

Visualização tri-dimensional interativa de informação demográfica mundial

Eduardo Duarte Pedro Bordonhos
DETI, Universidade de Aveiro
{emod,bordonhos}@ua.pt

Paulo Dias Beatriz Sousa Santos
DETI, IEETA, Universidade de Aveiro
{paulo.dias,bss}@ua.pt

Resumo

Este artigo propõe o Living Globe, uma aplicação de visualização de informação tri-dimensional que apresenta dados demográficos mundiais. A aplicação permite a visualização simultânea de vários dados demográficos, tais como a população, a densidade e o crescimento populacional, a esperança de vida, o saldo migratório, a taxa bruta de natalidade/mortalidade e o rácio entre estas duas taxas. O mapeamento dos dados para as estruturas visuais na aplicação é configurável pelo utilizador, introduzindo um aspeto de personalização que incentiva a obtenção somente visual e mais imediata de informação. Adicionalmente, os dados representados são normalizados conforme uma escala configurável, permitindo filtrar amostras muito dispare (como por exemplo na China e na Índia) e aumentando assim a visibilidade dos restantes dados. Foram realizados testes preliminares que indicam que a filtragem dos dados e a personalização no mapeamento incentivam uma leitura mais rápida de informação relacional e que a pesquisa e seleção de países tornam a navegação mais intuitiva comparado com as visualizações estado-de-arte.

Palavras-Chave

living globe, visualização de informação, tri-dimensional, demografia, população, filtragem, indicadores

1. INTRODUÇÃO

A visualização de dados demográficos é um método popular em estudos científicos para inferir e apresentar o crescimento populacional. Muitos destes estudos apresentam esta informação recolhida ao longo de vários anos em mapas bi-dimensionais e tri-dimensionais, para demonstrar os diferentes níveis de crescimento que ocorrem em várias localizações no globo terrestre.

Este artigo propõe o Living Globe, uma aplicação de visualização com o objetivo de incentivar a investigação e observação de dados demográficos de forma intuitiva e permitir aos utilizadores a comparação de dados de vários países em simultâneo. Esta aplicação implementa funcionalidades não-existentes nas visualizações estado-de-arte e que facilitam a navegação pelo globo, tais como o controlo e ampliação da câmara tri-dimensional, a seleção e a procura textual de países através de um sistema de pesquisa com preenchimento automático.

Nesta visualização estão representadas a população, a densidade e o crescimento populacional, a esperança de vida, o saldo migratório, a taxa bruta de natalidade e a taxa bruta de mortalidade. É possível filtrar esta informação conforme um valor máximo e um valor mínimo, podendo por isso retirar países com valores extremamente grandes ou pequenos, quando comparado com a moda estatística, e reconfigurar a visualização para melhorar a comparação

de dados entre diferentes países.

De modo a possibilitar a representação gráfica de dados diferentes em simultâneo, para inferir possíveis relações entre diferentes indicadores, o protótipo desenvolvido permite o mapeamento de dados sobre três componentes: altura de paralelepípedos/barras verticais em posições geográficas, cor destas barras e cor dos países. A natureza tridimensional da visualização permite o mapeamento de dois tipos de dados sobre as barras verticais, quanto à sua forma e quanto à sua cor. As cores das barras e dos países representam uma gama de valores normalizados entre 0 e 1, sendo a cor obtida através da seleção numa escala de cores.

Este artigo está organizado da seguinte forma: a secção 2 descreve as visualizações existentes no domínio da representação tri-dimensional de dados demográficos; a secção 3 descreve a visão por trás da aplicação proposta e aponta as funcionalidades ausentes de acordo com essa visão; a secção 4 aborda a implementação técnica da visualização, incluindo a fase inicial de recolha e tratamento de dados; a secção 5 descreve os testes realizados à aplicação para determinar a sua mais-valia quanto às visualizações estado-de-arte; e por fim, a secção 6 apresenta conclusões obtidas a partir dos testes da secção 5 e indica aspetos a desenvolver para trabalho futuro.

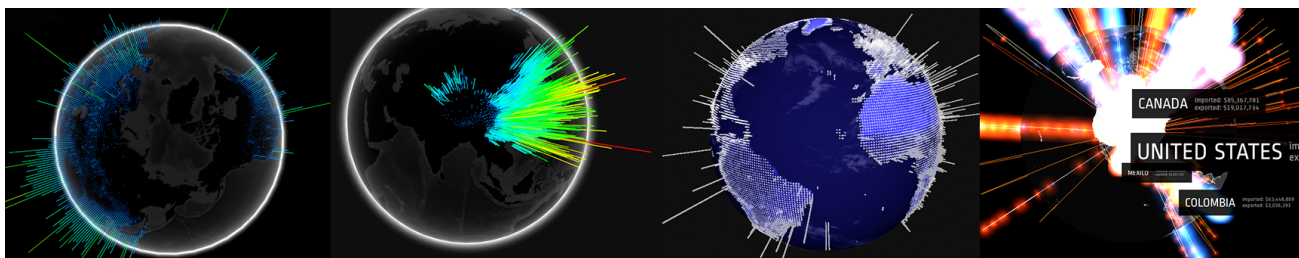


Figura 1. Visualizações estado de arte. Da esquerda para a direita: 'WebGL Globe - World Population', 'WebGL Globe - China and US Population', 'World Population Density - 2010' e 'Small Arms and Ammunition'.

