necessário, pois isto é conveniente, tanto a nível de utilizador final como a nível de programação.

A combinação do acesso e processamento a dados e de técnicas de visualização deve ser fácil e visual. A visualização pode ser *front-end* para a aplicação principal ou pode ser a aplicação em si. Ela também pode ser uma componente dentro de uma aplicação já existente. A infra-estrutura deve ser flexível o bastante de forma a se integrar a uma aplicação já existente, sem grande dificuldade, além de garantir vários níveis de acesso, dependendo do perfil do utilizador final. Nenhum tipo de programação deve nunca ser solicitado a utilizadores finais.

A PERIODIC TABLE OF VISUALIZATION METHODS

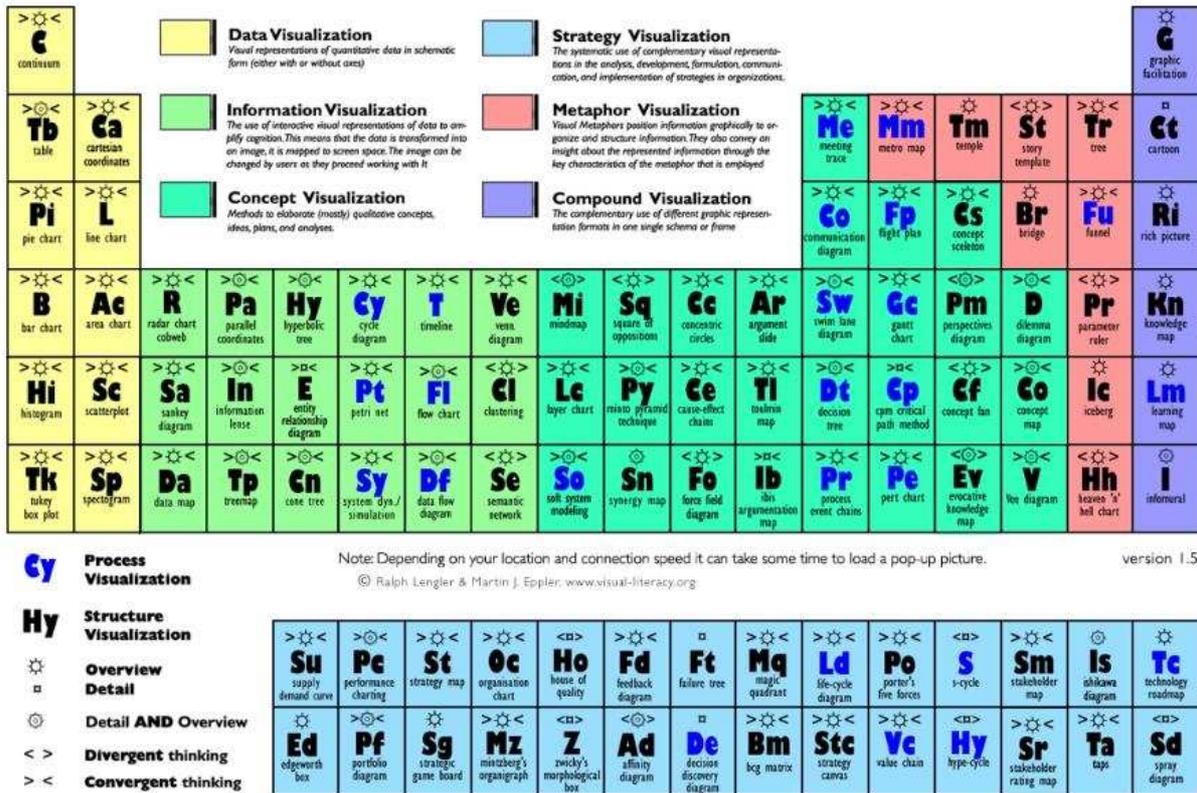


Figura 2.34. A periodic Table of Visualization Methods [134]

