

TIPOS DE DADOS VISUALIZÁVEIS

Aplicação no domínio da Engenharia Civil

João Duarte Cunha, LNEC
Ana Paula Cláudio, DICC/FCL
Maria Beatriz Carmo, DICC/FCL

RESUMO

A Visualização de Dados Científicos ou Visualização Científica envolve basicamente dois problemas distintos: converter dados correspondentes a conceitos abstractos em objectos visíveis (mapeamento) e gerar as imagens correspondentes (rendering).

O segundo problema pertence ao domínio da síntese de imagem mas reduz-se em muitos casos à geração de gráficos do tipo "business graphics". O estado actual dos conhecimentos relativos à síntese de imagem pode considerar-se satisfatório para as necessidades da visualização científica.

A fase de "mapping" pertence tradicionalmente ao foro das diversas disciplinas científicas, os "clientes" da visualização, tendo surgido ao longo dos anos soluções por vezes bastante imaginativas para conseguir gerar imagens sugestivas e informativas sobre os fenómenos em estudo.

Com o amadurecimento da disciplina Visualização Científica começaram a aparecer sistemas de visualização que pretendem ser genéricos e fáceis de configurar e usar pelos utilizadores finais. Por outro lado, reconhece-se que faz falta uma sistematização dos dados a visualizar e dos tipos de representação que lhes podem corresponder.

Nesta comunicação desenvolve-se o conceito de Tipo de Dados Visualizável e relatam-se algumas experiências de aplicação do apE, um dos sistemas de visualização da nova geração, a problemas do âmbito da engenharia civil.

1. Introdução

A Visualização de Dados Científicos ou Visualização Científica, precisa de se tornar acessível aos cientistas em geral, nas tarefas diárias de interpretação de resultados, sem necessidade de recorrerem a especialistas informáticos ou de se envolverem eles próprios na escrita de programas.

Nos últimos dois anos têm surgido no mercado comercial e académico vários sistemas que de alguma forma respondem a este requisito, mas estamos ainda longe do objectivo final.

É necessário construir um modelo de referência amplamente aceite para enquadrar os desenvolvimentos subsequentes e, para isso, há que encontrar formas de sistematizar os dados - ponto de partida da Visualização - e as suas representações possíveis - ponto de chegada.

O texto que se segue discute precisamente esta sistematização dos dados e, em particular, o conceito de Tipo de Dados Visualizável. Numa segunda parte, descreve-se uma curta experiência de utilização do sistema apE (Dyer, 1990) no âmbito da engenharia civil.

O apE foi escolhido por se tratar de um sistema facilmente acessível, aberto, bastante poderoso e usado por uma comunidade activa e comunicante.

2. Tipos de dados visualizáveis

Os dados são o objecto central da visualização: calculados ou adquiridos, eles contêm informação sobre os fenómenos em estudo. Quando são volumosos - e são-no quase sempre - a sua interpretação é difícil e a visualização, mesmo quando não é a solução definitiva, dá uma contribuição importante.

Na Workshop sobre "Data Structures and Access Software for Scientific Visualization" realizada por ocasião do SIGGRAPH'90 (Treinish, 1991) a situação actual no que toca à manipulação dos dados a visualizar foi considerada caótica. A maior parte dos sistemas de visualização ignora o problema, suportando apenas dados contidos em ficheiros com estruturas específicas, normalmente bastante simples, ou deixando a questão para o utilizador resolver "como exercício". Este é um dos obstáculos, normalmente o primeiro do ponto de vista cronológico, com que se debate o utilizador que quer tirar partido das capacidades de visualização sofisticadas disponíveis.

Este e outros obstáculos são responsáveis pelo facto de as formas mais sofisticadas de visualização serem usadas essencialmente para comunicação com o público e com as instituições que financiam a investigação, continuando a maioria dos cientistas a utilizar no dia a dia técnicas de visualização relativamente primitivas, desenvolvidas há mais de uma década, quando o volume e o ritmo de geração de dados eram muito inferiores aos actuais. Esta afirmação é feita frequentemente em workshops e reuniões de grupos de trabalho (p.ex. Treinish, 1991)

O caos na organização dos dados que interessa visualizar existe mesmo dentro da mesma disciplina científica e da mesma instituição, como podemos constatar no LNEC, gerando um enorme esforço de adaptação e conversão de programas e/ou de dados.

