

# LS3D: Pesquisa de Blocos LEGO em Ambientes Imersivos

Pedro B. Pascoal Daniel Mendes Diogo Henriques Isabel Trancoso Alfredo Ferreira  
Departamento Engenharia Informática

INESC-ID / IST / Universidade de Lisboa

pmbp@tecnico.ulisboa.pt, danielmendes@tecnico.ulisboa.pt, diogo.henriques@tecnico.ulisboa.pt,  
isabel.trancoso@inesc-id.pt, alfredo.ferreira@inesc-id.pt

## Resumo

A crescente popularidade dos ambientes imersivos tem aumentado o interesse em interações pessoa-máquina mais naturais. Contudo, no contexto de seleção de objetos e construção de modelos virtuais, o uso de ambientes imersivos tem sido pouco explorado. Neste trabalho apresentamos o LS3D, um protótipo que aborda este desafio, usando fala e gestos como forma de interação, desenvolvido para o cenário de construção com blocos LEGO. Neste âmbito, foram desenvolvidas novas abordagens para explorar resultados de uma pesquisa neste tipo de ambientes, que permitem uma melhor interpretação dos resultados e que oferecem uma interação direta aos utilizadores. Para avaliar a nossa solução, comparámos o LS3D com a aplicação comercial LEGO Digital Designer. Os resultados sugerem que a nossa proposta oferece uma forma mais simples e natural para procurar objetos 3D virtuais, garantindo também melhores desempenhos e percepção dos resultados do que abordagens tradicionais para a recuperação deste tipo de conteúdo. Os participantes afirmaram ainda que a nossa solução oferece uma experiência mais divertida e enriquecedora.

## Palavras-Chave

Interfaces multi-modais, interação natural, fala, gestos, descrição de objetos 3D, testes com utilizadores

## 1. INTRODUÇÃO

Os ambientes virtuais utilizados em aplicações de entretenimento estão a tornar-se vez mais imersivos [Trowbridge 09]. Recorrendo a CAVE [Jacobson 05] e *Heads-Mounted Displays* (por exemplo Oculus Rift<sup>1</sup> ou o recentemente anunciado Project Morpheus da Sony), este tipo de ambientes é cada vez mais comum. Diversas áreas podem tirar partido da capacidade de colocar o utilizador dentro do mundo virtual de uma forma mais imersiva em relação ao que as configurações tradicionais permitem, sendo possível oferecer interações mais naturais com o conteúdo virtual. Um dos exemplos destas aplicações é a criação de cenários virtuais [Cabral 11], que permitam ao utilizador posicionar objetos virtuais na cena, simulando interações do mundo físico.

No entanto, quando confrontados com o desafio de seleção de objetos em coleções, as soluções tradicionais baseiam-se normalmente em listas ou grelhas com vistas 2D dos objetos, o que pode não ser viável dada a mudança no paradigma de interação. Por outro lado, o uso de ambientes imersivos demonstra melhorias na visualização e exploração dos resultados de recuperação relativamente a abordagens tradicionais [Henriques 14, Pascoal 12]. Neste trabalho, focamos em formas de exploração de objetos 3D em grandes coleções através de ambientes imersivos,

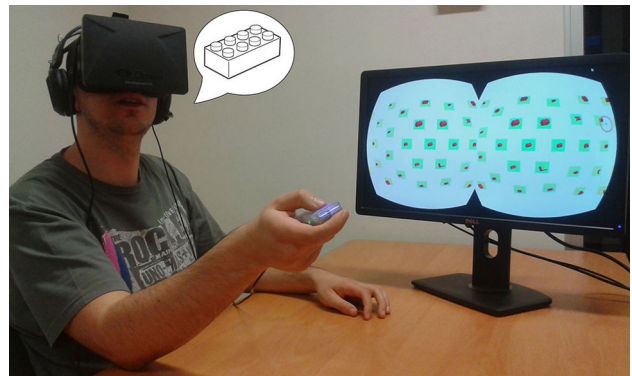


Figura 1. Utilizador a interagir com o protótipo LS3D.

com o propósito de selecionar um modelo específico. Para isso, usámos como cenário blocos de construção LEGO, que já provou ser um bom caso de estudo, que já originou aplicações de entretenimento interessantes [Santos 08, Mendes 11, Miller 12, Gupta 12]. Desenvolvemos um protótipo, o LS3D, que permite aos utilizadores pesquisar, selecionar, explorar e colocar blocos LEGO 3D num ambiente totalmente imersivo através do uso de interações multimodais, como ilustrado na Figura 1.

<sup>1</sup>Oculus Rift: <http://www.oculusvr.com/rift/>

Ao longo do documento iremos discutir o estado da arte relativamente a pesquisa de conteúdo multimédia, com foco especial em objetos 3D virtuais. Procederemos à descrição da solução implementada no nosso protótipo, assim como da avaliação onde o comparamos com uma aplicação comercial. Finalmente, apresentamos as nossas conclusões e indicamos algumas direções para trabalho futuro.

