

# Utilização de Motion Tracking para Navegação e Manipulação de Objectos frente a Ecrã de Larga Escala

José Pedro Dias

Instituto Superior Técnico

Av. Prof. Dr. Cavaco Silva

2744-016 Porto Salvo

jose.pedro.dias@gmail.com jose.goncalves@ist.utl.pt

José Gonçalves

Joaquim Jorge

INESC-ID

R. Alves Redol, 9

1000-029 Lisboa

jaj@inesc.pt

## Resumo

*Navegação em cenas tridimensionais em ambientes de visualização de larga escala constitui um desafio para a realização de interfaces simples, mas expressivas e fáceis de aprender. Este artigo apresenta uma solução baseada em interfaces multimodais, combinando captura de movimento, gestos e fala que permite a utilizadores navegarem na primeira pessoa, com as mãos livres num mundo virtual representado num ecrã multi-projecção de larga escala. Para tal utilizamos câmaras de infravermelhos que seguem marcadores fixados ao corpo dos utilizadores. O sistema utiliza a metáfora do Super-Homem para permitir voo controlado utilizando movimentos dos braços. De modo semelhante é possível manipular e transformar objectos. Este documento descreve os princípios e alguns detalhes da nossa abordagem, descrevendo testes efectuados com utilizadores e uma discussão dos resultados.*

## Palavras-Chave

*captura de movimentos, manipulação directa, navegação, manipulação de objectos, interacção por via de gestos.*

## 1 Introdução

A navegação numa cena tridimensional processa-se normalmente usando um computador de secretária, com rato e teclado como dispositivos de entrada e monitor como dispositivo de saída. Foi desenvolvido no laboratório Prof. Lourenço Fernandes[Araújo 05] um conjunto de modos de interacção que se têm provado eficazes para o efeito, recorrendo a equipamento de captura de movimentos e um ecrã de larga escala para navegar num ambiente virtual e manipular objectos na perspectiva da primeira pessoa.

Este trabalho pretende proporcionar uma interacção mais directa para navegação tridimensional, sem recurso a dispositivos intrusivos. Recorre a comandos de voz para seleccionar a acção a executar e a movimentos dos braços para controlar o desenrolar da mesma.

## 2 Modos de interacção

Segue-se a descrição dos diferentes modos de interacção suportados pela aplicação.

### 2.1 Navegação no espaço

A navegação no espaço faz uso dos braços estendidos (imitando o super-homem) para determinar a direcção e velocidade do voo.

O controlo da velocidade de voo é regulada pela distância entre mãos: quanto mais afastadas estiverem as mãos mais lento é o deslocamento, até a uma amplitude limite – cerca

de 90° – em que o utilizador pára no ar; a velocidade máxima é atingida com as mãos juntas (fig. 1a).

A direcção lateral do voo é afectada elevando um braço mais que o outro, sendo que quanto maior for a diferença de elevação, maior a rotação imprimida. Para rodar para direita deve levantar-se o braço esquerdo e baixar o direito (fig. 1b); vice-versa para rodar para a esquerda.

A alteração da altitude do voo efectua-se elevando ou baixando simultaneamente os braços (fig. 1c).

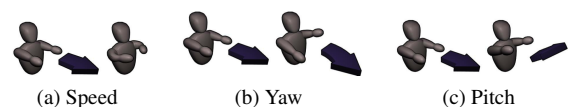


Figura 1: Modo Voar

### 2.2 Manipulação de objectos seleccionados

Para mover um objecto deve estender-se o braço direito. A sua direcção determina a translacção do objecto mapeada no plano horizontal do mundo (fig. 2).

Para rodar um objecto deve estender-se o braço direito em direcção ao cenário e fazer um movimento circular no plano paralelo ao ecrã imediatamente à frente do utilizador (como se estivesse a limpar um vidro). O sentido da rotação do braço é mapeado directamente no sentido da rotação do objecto (fig. 3). O eixo de rotação é previamente escolhido através do comando de voz que inicia o

modo.

Para alterar a escala do objecto seleccionado tem relevância a distância entre mãos. A posição inicial dos braços aquando a selecção do modo escala serve de ponto de partida para a alteração da mesma. Daí em diante, para aumentar a escala de um objecto deve afastar-se as mãos e para o reduzir aproximá-las (fig. 4). A escala é restringida a um eixo previamente escolhido através do comando de voz que inicia o modo.

