



INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Visualização de Indicadores de Negócio no Contexto de um Sistema de Suporte à Decisão
Baseado em *Balanced Scorecard: Spatial Dashboard*

Filipe Manuel Homem Miguel das Neves

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática e de Computadores

Júri

Presidente: Alberto Manuel Rodrigues da Silva

Orientador: Mário Serafim dos Santos Nunes

Co-Orientador: Miguel Leitão Bignolas Mira da Silva

Vogal: José Luís Brinquete Borbinha

Setembro de 2007

Resumo

Decisores e gestores de topo recorrem cada vez mais a ferramentas e metodologias de Business Performance como o *Balanced Scorecard* para enfrentar uma competitividade crescente. Nos seus ambientes de trabalho, os tempos de resposta a situações críticas constituem um aspecto crucial. No entanto, as ferramentas utilizadas pela maioria dos decisores não suportam características elaboradas de interacção, ainda que o processamento gráfico avançado seja um recurso amplamente acessível. O hardware gráfico actual permitiria facilmente que aplicações de *Business Intelligence* apresentassem de forma automática diagramas interactivos em tempo real, embora tabelas e gráficos bidimensionais sejam as interfaces mais comuns nessas aplicações.

Ainda que se defina o problema de forma genérica, a pesquisa foi conduzida a partir de um protótipo existente de uma aplicação de suporte à decisão com suporte à dimensão espacial, o *Spatial Dashboard*.

Acreditando nas mais-valias de técnicas avançadas de visualização encontradas na sequência de uma pesquisa nesse domínio, propomos um paradigma de visualização tridimensional para explorar o fluxo da estratégia pela empresa e indicadores de *scorecard* dispostos numa rede cujas ligações representam o fluxo de estratégia. Este paradigma baseia-se em três vistas onde indicadores pertencentes à mesma unidade podem ser apresentados de forma agregada ou isolado.

Desenvolveu-se um protótipo validado utilizando um caso de estudo de um aeroporto e através da colaboração na pesquisa de Doutoramento em que se insere o *Spatial Dashboard*. Os respectivos pormenores de implementação estão documentados. A avaliação preliminar sugere que, além do contributo inerente à concretização de pesquisa na intersecção de domínios tipicamente disjuntos, o paradigma proposto pode suportar a comunicação da estratégia empresarial e criação de sinergias entre unidades de negócio.

1.1. Palavras-chave

Visualização, Indicadores, Negócio, Balanced Scorecard, Desempenho

Abstract

Decision makers and top managers are increasingly making use of Business Performance Tools and methodologies like the Balanced Scorecard to face growing competitiveness. In their work environments, response times to critical situations is a crucial issue. However the tools most of them are using don't support enhanced interaction features, even though advanced graphical processing is nowadays widely available. Modern graphic hardware would easily allow Business Intelligence applications to automatically prepare interactive real-time and feature-rich diagrams, yet tables and bidimensional graphics are the most common interfaces found on such applications.

While stating the problem generically, the research was driven around an existing prototype of a Spatially Enabled Business Performance tool, the Spatial Dashboard.

Believing in the advantages of enhanced visualization techniques which we found on a visualization domain research, we propose a tridimensional visualization paradigm for exploring scorecard indicators and the strategy flow across the enterprise. Indicators are shown as network, in which links portray the strategy flow. This paradigm is based on three different views in which indicators belonging to the same business unit or department are either shown aggregated as single indicator or unaggregated.

A prototype was developed and validated using an airport case study and by means of cooperation with the Spatial Dashboard PhD research work. It's implementation details are hereby presented. Preliminary evaluation suggests that besides contributing by straightening the gap between 3D graphics and business intelligence, this paradigm may play an important role in supporting strategy communication and enabling synergy creation among business units.

1.2. Keywords

Visualization, Business, Indicators, Balanced Scorecard, Performance

Agradecimentos

Gostaria de agradecer a todos os que contribuíram para a concretização deste trabalho. Foi cerca de um ano em que tive a oportunidade de colaborar com bastantes pessoas com quem nunca tinha trabalhado. Aprendi com todas elas e considero que foi um ano extremamente enriquecedor.

Começo por agradecer em particular ao Professor Miguel Mira da Silva, de quem eu pouco mais conhecia além das virtudes de comunicador cativante, à partida para este projecto. Apesar da constantemente preenchida agenda, conseguiu sempre dispensar algum do seu tempo para nos ajudar e contribuir com críticas de grande valor construtivo. Demonstrou grande acessibilidade, promoveu o contacto entre os seus orientandos e correspondeu às expectativas que um aluno pode ter em relação a um Professor Orientador. Foi um prazer poder beneficiar de tal suporte.

Devo um agradecimento especial ao Professor Gabriel Pestana pela forma como ajudou a equipa de desenvolvimento ao longo destes dois semestres dando sempre o seu contributo de forma coerente e inovadora, esforçando-se por inculcar valor científico ao projecto. As tarefas acordadas permitiram-me sempre aprender algo de novo pois constituíam sempre novos desafios. Conseguiu como ninguém manter a equipa motivada e confiante no trabalho que desenvolvemos.

Agradeço a todos os elementos da equipa de desenvolvimento: Luís de Carvalho Campos, com quem partilhei muitos desafios, com quem aprendi imenso e cujo contributo durante as tarefas de desenvolvimento e para a presente tese foi absolutamente determinante; Pedro Alves Gomes e Ricardo Gomes Perleques com quem trabalhar chegou a parecer lazer, tal a boa disposição transmitida. Foi um prazer trabalhar com esta equipa, com quem estou certo de manter o contacto a nível profissional e pessoal.

Um agradecimento muito merecido à minha família mais próxima, Ana e Francisco que sempre me apoiaram e respeitaram os momentos particularmente atarefados.

Quero finalmente agradecer de forma especial à minha namorada, Maria João, cuja ajuda foi fundamental durante este ano, em especial na fase final de escrita da tese, e com a qual sei que posso contar sempre, ainda que este trabalho tenha implicado abdicar da partilha de muitos momentos.

2. Índice

FILIFE MANUEL HOMEM MIGUEL DAS NEVES	I
RESUMO.....	II
1.1. Palavras-chave.....	ii
ABSTRACT.....	III
1.2. Keywords.....	iii
AGRADECIMENTOS.....	V
LISTA DE FIGURAS.....	VIII
LISTA DE ABREVIATURAS.....	X

1. INTRODUÇÃO.....	1
1.1. Objectivo da Tese.....	1
1.2. Problema.....	1
1.2.1. Contexto.....	1
1.2.2. Descrição do Problema.....	2
1.2.3. Conclusão.....	2
1.3. Organização da Tese.....	2
2. ESTADO DA ARTE.....	5
2.1. Visualização - Motivação.....	5
2.2. Fundamentos de Visualização Multidimensional.....	6
2.2.1. Terminologia.....	6
2.2.2. Domínio de Aplicação.....	7
2.3. Caracterização do Estado da Arte.....	7
2.4. Técnicas de Visualização.....	8
2.4.1. Matrizes de Diagramas de Dispersão (Scatterplot Matrices).....	8
2.4.2. Coordenadas Paralelas.....	11
2.4.3. Iconografia.....	13
2.4.4. VisDB.....	16
2.4.5. GrandTour.....	18
2.4.6. Técnicas Hierárquicas.....	19
2.5. Bubble Charts.....	22
2.6. Floors And Walls.....	23
2.7. Técnicas Auxiliares / Boas Práticas.....	24
2.7.1. Filtragem.....	24
2.7.2. Brushing.....	24
2.7.3. Banking.....	25
2.7.4. Outras Técnicas Auxiliares.....	25
2.8. Conclusão.....	Erro! Marcador não definido.
2.9. Balanced Scorecard.....	27
2.9.1. Motivação.....	27
2.9.2. O que é o Balanced Scorecard?.....	28
2.9.3. Balanced Scorecard como Sistema de Gestão Estratégica.....	29
2.9.4. Implementação de Balanced Scorecard.....	30
3. PROBLEMA.....	33
4. PROPOSTA.....	36
4.1. Overview do Spatial Dashboard.....	37
4.2. Módulo Rede de Dendrogramas – Conceitos de Suporte.....	39
4.2.1. Conjunto de técnicas escolhido:.....	40
4.2.2. Mapeamento: Domínio de Dados → Domínio Gráfico.....	40
4.2.3. Espaço Virtual Tridimensional.....	41

4.2.4.	<i>Vistas e Navegação</i>	42
4.3.	Considerações adicionais	43
5.	IMPLEMENTAÇÃO	45
5.1.	Arquitectura.....	45
5.2.	Funcionalidades.....	49
5.3.	Constrangimentos	71
6.	CASO DE ESTUDO	52
6.1.	Contexto	52
6.1.1.	<i>Unidade de Negócio com desempenho abaixo do pretendido</i>	56
6.1.2.	<i>Relação de Cross-Business</i>	58
6.2.	Resultados	59
7.	CONCLUSÃO.....	61
7.1.	Principais Contribuições.....	61
7.2.	Trabalho Futuro	62
8.	BIBLIOGRAFIA.....	64
9.	ANEXOS	67
9.1.	Planificação	67
9.2.	Práticas	67
9.2.1.	<i>Prototipagem Rápida</i>	67
9.3.	Escolha de <i>Frameworks</i>	68
9.3.1.	<i>OpenGL em ambiente .NET (Framework TAO)</i>	68
9.3.2.	<i>Bibliotecas de Layout de Grafos (Lithium, PIGALE, yFiles.NET)</i>	69
9.3.3.	<i>Motores 3D</i>	69

Lista de Figuras

FIGURA 1 - PROPÓSITO DA VISUALIZAÇÃO: MAPEAMENTO DO ESPAÇO DE DADOS PARA O ESPAÇO VISUAL [1]	6
FIGURA 2 - MATRIZ DE DIAGRAMAS DE DISPERSÃO COM TRÊS VARIÁVEIS X, Y E Z.....	9
FIGURA 4 - SCATTERPLOT MATRIX DA APLICAÇÃO XMDVTOOL [11] COM UM SUBCONJUNTO DE DADOS SELECIONADO E REPRESENTADO A VERMELHO	10
FIGURA 5 - SCREENSHOT DO SOFTWARE ILOG DISCOVERY [13] REPRESENTANDO DADOS ESTATÍSTICOS DE DIFERENTES CIDADES AMERICANAS COM COORDENADAS PARALELAS.....	11
FIGURA 6 - PREVENÇÃO DE COLISÕES.....	12
FIGURA 7 – PROPORCIONALIDADE INVERSA	12
FIGURA 8 - PROPORCIONALIDADE DIRECTA.....	13
FIGURA 9 - CARACTERÍSTICAS FACIAIS DE UMA CARA DE CHERNOFF [14].....	14
FIGURA 10 - MAPEAMENTO DE DEZ DIMENSÕES DE UM CONJUNTO DE DADOS GERADOS PELA FERRAMENTA WOLFRAM MATHEMATICA [15], PARA UM ARRAY DE CARAS DE CHERNOFF.	14
FIGURA 11 - ESQUERDA: FIGURA DE PAUZINHOS COM IDENTIFICAÇÃO DAS VARIÁVEIS DE ORIENTAÇÃO; DIREITA: FAMÍLIA DE FIGURAS DE PAUZINHOS.....	14
FIGURA 12 - IMAGEM DE STICK FIGURES REPRESENTANDO 7 VARIÁVEIS.....	15
FIGURA 13 - DADOS DE RESSONÂNCIAS MAGNÉTICAS DE UM TUMOR, COM DIFERENTES MAPEAMENTOS.....	15
FIGURA 14 - ESQUEMA DA REPRESENTAÇÃO DE RESULTADOS EM TORNO DO PONTO FOCAL, DESCRREVENDO UMA ELIPSE.....	17
FIGURA 15 - ESQUEMA LEGENDADO DA UTILIZAÇÃO DE UM DISPLAY PARA CADA DIMENSÃO	17
FIGURA 16 - EXEMPLO COM DADOS REAIS (7000 ITENS COM 8 DIMENSÕES) DA UTILIZAÇÃO DE UM DISPLAY PARA CADA DIMENSÃO, PARA ALÉM DO PRINCIPAL (O DO CANTO SUPERIOR ESQUERDO).....	18
FIGURA 17 - ESQUERDA: REFERENCIAL ORTOGONAL TRIDIMENSIONAL; DIREITA: REFERENCIAL HIERÁRQUICO.....	19
FIGURA 18 - REPRESENTAÇÃO DOS TUPLOS $\{(0,0,1), (0,1,0), (0,2,1)\}$. AS VARIÁVEIS X E Y SÃO AS VARÁVEIS INDEPENDENTES E Z A VARIÁVEL DEPENDENTE	19
FIGURA 19 - DOIS MAPEAMENTOS DIFERENTES PARA O MESMO ESPAÇO DE DADOS	20
FIGURA 20 - VISUALIZAÇÃO DE DIMENSION STACKING COM MAPEAMENTO DE VARIÁVEIS DEPENDENTES NA FORMA DE INTENSIDADE NUMA ESCALA DE CINZAS	21
FIGURA 21 - VISUALIZAÇÃO DE DIMENSION STACKING SÓ COM VARIÁVEIS INDEPENDENTES, SENDO QUE CADA REFERENCIAL BIDIMENSIONAL É UM DIAGRAMA DE DISPERSÃO	21
FIGURA 22 - WORLDS WITHIN WORLDS	22
FIGURA 23 - SCREENSHOTS DA FERRAMENTA AUTOVISUAL.....	22
FIGURA 24 - BUBBLE CHART COM VARIÁVEIS DE 4 DIMENSÕES (X,Y,FORMA/COR E TAMANHO), CONSTRUÍDO COM A FRAMEWORK DUNDAS [17].	23
FIGURA 25 - TÉCNICA FLOORS AND WALLS [18].....	24
FIGURA 26 – PERCENTAGEM DE EMPRESAS QUE CONCRETIZA A SUA VISÃO E ESTRATÉGIA [22].	28
FIGURA 27 - PERSPECTIVAS DO BALANCED SCORECARD [23].....	29
FIGURA 28 - PROCESSOS DE GESTÃO INTRODUZIDOS PELO BSC.	30

FIGURA 29 – “CASCATA” DE SCORECARDS: PROCESSO DE ROLLOUT TÍPICO [26].	31
FIGURA 30 - FERRAMENTA DE SUPORTE A BSC HYPERION STUDIO DA ORACLE.	32
FIGURA 31 - SLIDE DE UMA APRESENTAÇÃO DE KARL SWEENEY, CAO DO MARRIOTT VACATION CLUB INTERNATIONAL, NA BRIGHAM YOUNG UNIVERSITY, SOBRE A IMPLEMENTAÇÃO DE BSC NA CADEIA DE HOTEIS MARRIOTT [29].	34
FIGURA 32 - ÁREAS CIENTÍFICAS DO SPATIAL DASHBOARD 1.0.	37
FIGURA 33 - ÁREAS CIENTÍFICAS DO SPATIAL DASHBOARD 2.0.	37
FIGURA 34 - HIERARQUIA DOS INDICADORES DE NEGÓCIO.	38
FIGURA 36 - ARQUITECTURA DA INTEGRAÇÃO DO MOTOR 3D IRRLICHT COM O SPATIAL DASHBOARD E COM O HARDWARE. ADAPTADO DE “HOW OPENGL FITS INTO VISTA” [47]	45
FIGURA 37 – DIAGRAMA DE CLASSES RELATIVO À INTEGRAÇÃO DO MÓDULO REDE DE DENDROGRAMAS COM OS PRINCIPAIS COMPONENTES DA APLICAÇÃO.	46
FIGURA 38 – DIAGRAMA DE CLASSES CORRESPONDENTE À IMPLEMENTAÇÃO DAS VISTAS DO MÓDULO REDE DE DENDROGRAMAS E RESPECTIVOS PADRÕES DE DESENHO.	47
FIGURA 39 – DIAGRAMA DE CLASSES DO DOMÍNIO DO MÓDULO REDE DE DENDROGRAMAS.	48
FIGURA 40 - VISTA REDE DE DENDROGRAMAS: ESTÁ SELECIONADO O CENÁRIO "AERPORTO DO PORTO"	49
FIGURA 41 - INDICADOR CLONADO AO FUNDO DO LADO DIREITO E CLONES À FRENTE DO LADO ESQUERDO. O PADRÃO DAS LINHAS É ANÁLOGO AO QUE PERMITE RECONHECER CLUSTERS EM COORDENADAS PARALELAS.	50
FIGURA 42 - PORMENOR NA VISTA DRILL ALL: INDICADORES DE DIFERENTES TAMANHOS DEMONSTRANDO QUE TÊM DIFERENTES PESOS.	50
FIGURA 43 – SEMANTIC MANAGER: INDICADORES DE PERFORMANCE DA ENTIDADE ORGANIZACIONAL “AEROPORTO DO PORTO”.	54
FIGURA 44 – CENÁRIO DE NEGÓCIO DA ENTIDADE ORGANIZACIONAL “AEROPORTO DO PORTO”.	54
FIGURA 45 – VISTA REDE DE DENDROGRAMAS COM CINCO DENDROGRAMAS.	55
FIGURA 46 – UNIDADE DE HANDLING REPRESENTADA A LARANJA DENOTANDO UMA QUEBRA NO DESEMPENHO.	56
FIGURA 47 - DRILL EXECUTADO NO DENDROGAMA “HANDLING”	57
FIGURA 48 - CENÁRIO "CONTROLO DE OPERAÇÕES EM TERRA" NO MÓDULO DASHBOARD.	58
FIGURA 49 – VISTA <i>DRILL ALL</i> APRESENTANDO O INDICADOR <i>FLIGHT DELAYS</i> E ALGUNS CLONES.	59
FIGURA 50 - ADAPTIVE ZOOMING NA APLICAÇÃO ONLINE GOOGLE MAPS.	63
FIGURA 35 - PLANEAMENTO DAS ACTIVIDADES.	67

Lista de Abreviaturas

API - Application Programming Interface

BI – Business Intelligence

BPM - Business Performance Management

BSC - Balanced Scorecard

ICAO - International Civil Aviation Organization

KPI - Key Performance Indicator

SIG - Sistemas de Informação Geográfica

UEN - Unidade Estratégica de Negócio

VIM - Visualização de Informação Multidimensional

1. Introdução

Este capítulo começa por introduzir a presente dissertação apresentando resumidamente as áreas de estudo, o problema a abordar, fazendo alusão à proposta de resolução. Será ainda apresentada a estrutura da dissertação.

1.1. Objectivo da Tese

Com o presente trabalho, pretende-se encontrar uma forma inovadora de visualizar e explorar indicadores de negócio, dando continuidade à filosofia que levou à criação da aplicação sobre a qual o protótipo será desenvolvido, o *Spatial Dashboard*: a mais-valia da representação gráfica sobre a representação textual e tabular. Na versão 1.0 (desenvolvida em 2005/06) essa orientação foi concretizada com recurso a cartografia digital para colmatar a lacuna da gestão do desempenho empresarial sob o ponto de vista geográfico. Para a versão resultante deste trabalho pretende-se colmatar lacunas identificadas em aspectos relativos à representação do desempenho em várias unidades empresariais, navegação e identificação de relações entre si.

1.2. Problema

1.2.1. Contexto

A visualização é considerada por vários autores a banda larga sensorial, permitindo processar grandes quantidades de informação em intervalos de tempo curtos, comparativamente com outras formas de *media* como texto ou som.

Os avanços recentes no domínio da computação gráfica permitem actualmente a preparação automática de diagramas de visualização elaborados, dinâmicos e interactivos em tempo real.

Devido à crescente vaga de sistemas de informação, bases de dados e *data warehouses*, cuja utilização se impôs como fundamental para a subsistência das empresas e organizações num contexto de competitividade, estas lidam hoje com uma quantidade e complexidade de dados sem precedentes. Trabalhar esta quantidade e complexidade de informação de forma útil é um desafio.

O contexto do Suporte à Decisão tem características que justificariam uma maior aposta em técnicas de visualização avançadas para as interfaces das aplicações usadas pelos decisores. Qualquer aspecto que contribua para os decisores tirarem ilações com maior brevidade será positivo, na medida em que a eficiência é tipicamente uma característica crítica das suas actividades profissionais. Providenciar informação com qualidade

acrescenta valor, visto que ajuda gestores a tomar melhores decisões, contribuindo para um melhor desempenho empresarial.

O BSC é uma metodologia de *Business Performance Management* (BPM) de grande divulgação e aceitação na comunidade especializada. Foi a metodologia seguida na implementação original do *Spatial Dashboard*. Os seus aspectos inovadores são a introdução de perspectivas não financeiras, sob as quais o desempenho empresarial deverá ser avaliado e também a identificação da pertinência dos gestores se focarem nos indicadores essenciais.

1.2.2. Descrição do Problema

O *software* de suporte a BPM e em particular de suporte a implementações de BSC não faz uso de interfaces elaboradas que permitam a um decisor obter a imagem geral do desempenho empresarial e navegar intuitivamente entre níveis de decisão, explorando relações entre entidades organizacionais. Ainda que esteja disponível uma panóplia de tecnologias gráficas de utilização comum em outros contextos, como em jogos, estas acabam por não ser utilizadas na prática em aplicações de suporte a BPM. Uma tal interface seria relevante pois:

Ajudaria os decisores a construir um modelo mental das relações através das quais a estratégia é traduzida em acções;

Contribuiria para uma agilização do processo de exploração de relações entre entidades organizacionais;

Poderia, tal como o BSC, desempenhar um papel importante na comunicação da estratégia empresarial.

1.2.3. Conclusão

Tendo como base o trabalho desenvolvido pelos colegas em trabalhos de final de curso anteriores, tentou-se conceber um protótipo que fosse uma interface e um paradigma de visualização que visa suprir necessidades de decisores empresariais ao nível da gestão de desempenho organizacional.

1.3. Organização da Tese

Introdução: apresentação resumida dos conteúdos descritos ao longo da dissertação, com especial enfoque no problema.

Estado da Arte: descrição de um conjunto representativo de técnicas de visualização multidimensional que serviu de base para a conceptualização do visualizador. É também apresentado o BSC, uma das mais divulgadas metodologias de BPM. Foi a metodologia de BPM adoptada na aplicação que serviu de ponto de partida à implementação do protótipo decorrente desta tese.

Problema: descrição das premissas que constituíram o ponto de partida para o presente trabalho.

Proposta: inclui uma descrição geral dos aspectos fundamentais da plataforma *Spatial Dashboard* necessários à compreensão do trabalho decorrente desta tese. É apresentada a proposta de solução ao problema disposto no capítulo anterior.

Implementação: descrição pormenorizada sobre o processo de desenvolvimento, opções de desenho e práticas seguidas durante a implementação do protótipo.

Caso de Estudo: apresentação das funcionalidades fundamentais do protótipo no contexto da ANA Aeroportos. O caso foi construído com base em dados reais e também fictícios, com suporte constante da bibliografia.

Conclusão: principais contribuições alcançadas por esta tese e também por indicações de trabalho futuro.

2. Estado da Arte

Os indicadores de *Balanced Scorecard* são inerentemente multidimensionais, em particular no contexto do *Spatial Dashboard*, pelo que neste capítulo se fará uma pesquisa sobre estado da arte em visualização de informação multidimensional. Será também descrita a metodologia *Balanced Scorecard*.

2.1. Visualização - Motivação

"Un bon croquis vaut mieux qu'un long discours"

Napoleão Bonaparte

O objectivo da seguinte pesquisa é apresentar o estado da arte no domínio da Visualização de Informação Multidimensional (VIM), um ramo importante da Visualização Científica, que trata a análise de dados com múltiplos parâmetros e relações entre si. A VIM visa auxiliar os processos de descrição, resumo e interpretação de dados e permite explorar a inferência cognitiva, reduzir a necessidade de memorização e alimentar as operações mentais de grandes quantidades de informação.

Os mecanismos de visualização tradicionais não são adequados à complexidade e quantidade de dados armazenados actualmente em bases de dados e *Data Warehouses*. A maior parte do trabalho desenvolvido na área da visualização foca-se em representações a duas ou três dimensões. No entanto, os avanços recentes em computação gráfica permitem a preparação automática de diagramas ou visualizações em tempo real, dinâmicos e interactivos que constituem ferramentas de interacção mais elaboradas e providenciam aos analistas intuição acerca de características dos dados, tais como padrões ou anomalias.

Vários métodos de visualização foram desenvolvidos neste âmbito com as suas qualidades e limitações. O seu objectivo é concretizar um mapeamento do espaço de dados, de alta dimensionalidade, para o espaço visual. Esta é uma tarefa não trivial e que deve coadunar-se com as capacidades e limitações cognitivas dos seres humanos.

A presente pesquisa pretende descrever algumas das principais técnicas de VIM e respectivas estratégias subjacentes. São apresentadas e discutidas as técnicas mais populares de VIM, como *ScatterplotMatrices*, *Coordenadas Paralelas*, *Stick Figures*, *Caras de Chernoff*, *VisDB*, *GrandTour*, entre outras. Serão ainda abordados a maturidade e o potencial deste domínio científico.

Foram consideradas apenas técnicas e não implementações das mesmas. É difícil distinguir os aspectos inovadores de uma nova abordagem quando uma técnica é apresentada sob a forma de uma ferramenta. As ferramentas não podem ser analisadas de forma isenta e imparcial, pois reúnem frequentemente características de diferentes técnicas e são afectadas pelo ruído introduzido por detalhes de implementação.

2.2. Fundamentos de Visualização Multidimensional

2.2.1. Terminologia

A literatura relativa a informação multidimensional sofre de alguma ambiguidade e inconsistência na notação. Este aspecto poderá derivar do facto de esta área ser objecto de estudo de profissionais de diferentes áreas. [1]. Assim, torna-se importante apresentar a terminologia adoptada no decurso do texto. Em seguida são definidos os principais termos específicos utilizados posteriormente, para uma correcta interpretação.

