

UNIVERSIDADE ESTADUAL PAULISTA

"JÚLIO DE MESQUITA FILHO"

CAMPUS UNIVERSITÁRIO DE BAURU

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE COMPUTAÇÃO

Azael de Melo e Sousa

**ÁREA DE TRABALHO DINÂMICA PARA O AMBIENTE EXPERIMENTAL E
DIDÁTICO COM ROBÔS MÓVEIS**

Bauru - SP

2015

Azael de Melo e Sousa

**ÁREA DE TRABALHO DINÂMICA PARA O AMBIENTE EXPERIMENTAL E
DIDÁTICO COM ROBÔS MÓVEIS**

Monografia apresentada junto à disciplina
Projeto de Implementação de Sistemas II,
do curso de Bacharelado em Ciência da
Computação, Faculdade de Ciências,
UNESP, campus de Bauru, como parte do
Trabalho de Conclusão de Curso.

Orientador:

Prof. Dr. Renê Pegoraro

Bauru - SP

2015

Azael de Melo e Sousa

**ÁREA DE TRABALHO DINÂMICA PARA O AMBIENTE EXPERIMENTAL E
DIDÁTICO COM ROBÔS MÓVEIS**

Monografia apresentada junto à disciplina
Projeto de Implementação de Sistemas II,
do curso de Bacharelado em Ciência da
Computação, Faculdade de Ciências,
UNESP, campus de Bauru, como parte do
Trabalho de Conclusão de Curso.

BANCA EXAMINADORA

Prof. Dr. Renê Pegoraro

DCo - FC - UNESP - Bauru

Orientador

Prof. Dr. Wilson Massashiro Yonezawa

DCo - FC - UNESP - Bauru

Prof. Dr. Marco Antônio Rahal Sacoman

DCo - FC - UNESP - Bauru

Bauru, 10 de Março de 2015

AGRADECIMENTOS

Agradeço primeiramente ao professor Renê Pegoraro pelo apoio, amizade, orientação e incansáveis reuniões e discussões de trabalho.

A todos os professores que me instruíram nas mais diversas áreas do conhecimento durante minha formação acadêmica.

À minha família, que me apoio e incentivou durante todos os anos de minha vida.

Aos meus grandes amigos, pelo companheirismo e todos os dias que passamos nos laboratórios

*"The needs of the many outweigh the
needs of the few, or the one."*

(Spock in The Wrath of Khan, 1982)

RESUMO

No AEDROMO (Ambiente Experimental e Didático com Robôs Móveis) robôs móveis e objetos passivos interagem numa área de trabalho. Esse ambiente tem o objetivo de permitir o controle de um robô de forma simples, na realização de atividades propostas. O propósito do sistema de visão do AEDROMO é detectar e classificar todos os robôs e objetos passivos na área de trabalho por meio de suas cores. Logo, a área de trabalho não pode ter cores conflitantes com as dos robôs e objetos. Este trabalho apresenta uma nova abordagem para a detecção dos robôs e objetos passivos através de seus tamanhos e cores, utilizando o sensor *Microsoft Kinect*. Isso trará maior versatilidade ao ambiente uma vez que as atividades podem ser melhor ilustradas e entendidas. Este trabalho resultou num sistema de visão onde as cores da área de trabalho não interferem na localização e classificação dos robôs e objetos, possibilitando a troca de estampas coloridas na área de trabalho.

Palavras-Chave: AEDROMO, Rastrear Objetos, *Kinect*, Calibração

ABSTRACT

In the AEDROMO (Experimental and Didactic Environment with Mobile Robots) mobile robots and passive objects interact in a workspace. This environment's goal is to allow simple control of robots, in the fulfillment of proposed activities. The purpose of AEDROMO's vision system is to detect and classify all mobile robots and passive objects in the workspace through its colors. Thus, the workspace cannot have conflicted colors with those of the robots and other objects. This work presents a new approach in the detection of robots and passive objects by means of size and color using the Microsoft Kinect sensor. This will make the environment more versatile once the activities can be better illustrated and understood. This work resulted in a system which the colors of the workspace doesn't interfere with the detection and classification of the robots and passive objects, allowing the exchange of colorful pictures in the workspace.

Key Words: AEDROMO, Object Tracking, Kinect, Calibration.

