

Manipulações no espaço sobre mesas interativas com estereoscopia

Fernando Fonseca Daniel Mendes Bruno Araújo
Alfredo Ferreira Joaquim Jorge
INESC-ID / IST / UTL

fernando.fonseca@ist.utl.pt, danielmendes@ist.utl.pt, brar@vimmi.inesc-id.pt,
alfredo.ferreira@inesc-id.pt, jaj@inesc-id.pt

Resumo

Mesas interativas com estereoscopia oferecem capacidades únicas de visualização, permitindo aos utilizadores perceberem objetos virtuais como se estes estivessem em cima da superfície. Com esta capacidade de visualização aparecem novas formas de interagir com os objetos virtuais, as quais representam desafios interessantes. Neste documento apresentamos um conjunto de cinco técnicas de interação, que permitem ao utilizador manipular objetos virtuais colocados acima da superfície. Das técnicas apresentadas, quatro utilizam o espaço tridimensional da visualização, recorrendo a tecnologias não intrusivas, enquanto outra recorre a multi toque.

Palavras-Chave

Manipulação de objetos 3D virtuais, Mesas interativas, Interações espaciais, Multi-toque, Estereoscopia

1. Introdução

A percepção em ambientes virtuais pode ser aumentada combinando visualização estereoscópica com seguimento da cabeça do utilizador, aumentando assim a sua imersão. Através da posição da cabeça do utilizador é possível criar um frustum de visualização personalizado para cada olho, criando a ilusão de que os objetos da cena estão acima da superfície. Esta ilusão torna-se ainda mais forte quando é dada a liberdade ao utilizador de se movimentar e a imagem é gerada segundo a sua perspetiva, sem que este necessite manipular qualquer câmara ou widget. Esta solução é facilmente aplicável a qualquer ambiente, desde ecrãs de larga escala a dispositivos móveis ou até a mesas interativas, onde não tem sido muito explorada. Este cenário interativo é especialmente apelativo para modelação 3D ou exploração e montagem. Por exemplo, um modelo arquitetural pode ser mostrado em cima de uma mesa sem necessitar que este seja reproduzido fisicamente, reduzindo assim os custos e aumentando a velocidade do processo de modelação, ou até mesmo servir para mostrar resultados a um cliente.

O avanço tecnológico que se tem vindo a verificar nos últimos anos tornou mais fácil o desenvolvimento deste tipo de cenários. Ainda há pouco tempo atrás a monitorização da cabeça do utilizador necessitava de uma sala repleta de câmaras de infravermelhos e marcadores colocados no utilizador, ou então de um sistema de fios invasivo. Hoje em dia este processo é possível utilizando uma única e acessível câmara de profundidade, tornando-o não intrusivo para o utilizador. Esta solução pode ser usada para definir a perspetiva do utilizador, assim como para se-

guir membros, mãos, os dedos do utilizador, possibilitando novos tipos de interação.

Interagir diretamente com mesas com estereoscopia tem sido alvo de investigação, usando superfícies multi-toque [5, 15] ou dispositivos vestíveis [2, 17]. Embora haja soluções robustas para interações multi-toque, estas estão limitadas a um espaço bidimensional, não podendo assim oferecer interações diretas no espaço de visualização. Através da utilização de dispositivos vestíveis consegue-se ultrapassar esta questão, mas estes são invasivos e restritivos. Outros investigadores [14, 16] propõem soluções para interação num espaço tridimensional, não necessitando de qualquer objeto adicional acoplado ao utilizador, mas, no entanto, estes não combinam as suas soluções com sistemas estereoscópicos, usando-as apenas como um cursor mais poderoso, mas na mesma indireto.

Neste documento apresentamos um conjunto de cinco técnicas para manipulações de objetos em três dimensões, desenvolvidas para mesas interativas com estereoscopia. Quatro destas técnicas são de interação espacial acima da superfície, enquanto outra usa toque na superfície. Estas técnicas proporcionam ao utilizador tanto manipulações diretas como indiretas.

2. Trabalho Relacionado

Várias aplicações requerem manipulações de objetos virtuais em três dimensões. Exemplos deste tipo de aplicações são exploração de modelos, desenho e montagem de componentes de engenharia, entre muitos outros. Sendo uma tarefa bastante comum, esta tem vindo a ser a motivo de investigação intensiva.