Dimensionalidade: Número de variáveis, atributos ou dimensões de um conjunto de dados.

Dados Multivariados: Conjunto de dados que tem muitas variáveis dependentes e que podem ser correlacionadas entre si. Normalmente estes conjuntos estão associados a modelos discretos de dados.

Dados Multidimensionais: Conjunto de dados que possui muitas variáveis independentes claramente identificadas e às quais está associada uma ou mais variáveis dependentes. Normalmente estes conjuntos estão associados a modelos contínuos de dados.

Alta-dimensionalidade: Um conjunto de dados diz-se de alta dimensionalidade se contiver mais do que três variáveis, dependentes e/ou independentes. Um método de visualização de alta-dimensionalidade terá de lidar com o problema de converter dados complexos de alta-dimensionalidade numa representação apropriada de baixa dimensionalidade. Associado a este conceito está a comumente denominada maldição da dimensionalidade: um espaço de alta-dimensionalidade cresce exponencialmente com o número de dimensões. Consequentemente, espaços multivariados de crescente dimensionalidade tendem a ser dispersos (e não densos), conduzindo a um aproveitamento deficiente do espaço de ecrã.

Baixa Dimensionalidade: Diz-se de um conjunto de dados cuja dimensionalidade é inferior ou igual a três.

Clustering: Processo de organização de itens em grupos de acordo com um determinado critério. É um dos frequentes objectivos da visualização multidimensional e multivariável. Pode ser considerado o problema mais importante de aprendizagem não supervisionada, ou seja, de estruturação de uma colecção de dados não rotulados. [2]

Formalmente uma técnica de visualização corresponde a um mapeamento do espaço de dados para o espaço visual.

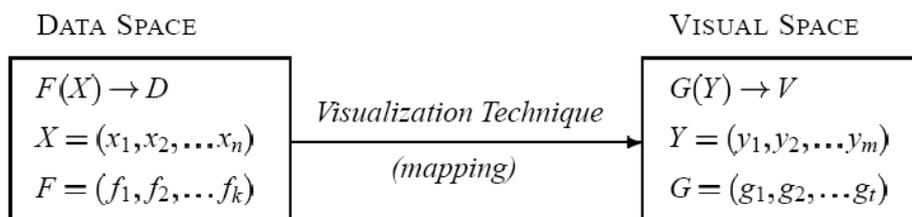


Figura 1 - Propósito da visualização: mapeamento do espaço de dados para o espaço visual [1]

No diagrama acima (Figura 1), X e Y (localização espacial de primitivas gráficas) representam as variáveis independentes do espaço de dados e do espaço visual, respectivamente, e F e G (propriedades gráficas das primitivas) representam as variáveis dependentes do espaço de dados e do espaço visual, respectivamente. No espaço de dados encontram-se os itens de dados e no espaço visual os itens mapeados.

Mapear n variáveis independentes do espaço de dados para m variáveis independentes do espaço de visualização é uma tarefa delicada sobretudo quando $n \gg m$, situação em que é necessária uma solução mais elaborada.

2.2.2. Domínio de Aplicação

São frequentemente identificadas na literatura de Visualização duas áreas principais de aplicação: Visualização Científica e Visualização de Informação [1, 3]. Santos [1] propõe as definições seguintes de cada uma destas áreas.

A Visualização Científica lida com conjuntos amostrais de dados e o objectivo é recriar a partir deste conjunto de dados o modelo $F(X)$ subjacente. Tipicamente o modelo de dados é contínuo. Quando o número de dimensões é pequeno (medição de temperaturas em intervalos de tempo regulares, imagiologia médica, etc.), a dimensionalidade do espaço de dados e do espaço visual coincidem, não sendo um problema encontrar uma representação visual adequada. No entanto, para problemas de alta-dimensionalidade, frequentes por exemplo em problemas de optimização de parâmetros, pretende-se visualizar uma função de um grande número de parâmetros.

A Visualização de Informação lida usualmente com dados multivariados de áreas como a análise estatística, bolsas de valores, dados de inquéritos, etc. Tipicamente o modelo de dados é discreto. Frequentemente estes dados são apresentados sob a forma de tabelas. O objectivo da visualização de informação depende do contexto do problema, mas normalmente consiste em pesquisa de padrões, *clusters*, tendências ou correlação entre atributos. A informação obtida será utilizada como suporte à elaboração de hipóteses acerca dos dados analisados. A visualização é uma ferramenta que aumenta o leque de abordagens analíticas para dados multivariados.

2.3. Caracterização do Estado da Arte

A visualização multidimensional e multivariada já era estudada por psicólogos e estatísticos muito antes da ciência da computação se constituir. O trabalho desenvolvido neste domínio científico teve maior actividade nos últimos 30 anos [3].

Wong e Bergeron [3] identificam quatro fases de características distintas: a fase da procura (antes de 1976), a fase do despertar (77-85), a fase da descoberta (86-91) e a fase da elaboração (92-presente).

A primeira fase caracteriza-se pela representação de dados de pequena dimensão e pela utilização de instrumentos essencial de papel e caneta. Os estatísticos lideravam os esforços. Uma importante contribuição

desta fase foi a das caras de Chernoff [4], uma técnica que consiste na representação de caras cuja expressão era determinada em função de características dos dados.

A segunda fase é caracterizada pelo uso de *displays* 2D e 3D e pelo estudo predominante de dados espaciais. Os principais acontecimentos foram a chegada do *PC* (a ferramenta científica mais poderosa até então) e a contribuição de Tukey o trabalho *Exploratory Data Analysis* [3, 5] que transformou a visualização numa forma de pensar e explorar os dados, não apenas uma forma de os expor.

A terceira fase caracterizou-se pela explosão de novos conceitos e técnicas a par com a evolução da capacidade de processamento gráfico. Um dos principais acontecimentos foi a declaração da necessidade de investigação em visualização espacial de objectos pela *National Science Foundation* [3].

Finalmente a quarta fase que se estende até hoje caracteriza-se pelo refinamento do trabalho efectuado nas fases anteriores. Uma questão essencial em aberto é a necessidade de avaliar a correcção, eficácia e utilidade das técnicas de visualização multidimensional e multivariada.

Nos tempos mais recentes têm sido desenvolvidas cada vez menos técnicas de VIM, sendo as mais recentes propostas aperfeiçoamentos de técnicas anteriores. O cerne da actividade centra-se agora na classificação das técnicas [6], pois não há ainda uma norma de categorização. As principais sugestões de classificação são a baseada no *display* de Wong e Bergeron [3], a baseada em entidades de Brodlie [7], a baseada em objectivos por Schneiderman [8] e a baseada em processos de Buja [9] e Keim[10]. A análise de cada uma destas propostas de taxonomia foge ao âmbito desta pesquisa, mas o seu contributo é de grande importância para a área e constituem um óptimo ponto de partida para o seu estudo.

2.4. Técnicas de Visualização

2.4.1. Matrizes de Diagramas de Dispersão (Scatterplot Matrices)

Esta é uma das técnicas de visualização de informação multidimensional e multivariada mais populares. É considerada uma norma e está disponível na quase totalidade de ferramentas de visualização de dados de alta-dimensionalidade.

Consiste na criação de uma matriz de referenciais em que cada referencial representa as relações entre um par de variáveis. A matriz contém $n(n-1)/2$ referenciais, sendo que n é a dimensionalidade do conjunto de dados. O intervalo de representação de cada eixo corresponde ao intervalo que compreende os dados dessa variável. Assim, uma matriz de diagramas de dispersão apresenta uma imagem geral de todo o conjunto de dados. Na Figura 2 e na **Erro! A origem da referência não foi encontrada.** são apresentados exemplos desta técnica.

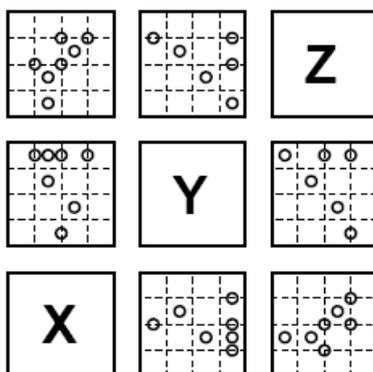


Figura 2 - Matriz de diagramas de dispersão com três variáveis X, Y e Z

A ideia base é associar visualmente características de um painel com características dos outros painéis. A redundância dos painéis simétricos em relação à diagonal (canto inferior esquerdo - canto superior direito) suporta a associação visual. Estruturas lineares denotam correlação entre os parâmetros envolvidos nessa projecção. Esta representação permite explorar a nossa capacidade perceptiva de identificação de clusters [6].

A diagonal principal da matriz pode ser utilizada de duas formas. Pode ser utilizada numa primeira abordagem para identificar a variável, que está representada nos demais referenciais dessa linha e dessa coluna nos eixos das ordenadas e das abcissas, respectivamente. Esta abordagem foi a adoptada na Figura 2. Numa abordagem alternativa pode ser utilizada para fornecer informação acerca dessa dimensão, como pode ser observado na **Erro! A origem da referência não foi encontrada.**

Na Figura 3, está representado um conjunto de dados com sete dimensões correspondentes a atributos de automóveis: consumo (em milhas/galão), cilindros, cavalos, peso, aceleração, ano de fabrico e país de origem. A representação das dimensões ao longo de si próprias permite inferir sobre a distribuição dos dados em cada dimensão (podemos observar que foram analisados carros com cinco cilindradas distintas e de três países diferentes). As implementações desta técnica normalmente oferecem mecanismos de auxiliares de interacção, como se pode observar na Figura 3 que corresponde a um *screenshot* da ferramenta *XMDVTool*.

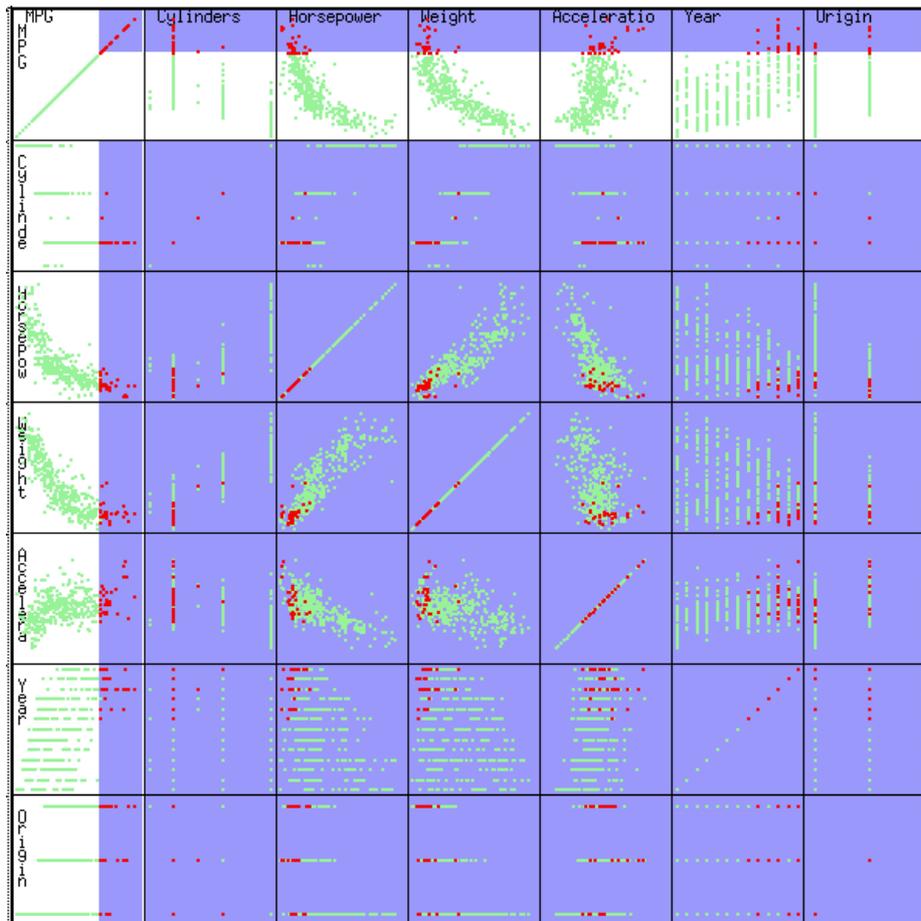


Figura 3 - ScatterPlot Matrix da aplicação XMDVTool [11] com um subconjunto de dados seleccionados e representados a vermelho

Na Figura 3 encontram-se seleccionados os automóveis com menores consumos (maiores valores de milhas/galão). O processo de selecção corresponde a desenhar um quadrado com o *mouse*, num dos referenciais (neste caso foi no primeiro) e todos os itens no interior do quadrado passam a ser representados a vermelho em todos os referenciais. Através desta selecção pode observar-se que a maior parte dos carros de baixo consumo têm quatro cilindros, pouco peso e baixa aceleração.

Têm sido propostas outras alternativas e melhoramentos para aumentar a quantidade de informação representada em cada referencial e facilitar a exploração. São exemplos a rotação da nuvem de pontos, a utilização de diferentes símbolos para agrupar classes de atributos e detectar sobreposições, e ainda a utilização de cores e sombreamento como dimensão extra em cada projecção [11].

Esta técnica está na origem de algumas outras. O princípio da utilização de adjacências de pares de variáveis é também basilar nas técnicas *HyperBox*, Eixos Hierárquicos e *HyperSlice* [3]. Outra técnica de visualização que representa uma extensão natural da apresentada é a Matriz de Diagramas de Dispersão Tridimensionais. Os conceitos são análogos. Os pontos são colocados num espaço tridimensional e posteriormente numa projecção bidimensional do mesmo. A aparência 3D é conseguida através de pistas de visuais que sugerem o efeito de profundidade como o sombreamento, nevoeiro e tamanho dos pontos. A tridimensionalidade possibilita a

identificação de relações entre três variáveis, mas levanta o problema da sobreposição de pontos, requerendo uma interacção mais complexa.

As principais vantagens desta técnica consistem na fácil interpretação e na aplicabilidade a grandes conjuntos de dados.

Uma limitação desta técnica consiste na ausência de características que permitam descobrir relações que envolvam mais do que dois atributos. Outra limitação consiste no facto de que as *ScatterPlot Matrixes* apresentam menor eficácia em contextos de muito alta dimensionalidade, pois será mais reduzido o espaço disponível para cada referencial visto que o espaço de ecrã é finito. Esta limitação poderá ser contornada efectuando a representação parcial da matriz e disponibilizando a funcionalidade de zoom [11].

2.4.2. Coordenadas Paralelas

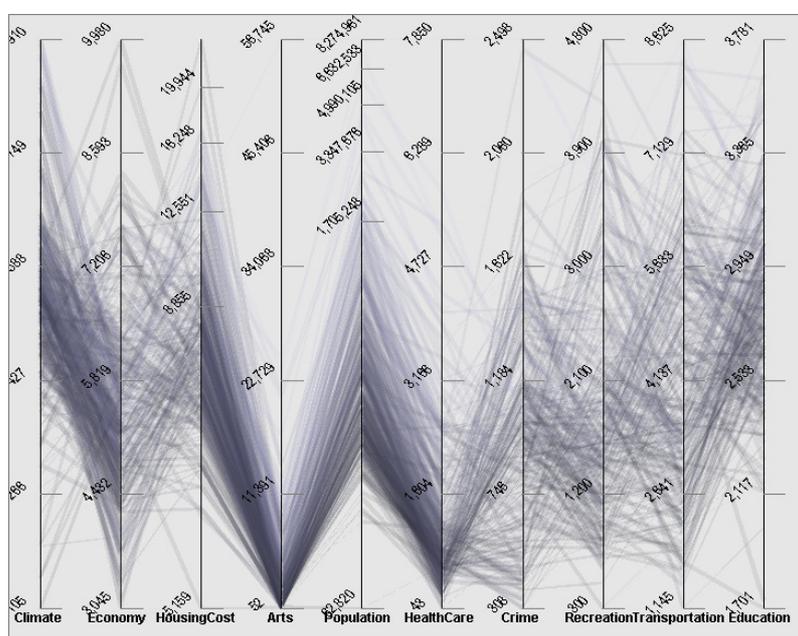


Figura 4 - Screenshot do software ILOG Discovery [13] representando dados estatísticos de diferentes cidades americanas com coordenadas paralelas

Num sistema de coordenadas paralelas, as N dimensões correspondem a N eixos verticais colocados paralelamente e igualmente espaçados. Representar um item do espaço de dados corresponde a marcar o valor de cada variável no respectivo eixo e a unir os pontos obtidos. Assim, cada item do espaço de dados corresponde a um conjunto de segmentos de recta que é usualmente referido como *poli-linha*. Na Figura 4 está um exemplo de dados estatísticos representados em coordenadas paralelas.

No conjunto de dados correspondente à representação em coordenadas paralelas da Figura 4 sobressai uma possível correlação entre os eixos dos custos de habitação, artes, população e saúde, evidenciada pela sobreposição dos segmentos de recta entre esses eixos. O clima e a economia não parecem correlacionados.

Uma grande vantagem do sistema de coordenadas paralelas face aos referenciais ortogonais cartesianos é que permite a fácil identificação de colisões e não está restrito ao limite de três dimensões.

Suponha-se que se pretende identificar colisões de dois aviões. Há que considerar as três dimensões espaciais e a dimensão tempo, correspondendo uma colisão a um par de itens com valores idênticos para cada dimensão. Isso seria difícil de confirmar recorrendo apenas a uma única vista de um referencial cartesiano.

Colocando num referencial tridimensional cartesiano dois aviões nos pontos (1,1,1) e (2,2,2), o ponto de vista em (3,3,3) e orientado para a origem, os aviões aparentariam estar em colisão. Para ajudar à prevenção de colisões podem ser desenhados paralelogramos onde estão incluídas as *poli-linhas*, como ilustrado na Figura 5. Esses paralelogramos representam uma margem de segurança, considerando-se que há colisão se 2 *poli-linhas* se encontrarem dentro do mesmo paralelogramo.

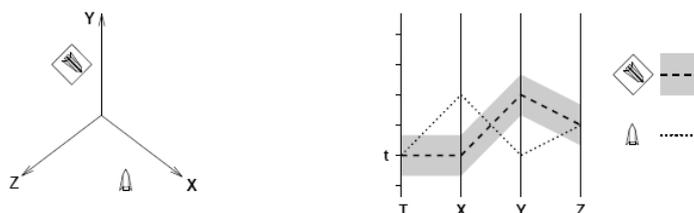


Figura 5 - Prevenção de Colisões

Uma aplicação relevante das coordenadas paralelas é o estudo de correlações entre variáveis através da localização dos pontos de intersecção. Repare-se que todas as *poli-linhas* correspondentes a duas variáveis directamente proporcionais irão intersectar-se no mesmo ponto, como se pode ver nas Figura 6 e 8.

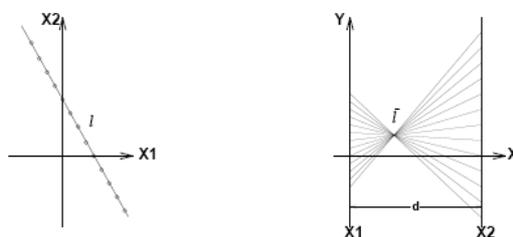


Figura 6 – Proporcionalidade Inversa

Na Figura 6 podemos observar que no referencial de coordenadas paralelas à direita são representados alguns pontos da recta representada no referencial cartesiano à esquerda. A intersecção dos pontos situa-se entre os eixos indicando proporcionalidade inversa.

A intersecção estará entre os respectivos eixos paralelos caso a proporcionalidade seja inversa (Figura 6) e fora caso seja directa (Figura 7).

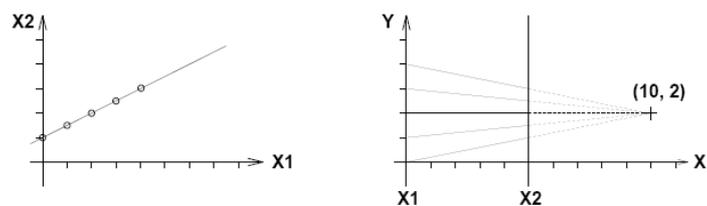


Figura 7 - Proporcionalidade Directa

Na Figura 7 observamos à esquerda: referencial cartesiano com a recta $X_2=1/2X_1+1$ e à direita: sistema de coordenadas paralelas que representa alguns pontos da mesma recta. A proporcionalidade directa é evidente pela ocorrência da intersecção no exterior dos eixos.

Na Figura 7, alguns pontos da recta $X_2=1/2X_1+1$ são representados num referencial cartesiano e num sistema de coordenadas paralelas. Os pontos de uma recta $y=mx+b$, que correspondem a segmentos de recta num referencial de coordenadas paralelas (com distância d entre eixos) é dado por $(d/1-m, b/1-m)$ [3]. Tal poderá ser verificado através de considerações geométricas triviais. Assim, o ponto de intersecção representado na Figura 7 é $(5/(1-0.5), 1/(1-0.5)) = (10, 2)$.

Como pode ser observado nas Figura 6 e 8, a fácil identificação de relações de proporcionalidade é um dos aspectos fortes desta técnica, bem como a identificação de clusters.

A principal limitação das coordenadas paralelas deriva do espaço limitado disponível para cada eixo paralelo. Em contextos com um numeroso conjunto de itens de dados, as *poli-linhas* geradas podem cobrir a maior parte do *display*.

Outra limitação consiste na maior facilidade de identificação de relações entre dimensões correspondentes a eixos paralelos adjacentes. Este último aspecto poderá ser contornado ao fornecer aos utilizadores mecanismos que permitam a rápida reordenação de eixos.

2.4.3. Iconografia

A iconografia engloba um conjunto de técnicas que assentam essencialmente nos mesmos princípios básicos. Mapeiam valores dos dados para atributos geométricos ou de cor de primitivas gráficas ou símbolos. O posicionamento dos ícones num *display* bidimensional poderá ser feito em função de duas dimensões do conjunto de dados, ou então poderá também fazer-se essa disposição sob a forma de *array*.

Caras de Chernoff

No ramo da iconografia, uma técnica que merece destaque pela inovação introduzida aquando da sua apresentação em 1973, é a técnica das caras de Chernoff [3, 11]. Os princípios são bastante simples. São construídas caras a partir dos atributos dos conjuntos de dados. Cada atributo é feito corresponder a uma característica facial (tamanho e forma da cara, forma e inclinação dos olhos, boca, nariz, etc.).

A Figura 8 representa um conjunto de características faciais aos quais se poderão fazer corresponder dimensões de um conjunto de dados.

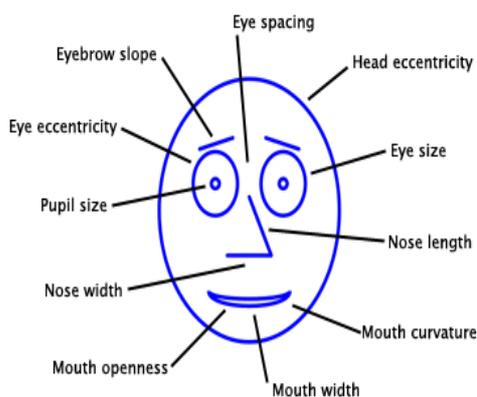


Figura 8 - Características faciais de uma cara de Chernoff [14].

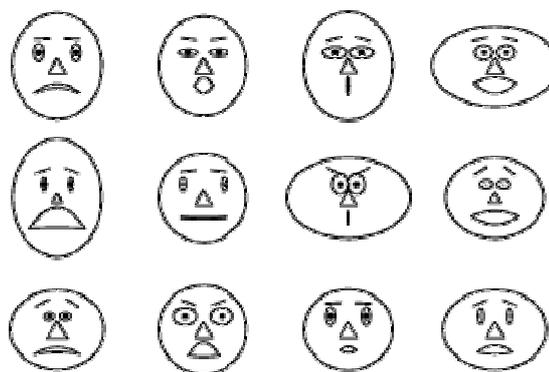


Figura 9 - Mapeamento de dez dimensões de um conjunto de dados gerados pela ferramenta Wolfram Mathematica [15], para um array de caras de Chernoff.

O reconhecimento facial é uma classe de estímulos particularmente desenvolvida dos seres humanos [12]. Algumas das caras do *array* da

Figura 9 são facilmente associáveis com expressões humanas.

Entre as principais vantagens desta técnica encontra-se o facto de o reconhecimento de caras e expressões ser uma capacidade humana inerentemente treinada e desenvolvida.

Entre as limitações, destacam-se a grande necessidade de espaço disponível para representar muitas caras e o facto de que caras podem não ser adequadas para grande número de contextos.

Stick Figure Icons

Aquando da sua apresentação, as figuras de pauzinhos [3, 13] consistiam num ícone, composto por cinco traços, correspondendo um ao tronco e os restantes aos membros. Esta família de ícones podia representar até cinco dimensões que eram mapeadas para as inclinações dos membros e do tronco. Mais membros permitem a adição de mais dimensões. Na Fig. 8 pode observar-se um exemplo de várias figuras de pauzinhos representando cinco variáveis.

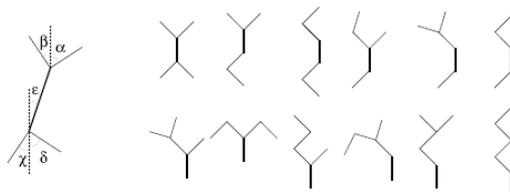


Figura 10 - Esquerda: figura de pauzinhos com identificação das variáveis de orientação; direita: família de figuras de pauzinhos

É possível mapear dimensões extra para o comprimento, espessura e cor dos traços. O alinhamento de ícones permite explorar a capacidade humana de discernimento de texturas em imagens complexas, como pode ser visto na Figura 10.