Bibliografia

- [1] "From Concepts to Data and Back Again: Connecting Mental Spaces with Data and Analysis Methods", Mak Gahegan et al., Proceedings of the 7th International Conference on GeoComputation, Setembro-2003, Southampton, UK.
- [2] "Information and Scientific Visualization: Separate but Equal or Happy Together at Last", Theresa Rhyne, IEEE Visualization 2003, Outubro-2003, pgs. 611-614.
- [3] "Frameworks in Support of InfoVis/Perceptualization", Stephen North et al., Proceedings of IEEE Symposium on Information 2004 Visualization (InfoVis), Outubro-2004.
- [4] "Visualizing Sequential Patterns for Text Mapping", Pak Wong et al., Proceedings IEEE Information Visualization 2000, Outubro-2000, Salt Lake City, Utah.
- [5] "Navigating Large Bodies of Text", D. Small, IBM Systems Journal, vol. 35, n^o 3 & 4, 1996.
- [6] "Extending the Utility of Treemaps with Flexible Hierarchy", Gouthami Chintalapani et al., Proceedings 8th International Conference on Information Visualization, 2004, pgs. 335-344.
- [7] "Topic Islands TM - a Wavelet-Based Text Visualization System", Nancy Miller et al., Proceedings of the Conference on Visualization '98, 1998, pgs. 189-196.
- [8] "Coordinating Overviews and Detail Views of WWW Log Data", Harry Hochheiser et al., Proceedings of New Paradigms in Information Visualization and Manipulation, Novembro-2000, Washington D. C., E.U.A.
- [9] "Dynamic Queries and Brushing on Choropleth Maps", Gunjan Dang et al., Proceedings of the 5th International Conference on Information Visualization, 2001, pgs. 757-761.
- [10] "Information Visualization and the Challenge of Universal Usability", Catherine Plaisant, Exploring Geovisualization, Ed. Elsevier, 2005.
- [11] "Knowledge Discovery in High Dimensional Data: Case Studies and a User Survey for an Information Visualization Tool", Jinwook Seo e Ben Shneiderman, The Institute for Systems Research, Technical Report n^o 2005-100, 2005.
- [12] "Inventing Discovery Tools: Combining Information Visualization with Data Mining", Ben Shneiderman, Proceedings of Discovery Science 2001 Conference, Novembro-2001, Washington DC, E. U. A.
- [13] "Domain Name Based Visualization of Web Histories in a Zoomable User Interface", Rajiv Gandhi et al., Proceedings of 11th International Workshop on Database and Expert Systems Applications, Setembro-2000, London, UK
- [14] "Dynamic Query Choropleth Maps for Information Seeking and Decision Making", Kent Norman et al., Proceedings of 10th International Conference on Human-Computer Interaction, 2003.
- [15] "Beyond Threads: Identifying Discussions in Email Archives", Adam Perer e Ben Shneiderman, Proceedings of InfoVis 05, Outubro-2005.
- [16] "The Challenge of Information Visualization Evaluation", Catherine Plaisant, Proceedings of Advanced Visual Interfaces 2004, Maio-2004. pgs. 109-116.
- [17] "A Fisheye Calendar Interface for PDAs: Providing Overviews for Small Displays", Benjamim Bederson et al., Proceedings of Conference on Human Factors in Computing Systems, 2003, pgs. 618-619.
- [18] "Image-based Highly Interactive Web Mapping for Geo-referenced Data publishing", Haixia Zhao e Ben Shneiderman, The Institute for Systems Research, Technical report n^o TR 2005-35, Dezembro-2002.
- [19] "A Knowledge Integration Framework for Information Visualization", Jim Seo e Ben Shneiderman, Springer LNCS, vol. 3379, Janeiro-2005.