Depois de desenvolver técnicas de controlo de seis graus de liberdade (DOF) usando um, dois e três toques, Hancock et al. [6] concluíram que um maior número de toques era melhor em termos de desempenho e de preferência dos utilizadores. Estes resultados sugerem que é desejável uma aproximação do mapeamento do input ao número de graus de liberdade do output. Os autores também definiram uma série de diretrizes para a criação de interfaces multi-toque, tal como a criação de um elo visual e físico com os objetos assim, como um retorno visual 3D apropriado. Mais tarde, os mesmos autores procederam a um aperfeiçoamento das técnicas propostas, acrescentando ferramentas de manipulação [7] a que eles chamaram de Sticky Tools.

Reisman et al. [13], considerando a norma *de facto* para manipulações a duas dimensões, o algoritmo Translate-Rotate-Scale (TRS), propuseram um método de interação usando vários pontos de contacto numa superfície multi-toque para manipular objetos 3D em seis graus de liberdade (DOF). A sua solução mantém os pontos de contacto constantes durante toda a interação, permitindo rotações e translações em simultâneo. O principal problema numa abordagem que permita diferentes transformações integradas prende-se sempre com o facto de poder provocar manipulações indesejadas, podendo dificultar tarefas mais precisas. Com isto em mente, foi sugerida a separação de DOFs [12], sendo seguida por vários autores. Martinet et al. [10] introduziram a DS3, uma técnica de manipulação 3D com separação de DOF e comparou-a com trabalhos anteriores [7, 13]. Esse estudo com utilizadores revelou que a separação de DOF trazia melhores resultados.

Cohé et al. [4], baseando-se em observações de utilizadores a manipular widgets, desenharam o seu próprio widget para transformações 3D, tBox. Este favorece o controlo direto e independente de DOF (translação, rotação e escala ao longo dos três eixos). LTouchIt [11], embora permita manipulações diretas para translações, recorre também a widgets para as rotações. Knoedel et al. [9] investigaram o impacto de manipulações diretas e indiretas em técnicas TRS. Os testes com utilizadores indicam que uma abordagem direta é melhor relativamente ao tempo de realização das tarefas, mas que uma abordagem indireta pode aumentar a eficiência e precisão.

Para manipular objetos virtuais em 9-DOF, Martinet et al. apresentaram Toucheo [5], uma configuração com visualização estereoscópica co-localizada, para evitar oclusões. O utilizador interage através de uma interface que combina widgets, a metáfora do balão [2] e o algoritmo TRS numa superfície multi-toque. O Mockup Builder [1] oferece um ambiente de modelação semi-imersivo no qual o utilizador pode manipular livremente objetos virtuais. Para isso os autores recorrem a dois Gametracks adaptados para detetar a posição dos dedos no espaço de interação.

Não se focando em estereoscopia, Song et al. [14] exploraram interações espaciais e propuseram a metáfora do espeto (Handle-Bar) como uma forma eficaz de manipular

objetos no espaço. Color Glove [17], um trabalho anterior dos mesmos autores, já o permitia, mas em vez de recorrer a um sensor de profundidade, recorria a luvas. Mais recentemente Wang et al. [16] introduziram uma nova forma de interagir usando apenas dois sensores de profundidade para rastrear mãos e dedos. Esta abordagem permite obter a posição e orientação de ambas as mãos em 6-DOF.

3. Técnicas de Interação

Num ambiente de mesa interativa com estereoscopia implementámos cinco técnicas diferentes de interação e manipulação 3D. Quatro destas técnicas consistem em interações no espaço acima da superfície, tanto indiretas como diretas, e uma usando apenas interações multi-toque na superfície. Todas as técnicas apresentadas possibilitam um controlo total em sete graus de liberdade (DOF): três graus para translação, três graus de rotação e um para escala uniforme.

Para permitir a implementação destas técnicas, a nossa configuração dispõe de duas câmaras de profundidade, uma colocada à frente do utilizador para lhe seguir a cabeça e outra colocada por cima da superfície, apontada para baixo de forma a conseguir seguir as mãos do utilizador. Para a deteção do toque recorreremos à tecnologia Laser Light Plane (LLP).