2. TRABALHO RELACIONADO

Nesta secção são apresentadas outras soluções de visualização de informação demográfica. Todas as visualizações do domínio, distribuídas de forma gratuita e *open-source*, são implementadas como aplicações web e utilizam a tecnologia WebGL ou outras APIs de representação tri-dimensional em Javascript, como *three.js*¹. O uso destas tecnologias permite: i) a criação e processamento gráfico de objetos tri-dimensionais, juntamente com as suas texturas, *canvas* e *shaders*, e manipuláveis em tempo-real; e ii) torna a solução compatível com um número elevado de dispositivos, necessitando apenas de conexão de rede e de um *browser* que suporte estas tecnologias, tal como os *browsers* mais populares no mercado. Todas as visualizações demográficas estudadas distribuem um tipo de dados, população total ou densidade populacional, por localização latitude-longitude ao longo de um mapa mundial, sob a forma de barras verticais posicionadas nessas localizações. Quanto maior a população ou densidade naquela localização, maior será essa barra.

Uma destas visualizações é o WebGL Globe², um globo implementado com a API nativa do WebGL e realizado no âmbito do programa "Chrome Experiments". O WebGL Globe é uma plataforma onde qualquer desenvolvedor pode fornecer um conjunto de dados com posições geográficas conforme uma API estabelecida neste e obter uma visualização num globo tri-dimensional, sendo assim totalmente abstraído da implementação deste globo e de elementos gráficos adicionais. Dos projetos WebGL Globe disponibilizados no website oficial, selecionamos dois projetos que apresentam dados demográficos: "World Population"[Arts 11] e "China and US Population"[Zhifang 13].

"World Population", que utiliza dados disponibilizados pelo Centro de Dados e Aplicações Socio-económicas (SE-DAC), é muito limitado em termos de usabilidade e estudo de informação, visto que a única funcionalidade disponível é o movimento do globo. A única componente representativa de dados é a altura das barras verticais, pelo que esta visualização não tira partido das suas três dimensões.

"China and US Population" utiliza um conjunto de dados mais limitado que o anterior, apresentando dados mais específicos para a China e os Estados Unidos da América. No que toca à interação e à usabilidade, estas são idênticas ao trabalho anterior, sofrendo por isso das mesmas falhas e tendo ausente o aproveitamento da oportunidade tri-dimensional.

Uma alternativa ao WebGL Globe é o "World Population Density - 2010"[Dirksen 12], um globo implementado em *three.js*. Identicamente ao "World Population", os dados utilizados são recolhidos do SEDA, mas em contraste com este, globo não é manipulável, rodando em torno do eixo de *yy* automaticamente. Deste modo, a usabilidade é menor pois um utilizador é forçado a esperar que o globo rode para observar os dados de um dado país e tem um tempo limitado para realizar esta observação. Mais uma vez, esta visualização não introduz benefícios na utilização de três dimensões.

Apesar de não fazer parte do domínio de visualização de dados demográficos, foi estudada a visualização "Small Arms and Ammunition - Imports & Exports"[Arts 14]. Os dados desta correspondem ao comércio de armas e munições de pequeno-porte entre os anos 1992 e 2010. Em contraste com as visualizações anteriores, este implementa a seleção e a pesquisa com preenchimento automático de países, para possibilitar a descoberta de países com localização desconhecida pelo utilizador. O tráfico entre dois países é representado com uma linha que interliga esses países. Estas linhas de ligação tiram partido da terceira dimensão para orbitar à volta do globo, não ocultando graficamente os países e o *feedback* de seleção. Deste modo, a tri-dimensionalidade da visualização introduz uma contribuição não apenas estética.

3. CONCEÇÃO DA VISUALIZAÇÃO

Nesta secção descrevemos o modelo da visualização proposta, referindo as funcionalidades que estão em falta de acordo com este modelo e os principais contributos que a aplicação vem trazer para o domínio em que se insere.

3.1. Motivação

As visualizações estudadas dentro do domínio de visualização de informação demográfica apresentam limitações ao nível da interação quando comparadas com

¹<http://threejs.org>

²<https://www.chromeexperiments.com/globe>

uma visualização fora deste domínio. O "Small Arms and Ammunition - Imports & Exports" implementa funcionalidades de seleção e pesquisa de países, funcionalidades que poderiam ser facilmente transpostas para as restantes visualizações, melhorando a usabilidade destas.

Adicionalmente, não são introduzidas mais-valias quanto ao uso de três dimensões, visto que os seus componentes tri-dimensionais poderiam ser substituídos por bi-dimensionais sem perda de legibilidade (por exemplo, substituir paralelepípedos, onde valores são representados pela sua altura, por pie-charts, onde valores são representados pelo seu preenchimento).