Enquanto na chamada "informática de gestão" o uso de sistemas de gestão de bases de dados se tornou regra, na área científica predominam os formatos ad-hoc, as soluções proprietárias e, nalguns casos, soluções institucionais. Este último caso, de todos o mais

favorável, é típico de grandes instituições como o CERN ou a NASA, e acaba por ter algum impacto para além das instituições de origem através da sua rede natural de influência.

Esta falta de estruturação dos dados impede a sua visualização de forma flexível, obrigando no mínimo a escrever “mais uma rotina” sempre que se quer visualizar algo que não estava inicialmente previsto ou alterar as formas de visualização.

O subgrupo que na Workshop referida se ocupou especialmente de Estruturas de Dados e Modelos considerou de extrema importância a criação de uma taxonomia de dados a visualizar, uma espécie de catálogo com contribuições do maior número possível de disciplinas. Cada item do catálogo deveria incluir, na opinião do grupo, a seguinte informação:

- Descrição do tipo de dados, sistema de coordenadas, propriedades da malha de referência e outras informações relevantes.
- Representação gráfica da malha.
- Definição dos termos específicos usados para descrever os dados, do “calão” próprio da disciplina em causa.
- Dimensão média e máxima (expectável) dos dados.
- Descrição das variáveis e respectivas estruturas de dados.
- Métodos de acesso, incluindo métodos de interpretação.
- Representações gráficas relevantes.

Em (Cunha, 1991) introduziu-se o conceito de “tipo de dados visualizável” e sugere-se que um tal conceito possa servir de base à construção de sistemas de visualização genéricos. Um tipo de dados é definido pelo conjunto de valores possíveis e pelas operações definidas sobre eles. Os tipos de dados viram a sua importância fortemente acrescida com a expansão do conceito de orientação por objectos que veio permitir, na prática, a definição e implementação de tipos de dados de complexidade arbitrária.

O que caracteriza um tipo de dados VISUALIZÁVEL é a existência de uma ou mais operações de mapificação, isto é, operações que geram objectos gráficos, susceptíveis de tratamento pelos métodos habituais da computação gráfica.

O conceito pode ser ilustrado recorrendo a um exemplo da área dos chamados “presentation graphics” que se podem considerar como o extremo inferior, em termos de requisitos tecnológicos, da visualização científica (Jern, 1989).

Encarando a visualização de uma tabela de uma entrada de acordo com o modelo de visualização corrente (Upton, 1989) teremos que identificar as fases de mapificação e de “rendering”. Concentrando-nos, para concretizar, em histogramas e gráficos tipo queijo, encontramos duas transformações de mapificação distintas: uma transforma a tabela num conjunto de paralelepípedos de altura proporcional aos valores da variável dependente e colocados em posições determinadas pelo valor da variável independente; a outra transforma a mesma tabela num conjunto de sectores cilíndricos justapostos de amplitude proporcional aos valores da variável dependente.