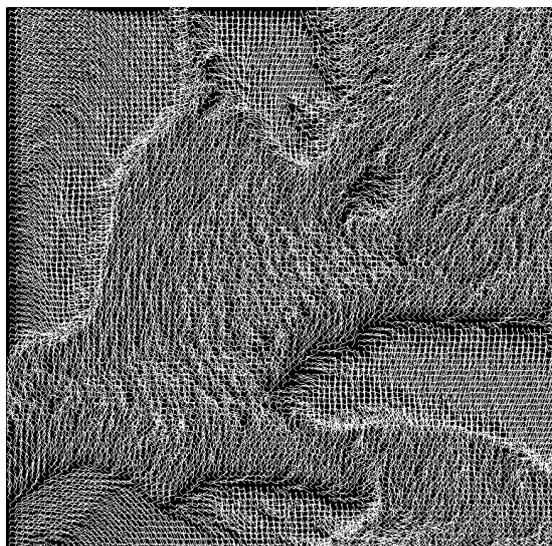


Figura 11 - Imagem de Stick Figures representando 7 variáveis

Em testes realizados no início da década de 90, esta técnica foi aplicada com sucesso com dados recolhidos a partir de ressonâncias magnéticas de tumores, permitindo visualizar um *hotspot* num tumor, ou seja, zona importante para o diagnóstico, que passaria despercebido com as técnicas de imagiologia médica tradicionais [13], Figura 11.



Figura 12 - Dados de ressonâncias magnéticas de um tumor, com diferentes mapeamentos

Na Figura 12 estão representadas duas composições iconográficas baseadas em dados de ressonâncias magnéticas de um tumor, com diferentes mapeamentos [13]. Esquerda: o tumor destaca-se a branco. Direita: o círculo vermelho assinala um *hotspot*, zona propícia à ocorrência de mutações, de difícil reconhecimento.

As vantagens desta técnica consistem na exploração da capacidade cognitiva de discernimento de texturas e na sua flexibilidade, permitindo configurações e ajustes a vários níveis.

As principais limitações consistem no comprometimento entre o número de dimensões representáveis e impacto de cada dimensão na imagem gerada e ainda no facto de que a facilidade de discernimento é altamente dependente de uma boa escolha para o mapeamento de parâmetros. É necessária bastante experimentação para obter um ajuste de parâmetros ideal [13], visto que o número de mapeamentos potenciais cresce factorialmente com o número de dimensões.

2.4.4. VisDB

A técnica VisDB [14] nasceu a partir da motivação de representar visualmente dados presentes em bases de dados e da observação de que as tabelas das bases de dados relacionais podem ser interpretadas como conjuntos de dados multidimensionais em que os atributos correspondem às dimensões.

As *queries* tradicionais foram consideradas insuficientes pelos autores devido a nem todos os dados poderem ser separados de forma clara (nomeadamente dados com intervalos contínuos de valores, desconhecidos *a priori*) e também ao facto de nem sempre os utilizadores sabem exactamente o que pretendem (desconhecem o intervalo e a distribuição de dados).

Em geral apenas se pode submeter uma query de cada vez, não sendo possível alterar ligeiramente uma query ou fazer queries vagas. Outra limitação importante da apresentação tradicional de resultados é a de que não proporciona qualquer tipo de *feedback* para além do resultado, que pode conter muitos dados para uma pesquisa eficiente ou muito poucos (ou nenhum), o que não dá pistas no sentido de continuar a explorar a base de dados.

Os objectivos do conceito são a representação visual de tantos resultados de uma query quantos os possíveis e mostrar quão distantes são os resultados em relação a um resultado exacto.

O utilizador introduz uma query que identifica um ponto focal no espaço multidimensional. Os dados são então processados utilizando uma função que representa o *factor de relevância* desse item em relação ao ponto focal. O cálculo destes factores de relevância é efectuado com o recurso a funções de distâncias, que dependem do tipo de dados em questão.

Alguns exemplos de funções de distâncias são a diferença numérica (para tipos métricos), matrizes de distâncias (para tipos ordinais e nominais), e diferença lexicográfica, de *character*, de *substring* e de *fonema* para *strings*. Cada item é representado no *display* por um *pixel* cuja cor depende do factor de relevância. A escala de cores utilizada é amarelo, verde, azul, vermelho e preto, denotando por esta ordem uma distância crescente para com as respostas correctas. O número de dimensões consideradas para a visualização é restrito ao número de dimensões presentes na query.

Uma vez calculados os referidos factores de relevâncias, são possíveis várias opções de disposição. A mais simples corresponde a organizar os itens de dados de forma que os mais relevantes fiquem no centro do *display* e os restantes em espiral à volta dessa região, como ilustrado na Figura 13. Quanto mais afastados do centro, menos relevantes são os itens correspondentes.

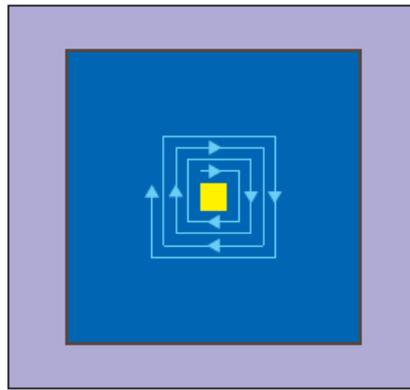


Figura 13 - Esquema da representação de resultados em torno do ponto focal, descrevendo uma elipse

Esta abordagem apresenta alguma redundância visto que a cor e a localização dos itens dependem ambas do factor de relevância. O padrão de variação de cor dá uma perspectiva de como todos os itens da base de dados se relacionam com o ponto focal, mas não cobre as relações entre essas variáveis. Tal poderá ser conseguido através da geração de *displays* adicionais visualizados simultaneamente, como apresentado na Figura 14. Esta abordagem consiste em definir um *display* para cada variável presente na *query* e que representa a sua contribuição para a *query*. A posição dos itens será a mesma do *display* principal, mas a cor será calculada com base na distância unidimensional correspondente a essa variável.

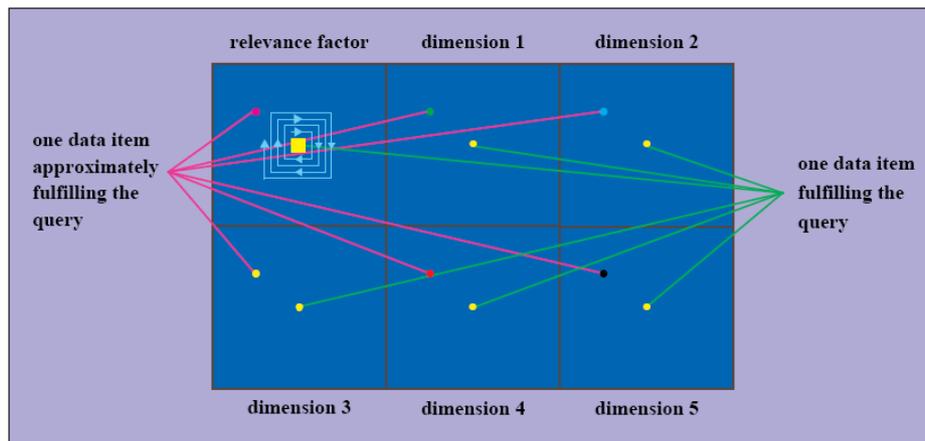


Figura 14 - Esquema legendado da utilização de um display para cada dimensão

A Figura 14 descreve o significado da representação em painéis separados para cada dimensão. O painel do canto superior esquerdo, corresponde à *query*, com a função de distância aplicada a todas as dimensões. Nos restantes painéis, a posição de cada item é a mesma que ocupa no primeiro painel, mas a sua cor é calculada a partir da função de distância apenas relativamente a uma dimensão. A Figura 15 apresenta uma concretização desta abordagem.

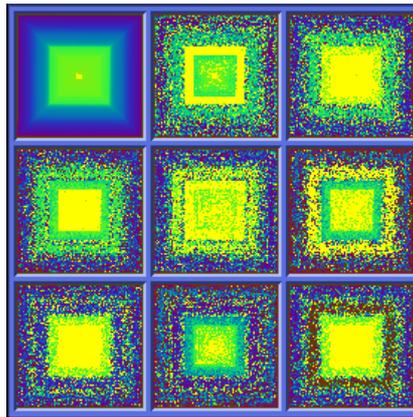


Figura 15 - Exemplo com dados reais (7000 itens com 8 dimensões) da utilização de um display para cada dimensão, para além do principal (o do canto superior esquerdo)

Padrões de cor semelhantes em várias dimensões indicam eventuais correlações. Analisando a Figura 15 e numerando os painéis de 1 a 9, sequencialmente da esquerda para a direita, de cima para baixo, pode discernir-se uma possível correlação entre as dimensões dos painéis 3 e 7, o que não se poderá dizer sobre os painéis 2 e 5.

Os aspectos de maior potencial desta técnica são o facto de o VisDB consistir numa extensão às bases de dados, uma tecnologia madura e de grande penetração, permite trabalhar em contextos de alta dimensionalidade e cardinalidade de dados e ainda a diferenciação da maioria das restantes técnicas por não representar os dados directamente mas sim através de factores de relevância calculados.

As principais limitações dos VisDB consistem no forte condicionamento em relação à resolução do *display* e no facto de a escolha de funções de distância não ser trivial para dados não métricos e ainda mais complexa para *queries* aninhadas com operadores booleanos [3].

2.4.5. GrandTour

O grand tour [15] é um processo de visualização, que gera uma animação ao invés das restantes técnicas descritas até agora que geram imagens estáticas. A ideia é a de mover um plano de projecção através de um espaço de alta dimensionalidade. Por outras palavras, se tivermos $p+1$ dimensões, tiramos 1 para representar o parâmetro tempo para a animação e as restantes p variáveis são projectadas em planos bidimensionais em rápida e suave sucessão de acordo com as alterações do parâmetro tempo. Assim, os dados são codificados em diagramas de dispersão providenciando, segundo os autores, Buja e Asimov, um rápido reconhecimento dos dados usualmente atingível apenas através de horas de análise de gráficos.

Como vantagem desta técnica identifica-se a análise de todo o espaço de dados em alguns instantes, havendo assim uma reduzida probabilidade de um utilizador não identificar uma estrutura importante e o recurso ao movimento que ajuda o sistema perceptivo humano a captar nuvens de pontos, linhas e superfícies [15].

Como desvantagem desta técnica os autores apontam a necessidade de utilizar outras técnicas ou conjugar o grand tour com métodos auxiliares quando os itens de dados formam estruturas que não sejam nuvens de

pontos, linhas ou superfícies. É também uma limitação a necessidade de que a animação gerada seja suave, pois a observação atenta de uma animação irregular requer maior esforço aos olhos, bem como o requisito de que o plano bidimensional que codifica a velocidade seja ortogonal em relação ao plano de projecção.

2.4.6. Técnicas Hierárquicas

Este conjunto de técnicas caracteriza-se pela inclusão de dimensões dentro de outras dimensões. As três variantes mais populares consistem na inclusão de espaços unidimensionais no interior de espaços unidimensionais (eixos hierárquicos), bidimensionais no interior de espaços bidimensionais (*dimension stacking*) e espaços tridimensionais no interior de espaços tridimensionais (*worlds within worlds*).

Eixos Hierárquicos

No caso unidimensional, o dos eixos hierárquicos [3, 11], cada dimensão deverá ser discretizada num conjunto de gamas de valores. Em seguida, as dimensões são ordenadas como na Figura 16 resultando que ficarão por ordem crescente de velocidades (num extremo ficará a dimensão mais lenta e no extremo oposto a dimensão mais rápida – a mais aninhada de todas as dimensões). Cada eixo hierárquico tem a sua velocidade, um conceito que caracteriza o grau de aninhamento da respectiva dimensão. Este processo permite definir um referencial hierárquico.

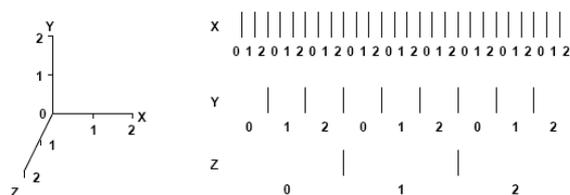


Figura 16 - Esquerda: Referencial ortogonal tridimensional; direita: Referencial hierárquico

Os valores de uma variável dependente podem em seguida ser traçados sob a forma de barras sobre o eixo mais rápido. Cada gama desse eixo corresponde a um conjunto de variáveis independentes. Na imagem seguinte está um exemplo simples correspondente à representação de um espaço tridimensional bivariado.



Figura 17 - Representação dos tuplos $\{(0,0,1), (0,1,0), (0,2,1)\}$. As variáveis x e y são as variáveis independentes e z a variável dependente

Na Figura 17, z é a variável dependente, ou seja, existe uma função f tal que $f(x,y)=z$. Neste caso tem-se $f(0,0)=1$, $f(0,1)=0$ e $f(0,2)=1$. A variável z está representada como a altura da coluna cinzenta.

Sem o recurso a mecanismos auxiliares, esta técnica suporta tipicamente cerca de vinte dimensões num ecrã. É possível representar mais dimensões recorrendo à técnica auxiliar de zoom de sub-espacos. No nível de abstracção correspondente a um determinado eixo hierárquico é possível ter uma representação aproximada dos dados contidos na dimensão de maior velocidade adjacente. Para tal basta aplicar uma função a todos os valores da variável independente (soma, média, máximo, mínimo, desvio padrão, etc.).

Como vantagens deste método identifica-se o facto de a representação favorecer a exploração iterativa da dimensão mais lenta para a mais rápida.

Entre as limitações destaca-se a dependência da discretização dos dados e a aplicabilidade restrita a funções (colocando sobre o eixo mais rápido barras com as variáveis dependentes e nos restantes eixos as variáveis independentes).

Dimension Stacking

Esta técnica é análoga à anterior, mas no contexto de duas dimensões a hierarquia corresponde a embutir imagens dentro de "pixels" de outras imagens de mais alto nível. Cada elemento é um gráfico bidimensional XY [3, 11, 16]. São construídos referenciais XY a partir de pares de variáveis. Esses referenciais são em seguida colocados recursivamente dentro das células dos outros referenciais.

Considere-se que se pretende utilizar esta técnica para representar dados correspondentes a quatro variáveis independentes x , y , z e w cujos tamanhos são respectivamente 2, 3, 5 e 6. Há que agrupá-las em pares, por exemplo $\{x,y\}$ e $\{w,z\}$. No caso do número de variáveis ser ímpar pode introduzir-se uma variável *dummy*. A partir de cada par de variáveis desenha-se uma grelha, com o número de células de cada lado igual ao tamanho da dimensão correspondente. Depois basta colocar uma grelha wz dentro de cada célula da grelha yz ou vice-versa. Ambas as possibilidades estão representadas na imagem seguinte.

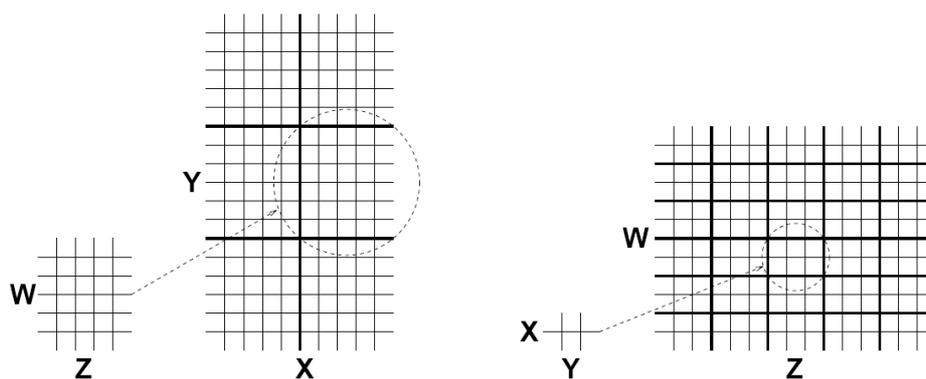


Figura 18 - Dois mapeamentos diferentes para o mesmo espaço de dados

No caso de existirem variáveis dependentes elas podem ser representadas como intensidades de cores (Figura 19), caso contrário cada referencial XY pode ser um simples diagrama de dispersão (Figura 20).

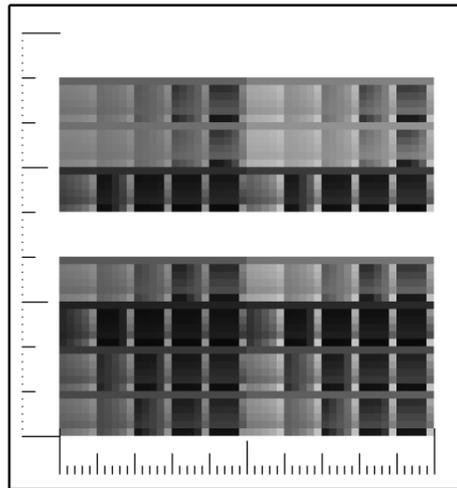


Figura 19 - Visualização de Dimension Stacking com mapeamento de variáveis dependentes na forma de intensidade numa escala de cinzas

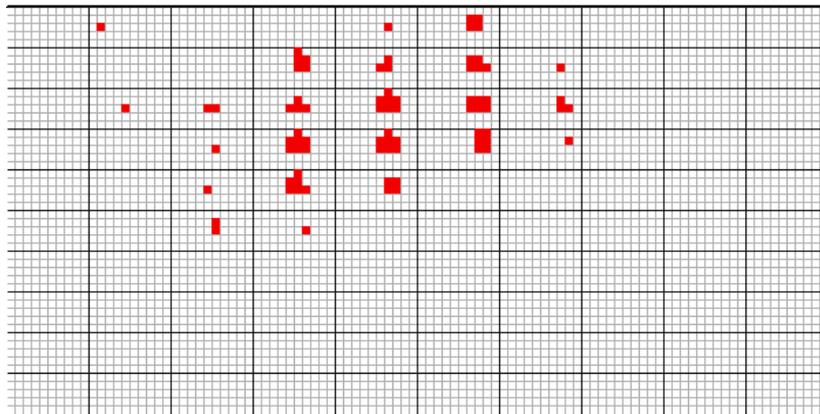


Figura 20 - Visualização de Dimension Stacking só com variáveis independentes, sendo que cada referencial bidimensional é um diagrama de dispersão

Worlds Within Worlds

Neste paradigma cada ponto de um espaço tridimensional contém outro espaço tridimensional. Em cada instante são apresentadas três variáveis no máximo. Esta técnica é necessariamente interactiva e requer uma luva de navegação tridimensional (*Power Glove*). Esta luva permite definir uma posição num espaço tridimensional onde será disposto um novo referencial tridimensional. Este processo repete-se até terem sido utilizadas todas as dimensões, sendo que o último referencial corresponde às últimas dimensões definidas.

A Figura 21 apresenta um aninhamento de dois mundos virtuais, no contexto de um espaço de dados com seis dimensões. Assim, primeiro foram traçados os eixos xyz , foi escolhido um ponto nesse espaço 3D e em seguida traçaram-se os eixos uvw .

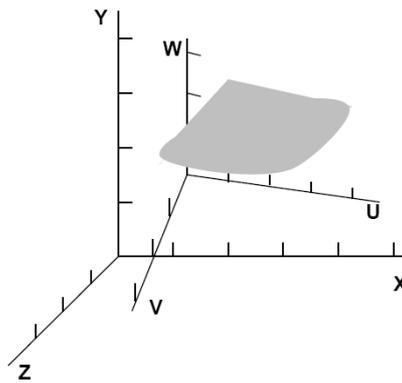


Figura 21 - Worlds Within Worlds

A utilização das mesmas variáveis por uma ordem diferente irá conduzir a diferentes vistas dos dados, originando muitas combinações possíveis. Existem ferramentas, como o AutoVisual [3], que geram automaticamente mundos virtuais com base em tarefas pretendidas pelo utilizador. Um dos aspectos diferenciadores mais importantes é a utilização de uma técnica de interação avançada baseada no tacto, o que também pode ser visto como uma contrapartida devido à sua reduzida divulgação e disponibilidade.

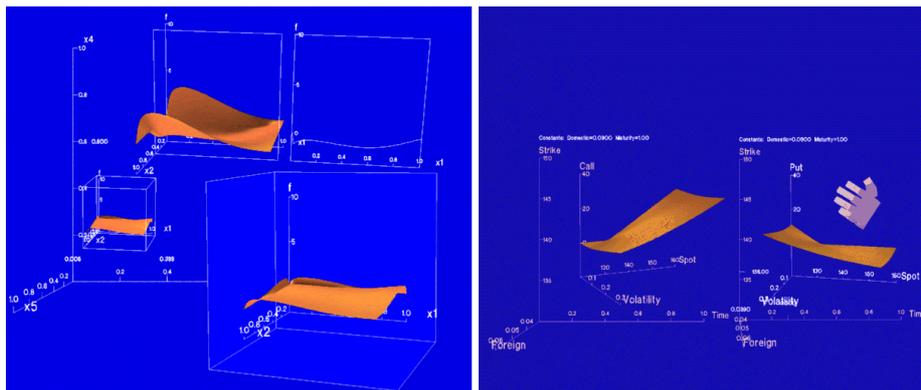


Figura 22 - Screenshots da ferramenta AutoVisual

As técnicas hierárquicas caracterizam-se por serem preferencialmente aplicáveis em espaços de dados densos, onde produzem bons resultados. Em espaços de dados dispersos há um maior desperdício de espaço de ecrã. São também técnicas de grande utilidade ao analisar dados categorizados, pois são de fácil hierarquização. É necessária uma rigorosa discretização de cada dimensão para garantir o mapeamento unívoco de pontos.

O maior problema intrínseco a esta classe de técnicas é a dificuldade de análise de dimensões não adjacentes. Tal como no caso das coordenadas paralelas, esta limitação pode ser atenuada permitindo a rápida reordenação de dimensões.

2.5. Bubble Charts

Os *bubble charts* ilustrados na Figura 23, constituem uma técnica de considerável divulgação que permite estender as potencialidades de um gráfico bidimensional permitindo representar variáveis com quatro ou mais dimensões de forma bastante intuitiva. Os valores são representados como bolas com os atributos ordenada, abcissa, dimensão, cor e eventualmente outros como a forma.

Ainda que bastante intuitiva, a principal desvantagem desta técnica consiste na incapacidade de representar um espaço de dados muito denso, que causaria uma saturação do *display*.

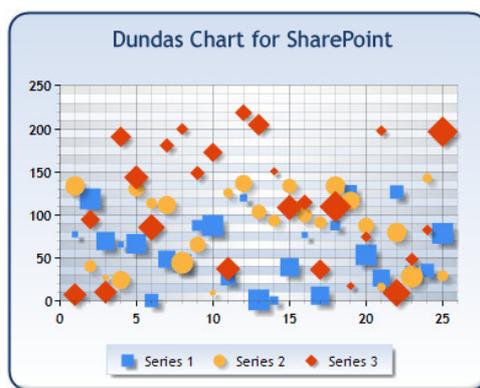


Figura 23 - Bubble Chart com variáveis de 4 dimensões (X,Y,forma/cor e tamanho), construído com a framework Dundas [17].

2.6. Floors And Walls

Esta técnica criativa baseia-se na metáfora da sala e faz recurso das paredes e do chão para representar informação. Permite que uma grande quantidade de informação seja apresentada numa cena 3D pequena. É flexível e permite o recurso a técnicas diversas através da utilização dos planos bidimensionais que constituem as paredes e chão do espaço virtual para mapear quaisquer gráficos bidimensionais, ou até mesmo tridimensionais como ilustrado no chão da Figura 24.

Esta técnica apresenta como principal vantagem a possibilidade de permitir ao decisor escolher a vista que compreende a informação que deseja analisar.

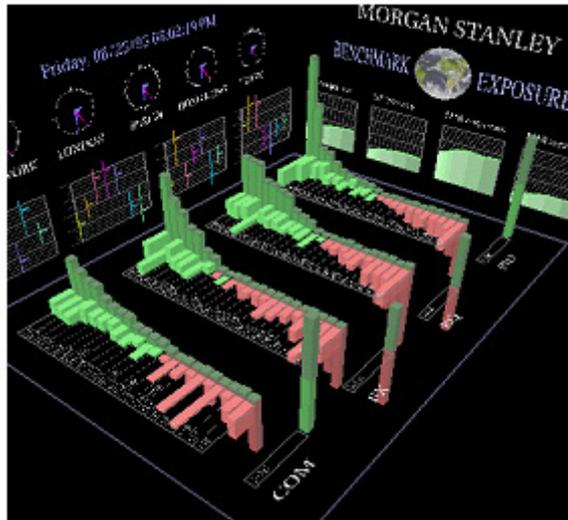


Figura 24 - Técnica floors and walls [18].

2.7. Técnicas Auxiliares / Boas Práticas

Porque só uma boa implementação pode permitir aproveitar as potencialidades das técnicas apresentadas, serão descritas em seguida algumas técnicas auxiliares de que podem ser aplicadas de forma genérica e complementar às técnicas anteriormente apresentadas.

2.7.1. Filtragem

A filtragem é uma técnica que permite apresentar uma grande quantidade de dados em conjuntos mais pequenos com características semelhantes.

Esta técnica torna-se fundamental quando nos apercebemos que a mente humana não está concebida para processar muita informação, sendo necessária a organização dessa informação em conjuntos tratáveis. Tipicamente conseguimos raciocinar com cinco a nove fragmentos de informação [19]. Selan Santos ao analisar técnicas auxiliares de visualização atribuí grande destaque à filtragem [1].

A grande desvantagem desta técnica consiste em que para se proceder à visualização de todo o espaço de dados, têm de ser apresentadas em múltiplas vistas, que terão de ser analisadas uma a uma. Esta desvantagem reside no facto de o ser humano apresentar limitações ao nível da concentração e da memória, capacidades fundamentais para reter e integrar os dados.

2.7.2. Brushing

O *Brushing* (pintar) é uma técnica de manipulação directa, sendo que existem duas formas de a aplicar em matrizes de diagramas de dispersão: rotulando as dimensões e estabelecendo ligações dinâmicas [3]. Rotular implica um ponteiro interactivo, usualmente um cursor de rato, que faça surgir rótulos com informação para determinado item (*mousetips*).