Todas as visualizações tri-dimensionais no domínio da demografia adotam o modelo de referência de visualização de informação [Chi 98][Mazza 09] para implicitamente transformar e mapear dados, dando ao utilizador a possibilidade de interação e configuração somente sobre a apresentação, ou seja, a última fase do modelo de referência. O mapeamento dos dados em estruturas visuais é feito implicitamente e sem possibilidade de configuração, o que implica que a leitura dos dados está cingida às opções de design estabelecidas pelos desenvolvedores. Numa perspetiva onde um utilizador utiliza estas ferramentas para estudo demográfico, onde este poderá procurar por relações entre diferentes tipos dados para inferir se o efeito de um indicador é a causa de perturbações noutra, seria interessante permitir uma maior flexibilidade na fase de mapeamento para apropriar a leitura e descoberta destas relações às preferências do utilizador. Juntamente com este aspeto, porque o utilizador teria a possibilidade de selecionar o tipo de dados que pretende consultar, seria benéfico poder observar outros tipos de dados para além da população total ou a densidade.

Por fim, todas estas visualizações apresentam a informação de forma estática, não permitindo realizar um *zoom* sobre a informação para melhorar a legibilidade dos dados relevantes. Em consequência, grande parte dos dados populacionais tornam-se pouco legíveis devido à existência de um pequeno conjunto de casos com valores extremamente elevados ou extremamente baixos quando comparados com os restantes dados. Isto verifica-se, por exemplo, devido ao facto da China e a Índia possuírem juntamente aproximadamente 36% da população mundial, ou do Mónaco possuir uma grande densidade populacional quando comparado com outros países, implicando que as alturas das barras verticais de países com população média e baixa sejam extremamente pequenas.

3.2. Contributos e Modelo

Como ponto de partida, propõe-se a implementação de funcionalidades de usabilidade existentes na visualização "Small Arms and Ammunition - Imports & Exports". Deste modo, a seleção permite a apresentação detalhada dos dados do país selecionado de modo textual e a pesquisa por nome de países permite a seleção de um país cuja localização é desconhecida pelo utilizador.

O Living Globe deverá naturalmente corresponder aos mo-

delos de visualização referidos, ao receber dados, transformar, mapear e apresentar estes. No entanto, apesar de fazer algum mapeamento por predefinição, o seu principal contributo está no facto de fornecer ao utilizador a possibilidade de realizar o seu próprio *mapping*. Para isto, conceberam-se três estruturas visuais, a que chamamos de Componentes, e cuja representação é configurável pelo utilizador: i) a altura de barras verticais, onde a menor altura é o valor mais baixo e a maior altura é o valor mais alto; ii) a cor das barras verticais, que varia entre um conjunto de cores de uma escala de azul a amarelo, onde azul é o valor mais baixo e amarelo é o valor mais alto; e iii) a cor dos países, que varia entre um conjunto de cores de uma escala de vermelho a verde, onde vermelho é o valor mais baixo e verde é o valor mais alto. O uso de três dimensões permite que um mesmo elemento gráfico, neste caso, a barra vertical, possa ser utilizado quanto à sua forma e a sua cor para representar dois tipos de dados.

A escolha dos pares cores azul-amarelo e vermelho-verde parte do princípio da harmonia de cores, visto que azul é cor complementar de amarelo e vermelho é cor complementar de verde (e vice-versa). As cores complementares, em termos de semiótica e perceção de cor, são ideais para representar ideias opostas, como por exemplo, direções cardinais, valores menores e maiores, entre outros [Pridmore 11]. Optou-se por dois pares de cores distintos em vez de um par para melhor representar Componentes (e portanto Indicadores) distintos.

Através desta funcionalidade, o utilizador deverá escolher qual o tipo de informação que cada Componente irá representar, entre um conjunto de tipos de informação possibilitados pelos dados originais recolhidos. Estes tipos de dados que denominamos Indicadores correspondem na aplicação desenvolvida a: i) população total; ii) densidade populacional; iii) crescimento populacional; iv) esperança média de vida; v) saldo migratório; vi) taxa bruta de natalidade (*Crude Birth Rate*, ou CBR); vii) taxa bruta de mortalidade (*Crude Death Rate*, ou CDR); viii) diferença entre o CBR e o CDR. O número de Indicadores que podem ser representados em simultâneo neste modelo é maior que o número de Indicadores disponíveis nas visualizações estado-de-arte mencionadas, viabilizando a procura de relações causa-efeito entre Indicadores diferentes.

Como referido anteriormente, um problema das visualizações analisadas é a dificuldade de visualização em situações em que existem valores máximos extremamente maiores (e/ou valores mínimos extremamente menores) do que a moda estatística da amostra. Para aliviar este problema, o Living Globe implementa uma funcionalidade de Filtragem onde utilizador estabelece o limite máximo e mínimo da amostra, definindo uma nova gama de valores de um Indicador e excluindo países com valores maiores ou menores que estes limites. Por exemplo, ao filtrar um país reduzindo o valor máximo, os valores dos restantes países são normalizados à nova escala de valores sem este país, aumentando por isso a altura das barras ou



Figura 2. "Living Globe" a apresentar dados demográficos de 2010 e com Reino Unido selecionado, mostrando detalhes textuais deste no canto inferior esquerdo.

aproximando as cores do amarelo, no caso das barras, e do verde, no caso dos países. Isto permite que representações de componentes semelhantes possam potencialmente ter uma maior diferenciação, sendo por isso mais facilmente comparáveis e legíveis visualmente.