LISTA DE FIGURAS

Figura 2.1 - AEDROMO com: (a) webcam; (b) robôs; e (c) área de trabalho. Fonte: (ALVES; et al. 2011).....	14
Figura 2.2 - Robô do AEDROMO (ALVES et al.2011).	14
Figura 2.3 - Posicionamento do robô reconhecido (ALVES et al., 2011).....	15
Figura 2.4 - Novo AEDROMO. Fonte: Autor.....	18
Figura 2.5 - TV projetando imagem de um campo de Futebol. Fonte: Autor.....	18
Figura 2.6 - Diagrama do processo de calibração e utilização. Fonte: Autor.	19
Figura 2.7 - Diagrama de Rota do Sistema de Identificação. Fonte: Autor	19
Figura 3.1 - Microsoft Kinect (SMISEK; JANCOSK; PAJDLA, 2013).	21
Figura 3.2 - Imagens tiradas do sensor RGB-D (HENRY et al., 2010).....	22
Figura 3.3 - Padrão de marcações capturado pela câmera IR. Fonte: Autor.	23
Figura 3.4 - Esquema representando a disparidade de profundidade (KHOSHELHAM, 2011).....	24
Figura 4.1 - Projeção dos pontos C, M, m (ZHANG, 2004, p. 8).	27
Figura 4.2 - Aplicação da matriz de Homografia H. Fonte: http://plus.maths.org/issue23/features/criminisi/homography.gif , acesso em: 24/11/2014.	29
Figura 4.3 - Inserção de um quadro em uma imagem através da homografia. Fonte: docs.opencv.org/_images/Feature_Homography_Result.jpg , acesso em: 24/11/2014.	29
Figura 4.4 - Diagrama de Atividade da Calibração do Posicionamento das Câmeras. Fonte: Autor.....	30
Figura 4.5 - Diagrama do funcionamento da Calibração das Cores. Fonte: Autor. ...	31
Figura 4.6 - Espaço RGB com a transformação SCT (UMBAUGH et al., 1989).	32

Figura 5.1 - Software de Calibração do Posicionamento das Câmeras. Área vazia (a), Captura dos pontos característicos (b) até (f). Fonte: Autor.	34
Figura 5.2 - À esquerda a área sem a proteção do vidro. No meio a imagem de profundidade. À direita o efeito físico da refração no AEDROMO. Fonte: Autor.	35
Figura 5.3 - Software de Calibração das Cores. (a) Com proteção, (b) Sem proteção e imagem sendo projetada na TV. Fonte: Autor.	36
Figura 5.4 - Gráfico da faixa de valores. Ângulos α e β e luminosidade do objeto de cor laranja. Fonte: Autor.	37
Figura 5.5 - Imagem colorida analisada. Processo de detecção de elevações e aplicação da homografia. Fonte: Autor.	38
Figura 5.6 - Imagem colorida indexada. Fonte: Autor.	39

LISTA DE SIGLAS

AEDROMO	Ambiente Experimental e Didático com Robôs Móveis
GISDI	Grupo de Integração de Sistemas e Dispositivos Inteligentes
DCo	Departamento de Computação
IR	<i>Infrared</i>
RGB	<i>Red, Green, Blue</i>
fps	<i>Frames per second</i>
ROI	<i>Region of Interest</i>
SCT	<i>Spherical Coordinate Transformation</i>
CIC	Congresso de Iniciação Científica

SUMÁRIO

Capítulo 1	INTRODUÇÃO	11
1.1	Objetivos.....	12
1.2	Justificativa	12
1.3	Estrutura da Monografia	13
Capítulo 2	O AEDROMO	14
2.1	Mudanças Estruturais no Novo AEDROMO	17
Capítulo 3	Microsoft Kinect	21
3.1	Medição de Profundidade	22
Capítulo 4	Calibração	26
4.1	Noções Matemáticas	27
4.2	Calibração do Posicionamento das Câmeras	28
4.2.1	Homografia	28
4.2.2	Aplicação da Metodologia de Calibração	30
4.3	Calibração das Cores	31
4.3.1	O SCT.....	32
Capítulo 5	Resultados e Discussões	34
5.1	<i>Software</i> de Calibração do Posicionamento das Câmeras.....	34
5.2	<i>Software</i> de Calibração das Cores.....	35
5.3	Utilização do Novo AEDROMO.....	37
Capítulo 6	Conclusão	40
	Referências Bibliográficas	40
	Apêndice A - Artigo apresentado no Congresso de Iniciação Científica da UNESP 2013	42
	Apêndice B - Artigo apresentado no Congresso de Iniciação Científica da UNESP 2014	44
	Apêndice C - Artigo apresentado no <i>Workshop Serious Games = Serious Business</i> 2015	46