## 2. TRABALHO RELACIONADO

Com o rápido aumento da informação multimédia disponível, a pesquisa e recuperação de um conteúdo específico tem-se apresentado como um desafio cada vez mais importante. As abordagens tradicionais permitem a recuperação recorrendo à utilização de pesquisas textuais. Contudo, este método não é trivial, visto que normalmente conteúdo multimédia não contém informação intrínseca suficiente. De forma geral, os motores de pesquisa fazem uso de anotações, tais como o nome dos ficheiros, legendas, ou mesmo referências associadas, que podem não estar diretamente relacionadas com conteúdo de média. Este conceito também se aplica na recuperação de imagens [Smith 97] e de objetos 3D [Funkhouser 03]. Funkhouser et al., faz uso de sinónimos para aumentar as palavras possíveis para descrever objetos 3D e minimizar a escassez do vocabulário.

Apesar dos trabalhos existentes, as descrições disponíveis para objetos 3D continuam a ser insuficientes, em especial relativamente à geometria dos objetos. Para facilitar, algumas das soluções propostas fazem uso da pesquisa-por-exemplo. O objetivo da pesquisa-por-exemplo é a obtenção de resultados semelhantes em termos do aspecto visual, tal como a cor [Paquet 99] ou forma [Laga 11]. Contudo, este método de interrogação exige que o utilizador tenha *a priori* um objeto semelhante ao procurado, o que geralmente não é o caso.

Uma forma de contornar essa limitação, é usar como interrogação um esboço do objeto pretendido, oferecendo aos utilizadores a possibilidade de pesquisarem objetos semelhantes aos esboços produzidos [Pu 05]. No trabalho de Santos et al. [Santos 08], os utilizadores realizam esboços que correspondem às dimensões de um determinado bloco LEGO desejado. Funkhouser et al. [Funkhouser 03] também apresenta um método de desenho, que permite aos utilizadores desenharem até três vistas 2D diferentes do modelo procurado, como forma de minimizarem o número de falsos positivos. Uma abordagem diferente é apresentada por Liu et al. [Liu 13] que tentou melhorar a pesquisa por esboços usando o perfil do utilizador, tendo em conta os hábitos do utilizador quando este desenha. Isso levou a melhorias nos resultados dos utilizadores que realizavam mais pesquisas.

Holz e Wilson [Holz 11] apresentam uma abordagem diferente, onde a descrição da interrogação é feita através de gestos. Este trabalho consiste na captura dos gestos dos utilizadores, usando a forma e a movimentação das mãos para criar esboços tridimensionais. Os autores concluíram que os participantes conseguiam manter as proporções dos objetos físicos relativamente corretas e que, em zonas com

mais detalhes, realizavam os gestos de forma mais lenta.

No entanto, a apresentação dos resultados nestes trabalhos foca-se principalmente no uso da abordagem tradicional de listas de vistas. Nakazako et al. [Nakazato 01] apresentaram o 3D MARS, que demonstra os benefícios no uso de ambientes imersivos no contexto da recuperação de conteúdo multimédia. O seu trabalho centra-se essencialmente na forma como são apresentados os resultados de pesquisas em sistemas de recuperação de imagens por conteúdo, usando uma CAVE. Estendendo a abordagem do MARS 3D a objetos 3D, Pascoal et al. [Pascoal 12] demonstram que alguns desafios na recuperação de objetos 3D podem ser superados ao apresentar os resultados em ambientes imersivos. Em ambas soluções, os resultados são distribuídos no espaço virtual de acordo com a sua semelhança face à interrogação. Após cada pesquisa, os utilizadores podem explorar os resultados, navegando no ambiente envolvente através do uso de dispositivos como *Head-Mounted Display's*. É também possível o uso diversificado de outros dispositivos de visualização e interação, que foram utilizados para testar vários paradigmas de interação para a recuperação de objetos 3D.

Recentemente, na área da recuperação de informação, temos testemunhado o aparecimento de um vasto leque de soluções com pesquisa-por-fala, especialmente em dispositivos móveis. Isto levou à preferência desta modalidade à tradicional de pesquisa por texto [Kamvar 10]. Na maioria dos casos, na procura por fala, a voz é convertida em texto, o qual é então utilizado como um parâmetro de pesquisa [Wang 11, Lee 10]. Mais recentemente, Lee e Kawahara [Lee 12] realizaram uma análise semântica das pesquisas por fala usadas para procurar livros, de forma a alcançar um maior entendimento sobre o elemento desejado.

Em conclusão, a pesquisa de conteúdos multimédia, nomeadamente de objetos tridimensionais, já foi alvo de estudos anteriores. Contudo, embora a maioria das soluções já comecem a explorar métodos mais naturais de interação, como gestos, ainda não exploram convenientemente todas as potencialidades da interação dos seres humanos e do seu poder descritivo. Em outras áreas, as descrições verbais têm sido utilizadas para recuperação de conteúdo. No entanto, as mesmas ainda não foram aplicadas no contexto dos objetos 3D nem complementadas com outros métodos multimodais. Algumas propostas usam ambientes imersivos, mas em alguns casos os resultados aparecem sobrepostos se forem muito semelhantes. As abordagens tradicionais não sobrepõem resultados, mas a sua visualização com base em vistas não tem a representação adequada dos objetos 3D e não permite interagir com os resultados da mesma forma que os sistemas imersivos permitem.