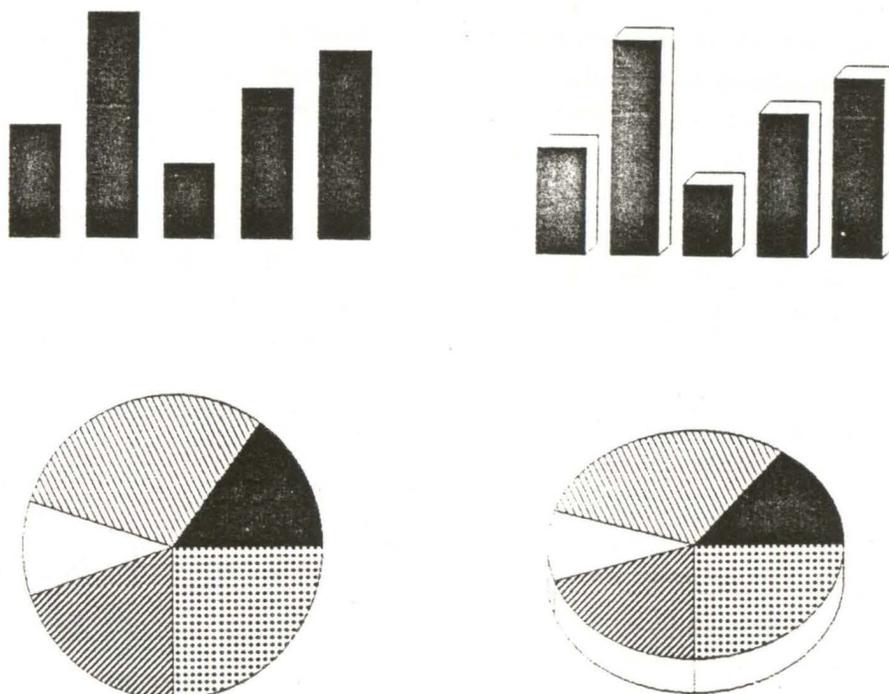


Fig. 1 - O tipo de mapeção define o tipo de gráfico: histograma ou queijo.
A parametrização do "rendering" define o aspecto do gráfico.

A decisão de apresentar gráficos normais ou 3D é, em qualquer das situações, tomada na fase de "rendering", dependendo em última análise do tipo de projecção efectuada. A fig. 1 ilustra esta situação.

Esta forma de encarar o processo de geração dum histograma, ou dum gráfico tipo queijo, ou de "business graphics" em geral, tem um interesse puramente conceptual mas, por isso mesmo, é útil para o fim em vista. Para definir o tipo de dados visualizável TABELA, para além de especificar os valores possíveis para as variáveis independente e dependente (p.ex., o conjunto dos inteiros positivos para a primeira e o conjunto dos reais não negativos para a segunda) e quaisquer operações de interesse (p.ex. normalização), teríamos que especificar as duas operações de mapeção que referimos.

3. O ambiente de visualização apE

O apE é um pacote de visualização desenvolvido a partir de 1987 por uma equipa do Ohio Supercomputer Center. A sua distribuição nos meios académicos é feita por um preço simbólico e inclui o código fonte. Recentemente foi criada uma empresa, a Tara Visual Corporation, com o objectivo de suportar o apE como um produto comercial. A versão actualmente distribuída é a 2.1.

O nome apE é um acrónimo da expressão inglesa *animation production Environment*, o que corresponde à ideia inicial dos seus autores (Dyer, 1990) mas não cobre

os desenvolvimentos posteriores. O apE pode ser descrito como um "toolkit" de visualização destinado sobretudo a cientistas não informáticos - e conseqüentemente sem conhecimentos nem interesse na área da Computação Gráfica - interessados em verem os dados das suas simulações ou medições representados numa imagem ou numa sequência animada de imagens. O utilizador do apE não necessita programar para produzir as imagens que pretende. A interface com o apE pode ser feita através de um programa próprio, o **Wrench** (work bench). Este programa tem uma janela própria, dividida em várias áreas com fins diversos, e um menu. Uma das zonas da janela destina-se à construção do **pipeline**. Por seu turno, uma das opções do menu permite visualizar os ícones que representam os diferentes módulos. O utilizador, por manipulação directa, começa por deslocar os ícones dos módulos que pretende utilizar para a zona da janela destinada ao "pipeline", ligando-os depois entre si pela ordem conveniente. Depois de bem definidos os parâmetros de entrada e saída dos vários módulos e os parâmetros de controlo dos que deles necessitem, o "pipeline" está pronto a executar, ou seja, a produzir a imagem pretendida.

As peças básicas são os módulos que funcionam como caixas pretas das quais só se conhecem as entradas, as saídas e a funcionalidade. Exemplos, são o módulo de normalização **Norm** e o de mudança de grelha, **Rezone**. Alguns destes módulos dispõem de janelas próprias. Os módulos ligados entre si vão, como já vimos, constituir um "pipeline". Quando este é posto a correr, cada módulo é executado por um processo distinto e os dados vão-se movendo ao longo do "pipeline". Trata-se pois de um sistema de "dataflow" do tipo "data-driven". Cada um daqueles processos pode ser executado numa máquina distinta pelo que o apE se adapta bem a ambientes distribuídos.

