

# Validação da biblioteca KinectProject através da implementação de uma interface natural ao Projeto da Trilha do Morro da Cruz

Sergio R. Gorniski Júnior, Ana Elisa F. Schmidt

Ciência da Computação - Universidade do Vale do Itajaí (UNIVALI) – Itajaí, SC –  
Brazil

sergiogorniski@univali.br, ana.schmidt@univali.br

**Abstract.** *This article presents a validation framework on KinectProject's library functionalities, which abstracts the Microsoft Kinect SDK and allows the integration between several applications that use natural interfaces. As a case of study, we proposed and implemented a natural interface for the Trail of Morro da Cruz's Project, a virtual trail that allows navigation through the environmental information related to the real Morro da Cruz trail, in Itajaí, SC. In addition, this article presents a usability evaluation for the natural interface proposed comparing two natural command sets.*

**Resumo.** *Este artigo apresenta a validação realizada sobre as funcionalidades da biblioteca KinectProject, a qual abstrai o SDK do Microsoft Kinect e permite a integração entre diversos aplicativos que utilizam interface natural. Como estudo de caso, foi proposta e desenvolvida uma interface natural para o Projeto da Trilha do Morro da Cruz, uma trilha virtual que possibilita acesso e visualização de informações ambientais relacionadas à trilha real do Morro da Cruz, em Itajaí, SC. Além disto, apresenta avaliações de usabilidade da interface natural proposta comparando dois conjuntos de comandos naturais.*

## 1. Introdução

O trabalho aqui apresentado faz uso de conceitos e ferramentas já desenvolvidas em dois outros trabalhos realizados por acadêmicos da UNIVALI, sendo estes o KinectProject desenvolvido pelo Bacharel em Ciência da Computação, Gabriel Schade Cardoso e o Projeto Trilha do Morro da Cruz desenvolvido pelo acadêmico em Ciência da Computação, Eduardo A. da Silva.

O KinectProject é uma biblioteca de funções que permite aplicações utilizarem as funções do Microsoft Kinect. A biblioteca captura e reconhece as poses e comandos de voz do usuário obtidos através do Kinect em tempo real e os compara com arquivos de entradas válidas previamente gravados [Schade, 2012].

O KinectProject é extensível para diversos tipos de aplicações através do paradigma de *publisher-subscriber*, onde o *publisher* é disponibilizado pela KinectProject, sendo necessário implementar o *subscriber*, que é o modulo da aplicação que recebe e interpreta os eventos dos comandos naturais reconhecidos e despachados pela biblioteca. A biblioteca já conta com uma implementação *publisher-subscriber* para integração ao Unity 3D [UNITY, 2013].

O projeto Trilha do Morro da Cruz implementa uma ferramenta de navegação virtual para a referida trilha [Silva, 2012]. Este projeto utilizou técnicas da área de Computação Gráfica e de Jogos para implementar a modelagem e navegação no ambiente virtual proposto, utilizando a *engine* Unity 3D como ferramenta de desenvolvimento [UNITY, 2013].

O projeto da Trilha consiste em uma modelagem 3D da Trilha do Morro da Cruz, situada em Itajaí, para a criação de um ambiente virtual, conforme mostra a figura 1. Semelhante a um jogo eletrônico, a interação com a trilha virtual é dada através da utilização de mouse e teclado, sendo o teclado para locomoção e o mouse para manipulação da câmera.



**Figura 1 – Passeando pelo Projeto da Trilha do Morro da Cruz**

Durante o desenvolvimento da biblioteca KinectProject não houve uma avaliação mais detalhada de suas funcionalidades em uma aplicação real, por este motivo não é possível identificar todas as suas limitações, características de desempenho e de usabilidade.

Levando-se em consideração que a falta de afinidade com tecnologias computacionais nas comunidades as quais o Projeto Trilha do Morro da Cruz será implantado, a forma de interação e a interface atual do projeto da trilha podem tornar um tanto restritivas e frustrantes a experiência de interação com a trilha virtual.