## 3. LS3D

Como solução para testar e entender quais os métodos mais naturais para a pesquisa de blocos LEGO, foi desenvolvido um protótipo, o LS3D (*LEGO Search combining Speech and Stereoscopic 3D*). Com o nosso protótipo, pretendemos validar algumas técnicas que ofereçam uma forma

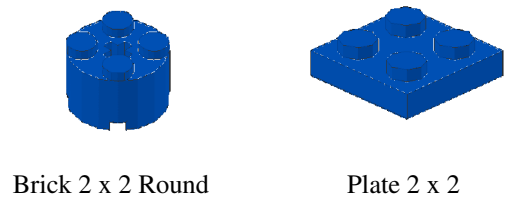
mais natural de pesquisa de objetos 3D, tirando partido de ambientes imersivos e de dispositivos de interação e visualização de última geração. No seu desenvolvimento, tivemos em consideração o estudo realizado por Mendes et al. [Mendes 13], onde é demonstrada a preferência dos utilizadores pela descrição verbal para peças LEGO. Adicionalmente, oferecemos também *feedback* auditivo ao utilizador através da síntese de voz (*text-to-speech*). À semelhança com o trabalho de Mendes et al. [Mendes 13], o nosso protótipo irá representar o papel de fornecedor, onde as pesquisas realizadas são interpretadas pelo sistema, sendo de seguida devolvida uma lista de resultados. Para os resultados retornados, é possível ao utilizador interagir com eles através de gestos ou usando um dispositivo apontador (por exemplo o *Space Point Fusion*). É também possível movimentar a câmara e controlar o ponto de vista do utilizador.

De forma a acelerar o processo de pesquisa das peças, foi necessário criar uma estrutura indexada de toda a coleção de modelos dos blocos de LEGO, que contivesse tanto informação textual como descrições de forma de todos os modelos. Para obter a informação das peças, recorremos ao *LDraw*<sup>2</sup>, uma biblioteca de código aberto que contém descrições textuais e geométricas. Desta biblioteca seleccionámos um subconjunto de 1.810 blocos de LEGO.

### 3.1 Indexação

Na biblioteca *LDraw*, cada bloco de LEGO tem associado descrições textuais, usando termos consistentes ao longo de toda a biblioteca, como ilustrado na Figura 2. Cada anotação é definida pelo designer após a criação do modelo virtual do bloco, e de seguida é validada pelos restantes membros. Através deste processo são garantidas a consistência e a qualidade das descrições textuais. Para cada bloco é também armazenada a sua informação geométrica, como número de linhas, triângulos, quadriláteros e cores. Para termos o acesso fácil e rápida a toda esta informação, criamos índices invertidos.

Considerando as soluções existentes para o texto, aplicámos uma abordagem proposta por Gennaro et al. [Gennaro 10] ao contexto de recuperação de objetos 3D. Como descritor de forma, usámos o algoritmo D2, que em trabalhos anteriores [Funkhouser 03] demonstrou oferecer resultados razoáveis, apesar do reduzido custo computacional associado. Após extraídas as características da forma para cada objeto, é criado um vetor de 128 características. Com base na ideia de que objetos semelhantes terão descritores semelhantes [Gennaro 10], seleccionamos um conjunto de objetos de referência (RO) da coleção. Para cada objeto da coleção é calculada a distância euclidiana entre ele e cada RO. Com base nestas distâncias, os RO são ordenados e, atribuindo um termo único a cada RO, repetimos este termo mais vezes quanto mais próximo o RO está do modelo. Desta forma criamos uma representação textual para a descrição de forma do objeto que pode ser processada de forma semelhante às descrições textuais que



**Figura 2. Exemplos de blocos LEGO e respectivas descrições LDraw.**

acompanham cada bloco, mantendo a relação entre objetos semelhantes.

Foram seleccionados 85 RO da coleção para representar os 1.810 objetos, com base na equação sugerida por Amato e Savino [Amato 08]:  $\#RO \geq 2 \cdot \sqrt{\#X}$ . Desta forma, alcançamos um compromisso aceitável entre tempo de consulta e devolução de resultados. Usando apenas os 30 RO mais próximos do objeto, é ainda possível reduzir significativamente o tamanho dos índices com as descrições. Adicionalmente, para cada RO, os objetos da coleção que estão mais próximos a ele estão associados num índice invertido.

Finalmente, quando na recuperação de objetos, baseamos-nos numa abordagem de TF-IDF para calcular a relevância dos resultados. Desenvolvemos duas modalidades de pesquisa: pesquisa-por-texto e pesquisa-por-exemplo. Na pesquisa-por-texto seguimos uma implementação direta semelhante à de um motor de pesquisa padrão para documentos de texto. Para cada termo da consulta, o sistema recolhe os objetos que contenham esse termo no seu índice invertido, e são combinados todos os resultados, e ordenados os resultados de acordo com a sua semelhança com a interrogação. Na pesquisa-por-exemplo combinamos a descrição textual com os descritores de forma. Uma vez que tanto a interrogação com base nas descrições textuais como usando a informação geométrica geram valores de similaridade contidos em intervalos [0,1], é possível calcular uma média de similaridade de todos os resultados, mantendo o valor final de semelhança dentro desse intervalo. Combinando estas duas abordagens, é possível contrariar alguns falsos positivos e falsos negativos que resultariam da aplicação de apenas uma delas.

### 3.2 Especificação da Interrogação

Como anteriormente indicado, para as pesquisas podem ser realizadas usando pesquisa-por-voz, através da integração de reconhecimento [Meinedo 10] e síntese [Paulo 08] de voz. Usando as expressões mais comuns, indicadas no estudo preliminar de Mendes et al. [Mendes 13], construímos uma gramática usando o formato GRXML. O vocabulário gerado inclui mais de 500 palavras diferentes, devido à inclusão de formas flexionadas das palavras.