No **Wrench** são apresentados os seguintes grupos de módulos: *Interface, Input & output, Utility, Editors, Data, Mapping, Geometry, Rendering, Image* e *Accad-contributed*. A título de exemplo do grupo *Interface* fazem parte os módulos *Panorama* e *Easel* em cujas janelas são visualizadas as imagens produzidas pelos "pipelines"; do grupo *Editors* faz parte o *Colormap*, um editor de tabelas de cores; do grupo *Image* faz parte um módulo de "zooming"; do grupo *Mapping* faz parte um módulo que produz gráficos de barras.

Existe uma interface de programação que permite a invocação directa do apE a partir de programas FORTRAN.

O apE utiliza um formato próprio para os dados que entram, circulam e saem do "pipeline": o formato **Flux**. Este formato, pelos vocábulos (ex: begin, end) e sintaxe que utiliza, recorda a linguagem Pascal. Os valores dos dados propriamente ditos podem ser especificados em ascii, xdr ou código binário nativo. O formato Flux é utilizado não só para ficheiros de dados mas para todos os ficheiros usados no apE. Por exemplo, quando no **Wrench** se guarda um "pipeline" ele é armazenado num ficheiro com este formato.

O apE define ainda um conjunto de **Data Groups**, cada um com a sua variante de formato Flux. Exemplos são *Variable, Colormap, Grid, Image, Object*. É possível associar vários tipos de **Data Groups**. Por exemplo, um **Variable Group** pode ter um **Grid Group** como sua referência espacial.

O apE está escrito em C e tem como suporte o sistema operativo Unix e um sistema de janelas próprio denominado **Face**. O **Face** foi implementado sobre vários sistemas de janelas, existindo versões de apE para X11R4, NeXSTep, SGI's GL e Sunview. Não foi adoptada nenhuma biblioteca gráfica por se considerar que nenhuma das soluções

normalizadas existentes na altura era suficiente para as necessidades da visualização.

Com estas opções os autores do apE pretenderam criar um sistema o mais possível independente de outros produtos de "software".

É um sistema extensível pois a adição de novos módulos está facilitada pela existência de ferramentas próprias que auxiliam esta tarefa. Um módulo existente pode ser alterado desde que se mantenham as suas entradas e saídas.

Existe um conjunto de tutoriais de demonstração que acompanham o sistema e facilitam grandemente a aprendizagem.

Para utilizar o sistema com os seus dados o utilizador começa por construir um "pipeline" com os módulos convenientes para que a imagem final seja da forma pretendida. Em seguida, tem de escolher o Data Group a que se adaptam os seus dados originais e construir o ficheiro de dados de entrada em formato Flux. Para tornar automática a filtragem dos dados pode construir-se um programa independente ou programar e adicionar um novo módulo ao conjunto de módulos do apE.

Para fazer a filtragem de imagens entre o formato Flux e outros formatos existem os módulos *Img-import* e *Img-export*. O primeiro executa a conversão de imagens de alguns formatos normalizados (p.ex. exemplo, a60, barney, hdf,...) para o formato Flux, o segundo executa a operação inversa. Nalguns casos só é possível a conversão num dos sentidos. É o caso do formato PostScript, que só pode ser utilizado pelo módulo *Img-export*.

Um mesmo "pipeline" que produza um determinado tipo de imagem pode ser utilizado facilmente para diferentes conjuntos de dados de entrada. Para adaptar um "pipeline" a um novo conjunto de dados basta alterar uns quantos identificadores (por exemplo, mudar o nome da variável de entrada) e ajustar os parâmetros de controlo usados por alguns módulos (por exemplo, o módulo de normalização usa como parâmetros de controlo os valores máximo e mínimo dos dados).

Uma das limitações sentidas ao produzir imagens no apE é a impossibilidade de lhes adicionar eixos ou legendas o que reduz a quantidade de informação que transmitem.

A informação contida nos manuais está dispersa e é em muitos casos insuficiente. Além disso, pelo facto de o apE ser ainda um produto em desenvolvimento, não é raro o utilizador encontrar no item *cautions* mensagens como a seguinte: "Although non-uniform grids are defined within apE, there are no apE módulos that use grids that are not uniform" (apE, 1990).