Ao estabelecer ligações dinâmicas, o ponteiro permite desenhar um rectângulo ajustável, usado para cobrir uma série de items num determinado painel. A informação retida no rectângulo é representada de forma realçada nesse e nos outros painéis.

Esta técnica torna-se uma ferramenta de manipulação directa muito poderosa, quando olhamos para os diferentes painéis e comparamos as diferentes perspectivas.

Num *display* interactivo e dinâmico, o *brushing* é mais eficaz. No geral, o *brushing* pode ser aplicado a muitas outras técnicas de visualização multidimensional e multivariada.

2.7.3. Banking

A percepção de propriedade de segmentos de linha pode ser melhorada ajustando as proporções do gráfico. Usando a técnica de *Banking* [3] as proporções do gráfico são definidas de forma a favorecer a percepção dessas propriedades. Um segmento de recta com um declive de 45º de inclinação é o mais adequado para observar a linearidade. A apresentação de um gráfico com proporções inadequadas distorce as linhas e implica assim um maior esforço para a identificação de características como a linearidade.

2.7.4. Outras Técnicas Auxiliares

- **Fitting:** Em visualização estatística, fitting significa encontrar uma curva suave que descreve um padrão subjacente, por recurso a técnicas tais como a interpolação polinomial.
- **Grelhas de Referência:** A utilização de grelhas em background, auxilia a percepção visual de padrões.
- **Animação:** Em teoria todas as técnicas de visualização baseadas em imagens estáticas podem ser estendidas a animações se os dados forem passíveis de representação na forma de série temporal.

Algumas técnicas poderão necessitar que os dados sejam submetidos a um pré-processamento. Normalmente estes procedimentos consistem em operações matemáticas cujos objectivos podem ser a redução da dimensionalidade ou a simplificação de algumas características dos dados.

Entre estas técnicas de pré-processamento encontram-se:

- **Multidimensional Scaling:** Encontrar um sistema de coordenadas que preserve as distâncias relativas entre itens de dados.
- **Análise dos Componentes Principais [20]:** Procedimento matemático que transforma um número de variáveis possivelmente correlacionadas num menor número de variáveis não correlacionadas. Pretende reduzir a dimensionalidade mas retendo a variabilidade presente no conjunto de dados original.
- **Clustering [2]:** o agrupamento de itens de dados ou de dimensões de acordo com determinado critério poderá ser realizado visualmente ou algoritmicamente.
- **Projection Pursuit:** técnica cujo objectivo é encontrar projecções de baixa dimensionalidade que descrevam apropriadamente potenciais clusters, através da optimização de um índice de projecção.
- **Redes Neurais:** técnica computacional baseada na arquitectura cerebral dos animais e que permite organizar padrões subjacentes aos dados preservando as relações entre esses padrões.

2.8. Sumário

A área da visualização multidimensional e multivariada apresenta ainda alguns sinais de imaturidade. São exemplo disso a inexistência de uma taxonomia de técnicas visuais consensual e o criticismo que desvaloriza o valor científico das técnicas mais elaboradas.

A investigação nesta área irá progredir no sentido de validar e avaliar rigorosamente as teorias subjacentes às técnicas de visualização para que a compilação de bons princípios seja efectuada. Assim será possível saber com maior exactidão que técnicas produzem de facto os melhores resultados e em que situações.

A reduzida divulgação das técnicas descritas é devida a factores como o treino necessário para a correcta interpretação dos dados em análise, a ampla divulgação das técnicas tradicionais e a resistência natural a técnicas com que os analistas estão menos familiarizados.

2.9. Balanced Scorecard

“One of the most dangerous forms of human error is forgetting what one is trying to achieve”

Paul Nitze

Sendo o âmbito desta tese a visualização de indicadores de BSC, torna-se imperativa a contextualização relativa a esta metodologia, ainda que breve. Será feita uma abordagem focando o seu intuito e os aspectos essenciais. O BSC é uma das mais divulgadas e conceituadas metodologias de BPM. É a metodologia utilizada no *Spatial Dashboard*.

2.9.1. Motivação

Durante a era industrial os indicadores de apoio à gestão eram unicamente financeiros, como o *Return On Investment* e o crescimento de vendas. Esses indicadores revelam-se insuficientes para inferir aspectos de performance como a inovação e outras competências que as empresas da actualidade necessitam de dominar. Numa pesquisa levada a cabo pela Deloitte em parceria com a *Economist Intelligence Unit*, em 2004, [21] concluiu-se que:

- 90% dos executivos tem a convicção de que os drivers de sucesso actuais (satisfação de clientes, qualidade do produto ou serviço, eficiência e eficácia de processos de negócio, empenho de colaboradores) não são financeiramente mensuráveis;
- 86% das empresas consideram monitorizar de forma “Boa” ou “Excelente” indicadores financeiros e apenas 34% têm a mesma convicção para os não financeiros. Justificam este aspecto com **carência de ferramentas** e **cepticismo** no seu impacto na performance geral.

Kaplan e Norton, os autores do BSC, identificaram a necessidade de encontrar mecanismos de suporte à gestão que complementassem os indicadores financeiros. Além de insuficientes, os indicadores financeiros apenas são úteis para reportar acerca das actividades passadas, não ajudando os gestores a melhorar a performance no futuro.

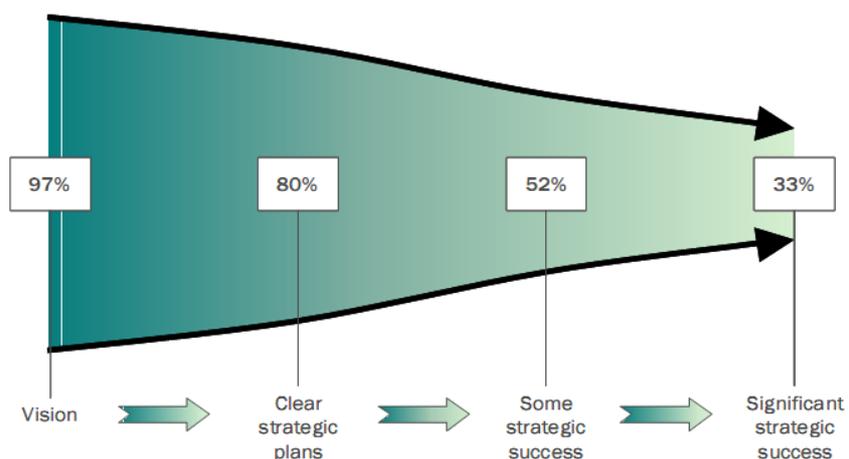


Figura 25 – Percentagem de empresas que concretiza a sua visão e estratégia [22].

Quase todas as empresas têm uma visão estratégica, no entanto apenas uma pequena percentagem a consegue concretizar, como ilustra a Figura 25.

Aquando da divulgação do BSC, as principais barreiras à eficaz gestão de performance eram as seguintes [22]:

- Elevado número de indicadores e conseqüente perda de foco;
- Utilização de sistemas complexos;
- Confiança injustificada em mecanismos de feedback informais;
- Objectivos vagos.

Estes aspectos levaram a que, na sequência de um estudo de um ano com doze empresas, Norton e Kaplan, autores do BSC, tirassem a ilação de que nenhuma métrica isolada poderá constituir um indicador claro de performance ou reflectir as áreas críticas do negócio [23].

2.9.2. O que é o Balanced Scorecard?

O BSC é uma das metodologias mais populares de monitorização de performance e comunicação estratégica. Foi considerado, por alguns autores, um dos 15 mais importantes conceitos introduzidos pela prestigiada Harvard Business Review, desde a sua criação [22]. A filosofia do BSC parte da premissa de que só se pode gerir adequadamente o que se pode medir [24].

Os seus autores definem-no como um conjunto de métricas que providencia aos gestores de topo, uma vista do negócio breve mas concisa. Inclui métricas financeiras que revelam os resultados de acções tomadas previamente que são complementadas com métricas operacionais, isto é, os drivers da performance financeira futura, e focam a satisfação dos clientes, processos internos e a aprendizagem organizacional [23]. Estas quatro categorias de métricas são denominadas de perspectivas sob as quais os gestores podem ver o negócio, ilustradas na Figura 26.



Figura 26 - Perspectivas do Balanced Scorecard [23].

Kaplan e Norton propuseram quatro perspectivas onde enquadrar as métricas. Ao obrigar os gestores a encarar as métricas operacionais em conjunto, o BSC deixa-os observar que melhoramentos numa área poderão comprometer outras áreas. A saturação de informação é prevenida restringindo o número de métricas a utilizar. A maior parte das empresas não tem falta de indicadores, apresentando uma panóplia de mecanismos de medida para actividades locais [23]. No entanto, estas medidas são frequentemente “*ad hoc*” e “*bottom up*” [24].

2.9.3. Balanced Scorecard como Sistema de Gestão Estratégica

O BSC introduz quatro processos de gestão [25] que contribuem para ligar objectivos estratégicos de longo prazo com acções a curto prazo: “Traduzir a Visão”, “Comunicar e Ligar”, “Planeamento de Negócio” e “Feedback e Aprendizagem”, descritos na Figura 27.

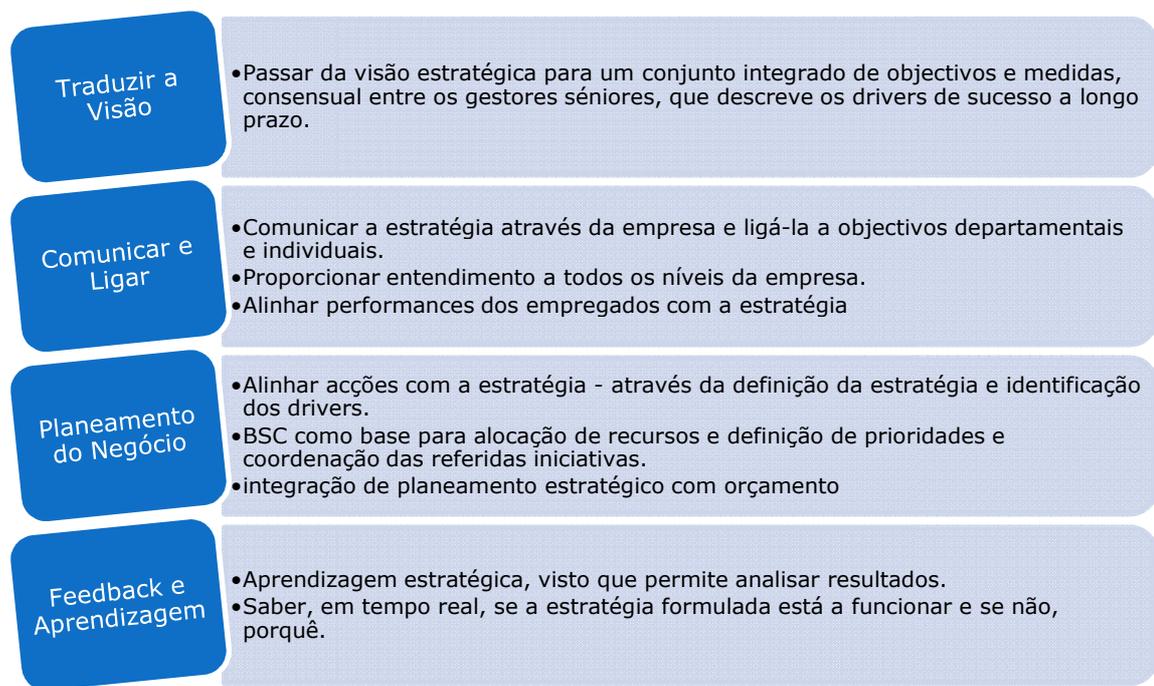


Figura 27 - Processos de Gestão introduzidos pelo BSC.

Ao utilizar o BSC para comunicar a investidores deve ter-se o cuidado de esconder informação sensível [25].

2.9.4. Implementação de Balanced Scorecard

Cada organização constitui um caso único e deverá seguir o seu caminho na implementação de *scorecard* [24]. A Figura 28 ilustra o perfil de uma implementação típica. Inicia-se a implementação desenhando os *scorecards* empresarial e o de uma unidade de negócio piloto (1). Em seguida (2) completam-se as unidades de negócio. A partir daí os *scorecards* são propagados num efeito de cascata para baixo pela hierarquia da empresa (3,4,5). Os *scorecards* de níveis inferiores alimentam os superiores. Observa-se que o *Corporate Scorecard* corresponde a uma reflexão estratégica *top-down* [24] mas será alimentado no sentido *bottom-to-top*.

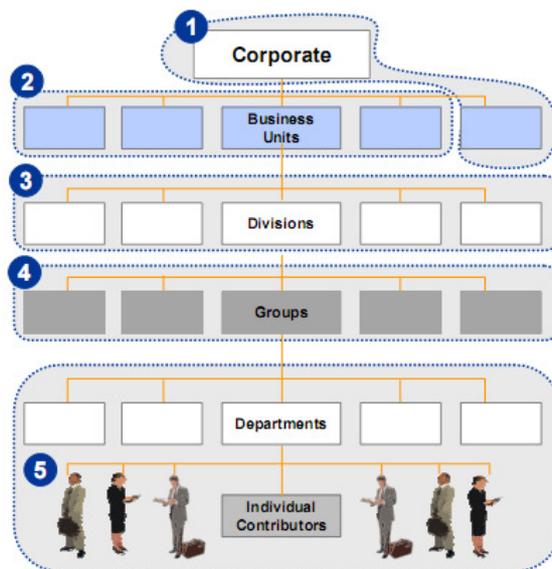


Figura 28 – “Cascata” de Scorecards: processo de Rollout típico [26].

Para desenhar o primeiro *scorecard*, deverá partir-se da visão estratégica e delinear objectivos. A partir destes medidas específicas e correspondentes valores *target*. Em cada passo os gestores responsáveis deverão partir do *scorecard* do nível superior para derivar os seus objectivos, métricas e *targets*.

Tipicamente as ferramentas de suporte à implementação de BSC em empresas, permitem observar os *scorecards* das diferentes unidades de negócio ou departamentos sob a forma de tabelas. Os paradigmas utilizados usualmente nas ferramentas tradicionais para organizar os *scorecards*, tais como a omnipresente metáfora das pastas e arquivos, não dão suporte a tarefas fundamentais para compreensão das relações entre *scorecards* distintos. Não permitem aos decisores navegar através dos *scorecards* de forma a explorar como a estratégia é traduzida em acções dos trabalhadores dos patamares inferiores da hierarquia da empresa. Veja-se o exemplo do software de referência *Hyperion Studio* da *Oracle* [27], na Figura 29. Ainda menos versátil é a adopção de folhas de cálculo como ferramenta de suporte à implementação de BSC.

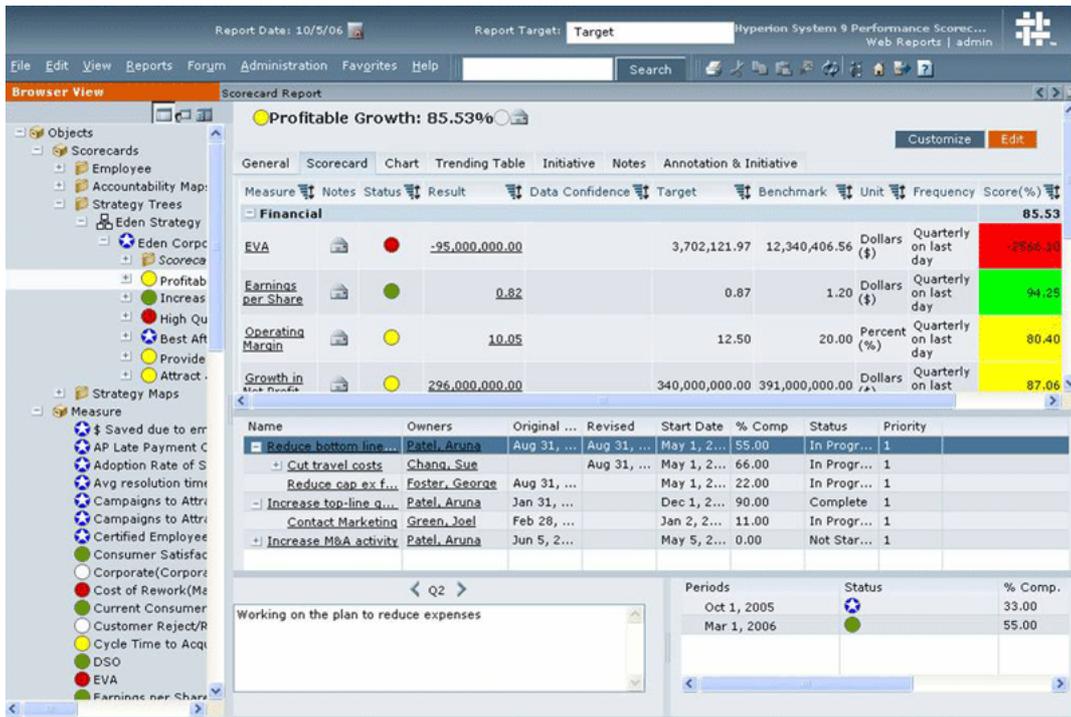


Figura 29 - Ferramenta de suporte a BSC Hyperion Studio da Oracle.

3. Problema

O problema a abordar nesta tese será considerado no contexto do *Spatial Dashboard*. Será apresentado primeiro de uma forma genérica, concretizando-se em seguida para a referida aplicação.

“Um elemento importante no processo de planeamento é identificar as ligações de uma unidade estratégica de negócio (UEN) com outras UEN na empresa e com actividades de suporte efectuadas ao nível empresarial. A ligação a outras UEN providencia oportunidades para o reforço mútuo de acções e partilha de melhores práticas. Estas oportunidades incluem o desenvolvimento e a partilha de conhecimento sobre tecnologias críticas e competências nucleares, coordenando esforços de marketing para clientes comuns, e partilhando recursos de produção e distribuição onde existam economias de escala ou de gama significativas. Uma das funções empresariais importantes é providenciar mecanismos onde tais oportunidades para sinergias através de unidades estratégicas de negócio descentralizadas possam ser identificadas e exploradas.”

Excerto de “The Balanced Scorecard: Translating Strategy into Action” [28]

Kaplan e Norton

Kaplan e Norton, os autores do BSC apontam a pertinência de identificar relações entre unidades de negócio. Referem também que o BSC providencia um mecanismo de suporte a essa tarefa: a comparação e análise de *scorecards* de unidades empresariais diferentes. Não providencia, no entanto uma forma explícita de as representar.

O software de suporte ao BSC, tipicamente permite visualizar apenas um *scorecard* de cada entidade organizacional de cada vez. No entanto há relações entre esses *scorecards*. Unidades de recursos humanos, de vendas, de manufactura fornecem indicadores de negócio que alimentam objectivos estratégicos. A imagem, extraída de uma apresentação sobre a realizada por Karl Sweeney, vice presidente e CAO da cadeia de hotéis Marriott, ilustra essas relações.

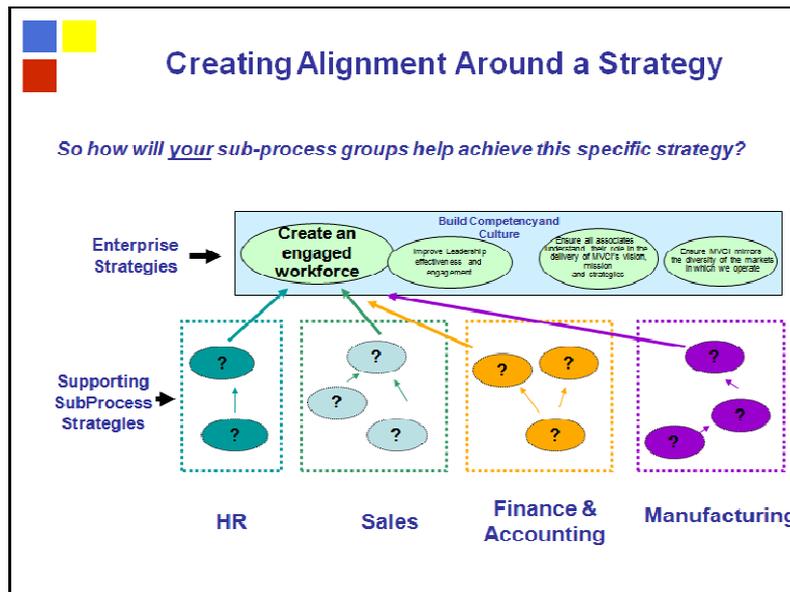


Figura 30 - Slide de uma apresentação de Karl Sweeney, CAO do Marriott Vacation Club International, na Brigham Young University, sobre a implementação de BSC na cadeia de hotéis Marriott [29].

Sendo a comunicação da estratégia empresarial uma das potenciais funções do BSC, não faria sentido representar os *scorecards* de uma forma que favoreça esse fim? As aplicações de suporte à implementação de BSC oferecem tipicamente interfaces tabulares, que permitem visualizar apenas um *scorecard* de cada vez. Essa abordagem é restritiva e não providencia mecanismos flexíveis de navegação. Dificulta imenso a possibilidade de representar toda a “cascata de *scorecards*” encadeados e interligados constantes de uma implementação de BSC. Uma tal forma de representar os *scorecards* contribuiria para facilitar a comunicação da estratégia empresarial aos colaboradores.

Os *scorecards* são um instrumento de suporte a tarefas de decisores. Pessoas com agendas preenchidas, com pouco tempo a despendar em actividades que não sejam críticas e cujo tempo de resposta a situações problemáticas seja crucial para o sucesso profissional. É de grande importância que os paradigmas de visualização com que os seus sistemas de informação lhes fornecem os dados sejam adequados às tarefas que desempenham.

Iremos agora contextualizar a problemática descrita anteriormente com o *Spatial Dashboard*.

A dimensão espacial dos dados é frequentemente subestimada ou mesmo ignorada em ferramentas de BI. A análise de informação espacial constitui um desafio devido à natureza do espaço geográfico, acarretando necessidades gráficas e analíticas específicas. No entanto, no contexto do suporte à decisão, toda a informação deverá poder ser analisada num ambiente homogéneo, suportado por uma integração transparente de tecnologias. Em particular a dimensão espacial, de grande importância na tomada de decisão. Partindo desta observação surgiu a ideia de desenvolver um sistema de suporte à decisão que, além da dimensão temporal, bem suportada pelas actuais ferramentas de BI, suportasse também a dimensão espacial: o *Spatial Dashboard*.

O *Spatial Dashboard* é uma ferramenta de BI baseada na metodologia do BSC. Nesta metodologia, a visão estratégica de uma empresa é traduzida em objectivos estratégicos, a partir dos quais são

derivadas métricas e *targets*, utilizadas para realizar a gestão da performance. O *Spatial Dashboard* utiliza a informação baseada em métricas para auxiliar o processo de tomada de decisão, através da monitorização de alterações graduais segundo a dimensão temporal e espacial. Essas métricas são representadas através de *Key Performance Indicators* (KPI) que são estruturados hierarquicamente em árvores de decisão, denominadas *dendrogramas*. Esta nomenclatura, derivada da palavra grega *dendron* – que significa árvore, será utilizada para evitar comparações com as técnicas de *data-mining* e também com o objectivo de manter a coerência entre os demais documentos produzidos no âmbito do projecto [30]. Em *data-mining*, as árvores de decisão são inferidas a partir de dados contínuos ou discretos e utilizadas como modelos preditivos. Nesta tese e no contexto do *Spatial Dashboard*, sendo distinto o propósito da sua utilização, utilizaremos as árvores de decisão como representações gráficas que simplificam a navegação em estruturas de dados complexas. Os dendrogramas são a estruturas que contêm os indicadores. Um dendrograma pertence a um cenário, que por sua vez pertence a um *branch* – o objecto do domínio do *Spatial Dashboard*, e que representa uma entidade organizacional (UEN, departamento, grupo, equipa, etc.).

O módulo do *Spatial Dashboard* que suporta a construção dos dendrogramas, o *Scenario Manager* permite observar apenas um dendrograma de cada vez. No entanto é possível importar KPI de outros dendrogramas, significando que um indicador de uma unidade de negócio poderá depender de um indicador de outra unidade de negócio. Chamamos a esta operação Clonagem. Esta operação cria relações de dependência entre as unidades de negócio/departamentos. À luz da filosofia do BSC, os indicadores de negócio apenas poderão ser copiados para níveis iguais ou inferiores da hierarquia empresarial, isto é, abordagem *top-down*, o que irá contribuir para a concretização da visão e estratégia da empresa. No entanto, estas relações de dependência não podem ser visualizadas na versão 1.0 do *Spatial Dashboard*, ou seja, a versão implementada no âmbito de três trabalhos de final de curso do ano lectivo 2005/06. Assim, a rede de dendrogramas é um conceito que não existe na versão 1.0 do *Spatial Dashboard*.

É pertinente a implementação de uma interface que providencie a um decisor mecanismos de rastreabilidade dos cenários que se encontram em produção, ou seja, que estão a ser utilizados para a medição da performance. Assim, no contexto do *Spatial Dashboard*, foram identificadas as seguintes lacunas, que poderão facilmente ser extrapoladas para contextos mais genéricos:

- *Scenario Manager* permite visualizar apenas um dendrograma de cada vez;
- Um gestor de topo, no contexto do *Spatial Dashboard*, não pode consultar uma visão geral correspondente à implementação de BSC na sua empresa;
- Não é possível saber, num determinado instante, que unidades de negócio ou departamentos não estão a corresponder aos níveis de performance estipulados aquando do desenho do *scorecard*;
- Não há um mecanismo navegável de suporte a uma análise de performance *enterprise-wide*, que permita identificar um departamento com um desempenho

abaixo do esperado e a partir daí explorar os indicadores cujas métricas se encontram distantes dos *targets* definidos.