Sendo assim, torna-se válido utilizar e validar a biblioteca KinectProject através da implementação de uma interface natural para uma aplicação real. Para tal avaliação foi escolhido o Projeto da Trilha do Morro da Cruz como caso de estudo devido a necessidade de aprimorar a forma de interação atual do mesmo.

Para possibilitar a avaliação e a validação do uso da biblioteca KinectProject no Projeto da Trilha do Morro da Cruz, este trabalho apresenta métricas voltadas a análise da facilidade do uso por parte dos desenvolvedores; a aplicabilidade em aplicações não relacionadas a jogos; a análises de performance no reconhecimento de poses e comandos de voz e avaliação de usabilidade na interação com a interface natural por parte dos usuários finais.

Através dos resultados obtidos nas métricas citadas acima pretende-se realizar a avaliação e validação da biblioteca KinectProject, agregando conhecimento para os futuros desenvolvedores que venham a utilizá-la.

## **2. KinectProject**

Com o lançamento do Kinect SDK (*Software Development Kit*) é possível integrar aplicações com os recursos oferecidos pelo dispositivo Kinect, desde que a aplicação tenha suporte ao .NET Framework 4 [Microsoft, 2013].

O KinectProject é uma biblioteca de funções que abstrai algumas das funcionalidades do Kinect SDK e acrescenta outras novas, permitindo utilizar diversas funções do dispositivo. Dentro das principais funções oferecidas pelo KinectProject, está a obtenção do esqueleto do usuário, reconhecimento de voz, reconhecimento de poses e reconhecimento de gestos, sendo estas duas últimas extensões existentes somente na KinectProject.

A biblioteca foi desenvolvida na estrutura filosofia *publisher-subscriber*, sendo a KinectScan DLL responsável por fazer a comunicação com o Kinect e o reconhecimento dos comando, despachando um evento a ser tratado na aplicação final (*publisher*). Desta forma, com a implementação da unidade consumidora de eventos (*subscriber*), as aplicações podem utilizar as funções do Kinect acrescentadas das novas funcionalidade de reconhecimento de comandos da KinectProject.

Para aplicações que não possuem suporte ao .NET Framework 4, como o caso da *engine* Unity 3D, a biblioteca traz também uma aplicação chamada KinectServer, que é utilizado como ponte entre a aplicação final e a DLL KinectScan, onde a comunicação é feita através do protocolo TCP via requisições de *socket*, conforme mostra a figura 2.

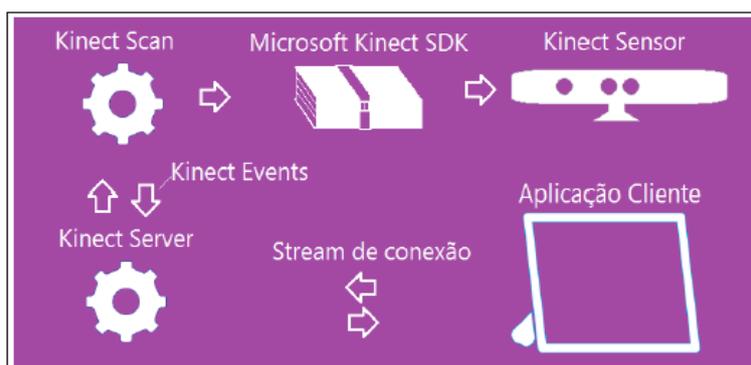


Figura 2. Estrutura do KinectServer - Fonte: SCHADE, Gabriel (2012).

O reconhecimento de poses é o principal diferencial da biblioteca. Através da aplicação InputRegisterApp, responsável pela a criação dos comandos de voz e poses, são escolhidas três juntas do esqueleto reconhecido do usuários para formar uma subpose de uma pose gravada. Através do cálculo da lei dos cossenos e do produto escalar, é efetuada a comparação destes dados com o esqueleto do usuário em tempo real; caso o resultado esteja dentro da margem de erro definido pelo usuário, a pose é reconhecida.