3.1 6-DOF Hand

Por forma a conseguir imitar interações com objetos físicos tão perto do real como possível, usamos a informação dos seis graus de liberdade fornecida pela aplicação que segue as mãos do utilizador¹. Nesta técnica o utilizador agarra os objetos diretamente com uma mão, tipicamente a mão dominante. Então, todo o movimento dessa mão é diretamente aplicado sobre o objeto, replicando assim a interação física com objetos, como exemplificado na Figura 1.

Arrastando o objeto no espaço irá fazer com que este se mova nessa direção, assim como qualquer rotação efetuada pelo pulso fará com que o objeto rode segundo o mesmo eixo e os mesmos graus que a mão. Usando a outra mão, e agarrando algures no espaço, não necessariamente no objeto, e variando a distância entre essa mão e a mão que agarra o objeto, provocará uma alteração uniforme da escala do objeto que o utilizador está a manipular. Ao afastar as mãos, a escala aumentará e ao aproximar a escala diminuirá. O ponto onde o utilizador agarra o objeto permanecerá constante durante toda a manipulação, até o objeto ser largado, sendo este o centro de todas as transformações.

3.2 3-DOF Hand

Tendo em conta a separação dos graus de liberdade com vista a evitar manipulações indesejadas, tal como sugerido em [12], separámos as rotações e as translações por ambas as mãos, criando assim esta técnica (Figura 2). Após agarrar um objeto com uma mão, o utilizador controla a translação nas suas três dimensões, usando apenas as

¹3Gear Systems, <http://www.threegear.com/>

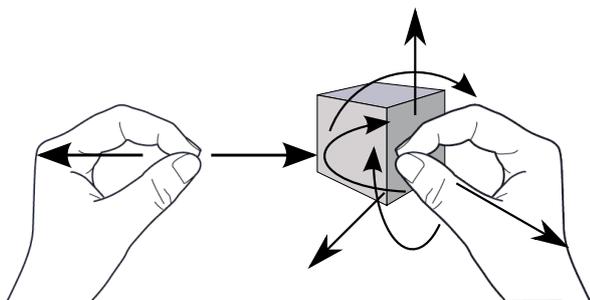


Figura 1. Técnica 6-DOF Hand.

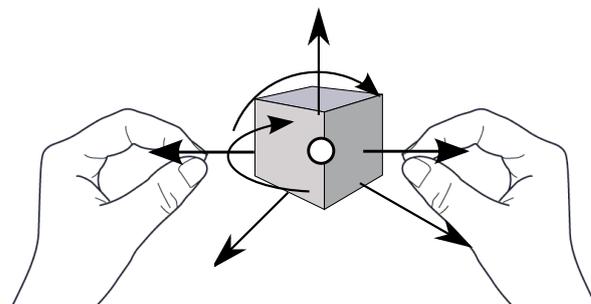


Figura 3. Técnica Handle-Bar.

movimentações dessa mão. Para controlar as rotações, o utilizador recorre às rotações do pulso da outra mão após agarrar algures no espaço e rodando a mão no eixo que pretende rodar o objeto, mantendo a primeira mão a agarrar o objeto. De forma semelhante à técnica anterior, a variação da distância entre ambas as mãos, enquanto estas estão a agarrar, escala uniformemente o objeto. O ponto onde o objeto foi agarrado mantém-se até ao final da interação, sendo este, mais uma vez, o centro de todas as transformações.

3.3 Handle-Bar

No seguimento dos trabalhos de Song et al [14], implementámos a metáfora da Handle-bar no nosso ambiente de mesa interativa com estereoscopia. Esta abordagem imita o uso de um espeto, tipicamente para assar frangos, em que o utilizador segura uma barra com as duas mãos e o frango se situa no seu ponto médio. Nesta técnica usamos o ponto médio entre as mãos do utilizador para manipular objetos (Figura 3). Ao introduzir o ponto médio dentro do objeto que o utilizador pretende manipular, o utilizador deverá efetuar o gesto de agarrar com ambas as mãos e mantê-lo.

O utilizador, ao estar a agarrar um objeto, pode deslocá-lo, movendo ambas as mãos na mesma direção, e rodar o objeto, movendo as mãos em torno do ponto médio. A distância entre as mãos, à semelhança das técnicas anteriores, controla a escala uniforme do objeto.