A possibilidade de observar dados de diferentes anos para poder acompanhar o seu crescimento ou decréscimo ao longo do tempo é um aspeto importante na comunidade científica para a descoberta de fenómenos e a inferência de comportamento futuro. Deste modo, o modelo aqui proposto assume a existência de valores para cada Indicador e para ARs (anos de recolha) diferentes, e por isso o protótipo que o suporta apresenta uma estrutura visual que permite ao utilizador definir o AR em visualização.

4. IMPLEMENTAÇÃO DO PROTÓTIPO

Nesta secção iremos descrever um protótipo que implementa o modelo descrito na secção anterior, utilizando a linguagem Javascript, a biblioteca *three.js*, e outras bibliotecas como *chroma.js*³ e *jQuery*⁴. Este protótipo foi testado com sucesso nos browsers Firefox 37.0.1, Chrome 42.0.2311.90 (64-bit) e Safari 8.0.5, e está disponível livremente em <https://github.com/eduardo/living-globe>.

4.1. Recolha e leitura de dados

Para conceber a visualização proposta, seria necessário recolher um conjunto de dados com os diversos Indicadores por país e por AR. Foram obtidos dados a partir do World Bank⁵, com *open license* para uso não-comercial. Estes dados estão estruturados como uma tabela, tendo por isso uma linha por cada par País-Indicador e uma coluna por cada AR. Os ARs destes dados vão desde 1980 até 2014, sendo que o Indicador de saldo migratório é apenas registado a cada 5 anos. Por fim, os países são identificados pelo nome oficial e pelo código ISO 3166 alpha-3⁶, um

³<http://old.driven-by-data.net/about/chromajs/>

⁴<https://jquery.com>

⁵<http://data.worldbank.org>

⁶http://www.iso.org/iso/country_codes

código que identifica unicamente cada país com três letras.

Tendo já os dados dos Indicadores para um período de tempo, resta obter a localização espacial dos países, representados por pares de valores latitude-longitude, para posicionar a barra vertical. Esta informação foi obtida a partir do *Sokrata Open Data Portal*⁷, um portal de dados open-source, também para uso não-comercial. Os dados obtidos correspondem a uma tabela com colunas para o código ISO 3166 alpha-3 e para a latitude e longitude, entre outros.

Todos os dados recolhidos são lidos assincronamente utilizando a tecnologia Ajax e analisado de modo a descobrir os valores máximos e mínimos dos Indicadores selecionados. Estes valores são utilizados para normalizar os restantes dados, como é descrito na sub-secção "Globo e Países" e "Barras Verticais".

4.2. Globo e Países

Outras funcionalidades propostas nesta visualização são a seleção e coloração de países. A implementação destas funcionalidades implica um conjunto de requisitos como o desenho e identificação dos países num objeto tri-dimensional, para possibilitar a deteção de seleção e a modificação da cor de cada país individual.

Como primeira solução para identificar os países no globo, utilizou-se um ficheiro CSV (*Comma Separated Values*)⁸ que contém um conjunto de pares latitude-longitude para cada país, sendo que o conjunto de pontos geográficos representa um polígono equivalente ao território de um dado país. O uso deste ficheiro implica o desenho programático dos países de modo a poder delimitá-los no globo, permitindo colorir os vários polígonos tal como requerido. Para identificar um país selecionado, a localização do cursor do rato era traduzida para latitude-longitude, sendo comparada com os limites do país. No entanto esta implementação mostrou-se como muito pouco eficiente pois o desenho dos múltiplos polígonos que se alteram quando se reconfiguram os Indicadores, o AR e/ou a Filtragem tem um custo computacional excessivo e torna a visualização demasiado lenta, afetando consideravelmente a experiência de utilização.

A segunda solução considerada consiste no uso de texturas, *canvas* e um *raycaster*, inspirado na implementação da visualização "Small Arms and Ammunition"⁹. São sobrepostas no globo cinco texturas diferentes: a) uma textura *satellite*, com uma fotografia da terra, utilizada apenas para efeitos estéticos; b) uma textura *specular*, que ajuda a manter a luminosidade dos países em relação ao mar; c) uma textura *ratio* inicialmente transparente, onde dadas áreas são coloridas com a cor correspondente ao valor normalizado do Indicador escolhido; d) uma textura *select* inicialmente transparente, onde dadas áreas são coloridas de branco, um de cada vez, para servir de feedback do país selecionado; e) uma textura *lookup*, que contém uma imagem

⁷<https://opendata.sokrata.com/dataset/Country-List-ISO-3166-Codes-Latitude-Longitude/mnkm-8ram>