Assim sendo, com a nossa gramática é possível utilizadores descreverem um bloco LEGO com mais de que uma

<sup>2</sup>LDraw: <http://www.ldraw.org/>

descrição. Desta forma, é possível descrever os blocos LEGO através das suas dimensões, forma, cor usando um conjunto diverso de adjetivos. Por exemplo, considerando o bloco na Figura 2 (à esquerda), que pertence à classe dos blocos de LEGO redondos, este poderia ser descrito usando 14 adjetivos diferentes (por exemplo: *redondo, circular, cilíndrica, circunferência*). O bloco na Figura 2 (à direita), que pertence à classe de placas, poderia por sua vez ser descrito com 21 adjetivos diferentes (por exemplo:  *fina, baixinha, espalhada, plana*).

No nosso protótipo, cada pesquisa começa com o uso da expressão *Acorda LEGO*, em que o sistema responde verbalmente com um *Sim*. Um exemplo de uma pesquisa pode ser: *Acorda LEGO. Quero uma peça fina, dois por dois*. No nosso trabalho, visto que cada bloco poder ter um vasto conjunto de cores, os utilizadores concentravam-se primeiro na descrição de forma e dimensão dos blocos. Uma vez identificado o bloco desejado, é fácil mudar a sua cor usando um novo comando de voz como *Acorda LEGO. Pinta de encarnado*. Dada a grande quantidade de blocos de LEGO, usamos três estratégias adicionais para restringir o número de peças mostradas. Os utilizadores podem filtrar verbalmente por característica (por exemplo: *Filtra por curva*), ou excluir por tipo (por exemplo: *Exclui DUPLO*) ou obter blocos visualmente semelhantes ao selecionado (por exemplo: *Dá-me semelhante a esta*).

Depois de o utilizador especificar uma pesquisa, o sistema responde imediatamente dependendo do que foi entendido. Por exemplo, na consulta *Acorda LEGO. Quero um bloco dois por dois*, o sistema irá sintetizar *Ilustrando peças dois por dois*. Após isso, o sistema exibe os blocos que correspondem à pesquisa do utilizador.

### 3.3 Exploração de Resultados

Para a exploração dos resultados das pesquisas, criamos uma nova abordagem que combina as visualizações cilíndricas e esféricas, apresentadas no trabalho de Henriques et al. [Henriques 14] como as preferidas dos utilizadores. Na nossa abordagem, mostramos os resultados distribuídos na metade frontal de um barril (Figura 3), situando o utilizador no interior do mesmo e colocando os blocos com mais relevantes na frente do utilizador. Adicionamos um quadrado colorido por trás de cada bloco para

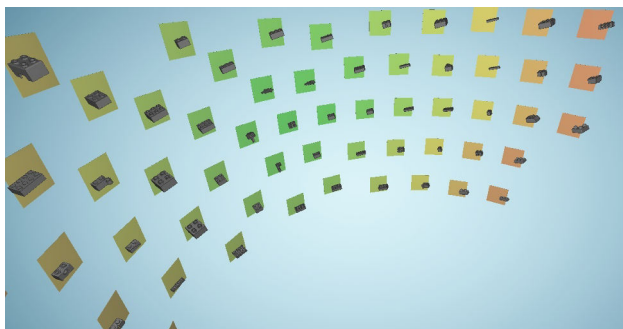


Figura 3. Modo de visualização barril.

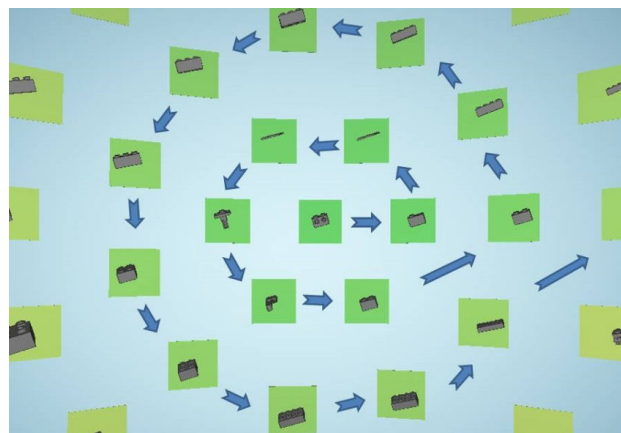


Figura 4. Disposição dos objetos por relevância.

ilustrar sua posição, com cor mais verde para os mais relevantes e vermelho os menos relevantes, conforme ilustrado na Figura 4. Além disso, para proporcionar uma melhor percepção dos objetos, todos os blocos vão girando lentamente. Isso permite aos utilizadores perceberem alguns detalhes adicionais. Quando o utilizador seleciona um objeto, os restantes param de rodar a fim de não chamar atenção desnecessariamente.

De forma semelhante aos modos desenvolvidos por Henriques et al. [Henriques 14], começámos por usar o dispositivo Leap Motion, mas depois de algumas experiências piloto apercebemo-nos que as pessoas muitas vezes perdiam a noção do volume onde poderiam interagir. Deste modo, exploramos diferentes interações, desta utilizando o dispositivo *Space Point Fusion*<sup>3</sup>, que não apresenta esta restrição. Após a integração deste dispositivo, foram comparadas as interações entre os dois dispositivos através de uma breve avaliação de usabilidade. A avaliação foi realizada por oito participantes, onde 50% deles não tinha qualquer experiência no uso de nenhum desses dispositivos, 5% já haviam experimentado usar ambos os dispositivos e os restantes 25% apenas um dos dispositivos. Verificou-se também que não só os utilizadores preferiram usar o *Space Point Fusion*, como este também foi mais rápido para realização dos testes do que o Leap Motion.