O KinectProject já oferece uma implementação de um *subscriber* consumidor de eventos para Unity 3D – o KinectUnity -, que foi utilizado para realizar a validação da biblioteca, através da implementação de uma interface natural ao Projeto da Trilha do Morro da Cruz.

### 3. Projeto Trilha do Morro da Cruz

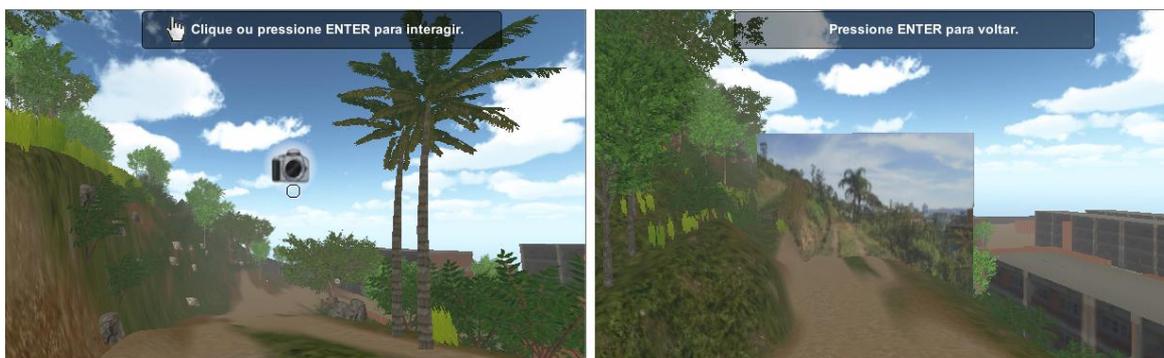
O Morro da Cruz, situado na cidade de Itajaí, é uma área de grande atrativo turístico na região, evidenciando a necessidade de sua preservação; contudo,. atualmente o morro encontra-se em estado de degradação da biodiversidade. Sendo assim, em 2009, a partir do projeto Restauração da Biodiversidade: Análise do processo de recuperação de uma

área do Morro da Cruz, iniciou-se um processo de recuperação. A Figura 3 mostra o início da trilha do Morro da Cruz acessível pelo campus da UNIVALI.



**Figura 3 - Trilha do Morro da Cruz em Itajaí**

O Projeto da Trilha Virtual consiste em um ambiente virtual 3D, no qual o usuário, através do mouse e teclado, locomove um personagem na referida trilha e interage com indicadores como para visualizar informações textuais e/ou visuais dos projetos ambientais desenvolvidos, conforme mostra a figura 4.



**Figura 4 - Interação Projeto da Trilha**

Este projeto da trilha foi escolhido como estudo de caso pois foi desenvolvido na *engine* Unity, que é a aplicação que será utilizada para efetuar a validação da biblioteca KinectProject, e também devido a possibilidade de modificar a forma de interação do projeto da trilha, tornando-a mais intuitiva através de uma interface natural baseada em comandos de voz e poses.