Capítulo 1 INTRODUÇÃO

O ensino de engenharia dentro das universidades enfrenta hoje novos desafios. Geralmente, as abordagens tradicionais que dão ênfase apenas aos aspectos teóricos cativam pouco a atenção dos alunos. Neste sentido, a introdução de novas tecnologias no processo de aprendizado pode trazer resultados positivos. Uma abordagem que tem apresentado êxito em diferentes níveis de ensino é a robótica. Como ferramenta pedagógica para o ensino de engenharia, a robótica apresenta vários aspectos interessantes. Pois se trata de uma área que é intrinsecamente multidisciplinar, capaz de estimular o trabalho em equipe e promove um retorno visual real de forma estimulante e motivador (AVANZATO, 2000).

Assim como os computadores, a robótica é um recurso tecnológico auxiliar utilizável no processo educacional que pode contribuir para o desenvolvimento cognitivo do aluno e de habilidades intelectuais específicas. Esta se oferece como uma ferramenta pedagógica interessante sobre vários aspectos e assim deve ser encarada e explorada (PAPERT, 1993).

Portanto, o uso da robótica em educação é, principalmente, motivado pelo aprendizado ativo que esta proporciona e isto acontece em dois aspectos. O primeiro, para o ensino de conceitos particulares ou gerais de disciplinas envolvidas com a própria robótica, e como bancada de testes para o ensino desta área. O segundo, como um elemento motivador e bancada de experimentos, para ensinar e verificar conceitos de disciplinas de áreas correlatas (BARKER; ANSORGE, 2007). Neste aspecto, a robótica é uma ferramenta pedagógica, um objeto didático.

Neste sentido, o Grupo de Integração de Sistemas e Dispositivos Inteligentes, (GISDI), do Departamento de Computação (DCo) da UNESP de Bauru desenvolve iniciativas de baixo custo, flexíveis e alinhadas para o ensino de Engenharia e da própria Robótica. O AEDROMO (ambiente experimental e didático com robôs móveis) é uma dessas iniciativas.

O AEDROMO pode apresentar diversas opções de aplicações. Possui a capacidade de reconhecer dois robôs (ou mais) pelas suas cores e vários outros objetos coloridos, sobre uma arena plana, através de uma câmera de vídeo

conectada a computadores responsáveis pelo controle dos robôs, possibilitando atividades que envolvam sistemas colaborativos e multiagentes.

Várias atividades, como os exemplos *Caça e Caçador*, *Colecionando Coisas* e *Resolvendo o Labirinto* podem ser desenvolvidas neste ambiente.

Em cada uma destas atividades, apesar das dimensões serem mantidas entre si, existe a necessidade de uma área de trabalho (arena), com desenhos coloridos específicos, caracterizados como estampas ou decoração, para cada uma das atividades, permitindo que os alunos, usuários ou clientes tenham uma representação mais concreta da tarefa e entendam melhor as finalidades das atividades. As cores dos desenhos não podem conflitar com as dos robôs e objetos. A ênfase deste trabalho é o desenvolvimento de um sistema de identificação de objetos para áreas de trabalho coloridas ou dinâmicas, ou seja, podem ser alteradas em tempo real, para o AEDROMO.

Durante este trabalho, o AEDROMO operando com o novo sistema de identificação será denominado Novo AEDROMO.

1.1 Objetivos

Desenvolver um novo sistema de identificação imune às cores da estampa, dos robôs e outros objetos na arena do AEDROMO.