4. Utilização do apE no âmbito da engenharia civil

No âmbito da engenharia civil existem problemas de visualização associados à utilização de modelos matemáticos, à utilização de modelos físicos e à observação de obras. No caso dos modelos matemáticos a visualização é importante quer na fase de preparação dos dados quer na de análise de resultados, embora esta segunda fase seja aquela em que o volume e complexidade dos dados a visualizar é maior.

Duma forma geral predomina nesta área o recurso a métodos clássicos de visualização - gráficos de linhas em terminal tipo Tektronix 4010 ou em "plotter" - produzidos por programas desenvolvidos sobre GKS ou outras bibliotecas gráficas com funcionalidade semelhante. Estes programas exploram dados em formatos específicos, sendo necessárias adaptações sempre que se pretende usá-los em contexto diferente, ainda que semelhante.

A utilização de programas específicos desenvolvidos sobre plataformas gráficas de baixo nível é incompatível com a exploração interactiva de grandes volumes de dados, onde a flexibilidade é uma exigência fundamental, e também com a utilização de técnicas sofisticadas de visualização. Este último problema, quando se trata da preparação de apresentações ou material para publicação, pode ser resolvido recorrendo a programas comerciais como o Data Visualizer da Wavefront Technologies, o UNIRAS, o DISSPLA ou outros. O processo de adaptação dos dados pode ser mais ou menos trabalhoso e requer quase sempre a intervenção de informáticos.

A substituição do GKS (ou equivalente) por PHIGS ou PHIGS+ não resolve nenhum dos problemas referidos. A utilização de bibliotecas especializadas como a Exponent Graphics da IMSL ou a da UNIRAS simplifica a programação, permite acesso a técnicas mais sofisticadas com um mínimo de esforço e garante, como aliás já o fazia a utilização do GKS ou PHIGS, a independência do "hardware". Contudo, não aumenta significativamente a flexibilidade na exploração dos dados nem a necessidade de adaptações.

O sistema apE usa uma abordagem diferente como vimos em 3., o mesmo acontecendo com o AVS (Upson, 1989). Se é certo que a possibilidade real de utilização autónoma pelos utilizadores finais, um dos objectivos visados com o desenvolvimento destes sistemas, é bastante questionável, não há dúvida que a sua arquitectura permite uma grande flexibilidade de utilização desde que estejam disponíveis os módulos necessários, incluindo os de importação de dados.

Para que tal seja possível, com esta arquitectura ou com qualquer outra, é necessário uma sistematização dos dados e das suas representações possíveis, que poderá ser feita seguindo, por exemplo, as linhas expostas em 2.

Nas experiências feitas com o apE no pouco tempo disponível desde que ficou operacional no LNEC procurou-se resolver dois problemas típicos de áreas distintas, numa óptica de avaliação de potencialidades e facilidade/dificuldade de utilização.

O primeiro problema diz respeito a dados originários do Núcleo de Águas Subterrâneas. Trata-se de valores dum conjunto de grandezas medidas mensalmente durante vários anos: quantidade de precipitação, humidade no solo, etc. Este tipo de dados é facilmente representável por métodos convencionais, podendo usar-se com toda a facilidade programas vulgares de representação gráfica.

Representações mais sofisticadas, nomeadamente envolvendo a representação combinada de várias grandezas para estudo das suas interrelações, podem ser obtidas com programas comerciais (UNIRAS, DISSPLA, etc).

A utilização do apE implica a conversão dos dados para o formato Flux. A obtenção de representações habituais como a da fig. 2¹ revelou-se simples, embora bastante menos do que usando programas comerciais vulgares. A falta de legendas e escalas, para o que não existem actualmente módulos adequados, é uma deficiência grave, sobretudo pensando em termos de resultados finais.

A exploração de outras possibilidades permitiu com relativa facilidade obter imagens como a da fig. 3, em que se combinam duas grandezas numa única representação. É possível gerar imagens correspondentes, por exemplo, a diferentes pontos de vista e criar com elas uma sequência animada. A flexibilidade na combinação de módulos e o controlo interactivo dos parâmetros respectivos permite testar com muita facilidade

¹A qualidade desta imagem e das seguintes é muito inferior à obtida no écran a cores. A existência no apE de um "drive" Postscript suportando cor virá resolver este problema.