As estruturas correspondentes às hierarquias de KPI construídas por *managers* e *decision makers* tendem a formar uma rede complexa de árvores de decisão com relações de dependência entre os seus nós, uma Rede de Dendrogramas. A compreensão dessas relações e da topologia da rede seria uma mais-valia para os decisores que analisam a performance empresarial e o impacto das relações de dependência entre indicadores de negócio de contextos distintos. Uma vez que os decisores não são, tipicamente, especialistas em ciências da computação, a representação dessa informação deverá ser feita de uma forma não técnica, com o suporte gráfico adequado e sincronização coerente entre as várias formas de representar a mesma informação. No contexto da filosofia subjacente à aplicação *Spatial Dashboard*, a dimensão espacial dos dados deverá ser suportada de forma integrada com os modelos de visualização a utilizar para representar a Rede de Dendrogramas.

Uma das qualidades reconhecidamente atribuídas ao BSC é o facto de facilitar a comunicação da estratégia empresarial aos seus empregados. Assim, a rede de dendrogramas deverá suportar essa característica do BSC, visto que a visualização é como que uma “banda larga” da comunicação, quando comparada com outras formas de comunicação, como a textual ou a sonora.

O objectivo de um *Dashboard* é essencialmente apresentar a “*Big Picture*” a quem compete decidir e gerir. O trabalho desenvolvido nesta tese parte da observação de que falta algo a essa imagem geral.

4. Proposta

O projecto *Spatial Dashboard* foi parte integrante de vários trabalhos de final de curso e também da tese de doutoramento do Professor Gabriel Pestana que acompanhou o trabalho desenvolvido no contexto desta tese. Assim, o trabalho foi iniciado partindo de uma visão clara dos objectivos pretendidos, o que foi de grande utilidade na delineação da proposta.

Para resolver as questões enunciadas no capítulo anterior, propomos um visualizador de redes de dendrogramas. As estruturas a visualizar serão construídas no *Scenario Manager*, outro módulo da aplicação *Spatial Dashboard*. As considerações relativas à modelação dos cenários utilizando árvores hierárquicas de indicadores de negócio saem do âmbito desta tese e foram o foco de outro trabalho de pesquisa desenvolvido em paralelo.

O *Spatial Dashboard* tinha, na versão a partir da qual partiu esta tese, as contribuições de três domínios principais: *DataWarehousing*, Sistemas de Informação Geográfica (SIG), e BPM, como ilustrado na Figura 31. O carácter inovador do projecto residia sobretudo na intersecção destes domínios, pouco explorada pela comunidade científica [30]. A contribuição para o *Spatial Dashboard* de cada um destes domínios é a seguinte:

- **BPM:** integra objectivos empresariais estratégicos com actividades e processos operacionais nucleares;

- **Data Warehousing:** suporta o repositório de dados empresariais recolhidos ao longo do tempo;
- **SIG:** exploração de correlações espaciais e topológicas e representação de KPI num *display* cartográfico.

Em sequência do problema descrito no capítulo anterior, propõe-se um novo módulo para o *Spatial Dashboard*, que lhe confira mais-valias do domínio da visualização de informação, como ilustrado seguidamente na Figura 32. As funcionalidades de visualização foram essencialmente conjugadas com conceitos de *DataWarehousing* e BPM, como induz a propositada colocação do domínio da visualização diametralmente oposto ao domínio SIG que sempre esteve fortemente ligado à visualização, ainda que numa orientação muito específica, a cartografia.

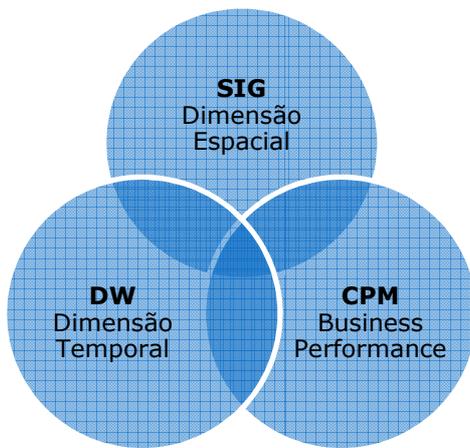


Figura 31 - Áreas científicas do Spatial Dashboard 1.0.

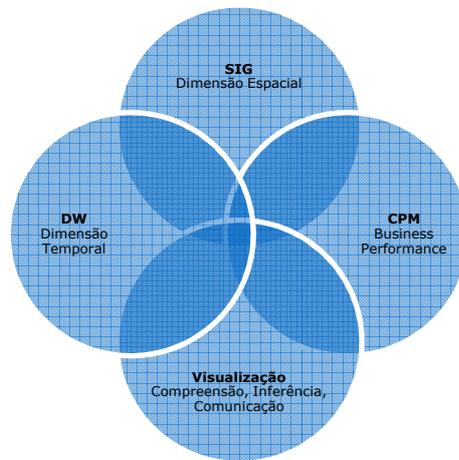


Figura 32 - Áreas científicas do Spatial Dashboard 2.0.

Para compreender um pouco melhor a contribuição deste trabalho para o *Spatial Dashboard*, será necessária uma breve explicação da aplicação essencialmente do ponto de vista do utilizador.

4.1. Overview do *Spatial Dashboard*

Sendo os indicadores as principais entidades do domínio do *Spatial Dashboard*, importa compreender as respectivas relações de especialização e conceitos relacionados.

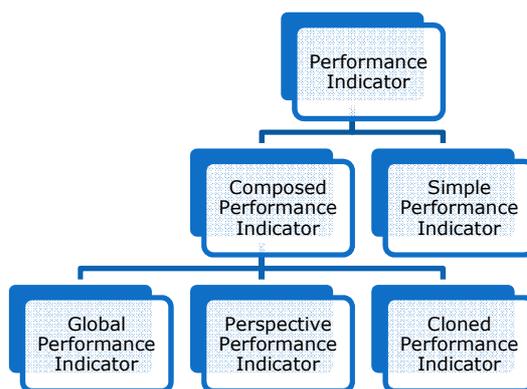


Figura 33 - Hierarquia dos indicadores de negócio.

Em seguida estão descritos os conceitos fundamentais do domínio do *Spatial Dashboard*:

- **Nível:** Nível de decisão. Por omissão pode ser estratégico, tático ou operacional. Podem ser adicionados outros níveis. Contém os *branches*.
- **Branch:** Entidade Organizacional, ou seja, uma unidade de negócio, um departamento, um grupo, ou no limite um individuo desde que faça sentido no contexto da empresa monitorizar o seu desempenho. Utiliza-se o termo *branch*, no sentido de ramo empresarial, fazendo alusão à forma como é representado na rede de dendrogramas, como um ramo de “*Corporate Performance*”. Cada *branch* contém um único cenário de produção a cada instante de tempo;
- **Indicadores de Performance ou Indicadores Simples (PI):** Indicadores alimentados directamente de um repositório de dados. Corporizam os conceitos de BSC métrica e *target*. São criados no *Semantic Manager*.
- **Indicadores Compostos ou Agregadores (CPI ou KPI):** Indicadores Compostos ou Agregadores. São alimentados por PI’s ou outros indicadores compostos. Mais do que um par métrica e *target*, representam um objectivo. São criados no *Scenario Manager*.
- **Clone:** Uma referência para um KPI. Transporta para um cenário diferente o estado do KPI que lhe deu origem.
- **Cenário de Negócio:** É a estrutura hierárquica de indicadores. No contexto da vista rede de dendrogramas, esta designação será também usada para nos referirmos a um nó da rede. Existem dois tipos: de produção e de teste. Os cenários de produção são os que estão activos, a monitorizar o desempenho empresarial. Em cada instante de tempo há só um para cada *branch*.
- **Dendrograma:** O mesmo que cenário de negócio.

O *Spatial Dashboard*, na sua versão 1.0, estava organizado em quatro módulos.

- **Business Metrics:** constrói as métricas;
- **Semantic Manager:** constrói os indicadores de negócio, associando *targets* às métricas;
- **Scenario Manager:** constrói os cenários de negócio;
- **Dashboard:** reporta o estado dos indicadores contidos nos cenários de negócio. Permite explorá-los, nomeadamente sobre uma base cartográfica, caso exista.

Um cenário pode ser comparado a um mapa estratégico, um conceito criado pelos mesmos autores do BSC. Num mapa estratégico as perspectivas são ordenadas ficando a perspectiva financeira no topo, permitindo observar as implicações entre indicadores e como estes acabam por contribuir para os indicadores da perspectiva financeira. Num cenário do *Spatial Dashboard* as perspectivas são tratadas de igual forma.

4.2. Módulo Rede de Dendrogramas – Conceitos de Suporte

Os indicadores de negócio, fragmentos elementares da informação a representar, são objectos intrinsecamente **multi-dimensionais**. Compreendem os seguintes atributos que se pretendem visualizar:

1. Dimensão temporal;
2. Nível hierárquico (Estratégico, Tático, Operacional ou outro);
3. Tipo de indicador (distinção entre PIs, indicador de baixo nível, próximo de um repositório de dados e KPIs, indicadores chave, de mais alto nível, próximo de um objectivo empresarial);
4. *Status* do indicador (a proximidade a que a métrica se encontra do *target value*);
5. *Peso do indicador*.

No contexto do *Spatial Dashboard* compreendem ainda:

6. Contexto espacial;
7. Identificação do indicador de negócio ou da entidade empresarial a que corresponde o nó.

Esta observação motivou a pesquisa sobre VIM apresentada no capítulo 2.

Os principais requisitos que se seguem ao problema descrito no capítulo anterior, são as necessidades de representar graficamente vários cenários de cada vez e de representar também as relações entre eles. É desejável representar os cenários de forma coerente entre os vários módulos do *Spatial Dashboard*. Desta forma, no módulo que propomos deveremos recorrer ao mesmo paradigma

da árvore utilizado no *Scenario Manager*. Tendo em conta a supracitada multi-dimensionalidade dos indicadores de negócio e as técnicas de VIM pesquisadas, há que realizar duas tarefas:

1. Escolher a técnica ou conjunto de técnicas adequado ao problema;
2. Definir o mapeamento de atributos dos dados para atributos gráficos, ilustrado na Figura 1, no capítulo 2.

4.2.1. Conjunto de técnicas escolhido:

Sendo um dos principais objectivos deste trabalho a representação de relações entre cenários que correspondem a relações entre entidades organizacionais, propomos uma vista em que as entidades atómicas em visualização correspondam a entidades organizacionais. Denominá-la-emos de vista *Dendrogram Network*. Cada entidade organizacional, que será representada nesta vista como uma esfera, contém um dendrograma, com a respectiva estrutura de indicadores de negócio. Assim, propomos uma funcionalidade de *Drill*, que permita explorar o dendrograma de cada entidade organizacional numa nova vista. Esta ideia é baseada na técnica *Worlds Within Worlds*, descrita no capítulo 2. A escolha desta técnica e algumas considerações que serão apresentadas num dos tópicos seguintes propiciaram a opção de uma abordagem baseada no recurso a um espaço virtual tridimensional.

A técnica *Floors and Walls*, consiste na utilização das paredes, base e topo de um espaço virtual para apresentação de informação. Neste caso será utilizada uma parede, para representar o contexto cartográfico associado a um determinado cenário ou indicador.

Da técnica *Bubble Charts*, a solução que propomos herda a utilização do raio e da cor das esferas para representar, respectivamente, o peso do indicador e o seu estado.

A elevada dependência de um determinado indicador poderá ser identificada de forma semelhante à que nos permite perceber um cluster usando coordenadas paralelas. Apesar de não haver uma aplicação pura da técnica de coordenadas paralelas, esse mecanismo de inferência é análogo ao obtido. Esta possibilidade, na técnica coordenadas paralelas, resulta do desdobramento de dimensões e sua representação em eixos paralelos. No âmbito da proposta, há também um desdobramento de dimensões. Na vista *Drill All* os cenários estão mesmo em paralelo, à semelhança do que sucede nas coordenadas paralelas. Na vista Rede de Dendrogramas os dendrogramas estão dispostos em circunferência. Em ambas as vistas, uma convergência de linhas indica elevada dependência de um indicador.

4.2.2. Mapeamento:

Domínio de Dados → Domínio Gráfico

Tabela 1. Representação Gráfica dos atributos dos Indicadores.

Atributos	Vistas		
	Dendrogram Network	Drill	Drill All
Dimensão temporal	Calendário	Calendário	Calendário
Nível hierárquico	Posição vertical no espaço virtual (coordenada Y)	Posição vertical no espaço virtual (coordenada Y)	Profundidade no espaço virtual (coordenada Z)
Tipo de indicador	- Não é representado -	Forma do nó (KPI- cubo, PI-esfera, CPI clonado – cubo com <i>border</i>)	Forma do nó (KPI- cubo, PI-esfera, CPI clonado – cubo com <i>border</i>)
Status do indicador / entidade empresarial	Cor do nó	Cor do nó	Cor do nó
Peso do nó	Tamanho do nó	Tamanho do nó	Tamanho do nó
Contexto Espacial	Textura de um plano colocado em background	Textura de um plano colocado em background	Textura de um plano colocado em background
Entidade Empresarial	Label do Nó	Label no topo da interface	Label do Nó raiz respectivo.
Identificação do Indicador	- Não é representado -	Label do Nó	Label do Nó

O nível hierárquico será representado através da posição vertical ou da profundidade dos nós. Para ajudar identificação dos níveis hierárquicos, estes serão limitados por planos. Todos os nós situados entre dois planos consecutivos irão corresponder ao mesmo nível. Os planos deverão ser ocultáveis, mediante uma opção na interface, para não interferirem na análise entre vários níveis.

4.2.3. Espaço Virtual Tridimensional

Um requisito fundamental a que a proposta deverá responder será a representação de ligações entre dendrogramas. Será portanto útil visualizar vários dendrogramas no mesmo *display*, o que será possível na vista **Drill All**. Para a representação simultânea de várias estruturas complexas no mesmo *display* convém uma forma que permita aumentar o detalhe nas estruturas em análise sem necessariamente esconder as demais. Neste contexto e considerando também a opção da técnica *Worlds Within Worlds*, a escolha de um espaço virtual tridimensional como base para o paradigma de visualização que propomos, foi uma escolha natural. Houve ainda um conjunto adicional observações a suportar esta decisão:

- A tridimensionalidade providencia um *look and feel* semelhante ao que é frequente encontrar nos SIG, portanto será familiar a utilizadores de SIG;
- Aumenta o espaço virtual disponível pelo que viabiliza e facilita a aplicação combinada de várias técnicas de visualização pesquisadas como, a *floors and walls*;

- Estudos realizados sugerem que para tarefas que requerem integração de informação, a performance dos utilizadores é superior em *displays* 3D do que em *displays* 2D [31];
- O desenvolvimento de modelos mentais é mais preciso se for suportado por informação representada em ambiente 3D [31];
- Cooper [32] sugere que indivíduos confrontados com projecções bidimensionais de objectos estruturais, constroem imagens mentais dos objectos originais;
- Wiegmann [31] evidencia a eficácia dos *displays* 3D para técnicas de visualização em sistemas de monitorização, controlo e análise.

Para navegar no espaço virtual, de entre as várias metáforas de navegação disponíveis, escolhemos a metáfora do voo livre [33] em que o utilizador controla posição e direcção com alcance total de posições espaciais. Como a entidade a visualizar será sempre a mesma estrutura colocada no centro do ambiente virtual, será imposta a restrição de que as rotações farão rodar toda a cena sobre o eixo dos YY – abordagem comum em aplicações de modelação 3D e CAD.

Existem também algumas desvantagens recorrentemente apontadas à tridimensionalidade, como as ambiguidades associadas à percepção de distância, tamanho e profundidade [31]. No entanto o impacto destas limitações será minorado no visualizador proposto através do recurso a pistas visuais bem conhecidas na área da computação gráfica, como o efeito de nevoeiro (*depth cueing*), a oclusão de objectos e a utilização de referências estáticas, que permitem enfatizar a sensação de distância e a noção de um espaço virtual coerente.

4.2.4. Vistas e Navegação

O Visualizador da Rede de Dendrogramas suporta três vistas distintas, para um determinado instante de tempo. A utilização de múltiplas vistas advém de um princípio de desenho de interfaces de visualização de informação a que alguns autores de referência dão grande importância, a filtragem [1]. Este princípio, já mencionado no capítulo 2, consiste em dividir o universo de dados a representar em conjuntos mais pequenos, com características semelhantes. Isto será implementado na vista *Drill*. Outra consideração que conduziu à opção das múltiplas vistas foi uma das conclusões de um estudo desenvolvido no *MIT Media Laboratory* que identifica a importância do recurso a formas complementares para observar uma rede de dados interligados [34].

As vistas propostas são:

- ***Dendrogram Network*** – Esta vista é a principal e deverá ser utilizada como ponto de partida para as outras duas. O dendrograma de cada cenário em produção é representado como uma esfera. Permite observar como as entidades

organizacionais se relacionam umas com as outras, sob a perspectiva da implementação de BSC na organização.

- **Drill** – Explora um dendrograma permitindo observar a sua estrutura de indicadores. A estrutura observada é similar à utilizada no Scenario Manager.
- **Drill All** – Corresponde a executar um Drill a todos os dendogramas.

Na vista *Drill All* observou-se ser mais vantajoso desenhar a raiz dos dendogramas com a mesma cota. Inicialmente tinha-se pensado desenhar de forma análoga à Rede de Dendogramas, ou seja, fazendo corresponder a cota ao nível hierárquico, mas tal iria implicar que as linhas que representam as relações entre os indicadores fossem muito maiores do que o necessário se fizermos corresponder os níveis hierárquicos à profundidade. Outra vantagem é o facto de se evitar que as linhas intersectem outros indicadores, em particular os estruturais "*Global Performance*" e os das Perspectivas.

Todos os nós contribuem para o estado do indicador raiz "*Corporate Performance*" o que é representado pelas ligações deste nó a todos os outros.

O fundo do volume de visualização deverá ser representado numa cor não pertencente ao leque de cores utilizado para representar os indicadores. Tons de azul ou cinzento apresentam um bom contraste com o gradiente de cores verde-amarelo-vermelho.

4.3. Considerações adicionais

A proposta apresentada tem ainda um enorme potencial de configuração, devido à considerável disponibilidade de atributos gráficos que poderão ser utilizados para mapear dimensões do espaço de dados. No entanto há que ter em consideração que essa configuração é um processo delicado e deverá ser feita com rigor visto que poderá causar uma saturação – *information overload*. Neste particular a referência a considerar será o trabalho clássico de Miller [19] que afirma que a nossa memória a curto prazo lida bem com sete “pedaços” de informação de cada vez (com uma margem de erro de 2). Alguns aspectos que acrescentam valor a esta proposta mas que não foram explorados encontram-se descritos no capítulo 7, como sugestão de trabalho futuro.

A representação dos indicadores em sistemas empresariais de monitorização de performance é actualmente feito com recurso a mecanismos bastante simples como semáforos, *gauges* e informação tabular. Não se conhecem à data da escrita desta tese esforços para a introdução de técnicas de visualização inovadoras neste domínio.

Uma vez que a interface a desenhar será utilizada para o suporte à decisão, será de grande importância que a linha de raciocínio dos utilizadores não seja interrompida devido a mecanismos de interação pouco eficientes. Simultaneamente deverá ser intuitiva. Assim, e seguindo uma *guideline* do conceituado perito em interfaces Jacob Nielsen, por motivos de flexibilidade e eficiência de interação devem haver mecanismos diferentes para utilizadores peritos e inexperientes [35]. Portanto, além da navegação através de uma *toolbar*, existirão também teclas de atalho. Outra recomendação do mesmo autor que será contemplada será a implementação de um mecanismo de informar os utilizadores acerca do estado da aplicação [35]. Fá-lo-emos através de uma barra indicativa do caminho percorrido semelhante às utilizadas frequentemente em interfaces Web. Finalmente, propomos para o visualizador implementação das funcionalidades de *Undo* e *Redo* como preconizado também por Jacob Nielsen [35].

5. Implementação

“We can't solve problems by using the same kind of thinking we used when we created them.”

Albert Einstein

5.1. Arquitectura

Uma grande vantagem da utilização de um motor tridimensional é a possibilidade de utilização de DirectX ou OpenGL, permitindo até alternar em *run-time*. A escolha de motor 3D resultou de uma análise comparativa entre frameworks diversas que se encontra em anexos. A Figura 34 ilustra como será feita a integração do motor Irrlicht .NET CP, o motor utilizado no protótipo, com o Spatial Dashboard e também com o hardware gráfico.

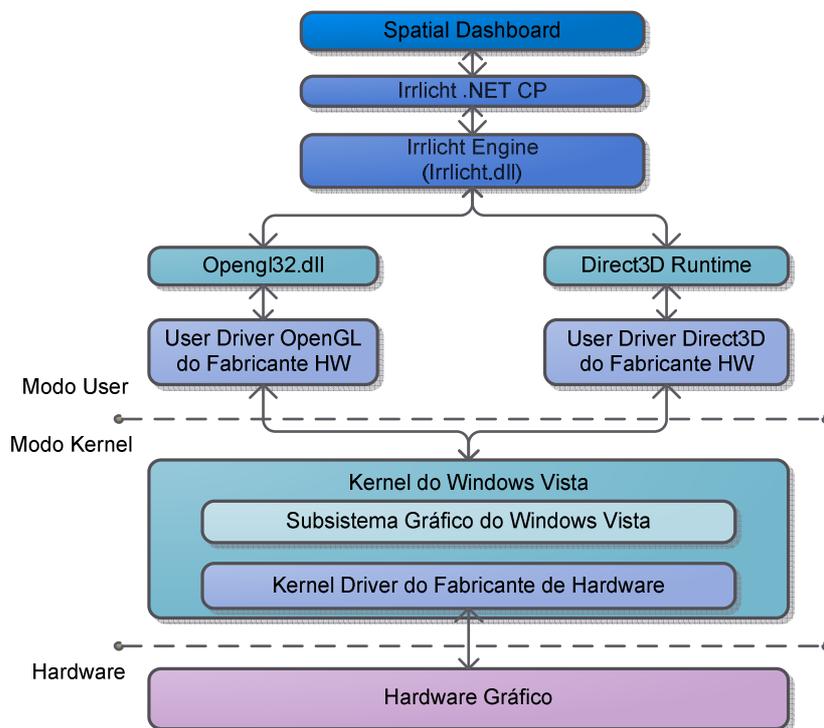


Figura 34 - Arquitectura da integração do motor 3D Irrlicht com o Spatial Dashboard e com o Hardware.

Adaptado de “How OpenGL fits into Vista” [47]

A classe *MainForm*, a classe de *start up* da aplicação, faz a inicialização dos contextos do *Spatial Dashboard*. Um dos contextos carregados corresponde à classe *Dendogram Manager Content* que tem

uma instância de *Dendogram Network App*, que por sua vez estende *IrrlichtApp*, fazendo assim a ligação com o motor *Irrlicht*. Como podemos ver representado na Figura 35. A transição entre contextos é feita com recurso a primitivas da *framework Weifen Luo* [48, 49].

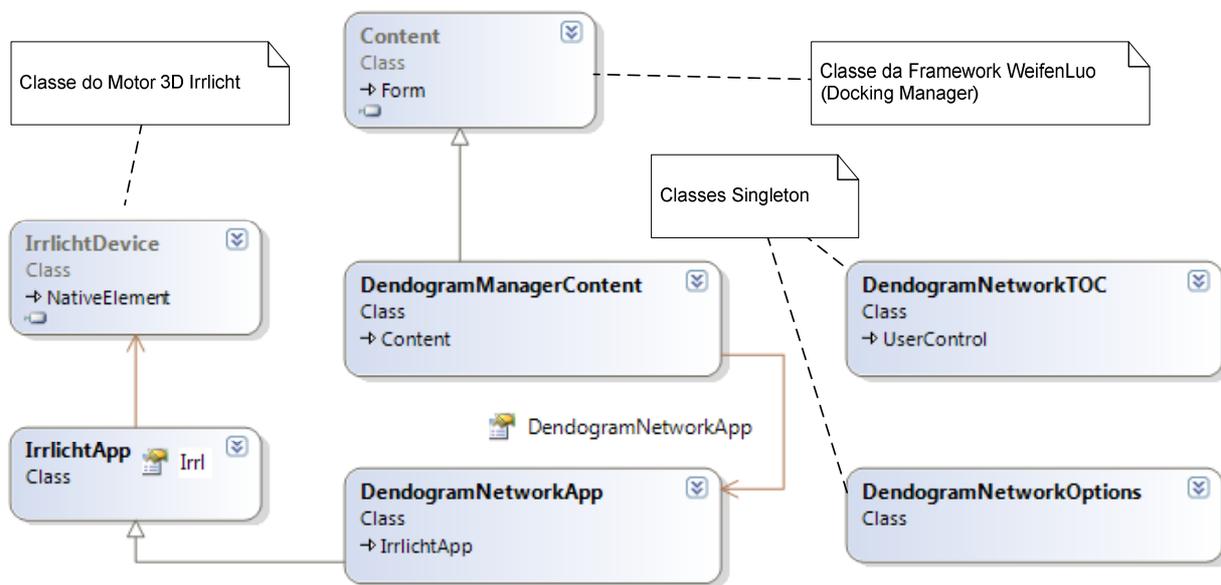


Figura 35 – Diagrama de classes relativo à integração do módulo Rede de Dendogramas com os principais componentes da aplicação.

As várias vistas podem ser encaradas como estados do visualizador de dendogramas. Esta observação conduziu à modelação dos objectos de acordo com o padrão de desenho *State* [50], cuja intenção é permitir que um objecto altere o seu comportamento quando o seu estado muda, como se fosse uma classe diferente. O comportamento do visualizador depende fortemente do seu estado e deve alterar o seu comportamento em tempo de execução dependendo do estado. Este padrão foi implementado aquando de uma refactorização motivada pelo recurso a instruções condicionais grandes em vários métodos cujas condições eram dependentes do estado do objecto, então representado por um enumerado. Com o padrão *State*, transferiu-se o código de cada ramo da estrutura condicional para uma classe separada que representa o estado.