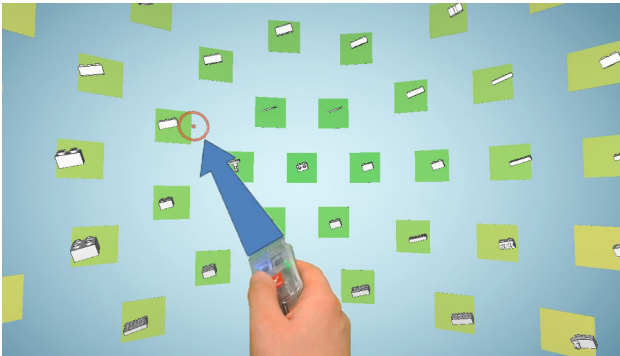
Recorrendo ao giroscópio do *Space Point Fusion*, os utilizadores podem apontar para um bloco específico (Figura 5). Após apontar para um bloco, o botão esquerdo do dispositivo pode ser premido para trazer o objeto para mais próximo do utilizador. Adicionalmente, o utilizador pode também rodar livremente o dispositivo, permitindo que o bloco seleccionado possa ser visto de todos os ângulos, conforme ilustrado na Figura 6.

### 3.4 Contrução do modelo de LEGO

Depois de navegar e encontrar o bloco desejado, o nosso protótipo permite ao utilizador construir um modelo LEGO

<sup>3</sup>PNI Sensor, <http://www.pnicorp.com/products/spacepoint-gaming>





**Figura 5. Selecionar um objeto usando o Space Point Fusion.**

virtual. Depois de selecionado o bloco, o utilizador pode esconder os resultados da pesquisa e alterar para o modo de construção, mantendo o bloco selecionado. Neste modo é exibida uma grelha de LEGO, como representado na Figura 7. O utilizador pode então apontar em qualquer posição para colocar o bloco na localização desejada.

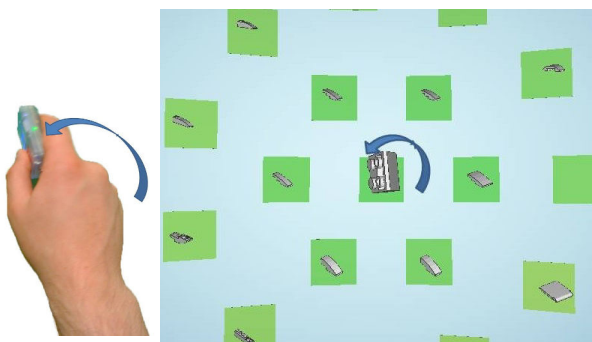
No modo de construção, é também possível manipular os blocos já colocados no modelo. Mantendo um bloco selecionado, o utilizador pode pesquisar por blocos semelhantes ou colori-los, usando as instruções de voz acima mencionados, ou mesmo descartá-lo.

#### 4. AVALIAÇÃO

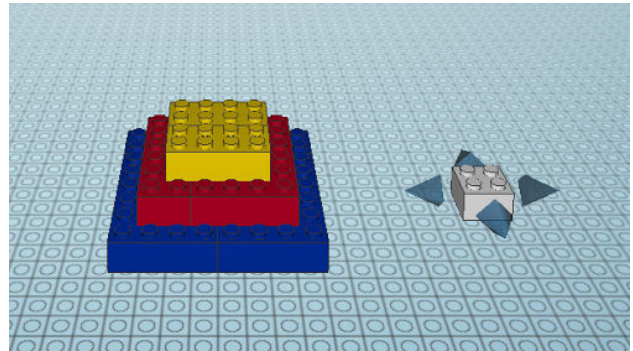
Para validar a nossa solução, comparámos LS3D contra uma aplicação comercial, Lego Digital Designer (LDD), através de um conjunto de testes com utilizadores. O LDD utiliza o paradigma *Windows Icons Menus and Pointing* (WIMP) apresentando uma coleção de blocos LEGO divididos em categorias numa grelha tradicional 2D, e cada objeto é representado por uma vista 2D.

##### 4.1 Participantes

A nossa avaliação foi realizada por 20 participantes (quatro do sexo feminino), com idades entre os 18 e 50 anos (70% entre 24 e 29). Todos os participantes eram falantes nati-



**Figura 6. Utilizador rodando um objeto usando o Space Point Fusion.**



**Figura 7. Construindo um modelo de LEGO virtual, com a peça selecionada.**

vos da língua portuguesa, sendo a maioria deles residentes da região urbana de Lisboa, e nenhum deles tinha anteriormente usado o LDD. No entanto, os participantes já tinham bastante experiência no uso de motores de busca, por exemplo de pesquisa de imagens, sendo usado numa base diária. Apenas 30% dos participantes já estavam familiarizados com motores de pesquisa de objetos 3D, onde em todos eles os objetos eram representados usando vistas 2D. 60% dos participantes já tinham usado sistemas de pesquisa por voz, e também 60% dos participantes já haviam experimentado algum tipo de visualização estereoscópica, contudo apenas 50% tinham anteriormente usado Head-Mounted Display's (HMD).