⁸<https://mledoze.github.io/countries/>

⁹<http://mflux.tumblr.com/post/28367579774/armstradeviz>

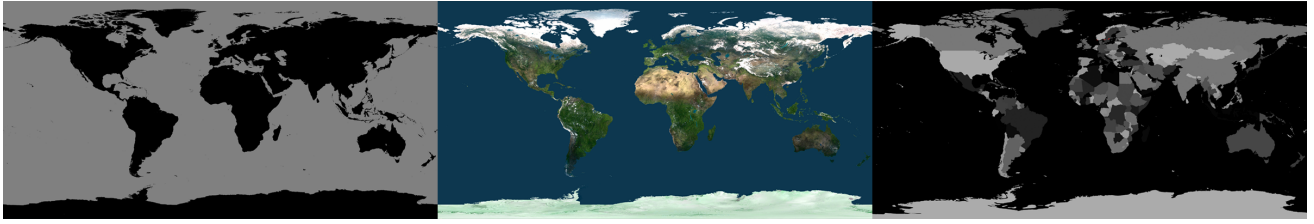


Figura 3. As três texturas utilizadas no material do globo. Da esquerda para a direita: textura *specular*, textura *satellite* e textura *lookup*

com o plano terrestre numa escala de cinzentos, onde cada país está colorido com um grau de luminosidade de cinzento diferente. Deste modo, esta textura *lookup* introduz as áreas que cada país cobre, e possibilita a sobreposição de cores sobre essas áreas. Porque não estamos a lidar com elementos gráficos vetoriais, isto é, de tamanho ajustável em tamanho real, o uso de texturas tal como descrito é apropriado.

Adicionalmente, adiciona-se um ficheiro JSON que mapeia o valor de luminosidade do cinzento num código ISO 3166-1 alpha-3 de países, gerado manualmente e baseado no ficheiro utilizado pela visualização "Small Arms and Ammunition", que continha códigos ISO 3166-1 alpha-2 (duas letras em vez de três). Como os dados tratados inicialmente possuem um código ISO 3166-1 alpha-3 por país, podemos assim associar os dados demográficos a cada nível de cinzento.

Para obter o cinzento selecionado na textura *lookup*, após um clique do rato, utiliza-se um *canvas* e um *raycaster*. O *raycaster* permite obter a interseção da posição do cursor do rato com o *canvas*, e o *canvas* permite associar a localização pressionada a uma localização na textura. Assim, o processo de seleção passa pelos seguintes passos: i) verificar qual o nível de cinzento intersetado pelo *raycaster*; ii) obter o código ISO 3166-1 alpha-3 correspondente ao cinzento intercetado; iii) determinar o país com o código associado; iv) identificar os dados demográficos deste, retirando qualquer cor na textura *select*; e v) pintar de branco a área cinzenta correspondente.

A cor associada a cada país é calculada a partir do valor normalizado do Indicador selecionado. Utiliza-se a biblioteca *chroma.js* para criar um gradiente de cor no modo RGB entre vermelho a verde, onde 0 é vermelho, 1 é verde e qualquer valor entre 0 e 1 é uma cor entre estas. Para obter a cor correspondente ao valor do Indicador, normaliza-se este entre 0 e 1 através da expressão 1 e obtêm-se a cor que identifica este valor normalizado. Na expressão 1, o valor V representa o valor do Indicador i escolhido para um dado País p no AR a .

$$V_{Normalizado}(i, p, a) = \frac{V(i, p, a) - \text{Min}(i, a)}{\text{Max}(i, a) - \text{Min}(i, a)} \quad (1)$$

4.3. Barras Verticais

Para aferir a altura das barras verticais, que deverão representar os valores de um Indicador sem exceder o plano de visão do *browser*, utilizou-se novamente a expressão 1 para normalizar os valores lidos entre 0 e 100 píxeis. Optou-se por uma altura máxima de 100 píxeis por esta ser visualmente confortável em relação à câmara da visualização.

A coloração das barras foi implementada de forma semelhante à coloração dos países (descrito na secção "Globo e Países"). Para este caso, criou-se através da biblioteca *chroma.js* um gradiente de amarelo a azul, onde 0 é amarelo, 1 é azul, e qualquer valor entre 0 e 1 é uma cor entre estas duas cores. Assim, de modo a obter a cor correspondente ao indicador escolhido, normaliza-se entre 0 e 1, através da expressão 1, e obtêm-se a cor mapeada com o valor normalizado.

4.4. Filtragem

Como referido na subsecção anterior, o uso de normalização evita o problema da altura excessiva das barras verticais no plano de visão. Ainda assim, por exemplo, para o Indicador de população total, esta solução leva a que as barras da China e Índia assumam alturas entre 90 e 95 píxeis e que as restantes barras assumam alturas entre 1 e 20 píxeis. No caso da coloração dos países, estes dois países terão cores próximas do verde e amarelo, e os restantes países terão cores avermelhadas e indistinguíveis a olho-nú.

Para poder filtrar os dados de países, utiliza-se a lógica de normalização referida anteriormente mas onde os valores mínimo e máximo são selecionados pelo utilizador. Se o valor for menor que o valor mínimo ou maior que o valor máximo, a barra ou cor no país deverá ser apagada, enquanto que as restantes barras deverão "crescer" ou "decrecer" e as restantes cores deverão progredir mais em direção a uma cor num dos extremos do gradiente.

Assim, a aplicação possui um elemento *slider* no topo da página, associado aos indicadores configurados, e basta ao utilizador escolher os indicadores que pretende analisar e escolher o seu valor mínimo e máximo no *slider*. A representação visual é dinamicamente renderizada à medida que o valor do *slider* é alterado.