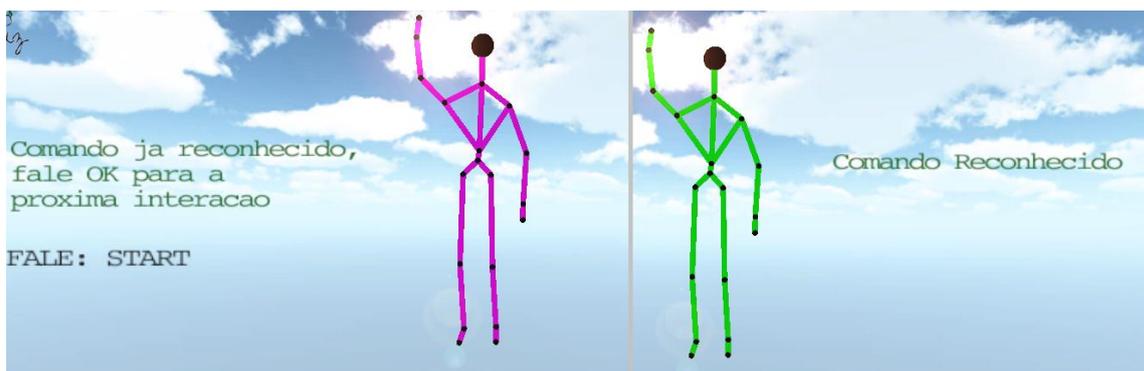
#### **4. Validação da KinectProject**

Para a validação da biblioteca KinectProject, foi proposta e desenvolvida uma interface natural para o Projeto da Trilha do Morro da Cruz; foram desenvolvidos e aplicados questionários com desenvolvedores para obter métricas de facilidade de uso da API do KinectProject, bem como questionários aplicados aos usuários finais da interface natural para obter métricas de usabilidade.

Para a implementação da interface natural no Projeto da Trilha do Morro da Cruz, algumas funcionalidades foram adicionadas à trilha virtual. Dentro das funcionalidades está um tutorial interativo para aprendizagem dos comandos naturais e

um menu de configuração, no qual é possível editar quais comandos naturais utilizar em cada interação.

O tutorial sobre a interface natural, acessível através do menu, é um recurso adicional para os usuários iniciais entenderem como interagir com o ambiente. Através deste, é informado aos usuários todas as poses, gestos ou comandos de voz disponíveis para cada tipo de interação com a trilha virtual. A figura 5 demonstra um usuário interagindo com o tutorial.



**Figura 5. Tela de tutorial**

A interface natural utiliza-se somente alguns recursos da biblioteca KinectProject, sendo eles: reconhecimento de poses; reconhecimento de gestos; e reconhecimento de voz. A obtenção da imagem RGB, das posições das juntas do esqueleto do usuário e a alteração do ângulo de visão do Kinect, foram somente testados *ad-hoc*, e não são utilizados pela interface da trilha virtual.

## 5. Resultados

### 5.1. Avaliação da funcionalidade dos recursos da KinectProject

A partir da implementação da interface natural, foi possível identificar o funcionamento correto e adequado de alguns dos recursos da biblioteca KinectProject:

- ✓ Os recursos de reconhecimento de gestos e de voz, funcionaram corretamente;
- ✓ Através do aplicativo InputRegisterApp foi possível criar e editar poses e comandos de voz corretamente;
- ✓ O reconhecimento de poses funcionou corretamente em situações específicas: onde o teste foi realizado com somente duas poses. A partir do momento que mais poses foram adicionadas, as mesmas deixaram de ser reconhecidas, evidenciando uma falha neste recurso.

Durante o processo de avaliação da funcionalidade dos recursos foram encontradas falhas no código que impediam a utilização destas, como por exemplo obter o esqueleto do usuário em tempo real. Nem todas as falhas foram corrigidas, pois não faziam parte do escopo original do projeto.

#### 5.1.1 Melhorias implementadas

Melhoria 1: Log do KinectServer não identifica qual o nome do comando reconhecido

Descrição: O log do KinectServer identifica corretamente o horário e o tipo do evento reconhecido pelo KinectScan, mas não apresenta o nome do comando.

Correção: Foi alterado o código para inserir no registro do log, o nome das poses ou comandos de voz reconhecidos, conforme mostra a figura 6.

```
The server sends to client a Message Response event at 23:49:41:391
Server receive a Voice Command from Kinect Scan Name: Start at 23:49:49:580
The server sends to client a Voice Command event at 23:49:49:627
```

**Figura 6. Log do KinectServer onde nome do evento é apresentado**

### Melhoria 2: Reconhecimento de poses simultâneas.

Descrição: O reconhecimento de poses só permite enviar um evento de pose reconhecida por vez, sendo necessário que nenhuma pose esteja sendo reconhecida para que se reconheça uma nova. Por exemplo, enquanto o usuário realizar a pose com o pé a frente, outras poses como um braço esticado não é reconhecida, somente reconhecerá esta última se o usuário retirar o pé a frente. O reconhecimento simultâneo necessário para cumprir um requisito de usabilidade interpretando assim vários comandos ao mesmo tempo.