1.2 Justificativa

O AEDROMO oferece inúmeras possibilidades de atividades, tendo cada uma delas um contexto específico para a situação proposta. Parte do contexto vem da decoração da arena, que pode conter qualquer cor ou estampa de forma a tornar mais concreto ao usuário o ambiente. No AEDROMO, os elementos, robôs e os objetos, são identificados por suas cores, que podem ser confundidas com as cores da decoração da arena. Assim, a identificação dos objetos levando em consideração outras características além das suas cores pode tornar o ambiente mais genérico e

permitir um maior número de atividades. Este trabalho se insere neste tipo de identificação.

1.3 Estrutura da Monografia

Esta monografia está estruturada da seguinte forma:

- Capítulo 2: O AEDROMO, apresenta a versão antiga do AEDROMO, as principais alterações feitas no projeto e a ideia do novo sistema de identificação dos objetos;
- Capítulo 3: O *Microsoft Kinect*, descreve o dispositivo *Kinect* e como a medição de profundidade é feita e calculada;
- Capítulo 4: Calibração, trata das descrições dos métodos de calibração para o posicionamento das câmeras e das cores;
- Capítulo 5: Resultados e discussões, mostra os *softwares* desenvolvidos para suprir as necessidades computacionais do projeto;
- Capítulo 6: Conclusões, apresenta a conclusão do projeto e os trabalhos futuros.

Capítulo 2 O AEDROMO

O AEDROMO é um Ambiente Experimental e Didático com Robôs Móveis, formado por uma área de trabalho¹, objetos passivos, dois robôs (ou mais), dois computadores, uma câmera global do tipo *webcam* e um transmissor (Figura 2.1). A flexibilidade de adaptação, uso por várias disciplinas, e o baixo custo são as premissas na concepção e desenvolvimento deste ambiente. Os experimentos, neste ambiente, podem ser motivados para fins de pesquisa, aprendizagem ou entretenimento (FERASOLI et al., 2006).

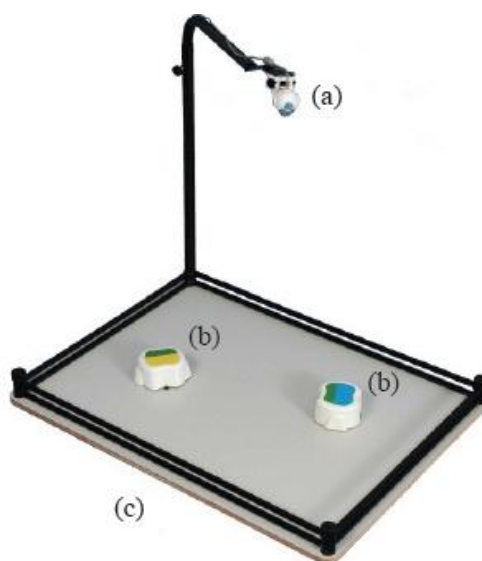


Figura 2.1 - AEDROMO com: (a) webcam; (b) robôs; e (c) área de trabalho. Fonte: (ALVES; et al. 2011).

O robô utilizado no AEDROMO pode ser visto na figura 2.2.



Figura 2.2 - Robô do AEDROMO (ALVES et al.2011).

¹ Neste trabalho, a área de trabalho será denominada arena.

Na arquitetura do AEDROMO há dois tipos de computadores: o Computador Servidor e o Computador Aluno. O computador servidor é responsável pela implementação do sistema de visão computacional e pela comunicação com os robôs. Por outro lado, no computador aluno é onde o software de controle dos robôs será desenvolvido pelo usuário (ALVES et al., 2011).

O uso de uma câmera global confere grande versatilidade ao sistema, uma vez que permite a pequenos robôs reagirem com o ambiente de forma complexa, sem a necessidade de inclusão de sensores sofisticados. Ao mesmo tempo, ela atribui ao ambiente uma ampla variedade de aplicações pedagógicas tanto para ensino quanto para pesquisa.

As imagens são capturadas pela câmera global e são enviadas ao computador servidor. Neste computador, as imagens são adquiridas e processadas. O processamento determina as posições cartesianas (x,y) e o ângulo θ dos objetos (robôs, bolas, blocos, etc.) inseridos na área de trabalho (ALVES et al., 2011).