A Filtragem de países possibilita a potencial diferenciação de barras e cores para uma melhor leitura dos dados, tal como se observa na figura 4.

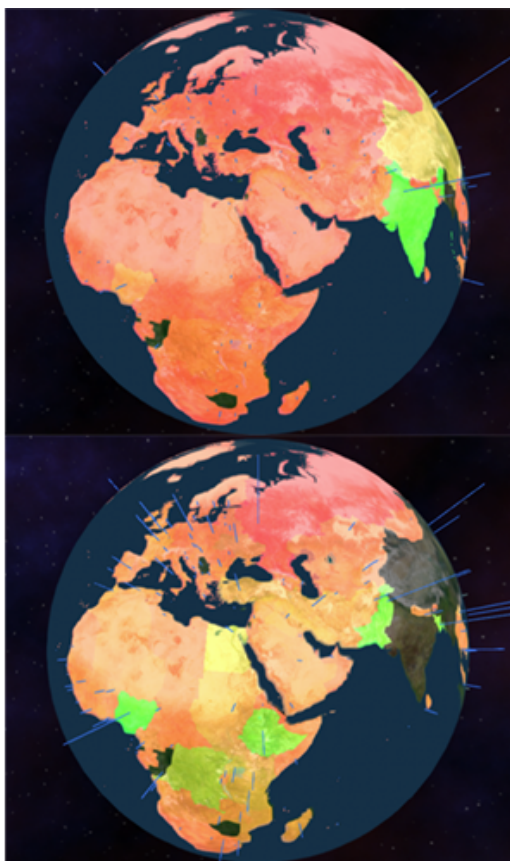


Figura 4. Indicador rácio natalidade-mortalidade mapeado na cor dos países. Quando não-filtrado (em cima) a cor da maioria dos países é uniformemente vermelho e indistinguível entre eles. Quando a China e Índia são filtrados (em baixo), a maioria das cores são distinguíveis e os países com melhor rácio (mais perto do verde) são mais facilmente detetáveis.

Durante o desenvolvimento consideraram-se diferentes formas de representação tri-dimensional de atributos que variam ao longo de um período de tempo, tal como a representação filtrada, onde os dados seriam filtrados de modo a se apresentar apenas um único ano de cada vez, selecionado pelo utilizador e a representação em lápis [CT05]. Esta segunda implicaria a introdução de um objeto adicional por cada país com a forma geométrica de um lápis, onde cada face é colorida desde ponta de carvão (localizado geograficamente no país) até ao cabo deste para representar os valores do primeiro até ao último AR. No entanto, esta representação iria gerar um congestionamento visual, tornando a leitura e comparação entre barras verticais ou entre lápis mais difícil, principalmente para países de pequena dimensão. Assim, optou-se pela representação filtrada, implementando um segundo *slider* horizontal para o utilizador "deslizar" de forma sequencial entre valores de AR.

4.5. Pesquisa

Para implementar a pesquisa, introduziu-se um campo de texto com sugestão de resultados dinâmica que se adapta ao conteúdo escrito no campo. Estas sugestões são adicionadas com a extensão de jQuery "DevBridge Auto-complete"¹⁰, uma extensão onde se especifica qual a lista de dados a dispor nas sugestões e que, por sua vez, os irá dispor automaticamente conforme o que o utilizador escrever no campo de pesquisa. Assim, apresentam-se nas sugestões os nomes de todos os países. Ao selecionar uma das sugestões, obtêm-se o código ISO 3166 alpha-3 e seleciona-se o país correspondente, utilizando o mesmo algoritmo especificado previamente em "Globo e Países".

5. AVALIAÇÃO

A avaliação da visualização implementada permite testar se o modelo concebido acrescenta alguma melhoria em relação às visualizações estado-de-arte no sentido da perceção e da leitura de informação. É necessário perceber se a leitura simultânea de um maior número de indicadores demográficos, representados sobre os Componentes conforme a preferência do utilizador, é intuitiva e facilita a compreensão de relações entre Indicadores.

Um aspeto importante desta fase da avaliação é perceber qual a audiência da visualização, ou seja, qual o contexto social ou profissional em que um dado indivíduo irá utilizar preferencialmente esta visualização em contraste com outras visualizações demográficas tri-dimensionais e até visualizações demográficas bi-dimensionais. Por esta razão, considerou-se que a audiência desta visualização deverá corresponder a dois tipos de utilizador: a) um utilizador com pouca literacia informática, com conhecimento moderado em Ciências Sociais e Estatística; ou b) um utilizador com moderada a alta literacia informática, com conhecimento alto em Ciências Sociais, Estatística ou até Visualização de Dados e Informação.

Tendo em consideração a audiência alvo, utilizaram-se dois tipos de avaliação para cobrir as diferentes necessidades e níveis de especialização dos utilizadores.

5.1. Teste de usabilidade

A primeira avaliação correspondeu a um teste de usabilidade, em que se pediu a dois utilizadores com pouca literacia informática, com idades de 23 e 59 anos e sem conhecimento nem experiência no domínio da Visualização de Informação para realizar um conjunto de tarefas e fornecer algumas observações finais sobre a facilidade de utilização da visualização e resolução destas tarefas. Cada indivíduo foi observado em separado e cada teste durou aproximadamente 7 minutos.