Correção: Foi alterado o código para que todas as poses continuem sendo verificadas, independente se já existe uma com o status de reconhecido.

### Melhoria 3: Verificar se as juntas do esqueleto da subpose que irá ser testada está sendo reconhecida.

Descrição: Os cálculos para reconhecimento de poses são efetuados mesmo que alguma junta da subpose a ser testada não esteja sendo reconhecida.

Correção: Foi alterado o código para que verificasse se todas as juntas das subposes do esqueleto do usuário, que serão comparadas com os arquivos de poses, estão sendo reconhecidos.

## **5.2 Facilidade uso da API do KinectProject**

Para a facilidade de uso da API da biblioteca, foi realizado um experimento com onze alunos de Ciência da Computação da UNIVALI de Itajaí, no qual os participantes deveriam implementar funções de reconhecimento de pose, voz e gestos em um minijogo e responder um questionário de avaliação logo após.

Ao início do teste, foi realizada uma breve apresentação sobre o Kinect, KinectProject e o Unity, assim como uma demonstração do que os participantes deveriam alcançar com suas implementações. Foi disponibilizado também a documentação da biblioteca e juntamente um guia com exemplos da utilização da biblioteca com o Unity.

Os resultados obtidos podem ser verificados na tabela 1. Os critérios utilizados para avaliação foram uma nota em relação a facilidade obtida para a implementação das funções e uma nota da documentação da biblioteca. Os resultados foram divididos pelo nível de experiência ou conhecimento em programação, no qual os próprios usuários identificam seu nível de experiência.

**Tabela 1. Resultados da avaliação da API do KinectProject**

Nível de Experiência		Facilidade	Nota da documentação
Básico	Nota	6,75	Péssima
	Desvio Padrão	2,872	0
Intermediário	Nota	7,20	Entre péssima e regular
	Desvio Padrão	1,483	0,57
Avançado	Nota	8,00	Não utilizaram
	Desvio Padrão	0	Não utilizaram
Total	Nota	7,18	Entre péssima e regular
	Desvio Padrão	1,888	0,50

Comparando os resultados de facilidade de uso da API da KinectProject obtidos entre os níveis de conhecimento em programação, percebe-se que não há muita dificuldade para a implementação das funcionalidades com um conhecimento básico em programação. Com apenas 3 comandos é possível iniciar um evento de reconhecimento, obter os eventos da fila e verificar se determinado comando foi reconhecido.

### 5.3 Usabilidade

Para avaliação da usabilidade da interface natural proposta, foram criados dois conjuntos de comandos naturais, um denominado padrão, no qual utiliza-se da posição do pé do usuário para se locomover, como pé a frente para locomover a frente, e o conjunto alternativo, que utiliza-se dos braços, sendo o braço levemente a frente para locomover a frente. A partir destes dois conjuntos foi solicitado que participantes utilizassem um dos conjuntos e preenchessem um questionário respondendo com uma nota de zero a dez para a facilidade na utilização e fluência na utilização dos comandos naturais, para a aprendizagem dos comandos e se houve alguma dificuldade na utilização.

Foram obtidos retornos de quatorze participantes e os resultados obtidos foram classificados quanto ao nível de conhecimento dos usuários em jogos e na interação com o dispositivo Kinect, no qual se auto definiam. Os usuários com mais familiaridade com o Kinect se adaptam a interação natural mais rapidamente por conhecerem as limitações do dispositivo, seja por questões de espaço ou da necessidade de precisão do comando natural a ser realizado.

Pelas anotações dos questionários e por observações durante os testes notou-se que participantes que utilizaram conjunto alternativo de comandos, tiveram mais dificuldade e desconforto na utilização da interface natural. Isso ocorre pelo fato da locomoção do personagem na trilha se tornar cansativa com a utilização dos braços em uma posição exaustiva.