- [20] "CounterPoint: Creating Jazzy Interactive Presentations", Lance Good e Ben Bederson, Technical Report HCIL-TR-2001-03, Março-2001.
- [21] "Broadening Access to Large Online Databases by Generalizing Query Previews", Egemen Tanis et al., Technical Report HCIL-TR-2000-14, Maio-2000.
- [22] "Temporal Visualization for Legal Case Histories", Chanda Harris et al., Technical Report HCIL-TR-1999-18, Junho-1999.
- [23] Lifelines, <http://www.cs.umd.edu/hcil/lifelines/> (visitado em Fevereiro de 2009)
- [24] "Interactive Information Visualization of a Million Items", Jean-Daniel Fekete e Catherine Plaisant, Proceedings of IEEE Symposium on Information Visualization 2002, Outubro-2002, pgs. 117-124.
- [25] "Visualizing Missing Data: Classification and Empirical Study", Cyntica Eaton et al., Proceedings of INTERACT 2005, Setembro-2005.
- [26] "Navigation Patterns and Usability of Overview+Detail and Zoomable User Interfaces for Maps", Kasper Hornbaek et al., ACM Transaction on Computer-Human Interaction, volume 9, Dezembro-2002, pgs. 362-389.
- [27] "Ordered Treemap Layouts", Ben Shneiderman e Martin Wattenberg, Proceedings of INFOVIS 2001, Outubro-2001.
- [28] "Data Exploration with Paired Hierarchical Visualizations: Initial Designs of Pair Trees", Bill Kules et al., Proceedings of 2003 National Conference Government Research, Setembro-2003.
- [29] "txtkit – Visual Text MiningTool", <http://sourceforge.net/projects/textkit> (visitado em Fevereiro de 2009).
- [30] "Interactive Pattern Search in Time Series", Paulo Buono et al., Proceedings of Conference on Visualization and data Analysis", Janeiro-2005, pgs. 175-186.
- [31] "Representing Unevenly-Spaced Time Series Data for Visualization and Interactive Exploration", Aleks Aris et al., Proceedings of INTERACT 2005, Springer, pg. 835-846.
- [32] "Browsing Large Online Data Tables Using Generalized Query Previews", Egemen Tanin et al., Proceedings of the NPIVM, Novembro-2000.
- [33] "A Rank-by-Feature Framework for Interactive Exploration of Multidimensional Data", Jinwook Seo e Ben Shneiderman,
- [34] "A Rank-by-Feature Framework for Unsupervised Multidimensional Data Exploration Using Low Dimensional Projections", Jinwook Seo e Ben Shneiderman, Proceedings of InfoVis 2004, Outubro-2004, pgs. 65-72.
- [35] "Dynamic Query Visualizations on World Wide Web Clientes: A DHTML Solution for Maps and Scattergrams", Evan Golub e Ben Shneiderman,
- [36] "SpaceTree:Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation", Catherine Planisant et al., Proceedings of InfoVis 2002, Outubro-2004, pgs. 57-64.
- [37] "Interfaces for Visualizing Multi-Valued Attributes: Design and Implementation Using Starfield Displays", Laurent Cailliteau, HCIL Technical Report nº 99-20, Setembro-1999.
- [38] "Browsers with Changing Parts: a Catalog Explorer for Philip Glass' Website", Harry Hochheiser, Proceedings of the Conference on Designing Interactive Systems: processes, practices, methods and techniques, Junho-2000, pgs. 105-115.
- [39] "TaxonTree: Visualizing Biodiversity Information", Cynthia Parr et al., HCIL Technical Report nº 2003-40, 2003.
- [40] "Visualizations for taxonomic and phylogenetic trees", Cynthia Parr et al., Bioinformatics, volume 20, Oxford University Press, Junho-2004.