A primeira fase do teste, com duração entre 30 segundos a 1 minuto, constituiu numa introdução breve sobre as funcionalidade básicas da aplicação, nomeadamente a possibilidade de seleção e pesquisa de países, os Componentes visuais, a personalização de Indicadores e a Filtragem.

¹⁰<https://www.devbridge.com/sourcery/components/jquery-autocomplete/>

A segunda fase do teste, com duração entre 5 a 6 minutos, correspondeu à resolução de um conjunto de cinco tarefas. Todas as tarefas incidiam na resolução de um dado problema, sendo por isso não só quanto à precisão da resposta ao problema como também quanto à rapidez na execução da tarefa e quanto à incidência de falsos positivos, onde o utilizador assume que concluiu a tarefa corretamente quando esta não o foi. As cinco tarefas corresponderam à resolução das seguintes questões, introduzidas pela mesma ordem ao utilizador:

1. Qual é a população total de Portugal?
2. Qual é o crescimento populacional de Porto-Novo?
3. Sem utilizar a seleção ou pesquisa, qual é o país com maior natalidade?
4. Sem utilizar a seleção ou pesquisa, entre o Brasil e a Argentina, qual o que apresenta uma maior razão de nados-vivos / óbitos?
5. Sem utilizar a seleção ou pesquisa e associando o Indicador "Taxa Bruta de Mortalidade" à cor dos países, qual o país com a maior taxa de mortalidade?

O objetivo principal das primeiras três tarefas, com um grau de dificuldade baixo, é colocar os utilizadores mais à vontade com a aplicação e incentivá-los a utilizar as funcionalidades, como a pesquisa. O objetivo das duas últimas tarefas, com um grau de dificuldade maior, é introduzir desafios para avaliar se estas funcionalidades são fáceis de utilizar para chegar às soluções.

É de salientar que a quinta tarefa tem como principal objetivo perceber se o uso das cores vermelho e verde para representar respetivamente valores baixos a valores altos de mortalidade iria confundir o utilizador na eventualidade de este assumir o significado das cores respetivamente pela sua conotação habitual negativa e positiva. Ou seja, o utilizador poderá inferir que um país com cor verde tem valores "melhores", logo irá ter um menor valor de mortalidade (e vice versa), e gerando um falso-positivo.

Os resultados dos testes realizados são apresentados na tabela 1, sendo: 'Tempo', o tempo que o utilizador utilizou para concluir a tarefa; 'Ajuda', se o utilizador solicitou ajuda; 'Confiança', se o utilizador tem confiança na resposta dada; e 'Correção', se o utilizador chegou ao resultado correto.

A última fase do teste, com duração de 1 minuto, passou por inquirir os utilizadores sobre que aspetos da visualização poderiam ter sido melhorados e que tarefas consideraram mais fáceis de executar. As observações obtidas indicam que a visualização foi fácil de utilizar e que a funcionalidade de Filtragem é extremamente útil para obter informação gráfica de certos países com mais detalhe. O segundo utilizador, que respondeu erradamente e com confiança à quinta tarefa, afirmou que momentaneamente não se recordou da escala de cores de vermelho a verde

Tabela 1. Resultados dos testes de usabilidade realizados.

| | | Utilizador 1 | Utilizador 2 |
|----------|-----------|--------------|--------------|
| Tarefa 1 | Tempo | 2 s | 8 s |
| | Ajuda | Não | Não |
| | Confiança | Sim | Sim |
| | Correção | Sim | Sim |
| Tarefa 2 | Tempo | 21 s | 16 s |
| | Ajuda | Não | Não |
| | Confiança | Sim | Sim |
| | Correção | Sim | Sim |
| Tarefa 3 | Tempo | 11 s | 47 s |
| | Ajuda | Não | Não |
| | Confiança | Sim | Sim |
| | Correção | Sim | Sim |
| Tarefa 4 | Tempo | 33 s | 39 s |
| | Ajuda | Não | Não |
| | Confiança | Sim | Sim |
| | Correção | Sim | Sim |
| Tarefa 5 | Tempo | 26 s | 43 s |
| | Ajuda | Não | Não |
| | Confiança | Sim | Sim |
| | Correção | Sim | Não |

usada, mas em vez disso da conotação negativa do vermelho. Ambos os inquiridos classificaram com 4 (numa escala do tipo Likert de 1 a 5) a satisfação de utilização da aplicação.

5.2. Avaliação Heurística

Foi também realizada uma avaliação heurística por dois analistas com conhecimentos no domínio de Visualização de Informação utilizando heurísticas apropriadas e à escolha do avaliador. Ambos os avaliadores utilizaram as heurísticas de Nielsen [Nielsen 90], dado o seu potencial para encontrar problemas de usabilidade [Nielsen 92]. As avaliações foram feitas em separado, e cada avaliador elaborou um relatório onde indicou os potenciais problemas de usabilidade, as heurísticas não cumpridas, bem como a gravidade de cada problema, indicando também aspetos em que as heurísticas eram cumpridas.