**Tabela 2. Resultados conjunto padrão**

Conhecimento		Aprendizagem	Facilidade	Fluência	Dificuldade
<b>Básico</b>	Nota	5,2500	5,7500	5,2500	Não se aplica

	Qtd	4	4	4	4
	Desvio Padr.	2,06155	1,89297	1,70783	0
<b>Intermed</b>	Nota	5,5000	7,0000	6,5000	Não se aplica
	Qtd	2	2	2	2
	Desvio Padr.	3,53553	0,00000	0,70711	0
<b>Avançado</b>	Nota	7,0000	7,0000	7,0000	Não se aplica
	Qtd	1	1	1	1
	Desvio Padr.	0	0	0	0
<b>Total</b>	Nota	5,5714	6,2857	5,8571	Não se aplica
	Qtd	7	7	7	7
	Desvio Padr.	2,14920	1,49603	1,46385	0

Tabela 3. Resultados conjunto alternativo

Conhecimento		Aprendizagem	Facilidade	Fluência	Dificuldade
<b>Básico</b>	Nota	5,6667	5,0000	4,3333	
	Qtd	3	3	3	3
	Desvio Padr.	0,57735	1,00000	0,57735	
<b>Intermed</b>	Nota	6,0000	5,0000	5,5000	
	Qtd	2	2	2	2
	Desvio Padr.	1,41421	0,00000	0,70711	
<b>Avançado</b>	Nota	5,0000	6,5000	6,5000	
	Qtd	2	2	2	2
	Desvio Padr.	0,00000	0,70711	0,70711	
<b>Total</b>	Nota	5,5714	5,4286	5,2857	
	Qtd	7	7	7	7
	Desvio Padr.	0,78680	0,97590	1,11270	

## 6. Conclusão

A partir dos dados exibidos na seção 5.3 sobre a usabilidade da interface, é possível perceber que o nível de conhecimento em computação, jogos ou no dispositivo Kinect,

influencia tanto na facilidade do uso, como na aprendizagem dos comandos naturais. Os usuários com mais conhecimento nas áreas citadas acima se adaptam mais rapidamente por conhecerem as limitações do dispositivo, seja por questões de espaço ou da necessidade de precisão do comando natural a ser realizado. O que não seria o ideal para uma interface natural, na qual esta deveria se adaptar ao usuário e não o usuário a ela.

A partir de observações de alguns participantes durante os testes, percebeu-se grande dificuldade na memorização dos comandos naturais e na realização destes comandos. Pelo fato da interface natural ter sido desenvolvida para suprir todos os comandos não naturais, uma grande diversidade de poses teve de ser criada, dificultando para a memorização do usuário. Além disto, surgiu a necessidade de restringir a margem de erro do reconhecimento das poses, para que as poses não conflitassem entre si, o que tornou difícil a execução dos comandos por usuário em um primeiro momento.

Outro empecilho para o reconhecimento é a falta de *feedback* da imagem do usuário no momento da interação com a trilha. O usuário não consegue saber se seu corpo está dentro da área de reconhecimento do Kinect, e nem comparar o esqueleto sendo reconhecido com a pose a ser efetuada. O *feedback* se torna indispensável, inclusive para os comandos de voz para que o usuário entenda que comando recém falado não foi reconhecido.

Pelas notas do questionário e por observações notou-se que participantes que utilizaram conjunto alternativo de comandos, tiveram mais dificuldade e desconforto na utilização da interface natural. Isso ocorre pelo fato da locomoção do personagem na trilha se tornar cansativa com a utilização dos braços em uma posição exaustiva.

Após a realização da avaliação da trilha com usuários, foi possível perceber que o cenário ideal para uma trilha virtual, seria a utilização da quantidade de comandos naturais mais reduzida. Desta forma poderia ser aumentada a margem de erro utilizada no reconhecimento da pose para facilitar o reconhecimento e dispensar a necessidade de um tutorial para memorização.