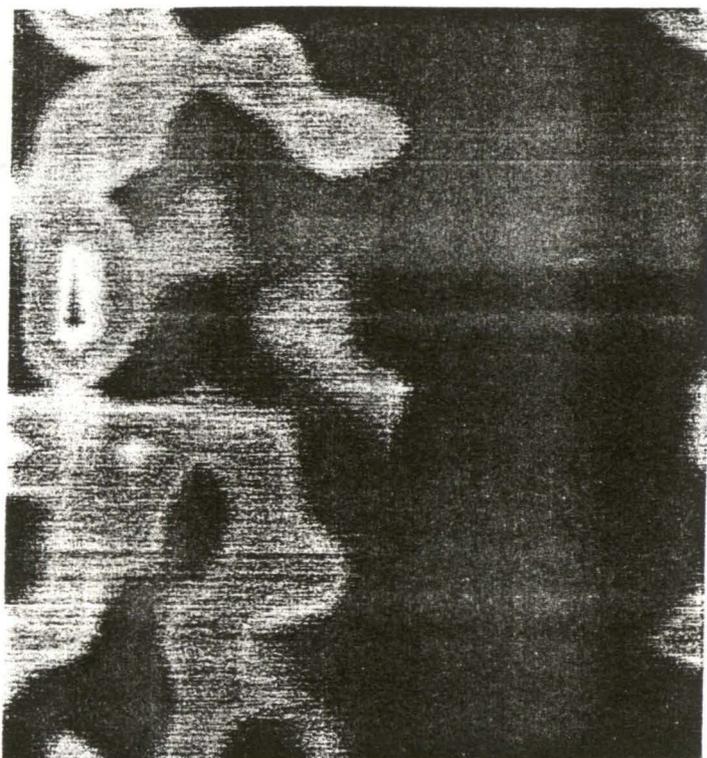


Fig. 2 - Quantidade de precipitação ao longo dos meses (abscissa), durante vários anos (ordenada).

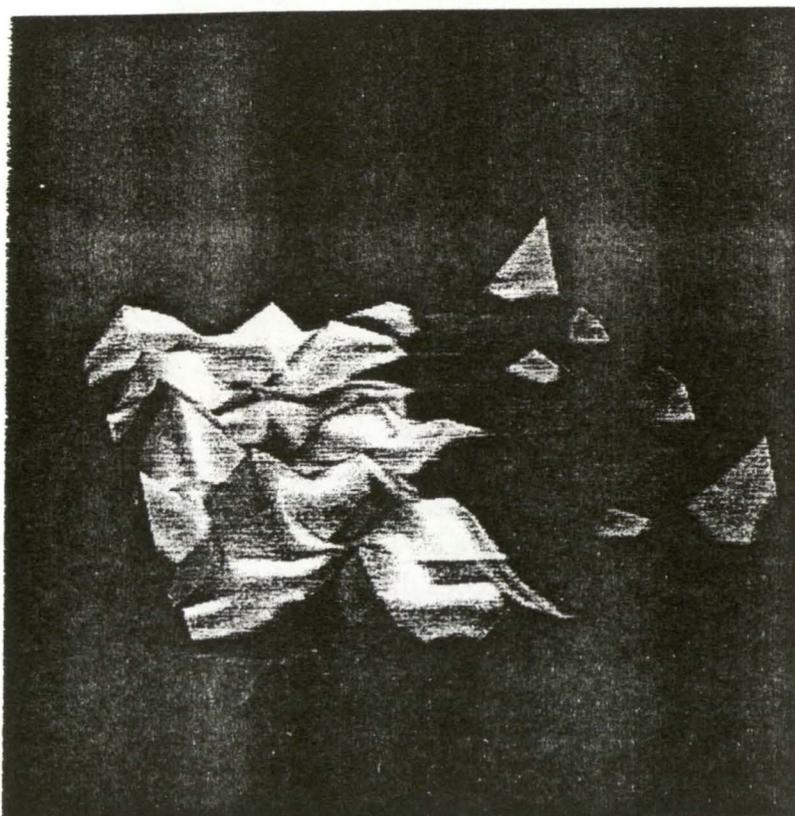


Fig. 3 - A precipitação foi associada ao relevo e a humidade no solo. à cor.

diversas alternativas.