De um modo geral, a visualização destacou-se positivamente nas heurísticas *Correspondência entre o sistema e o mundo real*, *Controlo e livre-arbítrio do utilizador*, *Consistência e Adesão a Normas*, *Reconhecimento em vez de lembrança*, *Flexibilidade e Eficiência* e *Desenho de ecrã estético e minimalista*. Isto significa que a representação gráfica de dados demográficos com escolha livre de Componentes parece não apresentar problemas severos de usabilidade. No entanto, ambos os avaliadores encontraram um aspeto negativo correspondendo ao não cumprimento das heurísticas *Visibilidade do estado do sistema* e *Ajuda e Documentação*, relacionado com a falta de documentação ou instruções que indiquem ao utilizador os próximos passos a tomar quando a visualização é iniciada.

6. CONCLUSÕES E TRABALHO FUTURO

Neste artigo propõe-se uma visualização de dados demográficos tridimensional sobre um globo e descreve-se a implementação dum protótipo que permite a exploração interativa desses dados de forma mais completa e flexível do que as aplicações com objetivos análogos atualmente disponíveis. A usabilidade deste protótipo foi avaliada através de avaliação heurística e testes de usabilidade. Um dos aspetos mais importantes desta visualização é a filtragem dos dados que facilita o processo de análise. Embora, os resultados dos testes de usabilidade pareçam indicar que a funcionalidade de filtragem não é evidente para os utilizadores, uma vez aprendida, é utilizada com frequência para facilitar a visualização de países com valores mais pequenos dos indicadores em estudo. Outro resultado da avaliação aponta para que a possibilidade dada aos utilizadores de associar cada tipo de indicador a uma forma de representação visual (por exemplo, cor), pode conduzir a associações menos intuitivas e provocar erros de interpretação. Por esta razão, propõe-se como trabalho futuro, facultar aos utilizadores a possibilidade de escolher não apenas a associação de indicadores a formas de representação visual (Componentes), mas também as escalas de cor utilizadas. É ainda de salientar que os utilizadores recorreram à pesquisa e seleção sem solicitar ajuda, o que sugere que estas funcionalidades são intuitivas e permitem realizar facilmente um *zoom* dos dados. Também os resultados da avaliação heurística foram geralmente positivos.

Em geral, o modelo proposto itera sobre os modelos estado-de-arte analisados e adiciona componentes que tornam a exploração visual e interativa de dados demográficos mais fácil e intuitiva.

Os conceitos e funcionalidade do modelo proposto neste artigo podem ser utilizados em contextos para além do domínio de estudo demográfico. Sendo assim, consideramos o trabalho atual como um ponto de partida para o desenho de uma arquitetura genérica que recebe dados, sob a forma de uma API (Application Program Interface) estruturada, e os representa no globo. Deste modo, a funcionalidade de seleção do país, seleção do ano de recolha, pesquisa, filtragem e navegação do globo estariam já implementadas e abstraídas, agilizando o processo de apresentação de dados num globo por parte de um desenvolvedor em qualquer página web. Esta API poderia também incluir ou esconder funcionalidades conforme a composição dos dados, por exemplo, se os dados não têm uma data de recolha, então o *slider* de AR é escondido.

7. AGRADECIMENTOS

Este trabalho é suportado por Fundos Nacionais através da FCT - Fundação para a Ciência e a Tecnologia, no contexto dos projetos UID/CEC/00127/2013 e Incentivo/EEI/UI0127/2014.

Os autores gostariam de agradecer a Luís Silva, Bruno Garcia, Pedro Miguel, Isabel Nascimento, Rui Simões e João Pedrosa pelos valiosos contributos que deram no teste e melhoria da aplicação.

Referências

- [Arts 11] Google Data Arts. World population. <http://data-arts.appspot.com/globe>, 2011.
- [Arts 14] Google Data Arts. Small arms and ammunition - imports & exports. <http://armsglobe.chromeexperiments.com>, 2014.
- [Chi 98] Ed H. Chi. An operator framework for information visualization systems. *Information Visualization Symposium*, 1998.
- [CT05] P. Schulze-Wollgast; H.Schumann C. Tominski. 3d information visualization for time dependent data on maps. *Proceedings of the International Conference on Information Visualisation*, 2005.
- [Dirksen 12] Jos Dirksen. World population density - 2010. <http://www.smartjava.org/examples/population>, 2012.
- [Mazza 09] Riccardo Mazza. Introduction to information visualization. *Springer-Verlag London*, 1, 2009.
- [Nielsen 90] Jakob Nielsen. Heuristic evaluation of user interfaces. *Proceedings of the ACM CHI'90 Conference*, páginas 249–256, 1990.
- [Nielsen 92] Jakob Nielsen. Finding usability problems through heuristic evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, páginas 373–380, 1992.
- [Pridmore 11] Ralph W. Pridmore. Complementary colors theory of colorvision: Physiology, color mixture, colorconstancy and color perception. *Color. Res. Appl.*, 2011.
- [Zhifang 13] Wang Zhifang. China and us population. <http://newnaw.com/pub/js/webglglobe/chinaanduspop>, 2013.