- [41] "ThemeRiver: Visualizing Thematic Changes in Large Document Collections", Susan Havre et al., IEEE Transactions on Visualization and Computer Graphics, volume 8, nº1, Janeiro-Março 2002.
- [42] "A User Interface for Coordinating Visualizations based on Relational Schemata: Snap Together Visualization", Christopher Loy North, Dissertação para obtenção do grau de doutor em Filosofia, Universidade de Maryland, Departamento de Ciências da Computação, 2000.
- [43] "Dynamic Query Tools for Time Séries Datas Sets: Timebox Widgets for Interactive Exploration", Harry Hochheiser e Ben Shneiderman, Information Visualization (2004) 3, 1-18. Palgrave Macmillian Ltd., 2004.
- [44] "Visualization and Analysis of Microarray and Gene Ontology Data with Treemaps", Eric Baehrecke et al., BMC Bioinformatics 2004, volume 5, Junho-2004.
- [45] Piccolo Toolkit, <http://www.cs.umd.edu/hcil/piccolo/> (visitado em Fevereiro de 2009).
- [46] JUNG – Java Universal Network Graph Framework, <http://jung.sourceforge.net/> (visitado em Fevereiro de 2009).
- [47] Prefuse Toolkit, <http://prefuse.sourceforge.net/> (visitado em Fevereiro de 2009).
- [48] IVC Software Framework, <http://iv.slis.indiana.edu/sw/> (visitado em Fevereiro de 2009).
- [49] InfoVis Toolkit, <http://ivtk.sourceforge.net/> (visitado em Fevereiro de 2009).
- [50] Snap-Together Visualization Toolkit, <http://www.cs.umd.edu/hcil/snap/> (visitado em Fevereiro de 2009).
- [51] Wilma Scope 3D Graph Visualization System, <http://wilma.sourceforge.net/> (visitado em Fevereiro de 2009).
- [52] VxInsightII, <http://www.cs.sandia.gov/projects/VxInsight.html> (visitado em Dezembro de 2006).
- [53] UVP – Universal Visualization Platform, <http://vw.indiana.edu/ivsi2004/grinstein/grinstein.pdf> (visitado em Fevereiro de 2009).
- [54] "Visualizing Multi-Dimensional Data", Stephen Eick, ACM SIGGRAPH, volume 34, nº 1, Fevereiro-2000.
- [55] "Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos", Ben Shneiderman e Hyunmo Kang, Proceedings of the International Conference on Information Visualization, 2000, pg. 88-90.
- [56] "Using Interactive Visualizations of WWW Log Data to Characterize Access Patterns and Inform Site Design", Harry Hochheiser e Ben Shneiderman, Journal of the American Society for Information Science Technology, Volume 52, nº 4, 2001, pgs. 331-343.
- [57] SpotFire, <http://spotfire.tibco.com/index.cfm> (visitado em Fevereiro de 2009).
- [58] ILOG, <http://www2.ilog.com/preview/Discovery/> (visitado em Fevereiro de 2009).
- [59] Fished Eyes Menus, <http://www.cs.umd.edu/hcil/fisheyemenu/> (visitado em Fevereiro de 2009).
- [60] Visible Human Project, <http://www.nlm.nih.gov/research/visible/> (visitado em Fevereiro de 2009).
- [61] "The WebBook and Web Forager: an information workspace for the WWW", Stuart Card et al., Proceedings of SIGCHI conference on Human factors in computing systems: common ground, ACM press, Abril-1996.
- [62] The Boost Graph Library, <http://www.boost.org/libs/graph/doc/index.html> (visitado em Fevereiro de 2009).
- [63] Graphvis, <http://www.graphviz.org/> (visitado em Fevereiro de 2009).
- [64] ILOG Discovery, <http://www2.ilog.com/preview/Discovery/> (visitado em Fevereiro de 2009).
- [65] Improvise, <http://www.personal.psu.edu/cew15/improvise/index.html> (visitado em Fevereiro de 2009).