3.4 Air TRS

Sendo possível obter as coordenadas espaciais de ambas as mãos, conseguimos trazer o algoritmo Translate-Rotate-

Scale (TRS) para a terceira dimensão. As mãos do utilizador são tratadas como dois pontos no espaço, utilizando-as de forma similar ao TRS [8] com escala, tal como está exemplificado na Figura 4.

A mão que agarra o objeto é usada para o movimentar, tal como nas técnicas anteriores. A outra mão, depois de agarrar algures no espaço, permite então manipular a rotação e a escala como no TRS. Estas duas transformações são centradas no ponto onde o objeto é agarrado pela primeira mão e o ângulo da rotação é definido pela variação de posição de uma mão em relação à outra. Para escalar o objeto, é novamente usada a distância entre mãos.

3.5 Touch TRS + Widgets

Embora interações que recorrem ao multi-toque apenas permitam manipulações de objetos virtuais em três dimensões no espaço acima da mesa de forma indireta, este é hoje em dia um método de entrada muito comum, estando presente na nossa vida quotidiana. A nossa técnica de interação usa o algoritmo TRS [8] combinado com um conjunto de três widgets para permitir os sete graus de liberdade.

Com esta decisão de implementação, seguimos a separação de graus de liberdade defendida por vários autores [6, 10, 7, 4, 3] permitindo ao utilizador deslocar objetos virtuais no plano paralelo à superfície, tocando diretamente por baixo do objeto e arrastando esse dedo na superfície. Enquanto este toque estiver ativo, são mostrados três widgets à esquerda ou à direita, dependendo da mão que o utilizador está a usar (Figura 5). Usando um segundo to-

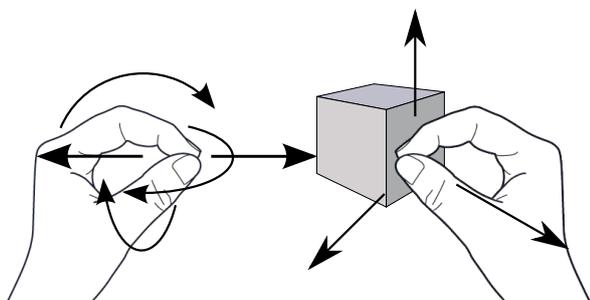


Figura 2. Técnica 3-DOF Hand.

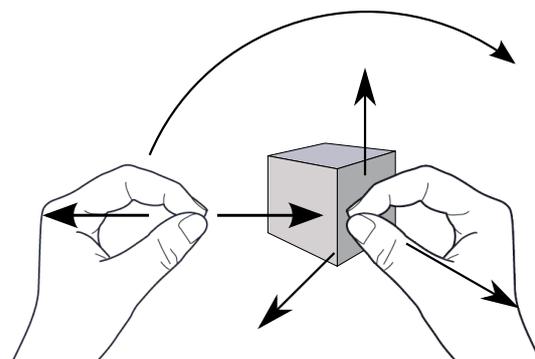


Figura 4. Técnica Air TRS.

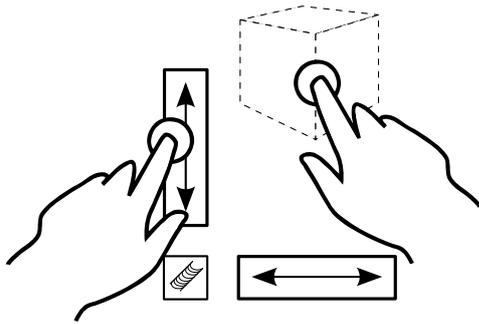


Figura 5. Técnica Touch TRS + Widgets.

que fora de qualquer widget, o utilizador pode rodar o objeto segundo o eixo vertical normal à superfície, e, ao mesmo tempo, escalar uniformemente. Se o segundo toque for num dos três widgets, o utilizador poderá rodar em volta de um dos dois eixos paralelos à superfície seguindo a metáfora da haste [5], ou alterar a altura do objeto, de forma similar à metáfora do balão [2].