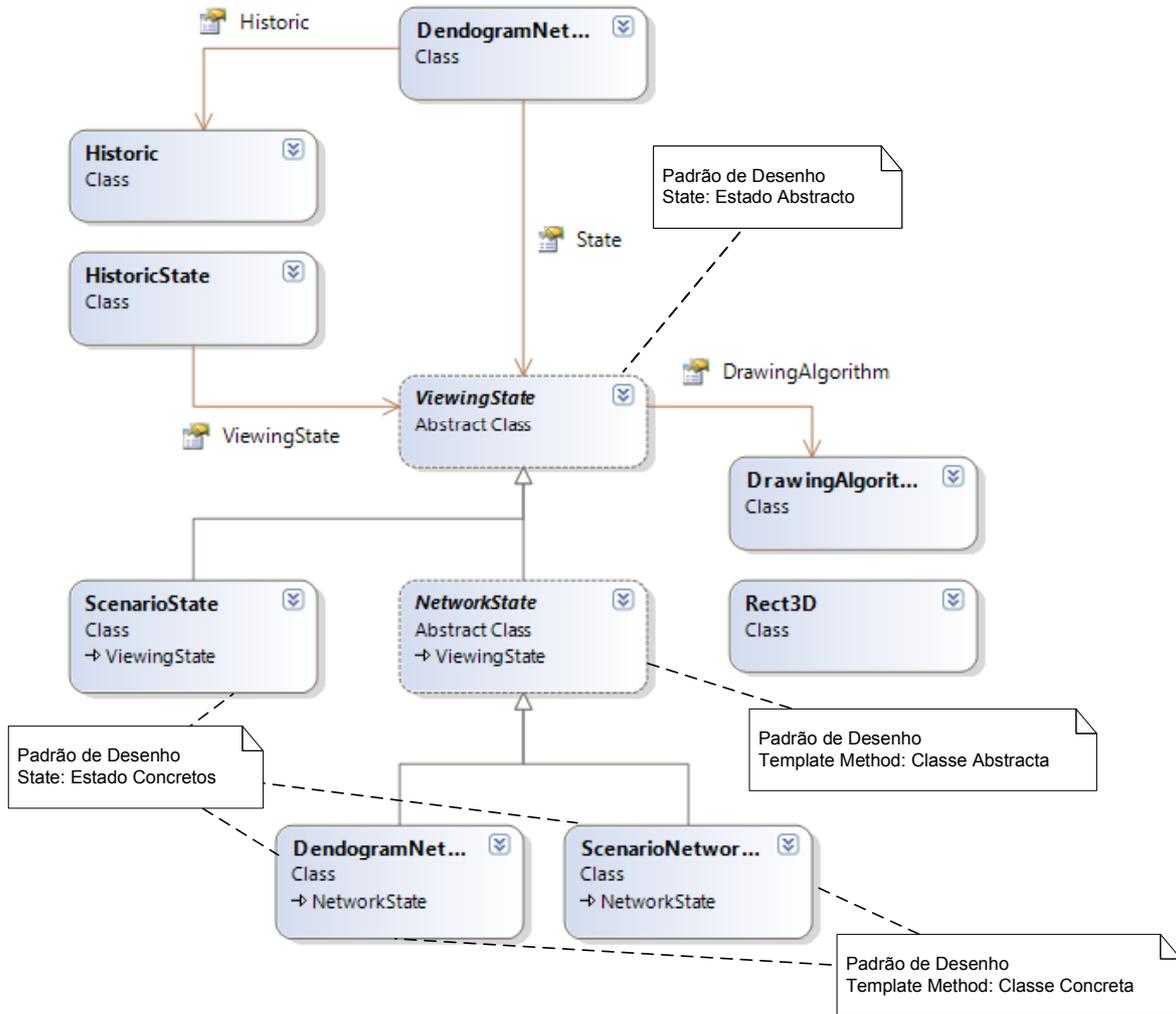


Figura 36 – Diagrama de classes correspondente à implementação das Vistas do módulo Rede de Dendogramas e respectivos padrões de desenho.

Obtiveram-se assim três objectos-estado: *ScenarioState*, *DendogramNetworkState* e *ScenarioNetworkState*. Observou-se que os dois últimos apresentavam muitas semelhanças a nível algorítmico. A estrutura de dados a percorrer eram essencialmente as mesmas, variando apenas a forma de representar os cenários e respectivo nível de detalhe. Fazia, portanto, sentido isolar o esqueleto do algoritmo principal e definir em objectos distintos os passos que apresentam as diferenças. Fazia sentido aplicar o padrão *Template Method* [50], o que, mais uma vez, foi concretizado numa refactorização. Este padrão consiste na utilização de classes derivadas para redefinir alguns passos de um algoritmo cuja estrutura está definida na classe pai. Evitou-se assim a duplicação de código e contribuiu-se para uma maior facilidade de manutenção e reutilização. A classe abstracta *NetworkState* possui o método *template* com a estrutura invariante do algoritmo de layout e os métodos abstractos variantes que são implementados nas classes filhas concretas *ScenarioState* e *DendogramNetworkState*.

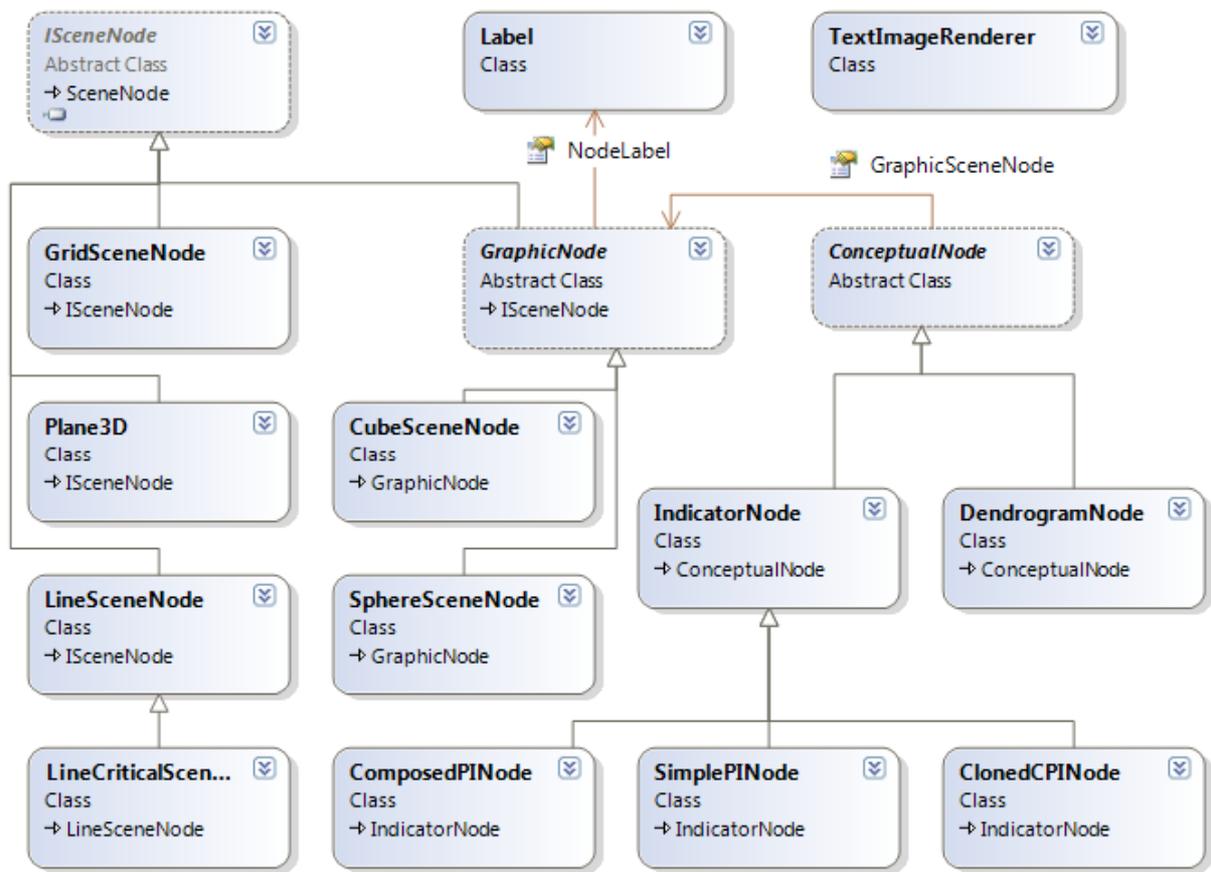


Figura 37 – Diagrama de classes do domínio do módulo Rede de Dendrogramas.

5.2. Funcionalidades

Faremos em seguida uma descrição geral das principais funcionalidades do visualizador implementado, tendo como ponto de partida uma *screenshot* do mesmo, a Figura 38.

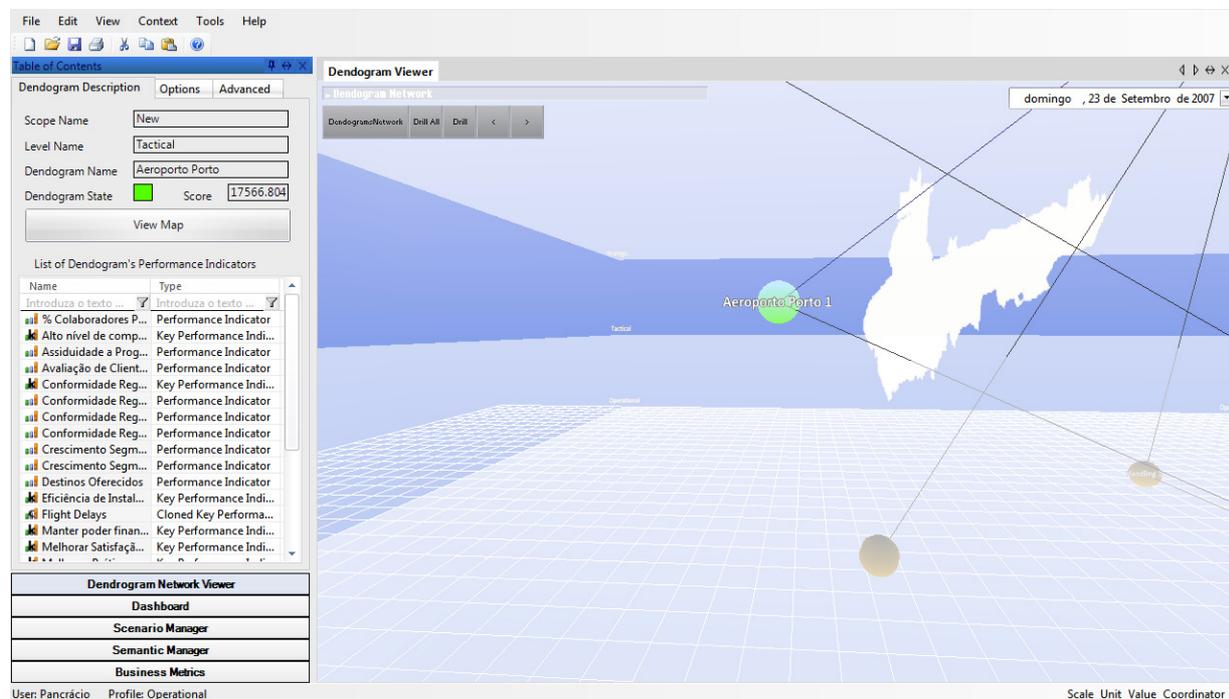


Figura 38 - Vista Rede de Dendrogramas: está seleccionado o cenário "Aeroporto do Porto".

Na *screenshot* anterior, podemos observar a interface da vista rede de dendrogramas. O nó seleccionado, no centro destaca-se por estar a transparente. Ao seleccionar um nó é possível consultar na *form* à esquerda um conjunto de dados a ele relativos: o nível, o nome, o estado, o *score* e a listagem de indicadores nele contidos – representados na *form* sob a label “List of Dendogram’s Performance Indicators”. Para pesquisar um indicador específico, basta introduzir o nome do indicador no campo “Name” sobre a lista de indicadores, que será feita uma pesquisa do tipo *search as you type*. O mapa representado no fundo do volume de visualização corresponde ao contexto cartográfico associado ao nó seleccionado.

A interface do visualizador apresenta uma *toolbar* no canto superior esquerdo. Permite alternar entre as diferentes vistas e possui ainda as funcionalidades de *back* e *forward*. Sobre a *toolbar* encontra-se a *path* que indica ao utilizador o caminho percorrido.

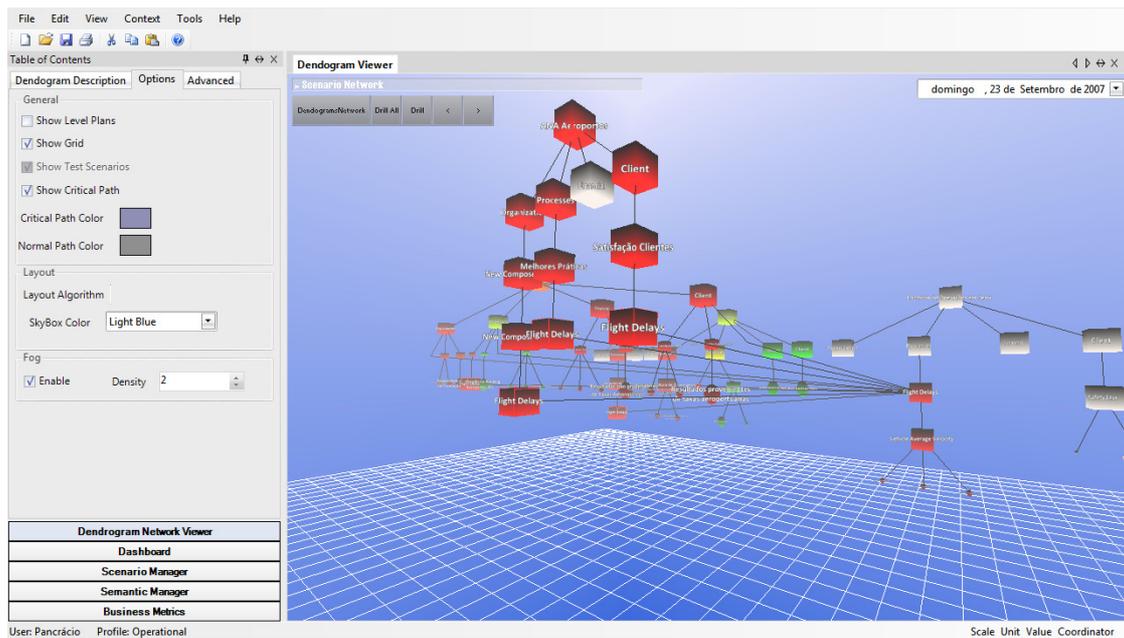


Figura 39 - Indicador clonado ao fundo do lado direito e clones à frente do lado esquerdo. O padrão das linhas é análogo ao que permite reconhecer clusters em coordenadas paralelas.

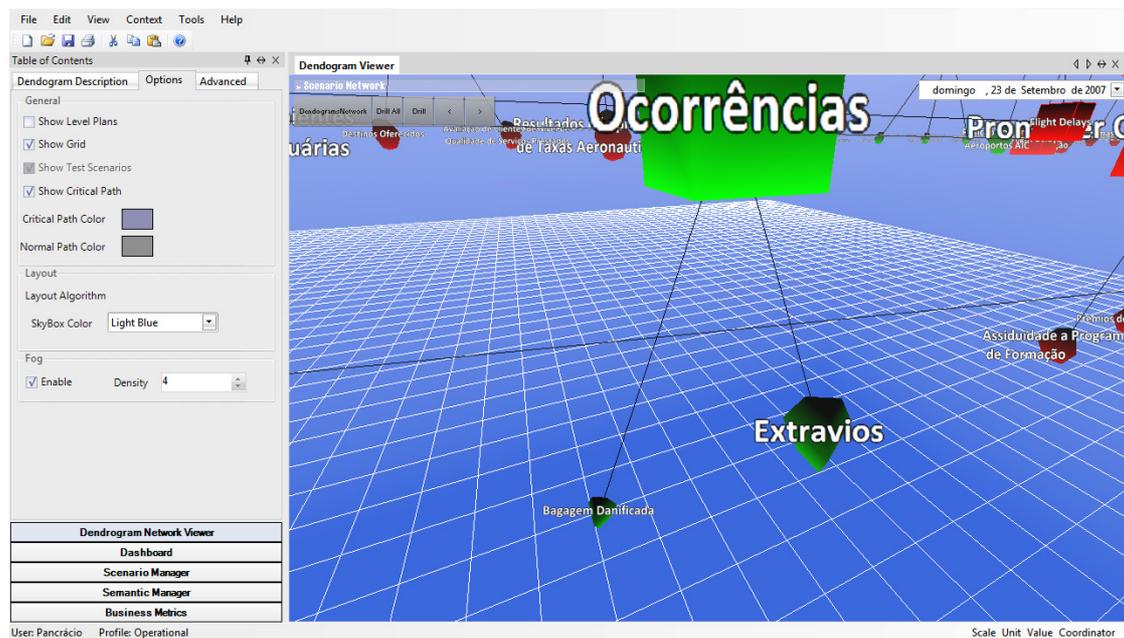


Figura 40 - Pormenor na vista Drill All: indicadores de diferentes tamanhos demonstrando que têm diferentes pesos.

A navegação, como referido no capítulo anterior segue a metáfora do voo livre, controlada com o rato, tendo os seguintes mecanismos de interação:

- Click + Arrastar com botão direito: Rotação da cena sobre o eixo dos YY.
- Click + Arrastar com botão esquerdo: Translação da câmara horizontal e vertical.
- Mouse-Scroll: Zoom.

- Click Simultâneo + Arrastar rato na horizontal: Zoom de pormenor (velocidade inferior à do Zoom com mouse-scroll).

Rotação da cena sobre o eixo dos YY, em vez de uma rotação da câmara sobre si própria confere maior usabilidade. Usando apenas esta funcionalidade de rotação de cena e o Zoom, a câmara fica sempre orientada para o eixo dos YY da cena. Uma vez que os dendrogramas são sempre desenhados no centro do volume de visualização, esses dois mecanismos serão suficientes para toda a navegação.

6. Caso de Estudo

6.1. Contexto

O caso de estudo utilizado neste projecto é o do Aeroporto Sá Carneiro. Foi também esse o caso de estudo da tese de doutoramento cujo âmbito abrange esta tese e demais projectos de fim de curso. Este conjunto de trabalhos herdou dados que foram cedidos no âmbito do projecto AirNet. O projecto AirNet tem o objectivo de desenvolver um sistema de controlo e monitorização de viaturas que circulam no lado terra de um aeroporto, sob uma base cartográfica. Conjuga informação de negócio do aeroporto (voos, aeronaves, tarefas, etc.) com informação recolhida dinamicamente de bases de dados.

Não dispusemos de informação estratégica relativa à totalidade do Aeroporto. Tipicamente as organizações protegem esse tipo de informação. Um *Scorecard* nas mãos de uma empresa concorrente poderá dar-lhe a conhecer vantagens competitivas ou informações sensíveis. Esta é, aliás, uma das razões que fazem desaconselhar a utilização do BSC como artefacto de comunicação do desempenho empresarial a investidores como discutido num artigo da Harvard Business Review dos autores do BSC [51]. Portanto, face a este constrangimento, complementaram-se os dados reais do projecto AirNet com dados fictícios, resultantes de uma recolha de informação sobre administração de aeroportos, em particular sobre a ANA Aeroportos. Criou-se assim um caso de estudo, baseado tanto quanto possível em dados reais, com o qual as funcionalidades implementadas foram testadas.

Da totalidade de dados relativos ao projecto AirNet que dispusemos, foram seleccionados os referentes a movimentos de veículos recolhidos ao longo de 21 dias nas diferentes áreas do aeroporto do Porto, nomeadamente: a *runway* (pista de aterragem e *take off*), a *taxiway* (ligação entre a *runway* e diferentes zonas do aeroporto que normalmente é assinalada por linhas amarelas) e a *apron* (zona destacada para cargas, descargas e manutenção de aeronaves) [52]. Estes dados foram utilizados para construir o *Scorecard* e o Cenário de Negócio de um departamento aeroportuário, “Centro de Operações em Terra”, responsável pelo controlo e monitorização de tráfego de veículos e aeronaves no lado terra do aeroporto. A informação de suporte para as restantes entidades organizacionais modeladas foi recolhida de um conjunto de fontes [29, 53, 54] pesquisado.

Um aspecto muito importante a considerar neste contexto é que as taxas pagas por uma companhia aérea à operadora do aeroporto são calculadas em função do tempo que o avião permanece em terra. Assim que o trem de aterragem toca o solo, o tempo começa a ser cronometrado e quaisquer operações a efectuar deverão ser feitas com a máxima celeridade e em conformidade com uma panóplia de critérios de segurança (definidos pela ICAO – *International Civil Aviation Organization*). É do interesse da companhia aérea e da operadora do aeroporto garantir que as operações se desenrolam com eficiência e segurança, para manter uma boa imagem, manter a confiança dos passageiros e

minimizar as filas de espera. Entre os critérios definidos pela ICAO destacam-se: velocidades máximas de circulação de veículos para as diferentes áreas do aeroporto, interdição de circulação de certos veículos em algumas áreas e distâncias mínimas entre veículos.

As razões mais significativas que podem contribuir para variações nos valores destes indicadores são a sobrecarga de utilização da pista (demasiados veículos) e a as condições meteorológicas. O controlo de tráfego no lado ar é de extrema importância nessas condições, pois a descoordenação entre os movimentos dos veículos propagar-se-á rapidamente aos demais sectores aeroportuários, num efeito bola de neve. Este cenário evocará certamente a qualquer pessoa que tenha feito algumas viagens de avião, recordações de longas horas de espera por um voo atrasado por motivos desconhecidos.

Para o caso de estudo, foram consideradas cinco entidades organizacionais da hierarquia da ANA Aeroportos, que foram colocados nos níveis estratégico, tático e operacional:

Entidade	Nível	Descrição
ANA Aeroportos	Estratégico	Gestão estratégica da ANA. Define o <i>Corporate Scorecard</i> .
Aeroporto do Porto	Tático	Direcção do Aeroporto do Porto.
Controlo de Operações em Terra	Operacional	Responsável pelo controlo de tráfego térreo.
<i>Handling</i>	Operacional	Responsável pela logística e processamento de mercadorias e bagagens.
Centro de Segurança Aeroportuária	Operacional	Responsável pela vigilância constante e protecção de pessoas e bens através de actividades de combate a iniciativas ilegais ou ilícitas.

Num exemplo real a modelação da hierarquia organizacional no *Spatial Dashboard* iria envolver um maior número de entidades, nomeadamente os restantes aeroportos da ANA (Faro, Lisboa, Ponta Delgada, etc.) e respectivos departamentos operacionais, que não foram contemplados para este caso de estudo. Traria uma complexidade adicional desnecessária para o teste das funcionalidades.

Suponha-se que a direcção da ANA identificou a necessidade de implementar um sistema de BPM. Foi criada uma equipa constituída por peritos do negócio, decisores das diversas entidades dos aeroportos da ANA, com o objectivo de modelar o negócio no *Spatial Dashboard*. O responsável de cada entidade organizacional (unidade de negócio, departamento, etc.) escolhe um conjunto de indicadores de performance, a partir dos quais construirá o respectivo cenário de negócio. Assim, a construção do projecto para o caso de estudo começa no *Semantic Manager*, onde são criados os indicadores de performance. Esta tarefa está ilustrada seguidamente, na Figura 41, onde podemos observar, pela cor verde dos indicadores, que os níveis de desempenho estão conforme as metas estipuladas.

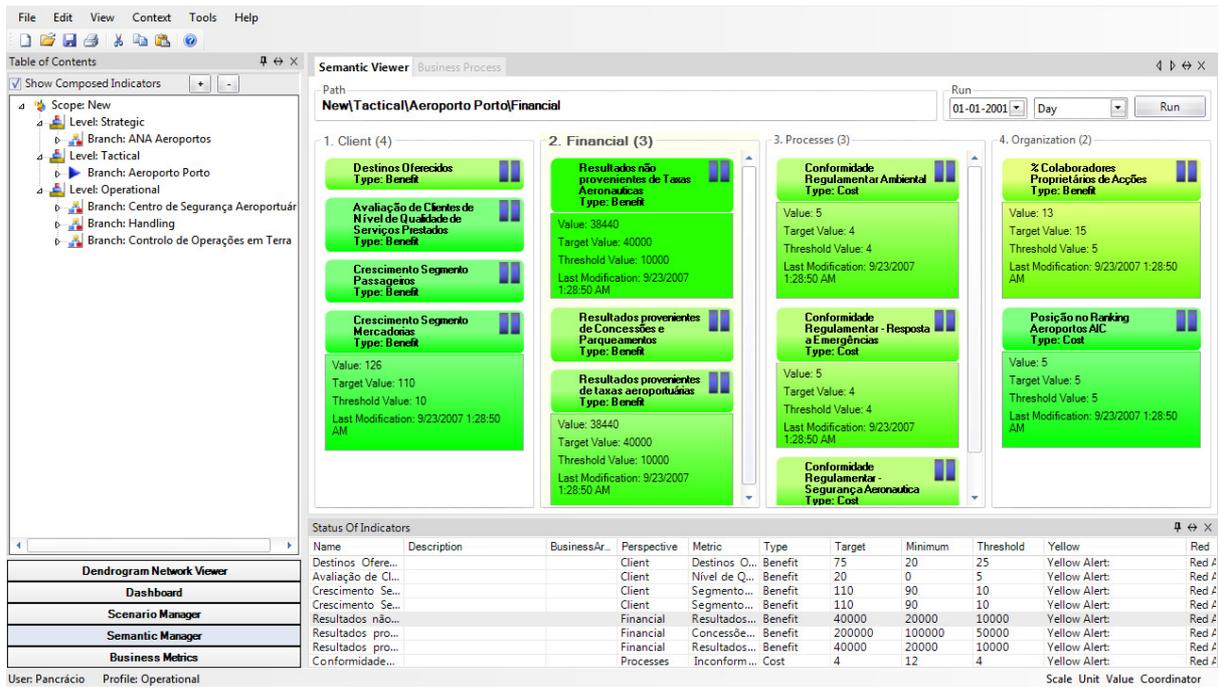


Figura 41 – Semantic Manager: indicadores de performance da entidade organizacional “Aeroporto do Porto”.