Outros tipos de tutoriais mais interativos também são possíveis, como por exemplo, a cada nova interação, uma pequena janela se abre informando como o usuário pode proceder, demonstrando a pose, comando de voz ou gesto a ser realizado.

Comparando os resultados apresentados na seção 5.1 sobre a facilidade de uso da biblioteca KinectProject obtidos entre os níveis de conhecimento em programação, percebe-se que não há muita dificuldade para a implementação das funcionalidades com um conhecimento básico em programação. Com apenas 3 comandos é possível iniciar um evento de reconhecimento, obter os eventos da fila e verificar se determinado comando foi reconhecido. Outra informação importante é que 63% dos participantes dos questionários não necessitaram recorrer à documentação, evidenciando uma boa arquitetura e estrutura da biblioteca.

O que também dificulta a aprendizagem na utilização da biblioteca é a falta de uma documentação mais completa, com exemplos e informações. A documentação é indispensável, para que os desenvolvedores tenham ao que recorrer quando necessitarem de apoio.

Uma contextualização mais profunda antes da realização do questionário com informações sobre a biblioteca, sua estrutura, qual a sua finalidade e o sobre o Unity,

ambiente no qual foram realizados os testes, poderia facilitar o entendimento dos participantes e assim consequentemente aumentar as notas de facilidade.

A biblioteca KinectProject possui várias funcionalidades úteis para a implementação de uma interface natural, como a possibilidade de pausar e retomar os eventos de reconhecimento para evitar processamento adicional indesejável. Entretanto outras se tornam necessárias para facilitar a adaptação das interfaces naturais a diversos tipos de aplicações, como discutido nos próximos parágrafos.

O fato do KinectServer ser uma aplicação e não um serviço, torna inviável para algumas aplicações utilizar esta biblioteca. Seria necessário modificá-lo como um serviço para que aplicações pudessem ser distribuídas normalmente, sem a necessidade da utilização de um segundo software.

Outra função interessante seria a adequação automática do ângulo do Kinect, baseando-se no esqueleto do usuário em reconhecimento. Essa função dispensaria que o usuário necessitasse configurar o ângulo manualmente.

Apesar de correções da biblioteca não fazerem parte do objetivo e escopo deste trabalho, grande parte do esforço realizado foi para correções das mesmas. As correções foram realizadas à medida que se tornaram necessárias e que não prejudicassem o cronograma deste trabalho, a fim de tornar possível a implementação da interface natural projetada.

## 7. Referências

- Microsoft. Kinect Documentatio & Resources. 2013. Disponível em: <<http://www.microsoft.com/en-us/kinectforwindows/develop/resources.aspx>>. Acesso em 09 mar. 2013.
- Schade, Gabriel Universidade do Vale do Itajaí. Biblioteca de Funções para utilização do Kinect em Aplicações GUI e Jogos Eletrônicos. 2012. TCC (graduação em Ciência da Computação) - Universidade do Vale do Itajaí, Itajaí, 2012. Disponível em : <[http://Siaibib01.univali.br/pdf/Gabriel\\_Schade\\_Cardoso.pdf](http://Siaibib01.univali.br/pdf/Gabriel_Schade_Cardoso.pdf)>. Acesso em: 21 fev. 2013.
- Silva, Eduardo A. Trilha virtual interativa em 3D do Morro da Cruz: Disseminação digital do conhecimento obtido no processo de restauração da biodiversidade do local. Itajaí, 2012. Universidade do Vale do Itajaí, CTTMAR – Centro de Ciências Tecnológicas da Terra e do Mar, Brasil. Disponível em: <[http://sbgames.org/sbgames2012/proceedings/papers/cultura/C\\_S10.pdf](http://sbgames.org/sbgames2012/proceedings/papers/cultura/C_S10.pdf)>. Acesso em: 08 mar. 2013.
- Unity Technologies. Unity Documentation. 2013. Disponível em: <<http://unity3d.com/company/support/documentation/>>. Acesso em: 09 mar. 2013.