##### 4.2 Equipamento

As sessões de testes, foram conduzidas num ambiente controlado, sem influências externas. Para executar LDD, foi usado um computador padrão com rato e monitor, como ilustrado na Figura 8 (esquerda). O LS3D foi testado usando o *Space Point Fusion* para interações, os Oculus Rift para a visualização e um conjunto com microfone e auriculares para as interação por voz, como pode ser visto na Figura 8 (direita).

##### 4.3 Metodologia

Os testes com os utilizadores foram estruturados em quatro etapas: questionário inicial para avaliar a experiência e perfil de cada utilizador; resumo dos objectivos dos testes; execução das tarefas nos dois sistemas; e questionário comparativo da usabilidade de ambos os sistemas. Para garantir uma distribuição uniforme dos testes, 50% dos utilizadores começaram por testar o LDD e os outros 50% começaram pelo nosso protótipo.

A execução das tarefas em cada sistema começou com um breve tempo de adaptação, onde os utilizadores experimentaram a interface. Após isso, foram convidados a procurar oito blocos, sendo mostrado exemplos de blocos de LEGO físicos. A ordem de apresentação dos blocos foi aleatoria, de forma a diferentes utilizadores não terem a mesma sequência, a fim de garantir uma distribuição uniforme. Depois de terminada a procura dos oito blocos, os



Figura 8. Participantes numa sessão de avaliação, usando o LDD (esquerda) e o LS3D (direita).

utilizadores foram convidados a responder a um pequeno questionário sobre sua experiência de utilização em relativamente aos dois sistemas.

#### 4.4 Resultados

Realizamos três perspectivas diferentes na análise dos resultados dos nossos testes com utilizadores. Em primeiro lugar, apresentamos uma análise quantitativa sobre as medidas objetivas, ou seja, tempo e erros. Em seguida, apresentamos também a análise qualitativa com base nas respostas do questionário. Finalmente, discutimos várias observações capturadas durante as sessões de teste.

##### 4.4.1 Análise Quantitativa

Durante os testes com utilizadores, registamos o tempo necessário para a realização da pesquisa de cada uma dos oito blocos, bem como o número de blocos errados selecionados. Estes valores estão representados no gráfico da Figura 9. Na análise dos resultados, na relação do tempo gasto na procura dos oito blocos, foi utilizado o teste de *signed-ranks* de Wilcoxon, com o qual podemos concluir

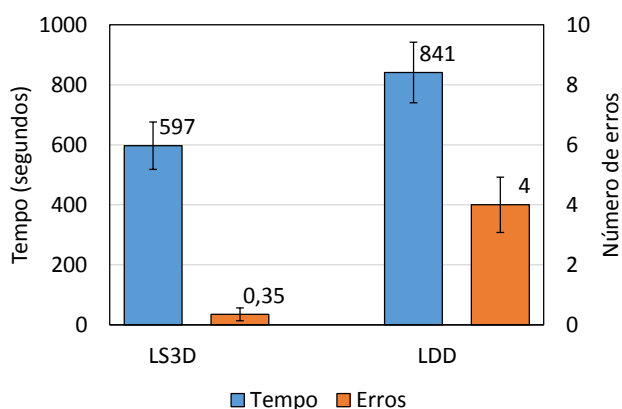


Figura 9. Tempo e número de erros na pesquisa dos oito blocos de LEGO (médias e intervalos de confiança a 95%).

que existem diferenças estatisticamente significativas: os utilizadores realizaram as pesquisas mais rápidas no LS3D do que com o LDD ( $Z = -3,509$ ,  $p = 0,000$ ). Quanto ao número de blocos errados, também foi utilizado o teste de *signed-ranks* de Wilcoxon, onde se conclui que também houveram diferenças estatisticamente significativas, sendo que a solução LS3D voltou a demonstrar melhor desempenho do que o LDD ( $Z = -6,697$ ,  $p = 0,000$ ), originando menos erros.

##### 4.4.2 Análise Qualitativa

Depois de completar as tarefas em ambos os sistemas, os utilizadores foram convidados a classificar as interfaces de pesquisa, usando uma escala de Likert de 4 valores, sobre: quão divertido foi de usar; quão fácil foi de ver os blocos; e quão simples foi de usar o sistema. Os resultados deste questionário são apresentados na Tabela 1. O teste de Wilcoxon foi novamente utilizado para encontrar diferenças estatisticamente significativas. Os utilizadores concordaram que LS3D foi o mais fácil de usar ( $Z = -2,441$ ,  $p = 0,015$ ). No nosso sistema, até mesmo os participantes que nunca usaram para pesquisas por fala, conseguiram concluir as tarefas com sucesso. Relativamente à visualização dos objetos, os utilizadores concordaram que LS3D permitiu melhor interpretação da representação 3D ( $Z = -2,780$ ,  $p = 0,005$ ). Finalmente, também concordaram que LDD foi menos divertido face ao LS3D ( $Z = -3,626$ ,  $p = 0,000$ ).