O segundo problema utilizado como teste é bastante mais interessante. Trata-se da refração de ondas de engenharia costeira calculadas de acordo com um modelo desenvolvido por A. Covas (Covas, 1984). Os resultados do modelo são normalmente apresentados sob a forma de diagramas de refração (fig. 4) contendo ortogonais (linhas perpendiculares à direcção de propagação das ondas) e cristas de onda.

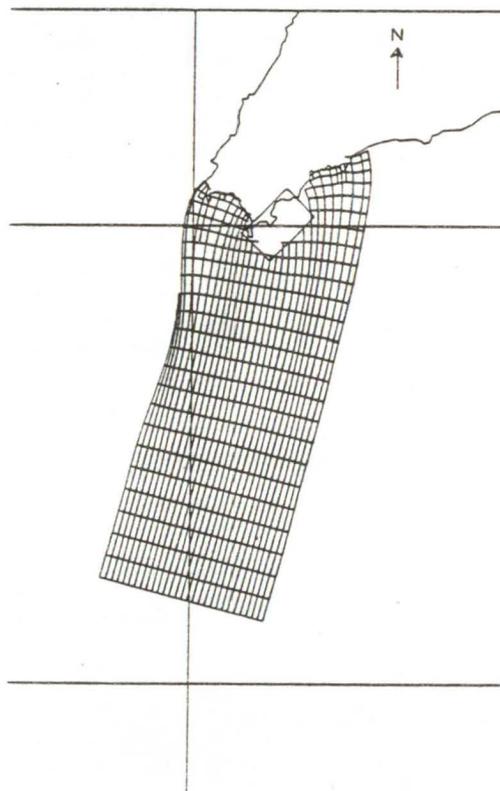


Fig. 4 - Diagrama de refração de ondas costeiras.

Utilizando resultados intermédios do modelo procurou-se gerar uma sequência animada mostrando a progressão das ondas. Trata-se de uma aplicação cujo interesse para a compreensão do fenómeno pelos especialistas é reduzida, sobretudo tendo em conta que o seu custo computacional será sempre elevado. Contudo, este tipo de visualizações tem frequentemente interesse quer em acções dirigidas ao grande público, quer em apresentações para entidades financiadoras ou potenciais clientes, constituindo aquilo a que poderemos chamar a “visualização espectáculo”.

Na figura 5 apresentam-se algumas imagens de uma sequência que pode ser animada no terminal. A batimetria (relevo do fundo) está representada através da cor e as ondas pelo relevo.

Neste caso a utilização do apE revelou-se já mais interessante dado que se trata de um tipo de representação que não está ao alcance dos programas vulgares.