Para facilitar o reconhecimento do robô, foram adotados marcadores coloridos que consistem em duas etiquetas de cores diferentes coladas sobre os robôs denotadas por C_0 e C_1 . Uma vez que processada a imagem da câmera, é possível reconhecer estas cores como regiões e determinar a posição e orientação do robô. Com as coordenadas cartesianas de C_0 e C_1 , o servidor calcula as coordenadas de C , dado pelo ponto médio entre C_0 e C_1 , e o ângulo θ do robô, obtido por $\alpha - 45^\circ$, sabendo que α é o ângulo formado entre a reta que passa por C_0 e C_1 e o eixo das abscissas (ALVES et al., 2011). A figura 2.3 mostra o reconhecimento do robô. Outros objetos passivos (bolas e blocos), que compõem o ambiente, também podem ser identificados na área de trabalho através de suas cores.

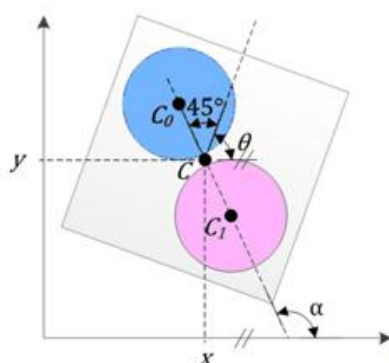


Figura 2.3 - Posicionamento do robô reconhecido (ALVES et al., 2011)

Durante a identificação das etiquetas, ruídos com as cores das etiquetas podem ser detectados. Para evitar que ruídos sejam considerados robôs, um sistema de identificação das melhores etiquetas e seus pares é utilizado. Esse sistema considera dois tipos de características: tamanho da etiqueta e a proximidade do par de etiquetas sobre o robô. Os valores de cada uma dessas características, quando obtidos pelo sistema de visão são comparados aos valores esperados ideais. As medidas para o tamanho da etiqueta são a largura, altura e o número de *pixels*, que devem ter seus valores comparados com seus respectivos valores ideais. A proximidade do par de etiquetas em um robô, leva em consideração a distância entre o centro de cada etiqueta. A soma das diferenças entre os valores observados e os ideais determinam a qualidade das etiquetas. O conjunto de etiquetas com a maior qualidade são as que identificam os robôs.

A atuação dos robôs é vista pela câmera e o cálculo se repete numa taxa equivalente à velocidade de captura da própria câmera, que em geral varia de 10 a 20 *fps (frames per second)*.

O objetivo das atividades propostas com o AEDROMO é demonstrar a programação do robô, habituar os alunos ao controle do mesmo (acionamento dos motores através do teclado e monitorar o estado dos sensores na tela) e experimentar aspectos mais complexos da robótica e da computação (ALVES et al., 2011).

Uma decoração com uma estampa colorida, pode ser colocada na superfície da arena. A decoração da arena possibilita aos alunos uma abstração da atividade proposta. Caso haja similaridade entre as cores da decoração com as dos robôs, poderá haver conflitos, com isso, o sistema de visão e identificação poderá detectar as etiquetas dos robôs em lugares inesperados. Assim, resultando em informações incorretas sobre os estados dos robôs, causando anomalias no controle e na realização das atividades destes robôs.

2.1 Mudanças Estruturais no Novo AEDROMO

A impressão e trocas da decoração para cada atividade no AEDROMO pode gerar dificuldades e descontinuidades em sequências de atividades. Portanto, a projeção de imagens coloridas e decorativas na arena pode facilitar a utilização do ambiente.

As projeções das imagens coloridas são feitas por meio de uma TV LCD que, no Novo AEDROMO, é a área de trabalho. Os robôs agem sobre um vidro posicionado diretamente sobre a TV. O propósito do vidro é o de proteger a TV de possíveis deformações e ranhuras causadas pelos usuários, robôs e objetos.

A partir das pesquisas bibliográficas, optou-se pela abordagem proposta por Wilson (2010) para a identificação dos objetos na área de trabalho. Nesta abordagem é proposta a utilização do sensor RGB-D *Microsoft Kinect* para obter informações da aproximação de mãos humanas sobre a superfície. Ainda no artigo, os testes apresentam a possibilidade da identificação da presença de outros objetos sobre a superfície, permitindo desta forma a verificação da presença e da posição de cada um dos robôs sobre a área de trabalho. Desta forma, permitindo a utilização da TV para criar os ambientes com os motivos dinâmicos desejados para a atividade proposta. O dispositivo *Microsoft Kinect* é instalado à uma altura de, aproximadamente, um metro do AEDROMO, que supera a distância mínima requerida pelo sensor.