- [66] "Interactive Simulation and Visualization", Christopher Johnson et al., IEEE – Computer, volume 32, nº 12, pg. 59-65, 1999.
- [67] "Case Study: WildFire Visualization", Ahrens, J.; McCormick, P.; Bossert, J.; Reisner, J.; Winterkamp, J., Visualization apos;97., Proceedings Volume 24 , no. 24, Outubro/1997 Pgs. 451 – 454.
- [68] "Guest Editors' Introduction: Visualization", Lawrence J. Rosenblum e Bruce E. Brown, IEEE - CG&A, pp. 18-20, volume 12, nº4, Julho/1992.
- [69] "The Generalized Detail-In-Context Problem", T. Alan Keahey, Advanced Computing Laboratory, Los Alamos National Laboratory, MMS B287, Los Alamos, NM 87545
- [70] "The Application Visualization System: A Computational Environment for Scientific Visualization", Craig Upson et al., IEEE - CG&A, volume 9, nº4, pp. 30-41, Julho/1989.
- [71] "Visualization Techniques", Scientific Visualization, K. Brodie, Eds. K. Brodie et al., Springer Verlag, 1992.
- [72] "Modeling of Scattered Multivariate Data", Gregory M. Nielson e Thomas A. Foley, pp. 38-55, Eds. C. Giertsen e P. Frevang, Eurographics Technical Report, 1994.
- [73] "Visualization of High-Dimensional Clusters Using Nonlinear Magnification", T. Alan Keahey, Advanced Computing Laboratory, Los Alamos National Laboratory, MMS B287, Los Alamos, NM 87545
- [74]] "User Interface Developments for the Nineties", Aaron Marcus e Andries Van Dam, IEEE CG&A, pp. 49-56, Setembro/1991.
- [75] "Research Issues in Data Modeling in Scientific Visualization", Gregory Nielson et al., IEEE - CG&A, pp. 70-72, volume 14, nº2, Março/1994.
- [76] "A Dataflow Toolkit for Visualization", D. Scott Dyer, IEEE - CG&A, volume 10, nº 4, pp. 60-69, Julho/1990.
- [77] "A Knowledge-Based System for Visualization Design", Hikmet Senay e Eve Ignatius, pp. 36-47, IEEE - CG&A, volume 14, nº 6, Novembro/1994.
- [78] "A Methodology for Choosing Data Representations", Philip K. Robertson, pp. 56-67, IEEE - CG&A, vol. 11, nº 3, Maio/1991.
- [79] "Techniques for Visualizing 3D Unstructured Meshes", Carole S. Gitlin & al., Departamento de Ciências da Computação, Universidade de Utah, Salt Lake City, UT 84112 USA, Janeiro/1995.
- [80] "Visualization of Scattered Meteorological Data", Lloyd A. Treinish, pp. 20-26, IEEE - CG&A, volume 15, nº 4, Julho/1995.
- [81] "New Visualization Techniques", Russel M. Taylor II, ACM SIGGRAPH, vol. 34, nº 1, Fevereiro/2000.
- [82] "Introductory Guide to Scientific Visualization", R. A. Earnshaw e N. Wiseman, Springer Verlag, 1992.
- [83] "Color versus Black and White in Visualization", Richard A. Holub et al., IEEE - CG&A, pp. 20-22, volume 12, nº 4, Julho/1992.
- [84] "Terrain Visibility", George Nagy, pp. 763-773, Computer & Graphics, volume 18, nº 6, Eds. Elsevier Science Ltda., 1994.
- [85] "Using Color Effectively in Computer Graphics", Lindsay W. MacDonald, IEEE – Computer Graphics and Applications, Julho/Agosto 1999, pg. 20- 34.
- [86] "Visualizing High-Resolution Accelerator Physics", Patrick S. McCormick al., IEEE – Computer Graphics and Applications, volume 19, nº 5, Setembro/Outubro 1999, pg. 11- 13.
- [87] "Perceptual Guidance for Visualization Design", Eve Ignatius et al., Thomson Technology Services Group, 1996.
- [88] "Interactive Visualization of Earth and Space Science Computation", William Hibbard, David A. Santek, M. Martinez, C. Dyer, A. Battaiola e Brian Paul, IEEE - Computer, pp. 65-72, vol. 27, nº 7, Julho/94.
- [89] "Scientific Visualization in the Next Millenium", Theresa Marie Rhyne, IEEE – Computer Graphic and Applications, volume 20, nº1, Janeiro/Fevereiro 2000.
- [90] "Web-Based Information Visualization", Randall Rohrer e Edward Swing, Computer Graphics IEEE, Julho/Agosto 1997.
- [91] "Texture Mapping for Scientific Visualization Environments", Fabrizio Cuccu e Laura Moltedo, IEEE - CG&A, volume 17, nº 2, pp. 131-136, 1993.
- [92] "Information Visualization Applications in the Real World", Nahym Gershon e Stephen Eick, Computer Graphics IEEE, Julho/Agosto 1997.