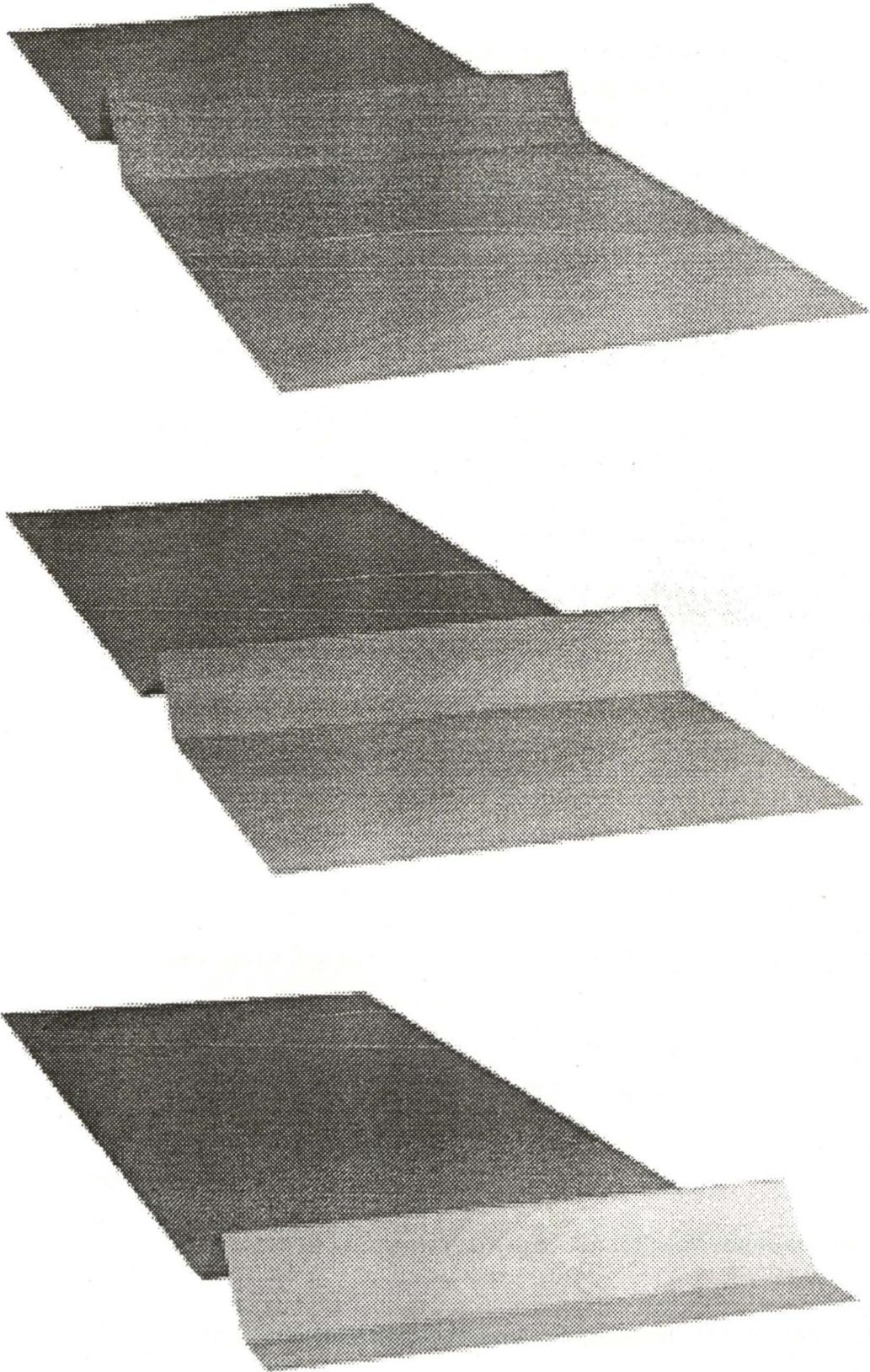


Fig. 5 - Imagens de uma sequência animada mostrando a progressão da onda de acordo com o modelo.

5. Conclusões

A generalidade dos cientistas só poderá beneficiar das capacidades oferecidas à visualização pela computação gráfica ("hardware" e "software") quando se dispuser de conceitos suficientemente gerais e amplamente aceites que permitam o desenvolvimento de sistemas realmente fáceis de usar, flexíveis e poderosos.

A utilização de sistemas abertos, baseados no paradigma do fluxo de dados e dotados de interfaces de programação visual pode ser muito útil na exploração de novas formas de visualização e de maneiras diferentes de olhar para as formas conhecidas.

6. Referências

apE Version 2.0 Reference Manual

The Ohio Supercomputer Graphics Project

Ohio Supercomputer Center

Covas, J. M. A.

"Modelos Matemáticos de Refracção das Ondas de Engenharia Costeira"

Tese de Doutoramento

LNEC, Lisboa, 1984

Cunha, J. D.

"Visualizable Data Types - General purpose mapping in the visualization pipeline"

Proceedings of the Second Eurographics Workshop on Visualization in Scientific Computing

Delft, The Netherlands, April 1991

Dyer, D. Scott

"A Dataflow Toolkit for Visualization"

IEEE Computer Graphics and Applications, Vol. 10, n.4 (July 1990)

Jern, Mikael

"Visualization of Scientific Data"

Eurographics Technical Report Series, 1990

Treinish, L.

"SIGGRAPH'90 Workshop Report - Data Structures and Access Software for Scientific Visualization"

Computer Graphics, Vol. 25, n.2 (April 1991)

Upson, C. et al.

"The Application Visualization System: A Computational Environment for Scientific Visualization"

IEEE Computer Graphics and Applications, Vol. 9, n.4 (July 1989)