##### 4.4.3 Observações

Durante as avaliações, notámos alguns aspectos relevantes no comportamento dos utilizadores, bem como alguns comentários. Ao longo da avaliação, vários utilizadores

System	LDD	LS3D
Utilização	2 (1)	3 (1)
Visualização	2 (1)	3 (0)
Divertimento	2 (1)	4 (0)

Tabela 1. Classificações de cada sistema (Média, amplitude inter-quartil).

mencionaram que a divisão por categorias do LDD era bastante útil. No entanto, essa divisão não ajudava em dois dos blocos a procurar, dado que esses blocos não aparentavam pertencer a nenhuma das categorias, e não eram semelhantes aos objetos representados. Isso obrigou esses utilizadores a abrir todas as categorias e navegar nelas. Contudo, este processo facilitou mais tarde a procura pelo blocos seguintes. Utilizadores que não começaram por esses dois blocos fizeram comentários que refletem frustração, enquanto procuravam no LDD, algo que não ocorreu no LS3D, mesmo em pesquisas mais longas.

No LDD, os utilizadores mostraram ter problemas em entender os blocos representados, o que não aconteceu no LS3D. Comentários sugeriram que o nosso protótipo oferece uma melhor percepção dos objetos 3D, do seu tamanho e detalhes, principalmente em graças à rotação dos blocos listados. Isso foi particularmente visível em blocos sem pinos, em que os utilizadores muitas vezes selecionavam um bloco com um tamanho diferente quando usando o LDD.

Além disso, alguns utilizadores comentaram que era mais fácil usar o LS3D para procurar objetos pela sua dimensão, apesar de alguns mencionarem dimensões erradas durante a pesquisa. Estes erros eram notados e corrigidas após serem mostrados os resultados da pesquisa e o sistema responder o que havia sido entendido. Notámos também que os utilizadores recorreram mais a metáforas e adjetivos nas últimas pesquisas. Isso deveu-se ao facto de não terem inicialmente o conhecimento de todas as palavras que poderiam ser usadas e apenas no final já tinham ideia do potencial do nosso protótipo.

Observamos que as funcionalidades disponibilizadas no LS3D ajudaram bastante os utilizadores a encontrar os objetos pretendidos. Além disso, foi também mencionado que o refinamento da exclusão foi o menos necessário, porque as outras possibilidades de refinamento, exemplo e filtragem, demonstraram ser suficientes para cumprimento das tarefas. Isso foi também observado na distribuição dos refinamentos, onde 55% da estes eram para filtrar, 36% para exemplos e apenas 9% para exclusões.

## 5. CONCLUSÕES E TRABALHO FUTURO

Com os ambientes virtuais imersivos sendo cada vez mais comuns, permitindo que utilizadores se coloquem de uma forma mais credível dentro de um mundo virtual, novos desafios têm surgido. Um exemplo disso é o seu uso na exploração de objetos de uma coleção dentro de um ambiente imersivo, ou de aplicações para a montagem de modelos virtuais. Com esta motivação, desenvolvemos um protótipo que permite a construção de modelos LEGO virtuais, a fim de melhor compreender os métodos mais naturais e simples para a recuperação de objetos em tais sistemas.

Com base no conhecimento obtido, desenvolvemos o nosso protótipo LS3D usando consultas faladas para recuperar os blocos de LEGO virtuais. Para ver a coleção de blocos, criámos um novo modo de visualização em barril,

no qual o utilizador é cercado por objetos virtuais.

Realizámos uma avaliação com 20 utilizadores, comparando o nosso protótipo com uma aplicação comercial da marca LEGO. Os resultados sugerem que o nosso protótipo é capaz de superar a aplicação tradicional, que usa uma grelha 2D com miniaturas, apesar de esta ser uma interface mais familiar aos utilizadores. Através dos resultados obtidos, podémos concluir que a nossa solução ofereceu uma melhor percepção dos objetos, sendo menos sujeita a erros do que a aplicação comercial. Além disso, através de comentários dos utilizadores e análise dos questionários, podémos concluir que a maioria dos participantes achou fácil usar a nossa solução de pesquisa, mesmo sem um período de treino extensivo.

Como trabalho futuro, consideramos que seria interessante maior investigação para o uso de modelos de linguagem mais sofisticados. Para mostrar resultados de consultas, melhorias poderiam ser feitas usando agrupamentos de objetos, como disponível no LDD, mas com mais de um objeto ilustrando cada grupo. Como mencionado, o nosso protótipo oferece uma melhor experiência ao utilizador, mas acreditamos que usando a nova versão dos Oculus Rift que possui seguimento da posição do utilizador, se conseguirá alcançar uma maior imersão do utilizador.

Embora nós exploremos blocos de LEGO, acreditamos que a nossa solução pode ser aplicada a outros cenários que utilizam modelos virtuais, como por exemplo no contexto de montagem nas indústrias automóvel e de construção civil.

## 6. AGRADECIMENTOS

O trabalho apresentado neste artigo foi parcialmente financiado pela Fundação para a Ciência e a Tecnologia (FCT) através do projeto TECTON-3D (PTDC/EEI-SII/3154/2012), Pest-OE/EEI/LA0021/2013 e através da bolsa SFRH/BD/91372/2012.