4. Conclusões e Trabalho Futuro

Através da combinação de mesas interativas com visualização estereoscópica dependente da perspetiva do utilizador é possível criar a ilusão de que os objetos estão em cima da mesa. Com esta abordagem aparecem desafios interessantes para novos tipos de interação. Neste documento apresentámos um conjunto de cinco técnicas de interação que permitem ao utilizador manipular objetos em sete graus de liberdade (três de translação, três de rotação e um de escala uniforme). Quatro destas técnicas são de interação espacial, recorrendo a tecnologia não intrusiva, outra recorrendo a multi-toque na superfície. Por forma a obter opinião dos utilizadores e perceber efetivamente qual a abordagem que melhor se adequa para este tipo de ambientes pretendemos realizar um estudo comparativo com utilizadores, usando um cenário prático.

5. Agradecimentos

O trabalho apresentado neste artigo foi parcialmente financiado pela Fundação para a Ciência e a Tecnologia (FCT) através dos projetos Alberti Digital, CEDAR e TECTON-3D, com as referências PTDC/AUR-AQI/108274/2008, PTDC/EIA-EIA/116070/2009 e PTDC/EEI-SII/3154/2012, respetivamente, e pelo projeto Pest-OE/EEI/LA0021/2013.

Referências

- [1] B. R. D. Araújo, G. Casiez, J. A. Jorge, and M. Hachet. Mockup builder: 3d modeling on and above the surface. *Computers & Graphics*, 37(3), 2013.
- [2] H. Benko and S. K. Feiner. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. In *3DUI*, 2007.
- [3] B. Bollensdorff, U. Hahne, and M. Alexa. The effect of perspective projection in multi-touch 3d interaction. In *Proceedings of GI'12*, 2012.
- [4] A. Cohé, F. Dècle, and M. Hachet. tbox: a 3d transformation widget designed for touch-screens. In *Proceedings of CHI'11*. ACM, 2011.
- [5] M. Hachet, B. Bossavit, A. Cohé, and J.-B. de la Rivière. Toucheo: multitouch and stereo combined in a seamless workspace. In *Proceedings of UIST'11*, UIST '11. ACM, 2011.
- [6] M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *Proceedings of CHI'07*. ACM, 2007.
- [7] M. Hancock, T. ten Cate, and S. Carpendale. Sticky tools: full 6dof force-based interaction for multi-touch tables. In *Proceedings of ITS'09*. ACM, 2009.
- [8] M. Hancock, F. Vernier, D. Wigdor, S. Carpendale, and C. Shen. Rotation and translation mechanisms for tabletop interaction. In *Horizontal Interactive Human-Computer Systems, TableTop 2006*.
- [9] S. Knoedel and M. Hachet. Multi-touch rst in 2d and 3d spaces: Studying the impact of directness on user performance. In *Proceedings of 3DUI'11*. IEEE, 2011.
- [10] A. Martinet, G. Casiez, and L. Grisoni. The effect of dof separation in 3d manipulation tasks with multi-touch displays. In *Proceedings of VRST'10*. ACM, 2010.
- [11] D. Mendes, P. Lopes, and A. Ferreira. Hands-on interactive tabletop lego application. In *Proceedings of ACE'11*. ACM, 2011.
- [12] M. A. Nacenta, P. Baudisch, H. Benko, and A. Wilson. Separability of spatial manipulations in multi-touch interfaces. In *Proceedings of GI'09*, 2009.
- [13] J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2d and 3d direct manipulation. In *Proceedings of UIST'09*. ACM, 2009.
- [14] P. Song, W. B. Goh, W. Hutama, C.-W. Fu, and X. Liu. A handle bar metaphor for virtual object manipulation with mid-air interaction. In *Proceedings of CHI'12*. ACM, 2012.
- [15] S. Strothoff, D. Valkov, and K. Hinrichs. Triangle cursor: interactions with objects above the tabletop. In *Proceedings of ITS'11*. ACM, 2011.
- [16] R. Wang, S. Paris, and J. Popović. 6d hands: markerless hand-tracking for computer aided design. In *Proceedings of UIST'11*. ACM, 2011.
- [17] R. Y. Wang and J. Popović. Real-time hand-tracking with a color glove. In *Proceedings of SIGGRAPH'09*. ACM, 2009.