- [93] "Zooming of Terrain Imagery Using Fractal-Based Interpolation", M. A. Pumar, pp. 171-176, Computer & Graphics, volume 20, nº 1, Eds. Elsevier Science Ltda, 1996.
- [95] "Theories of Visual Perception", I. E. Gordon, John Wiley & Son, 1989.
- [96] "Visualizing Multi-Dimensional Data", Stephen Eick, SIGGRAPH Computer Graphics Newsletter, february/2000.
- [97] "HD-Eye: Visual Mining of High Dimensional Data", Alexander Hinneburg al., IEEE – Computer Graphics and Applications, volume 19, nº 5, Setembro/Outubro 1999, pg. 22- 31.
- [98] "Fourteen Ways to Say Nothing with Scientific Visualization", Al Globus e Eric Raible, IEEE - Computer, pp. 86-87, vol. 27, nº 7, julho/1994.
- [99] "Scientific Visualization and Semiotic Morphisms", Joseph Gogien, <http://www-cse.ucsd.edu/users/goguen/papers/sm/vzln.html>, Junho/2000 (visitado em Fevereiro de 2009).
- [100] "On the Use of Perceptual Cues and Data Mining for Effective Visualization of Scientific Datasets", Christopher G. Healey, EECS Department, University of California at Berkeley, 1997.
- [101] "Two Stepping Information Technology with Visualization", Theresa Maria Rhyne e Lockheed Martin, ACM SIGGRAPH, volume 34, nº 1 Fevereiro/2000.
- [102] "Enhanced Line Integral Convolution with Flow Feature Detection", Arthur Okada e David Lane, NASA Ames Research Center, 1996.
- [103] "Selective Visualization of Vector Fields", Theo Walsum e Frits Post, pp. 339-347, Eurographics '94, Eds. Daehlen e Kjeldahl, Blackwell Publ., 1994.
- [104] "Understanding Quality in Conceptual Modeling", Odd Lindland, G. Sindre e A. Solvberg, pp. 42-49, IEEE - Software, vol. 11, nº 2, Março/1994.
- [105] "Visual Cues", Peter Keller e Mary Keller, IEEE Society Press, 1993.
- [106] "Interface Design When You Don't Know How", Marc Rettig, Communications of the ACM, pp. 29-34, vol. 35, nº. 1, Janeiro/1992.
- [107] "Using Sound to Extract meaning from complex Data", Carla Scaletti e Alan Craig, University of Illinois, 1991.
- [108] "Principles of Information Display for Visualization Practioners", Edward Tufte, NASA Ames Research Center, Novembro/1994.
- [109] "Tools and Techniques for Scientific Visualization", R. A. Earnshaw, New Trends in Animation and Visualization", Eds. Nadia Thalmann e Daniel Thalmann, 1991.
- [110] "Selected New Trends in Scientific Visualization", Werner Purgathofer e Helwig Löffelmann, Vienna University of Technology, 1997.
- [111] "Visualization Modeling: Making Visualization a Creative Discipline", Tosiyasu L. Kunni e Yoshihisa Shinagawa, New Trends in Animation and Visualization, Eds. Nadia Thalmann, Daniel Thalmann, Wiley, 1991.
- [112] "Applications of Multivariate Visualization to Behavioral Sciences", Chong Ho Yu e John T. Behrens, Arizon State University, 325 Payne Hall, ASU, Tempe AZ 85287-0611
- [113] "Glyphmaker: Creating Customized Visualizations of Complex Data", William Ribarsky et al., IEEE - Computer, volume 27, nº 7, pp. 57-64, Julho/1994.
- [114] "Rules and Principles of Scientific Data Visualization", Hikmet Senay e Eve Hignatius, The George Washington University, Agosto/1996.
- [115] "Real Time exploratory Scientific Visualization and Virtual Reality", Steve Bryson, Scientific Visualization Advances and Challenges, Academic Press, Ed. L. Roseblum et al., 1994.
- [116] "Feature Extraction and Iconic Visualization", Theo van Waltsum et al., pp. 111-119, IEEE - Transactions on Visualization and Computer Graphics, volume 2, nº 2, Junho/1996.
- [117] "Discovery Visualization Using fast Clustering", William Ribarsky et al., IEEE – Computer Graphics and Applications, volume 19, nº 5, Setembro/Outubro 1999, pg. 32- 39.
- [118] "Interdisciplinary Techniques, Toolkits and Models for Scientific Visualization", R. A. Earnshaw, pp. 1-21, Graphics Modeling and Visualization in Science and Technology, Eds. M. Gobel e J. C. Teixeira, Springer Verlag, 1993.
- [119] "Visualizing and Modeling Scattered Multivariate Data", Gregory M. Nielson, Thomas Foley, B. Hammann e D. Lane, IEEE - CG&A, pp. 47-54, vol. 11, nº 3, Maio/1991.
- [120] "Screen Design", Thomas S. Tullis, Handbook of Human Computer Interaction, North- Holland, 1990.