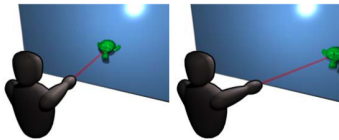


Figura 2: Modo mover

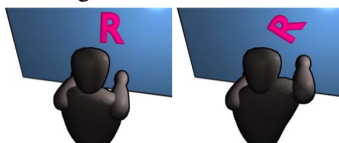


Figura 3: Modo rodar

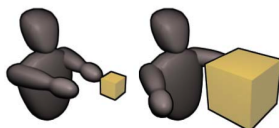


Figura 4: Modo escala

### 3 Configuração Física

A captura de movimentos efectuou-se recorrendo à fixação de marcadores reflectores nos braços do utilizador e à instalação de um conjunto de 4 câmaras com lâmpadas de infra-vermelhos, cuja imagem é processada por um computador que identifica em cada imagem os marcadores visíveis no espaço-imagem, faz a computação da posição espacial referente a cada marcador pelo cruzamento dos dados das 4 câmaras e efectua uma etiquetagem dos marcadores segundo uma posição inicial dos braços.

Tivemos à nossa disposição o Microsoft Speech API 5.1, versão american english. A captura de comandos de voz foi efectuada com um microfone wireless da Microsoft, após a comparação do reconhecimento do mesmo face a um microfone fixo e um Bluetooth.

#### 3.1 Gramática

Foi projectada uma pequena gramática que permitisse a activação dos diferentes modos multimodais:

**Fly:** begin flying , stop flying

**Drag:** begin drag, stop drag

**Rotate:**begin rotate ( X | Y | Z ) , stop rotate

**Scale:** begin scale ( X | Y | Z | Overall ) , stop scale

O reconhecedor mostrou-se algo limitado, tendo sido necessário adaptar a gramática a palavras não ambíguas entre si ou que gerassem falsos positivos.

### 4 Testes com utilizadores

Em Julho de 2007 foram efectuados testes de usabilidade na cidade de Glasgow, na Escócia, no âmbito do programa IMPROVE, em que o ImmiView se insere.

Participaram 10 utilizadores, estudantes de arquitectura e arquitectos. Os mesmos foram distribuídos em grupos de 2, realizando cada um um conjunto de tarefas em separado e uma tarefa colaborativa, estando um deles munido dos marcadores.



Figura 5: utilizador em modo voo

### 5 Avaliação e Discussão

A leitura dos marcadores mostrou-se problemática em indivíduos de baixa estatura. Supomos que a origem esteja numa maior frequência de oclusões.

A necessidade do sistema de captura de movimentos ter início com uma posição predefinida do utilizador e a pouca robustez do mesmo face a oclusões dificulta a presença de outros indivíduos no espaço capturado.

Foi necessário alterar a gramática de voz para desambiguar alguns comandos.

Os utilizadores apreciaram as funcionalidades fornecidas, não se mostrando inibidos na sua aplicação.

Os modos alternativos de navegação e manipulação aqui descritos mostram-se não só viáveis como vantajosos face a interfaces WIMP. A navegação mostrou-se particularmente poderosa, podendo o utilizador controlar de forma contínua diversos parâmetros do voo recorrendo apenas à movimentação dos seus braços para o fazer.

### 6 Conclusões

Neste artigo apresentámos uma interface multimodal simples mas poderosa para suportar as acções mais comuns de manipulação e visualização tridimensional em ecrãs de larga escala, combinando captura de movimentos, fala e gestos numa interface natural.

Testes realizados com utilizadores permitiram verificar a eficácia e virtudes da nossa abordagem mãos-livres que substitui com vantagem técnicas mais convencionais.

### Referências

- [Araújo 05] Bruno Araújo, Tiago Guerreiro, Ricardo Costa, e Joaquim Jorge. Leme wall: Desenvolvendo um sistema de multi-projecção. Em *13o Encontro Português de Computação Gráfica, Universidade Trás-Os-Montes e Alto Douro, Vila Real, Portugal*, 10 2005.