O passo seguinte será, para cada uma das cinco áreas a modelar, criar o respectivo cenário de negócio, onde os indicadores de performance serão combinados de forma a obter indicadores de performance compostos (agregadores de indicadores simples), conforme a Figura 42. As ilustrações correspondem à entidade “Aeroporto do Porto” mas os procedimentos serão análogos para as restantes.

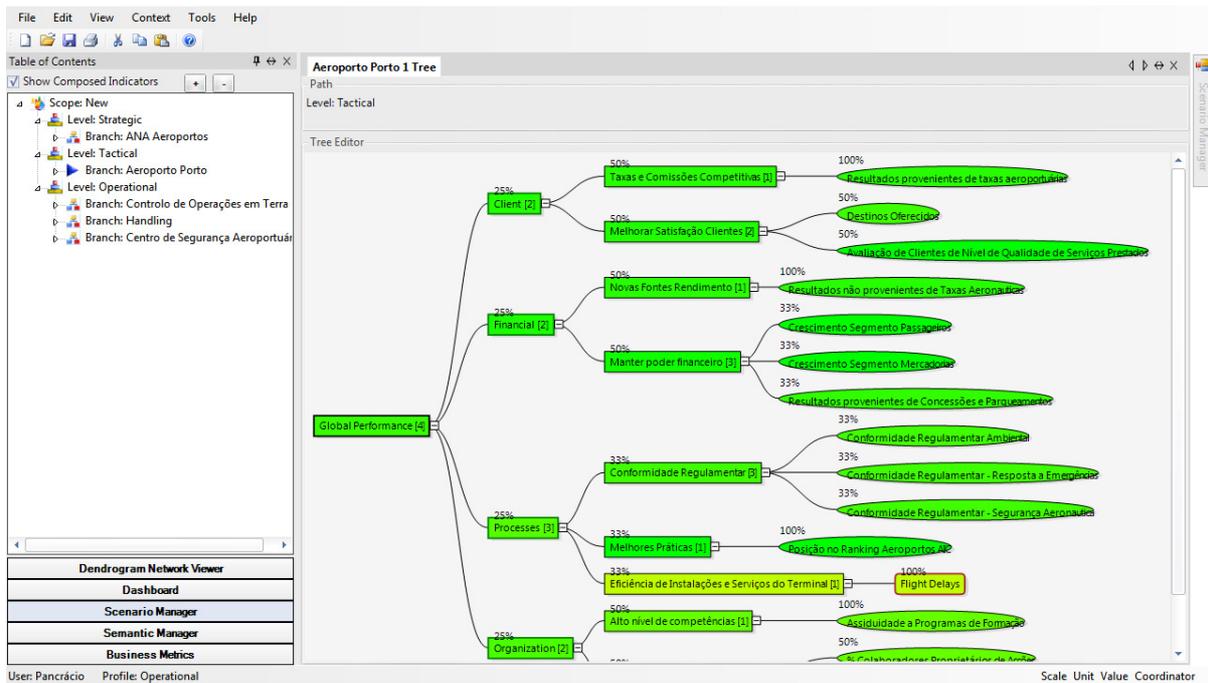


Figura 42 – Cenário de Negócio da entidade organizacional “Aeroporto do Porto”.

Depois de todos os elementos da equipa darem o seu contributo, ou seja, de modelarem no *Spatial Dashboard* a área de negócio/departamento que dominam, a Rede de Dendrogramas é gerada dinamicamente pela aplicação e foi analisada pelos administradores de topo da ANA.

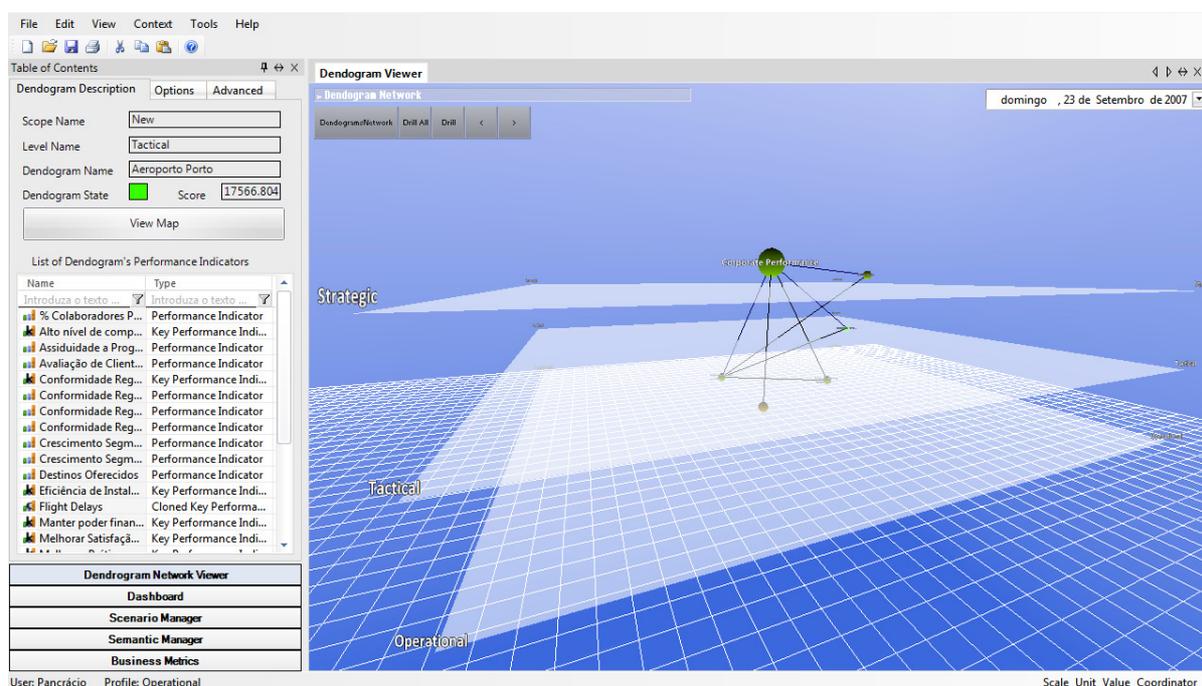


Figura 43 – Vista Rede de Dendrogramas com cinco dendrogramas.

Podemos observar, na figura acima, a rede de dendrogramas com os cinco nós que representam as respectivas unidades da ANA e também o nó “Corporate Performance”, o nó maior cujo estado resulta da contribuição ponderada aritmeticamente dos estados de todos os outros nós.

Existem dois aspectos que sobressaem à partida:

- um dos indicadores está representado a laranja, indicando assim uma quebra de performance – unidade empresarial com quebra no desempenho;
- existem vários nós ligados entre si, o que significa os cenários de algumas unidades dependem de indicadores presentes noutras unidades – relação de *cross-business*.

O decisor poderá realizar agora um *zoom*, como representado na imagem seguinte, para observar os nós em questão, identificando a entidade em sub-desempenho, a unidade de “handling”, e a entidade que tem as associações referidas, a unidade de “Controlo de Operações em Terra”.

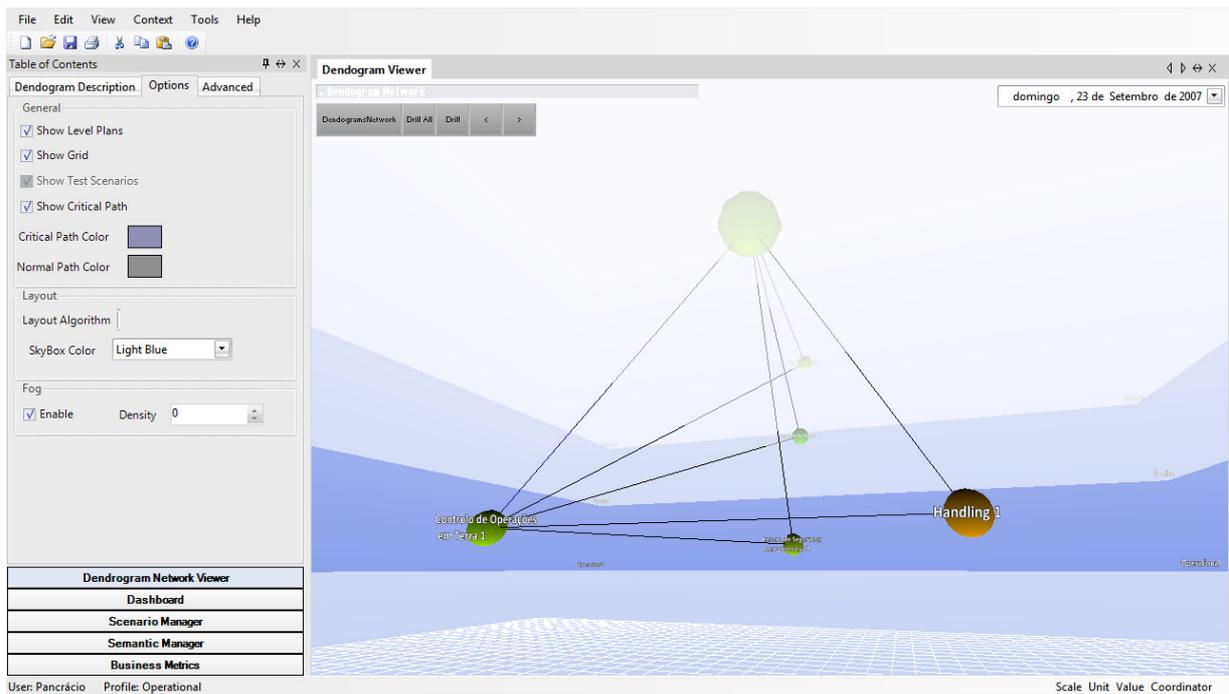


Figura 44 – Unidade de Handling representada a laranja denotando uma quebra no desempenho.

6.1.1. Unidade de Negócio com desempenho abaixo do pretendido

Para identificar a causa do estado alaranjado do nó “handling”, o decisor poderá seleccioná-lo e proceder à operação de *Drill*. Assim irá observar a estrutura de indicadores do dendrograma respectivo.

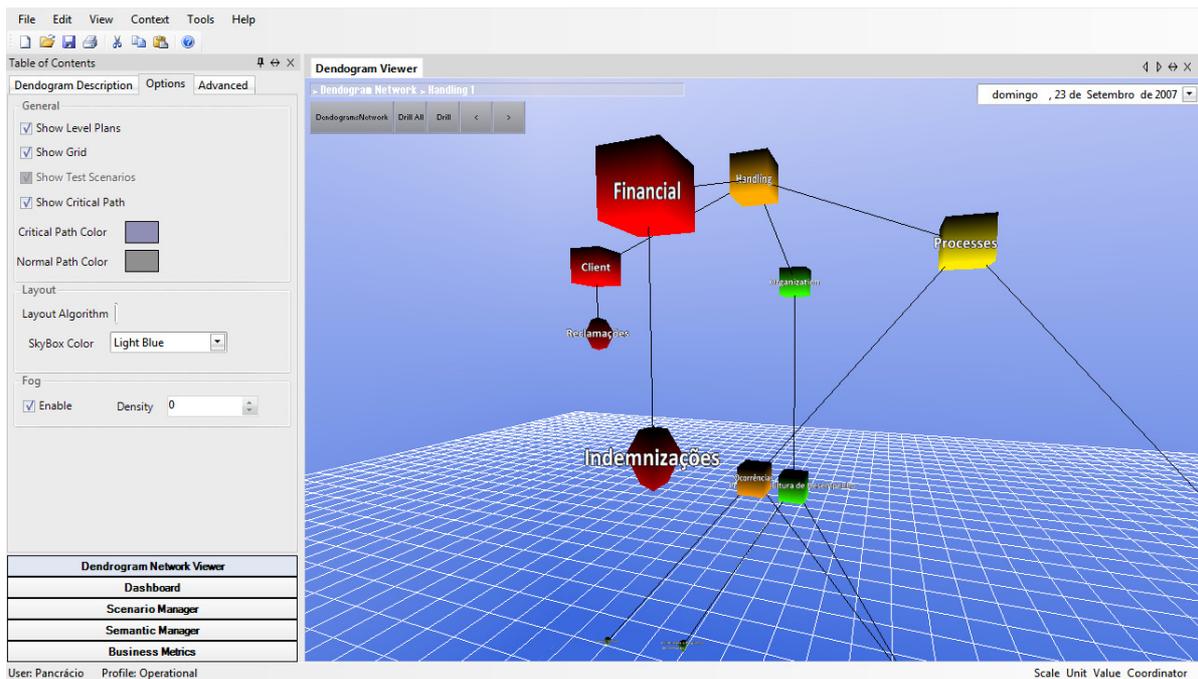


Figura 45 - Drill executado no dendrograma "Handling".

Rapidamente são identificadas duas perspectivas com um mau desempenho: *Client* e *Financial*. Os indicadores directamente responsáveis são respectivamente as "indemnizações" e as "reclamações". Outro nó merece a atenção do decisor, por se encontrar representado a laranja: "ocorrências", que é um indicador composto alimentado pelo indicador simples "extravios". Concluiu-se que este último estava na causa do sub-desempenho dos restantes. Em períodos de maior afluência verifica-se um grande número de bagagens perdidas que fez aumentar as "indemnizações" e as "reclamações". A situação levou à consideração de aquisição de um sistema de leitura óptica.

Considere-se uma outra situação possível de sub-desempenho de um indicador, por exemplo o indicador *Flight Delays* do cenário do "Controlo de Operações em Terra". Este caso será diferente visto que os indicadores simples que alimentam o indicador composto *flight delays* têm componente espacial. Merece menção também para destacar a integração da Rede de Dendrogramas com o módulo *Dashboard*: ao seleccionar um dendrograma na vista Rede de Dendrogramas e seguidamente mudar de contexto para o módulo *Dashboard*, carrega o cenário correspondente. Na Figura 46 vemos que o cenário foi carregado no módulo *dashboard*, estando os indicadores com componente espacial representados no mapa, os pontos a amarelo.

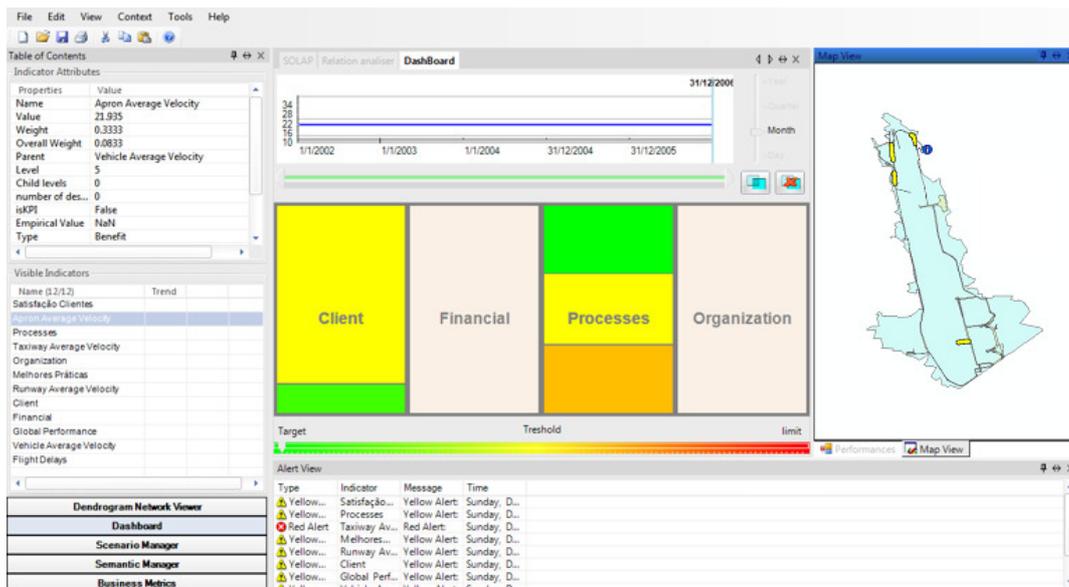


Figura 46 - Cenário "Controlo de Operações em Terra" no módulo dashboard.

Pode assim dedicar-se uma análise complementar (espacial) à quebra no desempenho. A razão foi então identificada e prendia-se com motivos de obras num troço da *taxiway*, que impedia os veículos de circular à velocidade desejável.

6.1.2. Relação de Cross-Business

Para ilustrar as relações de *cross-business* que podem ser identificadas com o paradigma de visualização proposto, inspirámo-nos num exemplo que David Norton e Robert Kaplan usaram para descrever a o papel do BSC para a identificação destas relações [29, 55] e adaptámo-lo para o nosso caso de estudo. O exemplo refere a cadeia de hotéis Marriott, onde o BSC ajudou a compreender como o desempenho da unidade de negócio "*Land, Development and Construction*" tinha influência directa no desempenho das restantes unidades da empresa: "*Marketing and Sales*", "*Mortgage Banks*" e "*Resort Management*". Quando a primeira unidade de negócio está em dificuldades as outras três unidades de negócio têm de tomar conhecimento e reagir de acordo com as necessidades da situação. O BSC contribui assim para criar sinergias entre diferentes unidades de negócio.

De forma análoga à do exemplo citado, o desempenho da unidade "Controlo de Operações em Terra" influencia directamente o desempenho das unidades às quais se encontra ligada: "*Handling*", "*Centro de Segurança Aeroportuária*" e "*Aeroporto do Porto*". Para melhor compreender a natureza destas ligações o decisor poderá utilizar o visualizador para identificar o indicador de negócio. Para tal, poderá realizar uma operação de *Drill All*, que irá expandir os cenários subjacentes a cada uma das entidades. Irá obter uma representação de acordo com o ilustrado na imagem seguinte.

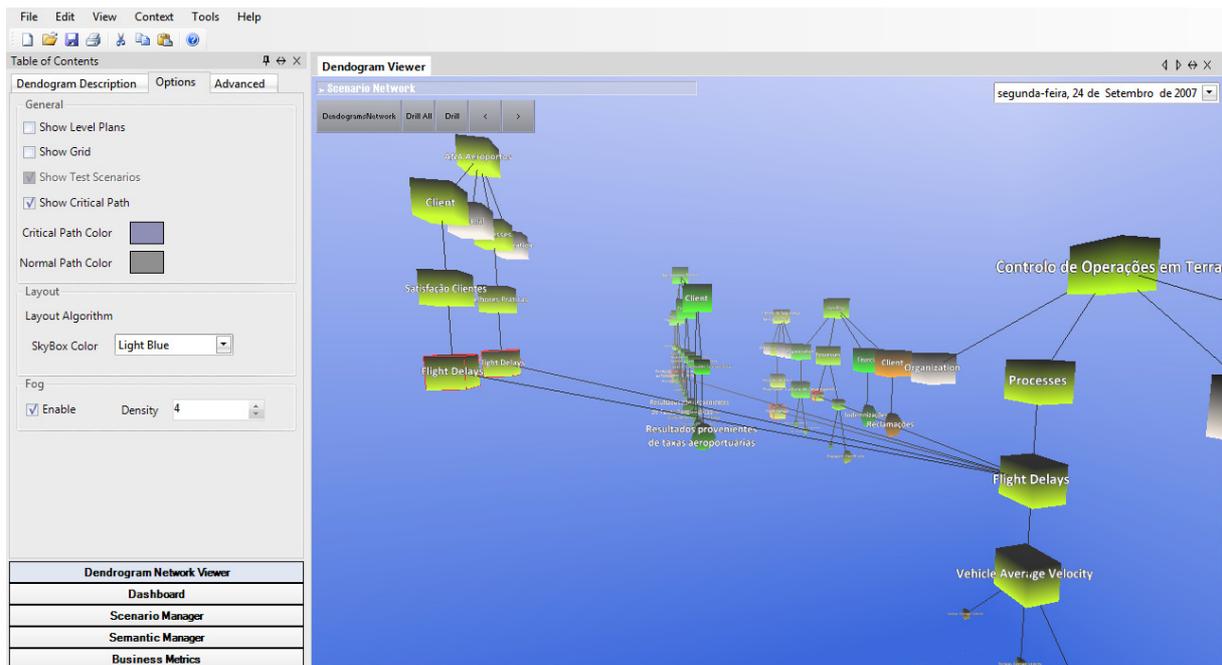


Figura 47 – Vista *Drill All* apresentando o indicador *Flight Delays* e alguns clones.

Conclui-se assim que o indicador responsável pelas dependências observadas na vista Rede de Dendrogramas é o *Flight Delays*. A Figura 47 esclarece ainda que existem dois clones do indicador *flight delays* no cenário de negócio do *branch* “Ana Aeroportos”, além dos clones presentes nos outros cenários mais distantes no ponto de vista da cena. Agora é uma questão de seguir as linhas e observar que indicadores estão a ser alimentados por estes clones e em que *branches*. Este conhecimento dará ao decisor uma boa compreensão de como a estratégia empresarial está a ser monitorizada pela empresa. Poderá agora fazer uso desta informação para criar sinergias entre estas unidades, por exemplo melhorando a eficiência de comunicações ou a partilha de recursos.

Os exemplos descritos ilustram uma das grandes intenções do BSC: traduzir estratégia em acções. O papel do *Spatial Dashboard*, em particular do módulo implementado no contexto desta tese, para além de suportar a implementação de BSC na empresa, consistiu em providenciar mecanismos navegáveis de análise para reduzir os tempos de resposta dos decisores a situações que necessitavam de intervenção.

6.2. Resultados

O caso de estudo utilizado permitiu ilustrar e demonstrar a praticabilidade dos conceitos propostos. As funcionalidades implementadas e a aplicação a casos de estudo foram analisadas e validadas pelo Professor Acompanhante do projecto, Gabriel Pestana, que é também o autor da tese de Doutoramento que criou as bases onde se fundamenta o protótipo *Spatial Dashboard* que aloja o visualizador de dendrogramas.

Seria agora de grande interesse o teste do protótipo desenvolvido num contexto empresarial complexo. Será provavelmente esse o objectivo de uma futura tese a realizar por um colega nos próximos anos. Foram já feitos esforços nesse sentido. Em particular a ANA Aeroportos e a Marinha Portuguesa foram contactadas com o intuito de testar o *Spatial Dashboard*. Relativamente à Marinha Portuguesa, chegou mesmo a estar agendada uma apresentação que não se concretizou por motivos a que fomos alheios.

7. Conclusão

7.1. Principais Contribuições

O trabalho apresentado reveste-se de uma considerável ambição na medida em que se baseia na intersecção de áreas científicas tipicamente disjuntas e portanto relativamente infértil em publicações científicas. No entanto este pretende ser um trabalho com características inovadoras e o cruzamento de conceitos de áreas científicas é um potenciador para a colocação de novas hipóteses e afirmação de novos paradigmas.

A crescente complexidade e quantidade dos parâmetros a equacionar para realizar um crescente número de tarefas do mundo actual implica a necessidade de mecanismos que permitam ver e raciocinar acerca desses mesmos parâmetros. O conjunto de mecanismos de visualização utilizados em sistemas de suporte à decisão tem vindo a crescer. Acreditamos que técnicas de grande criatividade constituem soluções para problemas de visualização específicos e permitem contornar as limitações cognitivas dos seres humanos. Esperamos que as constantes da proposta descrita anteriormente possam representar um contributo nesse sentido.

Esperamos ter demonstrado a aplicabilidade e utilidade da tecnologia de computação gráfica no âmbito da BI, cujas interfaces são tipicamente pautadas pela sobriedade de tabelas, *gauges* e gráficos 2D. Este terá sido o principal contributo alcançado. A apresentação profissional e imagem robusta que as aplicações de BI pretendem transmitir parecem restringir a criatividade e introdução de paradigmas mais inovadores nas mesmas.

Uma barreira que parece óbvia que paradigmas de visualização inovadores irão ter de enfrentar para vingar será a grande habituação dos utilizadores às representações tradicionais.

Esta conclusão foi tirada aquando da pesquisa sobre o estado da arte em técnicas de visualização multidimensional. Novos paradigmas de visualização requerem treino. Seja de alguns dias, horas ou minutos, isso constituirá sempre uma desvantagem visto que ninguém tem de aprender a consultar uma tabela. Como foi referido no Capítulo 2, a visualização encontra-se num estágio de desenvolvimento relativamente precoce. Procura-se actualmente uma taxonomia consensual entre a comunidade científica. Ultrapassada esta fase, estamos convictos que poderemos assistir a uma maior divulgação e penetração de técnicas de visualização mesmo em âmbitos atípicos como o de BI.

O protótipo, artefacto que corporiza os aspectos propostos, foi implementado de acordo com boas práticas desde o desenho da interface ao desenho arquitectural da aplicação que suporta os conceitos de alto nível do domínio do problema. Acreditamos que as considerações descritas no Capítulo 5 serão importantes num eventual redesenho para passagem a produção da aplicação.

Encaramos, no entanto, com reservas a apreciação do paradigma de visualização proposto e do protótipo implementado. Estamos conscientes que uma avaliação concludente apenas poderá seguir-se

a uma aplicação num contexto empresarial real envolvendo testes com utilizadores. Tal avaliação não fazia parte dos objectivos desta tese. Procurou-se suprir tal lacuna com um caso de estudo de relativa dimensão e complexidade. Permitiu explorar transversalmente as funcionalidades implementadas no protótipo correspondendo às expectativas do Professor Acompanhante.