- [121] "Visualization Idioms: A conceptual Model for Scientific Visualization Systems", R. B. Haber e D. A. McNabb, Visualization in Scientific Computing, R. B. Haber e D. A. McNabb, Eds. G. Nielson e B. Shriver, IEEE Press, 1990.
- [122] "Volume Models for Volumetric Data", Vishwa Ranjan e Alain Fournier, IEEE - Computer, pp. 28-36, vol. 27, nº 7, Julho/1994.
- [123] "Interactive Simulation and Visualization", Christopher Johnson et al., IEEE – Computer, volume 32, nº 12, pg. 59-65.
- [124] "Sieve: A Collaborative Interactive Modular Visualization Environment", Philip Isenhour et al., Virginia Tech, 1998.
- [125] "Virtual Reality in Scientific Visualization", Steve Bryson, Computer & Graphics, pp. 679-685, volume 17, nº 6, 1993.
- [126] "Evaluation of Some New Techniques for Visualization of Highdimensional Dynamical Systems", Peter Steiger, www.cg.tuwien.ac.at/~steiger/eval/index.html, Maio/1998. (visitado em Fevereiro de 2009)
- [127] "MagicSphere: an insight tool for 3D data visualization", C. Montani e R. Scorpigno, pp. 317-328, Eurographics '94, Eds. Daehlen e Kjellidahl, Blackwell Publ., 1994.
- [128] "Collaborative Visualization", Ken Brodie et al., University of Leeds, 1997.
- [129] "Research Issues in Foundations of Visualization", Philip Robertson et al., IEEE - CG&A, pp. 73-76, volume 14, nº2, Março/1994.
- [130] "Visualization of Damaged Structures", M. Pauline Baker et al., NCSA, University of Illinois, Março/1988.
- [131] VTK – Visualization Toolkit, <http://www.vtk.org> (visitado em Fevereiro de 2009).
- [132] "Visualizing author attribution using Blobby objects", Mala, T.; Geetha, T.V. Computer Graphics, Imaging and Visualisation, Volume 14, no. 17, pgs.:460 – 464, Agosto/2007.
- [133] VisualComplexity.com, <http://www.visualcomplexity.com/vc/> (visitado em Fevereiro de 2009).
- [134] A Periodic Table of Visualization Methods, http://www.visual-literacy.org/periodic_table/periodic_table.html (visitado em Fevereiro de 2009).



Centro de Computação Gráfica

Publicações CCG

Elizabeth Simão Carvalho
Adérito Fernandes Marcos
(autores)

Publicação:
Associação CCG / ZGDV – Centro de Computação Gráfica
Campus de Azurém
4800-058 Guimarães
Portugal
tel.: +351 253 510 580
fax: +351 253 510 581
e-mail: info@ccg.pt
Relatório Interno CCG

Edição digital
ISBN 978-972-99062-5-1