## Referências

- [Amato 08] Giuseppe Amato e Pasquale Savino. Approximate similarity search in metric spaces using inverted files. Em *Proceedings of the International Conference on Scalable Information Systems*, 2008.
- [Cabral 11] M. Cabral, P. Vangorp, G. Chaurasia, E. Chapoulie, Martin Hachet, e G. Drettakis. A multimode immersive conceptual design system for architectural modeling and lighting. Em *3D User Interfaces (3DUI), 2011 IEEE Symposium on*, páginas 15–18, March 2011.
- [Funkhouser 03] Thomas Funkhouser, Patrick Min, Michael Kazhdan, Joyce Chen, Alex Halderman, David Dobkin, e David Jacobs. A search engine for 3D models. volume 22, páginas 83–105, Janeiro 2003.
- [Gennaro 10] Claudio Gennaro, Giuseppe Amato, Paolo Bolettieri, e Pasquale Savino. An approach to content-based image retrieval based on the lucene search engine library. Em *Research and Advanced Technology for Digital Libraries*, páginas 55–66. Springer, 2010.

- [Gupta 12] Ankit Gupta, Dieter Fox, Brian Curless, e Michael Cohen. Duplotrack: a real-time system for authoring and guiding duplo block assembly. Em *Proceedings of the 25th annual ACM symposium on User interface software and technology*, páginas 389–402. ACM, 2012.
- [Henriques 14] D. Henriques, D. Mendes, P. Pascoal, I. Trancoso, e A. Ferreira. Poster: Evaluation of immersive visualization techniques for 3d object retrieval. Em *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, páginas 145–146, March 2014.
- [Holz 11] Christian Holz e Andrew Wilson. Data miming: inferring spatial object descriptions from human gesture. páginas 811–820, 2011.
- [Jacobson 05] Jeffrey Jacobson, Marc Le Renard, Jean-Luc Lugin, e Marc Cavazza. The caveat system: Immersive entertainment based on a game engine. Em *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACE '05, páginas 184–187, New York, NY, USA, 2005. ACM.
- [Kamvar 10] Maryam Kamvar e Doug Beeferman. Say what? why users choose to speak their web queries. Em *Interspeech*, 2010.
- [Laga 11] H Laga, T Schreck, A Ferreira, A Godil, I Pratikakis, e R Veltkamp. Bag of words and local spectral descriptor for 3d partial shape retrieval. 2011.
- [Lee 10] C Lee, Alexander Rudnicky, e GG Lee. Let's buy books: finding ebooks using voice search. páginas 85–90, 2010.
- [Lee 12] Cheongjae Lee e Tatsuya Kawahara. Hybrid vector space model for flexible voice search. Em *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, páginas 1–4. IEEE, 2012.
- [Liu 13] Yong-Jin Liu, Xi Luo, Ajay Joneja, Cui-Xia Ma, Xiao-Lan Fu, e Dawei Song. User-Adaptive Sketch-Based 3-D CAD Model Retrieval. páginas 1–13, 2013.
- [Meinedo 10] Hugo Meinedo, Alberto Abad, T Pellegrini, I Trancoso, e J Neto. The L2F Broadcast News Speech Recognition System. páginas 93–96, 2010.
- [Mendes 11] Daniel Mendes, Pedro Lopes, e Alfredo Ferreira. Hands-on interactive tabletop LEGO application. Em *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology - ACE '11*, Novembro 2011.
- [Mendes 13] D. Mendes, D. Henriques, I. Trancoso, e A. Ferreira. Qual peça? identificação de métodos naturais para a descrição de objectos 3d. Em *INTERAÇÃO 2013*, 2013.
- [Miller 12] Andrew Miller, Brandyn White, Emiko Charbonneau, Zach Kanzler, e Joseph J LaViola. Interactive 3d model acquisition and tracking of building block structures. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4):651–659, 2012.
- [Nakazato 01] M. Nakazato e T.S. Huang. 3D MARS: immersive virtual reality for content-based image retrieval. páginas 44–47. Ieee, 2001.
- [Paquet 99] Eric Paquet e Marc Rioux. Nefertiti: a query by content system for three-dimensional model and image databases management. volume 17, páginas 157–166, 1999.
- [Pascoal 12] Pedro Pascoal, Alfredo Ferreira, e Joaquim Jorge. Towards an immersive interface for 3d object retrieval. Em *Proceedings of the 5th Eurographics conference on 3D Object Retrieval*, EG 3DOR'12, páginas 51–54, Aire-la-Ville, Switzerland, Switzerland, 2012. Eurographics Association.
- [Paulo 08] Sérgio Paulo, Luís C Oliveira, Carlos Mendes, Luís Figueira, Renato Cassaca, Céu Viana, e Helena Moniz. Dixi—a generic text-to-speech system for european portuguese. Em *Computational Processing of the Portuguese Language*, páginas 91–100. Springer, 2008.
- [Pu 05] Jiantao Pu, Kuiyang Lou, e Karthik Ramani. A 2 d sketch-based user interface for 3 d cad model retrieval. *Computer-Aided Design and Applications*, 2(6):717–725, 2005.
- [Santos 08] Tiago Santos, Alfredo Ferreira, Filipe Dias, e Manuel J. Fonseca. Using sketches and retrieval to create lego models. 2008.
- [Smith 97] JR Smith e SF Chang. Visually searching the web for content. 1997.
- [Trowbridge 09] S. Trowbridge e C. Stapleton. Melting the boundaries between fantasy and reality. *Computer*, 42(7):57–62, July 2009.
- [Wang 11] Robert Wang, Sylvain Paris, e Jovan Popović. 6d hands: markerless hand-tracking for computer aided design. Em *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, páginas 549–558, New York, NY, USA, 2011.