7.2. Trabalho Futuro

No decurso do trabalho desenvolvido, foram identificadas algumas abordagens que seriam interessantes de explorar e que não o foram por se desviarem do âmbito ou por se secundarizarem perante outras de maior prioridade. Assim, caso se dê continuidade ao trabalho, recomendaríamos a utilização de uma versão do motor 3D *Irrlicht* desenvolvida em *C#* puro, utilizando uma *framework* que não seja um *binding* mas sim um *port* – ou seja uma versão completamente reescrita em *C#*, como o projecto *Irrlicht Sharp* [56] que disponibiliza um *assembly .NET* puro. Isso iria ter um impacto relevante na estabilidade da aplicação e esforço de desenvolvimento por permitir contornar algumas das condicionantes descritas em 5.8.

Outra abordagem de potencial valor seria a utilização de modelos tridimensionais construídos numa aplicação de modelação tridimensional como nós da rede de dendrogramas de forma a incutir semântica nos mesmos. A utilização de um motor 3D facilita o carregamento de malhas poligonais (*meshes*), ou seja colecções de vértices e polígonos que definem a forma de um objecto tridimensional complexo. As *meshes* são de utilização comum para exportação de objectos em aplicações de modelação gráfica tridimensional. O motor *Irrlicht* suporta os formatos de ficheiros de *meshes* mais comuns incluindo os do software de referência Maya e 3d Studio. Poder-se-ia assim substituir algumas das esferas por objectos mais gráficos representativos, como um cifrão para o nó da perspectiva financeira, um logótipo da empresa para o nó *Corporate Performance*, uma silhueta para o nó de Recursos Humanos, etc. Outra alternativa para introdução de semântica na rede de dendrogramas seria a utilização de imagens como texturas para os nós. Alguns estudos sugerem que a utilização de representações gráficas realistas em sistemas de suporte à decisão pode contribuir para que sejam tomadas decisões mais exactas [57].

Seria uma mais-valia a substituição do calendário por um *time slider* ou implementação da funcionalidade de animação (à semelhança da técnica *GrandTour*) para melhorar o suporte da dimensão temporal. Na versão actual do protótipo apenas se pode visualizar um instante de tempo de cada vez. A animação poderia abreviar para alguns minutos, uma tarefa de que tipicamente demoraria horas se feita com recurso a análise de gráficos e tabelas, como referem os autores do *GrandTour*, Buja e Asimov [15]. Um estudo que procurou responder à questão “*Does animation in user interfaces improve decision making?*” [57] concluiu que a introdução nos mecanismos de interacção de animações suaves, simples e adequadas ao modelo mental do utilizador para a tarefa pode contribuir para uma tomada de decisões mais adequadas e melhorar tempos de resposta.

O Adaptive Zooming, ilustrado na Figura 48, é um princípio de utilização comum em SIGs, que consiste no rendering automático de conteúdos e simbologia diferentes em função da escala na sequência de uma operação de zoom. Difere do zoom tradicional que apenas amplia sem alterar a informação representada. A Figura 48 ilustra a diferença entre o Zoom tradicional (à esquerda) e o adaptivo (à direita) de uma mesma selecção de um mapa (a região no interior do rectângulo, acima). A utilização desta técnica é defendida por alguns autores para a exploração de redes complexas de informação [34], como redes sociais.



Figura 48 - Adaptive Zooming na aplicação online Google Maps.

8. Bibliografia

1. Santos, S., *A Framework for the Visualization of Multidimensional and Multivariate Data*, in *School of Computing*. 2004, University of Leeds: Leeds.
2. *Tutorial on Clustering Algorithms*. [cited 2007 2007/02/01]; Available from: http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial_html/.
3. Bergeron, R.D. and P. Wong, *Thirty Years of Multidimensional Multivariate Visualization*, in *IEEE Computer Society*. 1997.
4. Butavicius, M.E.L., M. D.; Reilly, R. E., *An Empirical Evaluation of Chernoff Faces, Star Glyphs and Spatial Visualizations for Binary Data*, in *Proceedings of the Asia-Pacific Symposium on information Visualisation*, T.P.a.B. Thomas, Editor. 2003: Darlington. p. 1-10.
5. Datamology. *Exploratory Data Analysis*. [cited 2007 2007/02/01]; Available from: <http://www.datamology.com/eda.shtml>.
6. Ware, C., *Information Visualization: Perception for Design*. Interactive Technologies, ed. M. Kauffman. 2000: Elsevier.
7. Brodlie, K., *A Classification Scheme for Scientific Visualization*, in *Animation and Scientific Visualization*, R.E.E.a.D.W. Eds, Editor. 1993.
8. Shneiderman, B., *A Task by Data Type Taxonomy for Information Visualizations*, in *Proceedings IEEE Visual Languages*, I.C. Society, Editor. 1996.
9. Buja, A., [et al.], *Interactive High-Dimensional Data Visualization*, in *Journal of Computational and Graphical Statistics*, I.C. Society, Editor. 1996.
10. Keim, D.A., *Designing Pixel-Oriented Visualization Techniques: Theory and Applications*, in *IEEE Transactions on Visualization and Computer Graphics*, I.C. Society, Editor. 2000.
11. Ward, M.O., *XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data*, in *Proceeding IEEE Visualization*, R.D.B.a.A.E.K. Eds, Editor. 1994.
12. Davidoff, J.B., *The Specificity of Face Perception: Evidence from Psychological Investigations*, in *The Neuropsychology of Face Perception and Facial Expression*, R. Bruyer, Editor. 1986. p. 147-162.
13. Bergeron, R.D., G.G. Grinstein, and S. Smith, *Interactive Data Exploration With a Supercomputer*, in *Proceedings IEEE Visualization, San Diego*, G.M.N.a.L.R. Eds., Editor. 1991.
14. Keim, D.A. and K.H. P., *VisDB: Database Exploration Using Multidimensional Visualization*, in *IEEE Computer Graphics and Applications*, I.C. Society, Editor. 1994.
15. Asimov, D., Buja, A., *Grand Tour Methods: An Outline*, in *Proceedings of the 18th Symposium on the Interface*, A.S. Association, Editor. 1986.
16. Sachinopoulou, A., *Multidimensional Visualization*, in *VTT Electronics*. 2001: Espoo.
17. Dundas. *3D Bubble Chart*. Dundas Chart Screenshot Gallery 2007 [cited 2007 2007/08/20]; Available from: <http://www.dundas.com/Gallery/Chart/SharePoint/index.aspx?Img=BubblePoint13>.
18. Tegarden, D., *Business information visualization*. Communications of the Association for Information Systems, 1999. **1**(1es): p. 4.
19. Miller, G.A., *The Magical Number Seven, Plus or Minus Two*, in *The Psychological Review*, A.P. Association, Editor. 1956: Philadelphia. p. 81-97.
20. Smith, L. *A tutorial on Principal Components Analysis*. 2002 February 26,2002 [cited; Available from: csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf].
21. Melo, P. *Aplicação Prática do Balanced Scorecard*. in *Fórum APCER VII - Ferramentas de Gestão em Busca de uma Vantagem Competitiva*. 2004. Europarque, Santa Maria da Feira: Deloitte.
22. 2GC, *Introducing The Balanced Scorecard*, G.A. Management, Editor. 2001, 2GC Active Management.
23. Kaplan, R.S. and D.R. Norton, *The Balanced Scorecard: Measures That Drive Performance*. [cover story]. *Harvard Business Review*, 2005. **83**(7/8): p. 172.
24. Kaplan, R.S. and D.P. Norton, *Putting the Balanced Scorecard to Work*. *Harvard Business Review*, 1993. **71**(5): p. 134.
25. Kaplan, R.S. and D.P. Norton, *Using the Balanced Scorecard as a Strategic Management System*. *Harvard Business Review*, 1996. **85**(7/8): p. 150.
26. Norton, D., *Strategic Organization Alignment*. 2004, Balanced Scorecard Colaborative: Chicago.

27. Oracle. *Hyperion Performance Scorecard—System 9*. 2007 [cited 2007/08/10]; Available from: <http://www.oracle.com/appserver/business-intelligence/hyperion-financial-performance-management/hyperion-performance-scorecard-system9.html>.
28. Kaplan, R.S. and D.P. Norton, *The Balanced Scorecard: Translating Strategy Into Action*, ed. H.B.S. Press. 1996, Boston.
29. Sweeney, K. *Give Me Great Strategy Execution, Or Give Me Death (at least fire me)*. BYU-Idaho Business Summit 2007 15/2/2007 [cited 2007 2007/08/12]; Available from: <http://www.byui.edu/busandcomm/summitwin07.htm>.
30. Pestana, G., *Spatial Dashboard: Using the Spatio-Temporal Context to Analyse Business Performance*, in *Departamento de Engenharia Informática*. 2007, Universidade Técnica de Lisboa: Lisboa.
31. Wiegmann, D.A., et al. *Human factors aspects of three-dimensional visualization of power system information*. in *Power Engineering Society General Meeting, 2006. IEEE*. 2006.
32. Cooper, L.A. *Mental representation of three-dimensional objects in visual problem solving and recognition*. *PsycARTICLES* 1990 [cited 2007; Available from: <http://content.apa.org/journals/xlm/16/6/1097>.
33. Baker, P. and C. Wickens, *Human Factors in Virtual Environments for the Visual Analysis of Scientific Data*. 1995, unknown.
34. Donath, J. and F. Viègas, *Social Network Visualization: Can We Go Beyond the Graph?*, in *Computer Supported Cooperative Work 2004*. 2004, MIT: Massachusetts.
35. Nielsen, J. and R. Molich, *Improving a human-computer dialogue*. *Commun. ACM*, 1990. **33**(3): p. 338--348.
36. Overmyer, S.P. *Revolutionary vs. Evolutionary Rapid Prototyping: Balancing Software Productivity and HCI Design Concerns*. in *4th International Conference on Human-Computer Interaction*. 1991. Stuttgart, Germany.
37. Sommerville, I., *Software engineering*. International computer science series. 2004, Boston: Pearson/Addison Wesley.
38. Loach, R. *The Tao Framework - Various .NET and Mono Bindings*. 2007 [cited 2007/08/17]; Available from: <http://www.taoframework.com/>.
39. Mendez, H.d.F.P.O.d. *PIGALE - Public Implementation of a Graph Algorithm Library and Editor*. 2007 [cited 2007/08/17]; Available from: <http://pigale.sourceforge.net/>.
40. yWorks. *yFiles.NET - Graph Layout and Visualization Library*. 2007 [cited 2007/08/17]; Available from: http://www.yworks.com/en/products_yfilesdotnet_about.htm.
41. Bossiere, E. and Y. Yashchenko. *Irrlicht .NET CP*. 2007 [cited 2007/08/17]; Available from: <http://irrlichtnetcp.sourceforge.net/>.
42. Gebhardt, N. *Irrlicht.net*. 2007 [cited 2007/08/17]; Available from: <http://irrlicht.sourceforge.net/docu.net/index.html>.
43. McGuiirk, C. *Axiom 3D Engine*. 2007 [cited 2007/08/17]; Available from: <http://axiomengine.sourceforge.net>.
44. Mark, J.J. *OgreDotNet Binding Set and CIL Wrapper dlls for Ogre3D*. 2007 [cited 2007/08/17]; Available from: <http://www.ogre3d.org/wiki/index.php/OgreDotNet>.
45. *MOGRE Wiki*. 2007 [cited 2007/08/17]; Available from: <http://www.ogre3d.org/wiki/index.php/MOGRE>.
46. Hart, J. *3D Engine and Game Development SDK*. 2007 [cited 2007/08/17]; Available from: <http://www.truevision3d.com/>.
47. Maher, K. *The state of OpenGL on Windows Vista*. 2007 [cited 2007/08/16]; Available from: <http://www.beyond3d.com/content/articles/48/2>.
48. Luo, W. *DockManager Control*. 2007 [cited 2007/06/20]; Available from: <http://weifenluo.bravehost.com/DockManager.htm>.
49. Luo, W. *DockPanel Suite*. 2007 [cited 2007/06/20; Available from: <http://sourceforge.net/projects/dockpanelsuite/>.
50. Gamma, E., et al., *Design patterns: elements of reusable object-oriented software*. 1995: Addison-Wesley.
51. Kaplan, R.S. and D.P. Norton, *Using the Balanced Scorecard as a Strategic Management System*. *Harvard Business Review*, 2007. **85**(7/8): p. 150.
52. PLC, S.M. *Airport Technology*. 2007 [cited 2007/09/02]; Available from: <http://www.airport-technology.com/glossary/>.

53. ANA, A.d.P.-. *ANA - Área de Empresa*. 2007 [cited 2007/09/02]; Available from: http://www.ana.pt/portal/page/portal/ANA/ANA_EMPRESA.
54. Winnipeg, A.A.I.-. *Winnipeg Airpor. ts. Authority Inc .2005 Business Plan*. 2007 [cited 2007/09/02]; Available from: www.waa.ca/UserFiles/File/WAA_BP2005-abridged.pdf.
55. Kaplan, R.S. and D.P. Norton, *Alignment using the balanced scorecard to create corporate synergies*. 2006, Boston, Mass.: Harvard Business School.
56. Hill, K. *IrrlichtSharp Engine*. 2007 [cited 2007/08/23]; Available from: http://irrlichtsharp.sourceforge.net/wiki/index.php/Main_Page.
57. Cleotilde, G. *Does animation in user interfaces improve decision making?* in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1996. Vancouver, British Columbia, Canada: ACM Press.

9. Anexos

9.1. Planificação

O ponto de partida para a implementação dos requisitos pretendidos foi a versão desenvolvida pelos colegas envolvidos no projecto no ano passado. Assim, as primeiras tarefas consistiram numa análise funcional da aplicação na perspectiva do utilizador, leitura da documentação existente e análise da arquitectura da aplicação. O planeamento detalhado pode ser consultado na Figura 49.

ID	Task Name	Start	Finish	Duration	2006				2007											
					Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago				
1	Instalação da Plataforma de Desenvolvimento e Análise da Documentação Técnica	01-09-2006	07-09-2006	1w																
2	Desenvolvimentos sobre a aplicação existente: nova versão do algoritmo treemap	08-09-2006	29-09-2006	3,2w																
3	Desenvolvimentos sobre a aplicação existente: melhorias na TOC de Scenario e Semantic Manager	25-09-2006	13-10-2006	3w																
4	Especificação de Requisitos	03-10-2006	09-10-2006	1w																
5	Testes para escolha de framework	10-10-2006	15-11-2006	5,4w																
6	Desenvolvimentos	02-11-2006	20-11-2006	2,6w																
7	Meta: 1ª Versão do Protótipo (20 Nov 2006) Independente do Spatial Dashboard	20-11-2006	24-11-2006	1w																
8	Desenvolvimentos e Testes	24-11-2006	09-03-2007	15,2w																
9	Meta: 2ª Versão do Protótipo (9 Mar 2007) Integrado com Spatial Dashboard	09-03-2007	15-03-2007	1w																
10	Desenvolvimento e Testes	16-03-2007	15-06-2007	13,2w																
11	Meta: Versão Final do Protótipo	15-06-2007	20-06-2007	,8w																

Figura 49 - Planeamento das actividades.

9.2. Práticas

Cientes do impacto que o recurso a boas práticas de desenvolvimento poderá ter na qualidade do protótipo a implementar, utilizámos o protótipo com recurso a um repositório de código com controlo de versões (*Sub Version*), procurámos seguir alguns aspectos da filosofia *extreme programming* como a integração contínua e a refactorização. Utilizámos ainda aplicações de suporte à comunicação e ao trabalho colaborativo como o *Basecamp* e um pequeno e simplificado repositório de bugs que consistia numa *spreadsheet* partilhada e formatada para o efeito, disponibilizada através do Google Docs.

9.2.1. Prototipagem Rápida

Um passo importante de um projecto que envolva o desenho de uma interface deve ser a construção de um protótipo de baixa fidelidade. Overmyer [36] considera a prototipagem rápida uma forma eficaz de lidar com os requisitos contribuindo para um aumento de produtividade que resulta de se ignorarem aspectos como a reutilização, a escalabilidade e a facilidade de manutenção. Em particular foi utilizado um protótipo do tipo *throw-away*, ou seja, descartável, para validação dos requisitos relativos à interface, adequado para consolidar requisitos, reduzindo assim o risco de se investir tempo e esforço em requisitos incertos [37]. Assim, pretende-se realizar uma análise aos requisitos, ou seja derivar os requisitos do sistema (*get the right requirements*). Em sequência dessa validação, prosseguiu-se com a implementação de um protótipo evolutivo, abordagem de prototipagem adequada para sistemas que não se podem especificar à partida, como sistemas de inteligência artificial e interfaces de utilizador, como era o nosso caso. A validação da adequação do sistema teve várias iterações realizadas em conjunto com o Professor Gabriel Pestana, no papel de utilizador. Com a validação pretendeu-se implementar correctamente os requisitos anteriormente definidos (*get the requirements right*).

9.3. Escolha de *Frameworks*

Relativamente à escolha da plataforma de desenvolvimento, tivemos total liberdade, não nos tendo sido imposta nenhuma *framework*. O principal condicionalismo consistiu no facto de que a grande maioria das soluções para gráficos interactivos em tempo real são escritas na linguagem C++, enquanto o *Spatial Dashboard* está escrito totalmente em C#.

Foi analisado um conjunto de *frameworks* que incluiu API gráficas e de *layout* de grafos. As considerações daí decorrentes serão seguidamente apresentadas.

9.3.1. *OpenGL* em ambiente *.NET* (*Framework TAO*)

O *OpenGL* (Open Graphics Library) é uma especificação de uma *Application Programming Interface* (API) multi-plataforma para a concepção de aplicações gráficas 2D e 3D. É amplamente utilizada em CAD, visualização de informação, visualização científica, simulação e em jogos, competindo com a solução da Microsoft, o *DirectX*.

A *framework* TAO [38] permitir-nos-ia aceder às primitivas básicas do *OpenGL*. Não é um motor 3D, assim, todas as funcionalidades de alto nível (gestão da câmara, gestão de eventos do rato, gestão do volume de visualização, etc.) teriam de ser escritas por nós. Numa primeira abordagem considerámos a conversão de C++ para C# de uma biblioteca (AVTGL) que utilizámos no âmbito de um projecto de Animação e Visualização Tridimensional e que serve de suporte a algumas funções de alto nível. A principal vantagem seria o facto de já termos anteriormente trabalhado com essa biblioteca. No

entanto, ao rever a referida biblioteca considerámos que o nível de abstracção não era o ideal e a conversão de linguagens poderia propiciar *bugs* no futuro, além do esforço considerável que implicaria.

9.3.2. Bibliotecas de *Layout* de Grafos (Lithium, PIGALE, yFiles.NET)

A *framework* PIGALE [39] é uma ferramenta de edição de grafos e biblioteca de algoritmos de *layout* escrita em C++ e *OpenGL*, essencialmente dedicada a grafos planares, mas com algum suporte a *layout* tridimensional. Está vocacionada para investigação em teoria de grafos. Visto que está escrita em C++ e *OpenGL*, a sua utilização iria implicar a reescrita das componentes necessárias em C# com o recurso à *framework* TAO, o que representaria um grande esforço, desnecessário com outras soluções possíveis.

A *framework* yFiles.NET [40] é uma biblioteca para o ambiente .NET que disponibiliza estruturas de dados, algoritmos e controlos de visualização para a análise de grafos, diagramas e redes.

Pretendíamos uma solução mais flexível. As bibliotecas de *layout* de grafos que encontrámos, se utilizadas, iriam condicionar determinadamente o *look and feel* do visualizador.

9.3.3. Motores 3D

Um motor 3D, um componente nuclear de aplicações gráficas interactivas em tempo real, acarretaria as principais vantagens:

- Suportar funcionalidades fundamentais para aplicações de gráficos 3D;

- Abstrair detalhes inerentes a bibliotecas de sistema (*OpenGL*, *DirectX*), possibilitando que o visualizador seja suportado em contextos de *hardware* heterogéneos.

Entre as referidas funcionalidades fundamentais a que um motor gráfico dá suporte, destacam-se:

- Back Face Culling* (remoção de faces ocultas): realiza um teste para determinar se a face visível de um polígono está orientada para a câmara e se o teste der verdadeiro, o vector normal ao polígono aponta para fora em relação à câmara e não deve ser desenhado;

- Clipping*: corresponde a não desenhar os objectos gráficos que se encontram no exterior do volume de visualização. É uma operação não trivial, crítica para o desempenho em animações tridimensionais;

- Picking*: selecção de objectos gráficos, tipicamente através de um *mouse-click* num ponto da cena. Faz recurso de volumes envolventes aos objectos denominados *bounding boxes*.

Ao descobrir as potencialidades de um motor 3D de alto nível e o impacto que poderia ter para o desenvolvimento de uma solução estável e robusta, pesquisámos alternativas para a referida biblioteca AVTGL, tendo encontrado os motores:

- Irrlicht*, nas versões compatíveis com .NET: *Irrlicht .NET CP* [41] e *Irrlicht API .NET* [42];

- OGRE*, nas versões compatíveis com .NET: *Axiom* [43], *OgreDotNet* [44] e *MOGRE* [45];

- Truevision* [46].

Os dois primeiros são as principais referências de motores 3D *open source*. Será necessário o recurso a *wrappers* (acima indicados), que permitem aceder em ambiente .NET às dlls dos motores que foram escritas e compiladas em C++.

Comparando funcionalidades verificou-se que o motor Truevision era menos poderoso e versátil que os outros dois. Excluída esta hipótese e após alguma pesquisa concluímos que o Motor Ogre apresentava duas desvantagens principais em relação ao Irrlicht:

as soluções para .NET encontravam-se num estado de desenvolvimento menos maduro que as do Irrlicht;

maior curva de aprendizagem.

Optámos portanto pelo motor 3D *Irrlicht* que nos agradou pois é *open source*, tem uma grande comunidade activa, é descrito como sendo bastante compatível tanto com hardware recente como antigo e por ter uma curva de aprendizagem suave, sem comprometer as funcionalidades.

A principal e importante desvantagem consistia no facto de ter sido escrito em C++. No entanto, como já referimos existem *wrappers*. Um deles, seria à partida o preferível visto que é o *wrapper* oficial do *Irrlicht* para as liguagens .NET: o Irrlicht API .NET. No entanto encontrámos-lhe uma limitação grave, que era a impossibilidade de estender objectos gráficos, os *Scene Nodes* do *Irrlicht*.

A opção que tomámos pelo projecto Irrlicht.NET CP, em detrimento do referido *wrapper* oficial parece ter sido adequada, visto que este deixou recentemente de ser suportado, podendo ler-se na página oficial do *Irrlicht*: “*Support for Irrlicht.NET (...) will be dropped with the next release. Instead, we encourage to use Irrlicht.NET CP*”.

Framework	Curva de Aprendizagem	Comunidade, Suporte e Documentação	Operações Gráficas	Licença	Tipo de Solução
Netron Lithium	*****	**	*	Open Source	Biblioteca .NET para layout de grafos
yFiles.NET	***	***	***	Comercial	Biblioteca .NET para layout de grafos
PIGALE	***	****	**	Open Source	Biblioteca C++ para layout de grafos
TAO Framework	***	**	***	Open Source	<i>Binding</i> de OpenGL para .NET
TrueVision	***	**	****	Comercial	Motor 3D
Irrlicht .NET CP	****	****	*****	Open Source	<i>Wrapper</i> de Motor 3D C++ para C#
Irrlicht API .NET	****	**	*****	Open Source	<i>Wrapper</i> de Motor 3D C++ para C#
OGRE (wrappers)	**	***	*****	Open Source	<i>Wrapper</i> de Motor 3D C++ para C#

9.4. Constrangimentos

O suporte de texto e fontes do *Irrlicht* é desenhado para a utilização em *Graphic User Interfaces*, em 2D. Apesar de permitir a utilização simultânea de gráficos 2D e 3D no mesmo grafo de cena o texto 3D suportado nativamente não pôde ser utilizado pois o *Irrlicht* não calcula a oclusão para objectos 2D.

Outro constrangimento relevante que confrontámos deveu-se ao facto de se terem utilizado *frameworks open-source* cuja robustez para a implementação de funcionalidades em ambiente .NET puro não nos mereceu qualquer reparo, ficando, no entanto, um pouco aquém das expectativas no capítulo da versatilidade. De facto surgiram alguns problemas quando as funcionalidades a implementar implicavam a o funcionamento integrado das *frameworks*, nomeadamente do motor 3d *Irrlicht .NET CP* e do *docking manager Weifen Luo*.

O *Irrlicht .NET CP* é um *wrapper* o que levantou alguns conflitos com o modelo de eventos .NET utilizado no projecto *Spatial Dashboard*. Em particular, um constrangimento subtil, mas de grande impacto no desempenho e estabilidade da aplicação, com que nos deparámos foi relativo às mudanças de contexto para o contexto do Visualizador de Dendrogramas. Como referido anteriormente o *Spatial Dashboard* está organizado por contextos que são geridos com recurso à *framework Weifen Luo*. O módulo implementado no contexto desta tese corresponde a um desses contextos. Como também foi referido anteriormente, fez-se uso do motor 3D *Irrlicht*, numa versão *wrapper*, ou seja, não era um *assembly* de C# .NET, pelo que o modelo de eventos .NET funcionava de forma limitada enquanto o *wrapper* se encontrava em execução. Mais concretamente, ao efectuar a mudança de contexto, assim que o método de inicialização do motor *Irrlicht* era invocado, não iria correr mais instrução nenhuma relativa à mudança de contexto. Como solução temporária, colocou-se na interface um botão para inicializar o *Irrlicht*, que o utilizador deveria pressionar após a mudança de contexto. Após uma detalhada análise da *framework Weifen Luo*, o problema acabou por ser contornado com uma refactorização no *core* da *Framework*, que permitiu que a última instrução a ser executada na mudança de contexto fosse precisamente a de inicialização do motor 3D.