A TV é suportada mecanicamente por uma estrutura metálica. Nessa mesma estrutura é instalado um microcomputador de mesa que funciona como o gerenciador dos dados do *Kinect*. Após o tratamento destes dados, como a identificação dos elementos na arena (robôs e objetos passivos), as coordenadas dos objetos são disponibilizados à uma máquina cliente, que definirá os comandos a serem executados pelos robôs.

A figura 2.4 apresenta a aparência da nova estrutura do AEDROMO.



Figura 2.4 - Novo AEDROMO. Fonte: Autor.

A figura 2.5 mostra a área de trabalho do AEDROMO pronto para uma partida de futebol. Na versão antiga seria necessário construir a área de trabalho evitando a utilização de cores conflitantes com às dos robôs ou objetos passivos.



Figura 2.5 - TV projetando imagem de um campo de Futebol. Fonte: Autor.

O sensor de profundidade do *Kinect* é utilizado para encontrar objetos na área de trabalho por meio de elevações, ou seja, regiões acima do plano da arena. Uma vez identificada as regiões de elevação, somente estas regiões tem importância na imagem colorida do *Kinect*. Caso, na imagem colorida, ao se analisar a região de elevação, seja encontrada a etiqueta de um robô, então esse objeto é um robô, caso contrário é algum objeto passivo (bola, cubo, etc.). Este procedimento é feito no momento da utilização do AEDROMO.

Para que, no momento da calibração, se possa analisar as regiões de elevação, é necessário sobrepor ambas imagens coloridas e de profundidade e

conhecer as cores de cada objeto. Isto é feito por meio de duas calibrações: a calibração do posicionamento das câmeras e a calibração das cores. As calibrações são apresentadas com mais detalhes posteriormente. O diagrama da figura 2.6 mostra como os procedimentos de calibração e utilização se relacionam.

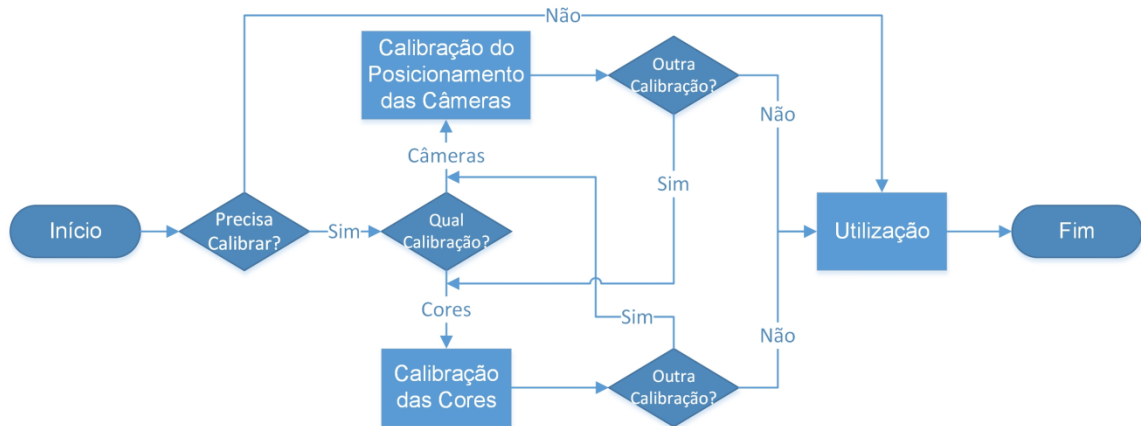


Figura 2.6 - Diagrama do processo de calibração e utilização. Fonte: Autor.

Cada uma das etapas de *Calibração do posicionamento das câmeras*, *Calibração das cores* e *Utilização*, possuem seus próprios processos a serem executados. Porém não há a necessidade desses processos serem sempre feitos seguindo a mesma ordem. Caso algum conjunto de parâmetros obtidos nas calibrações ainda sejam válidos, ou seja, todos os elementos na arena podem ser vistos claramente no momento da utilização, não há a necessidade de refazer essa etapa específica para usar o ambiente. A figura 2.7 mostra cada processo individualmente.

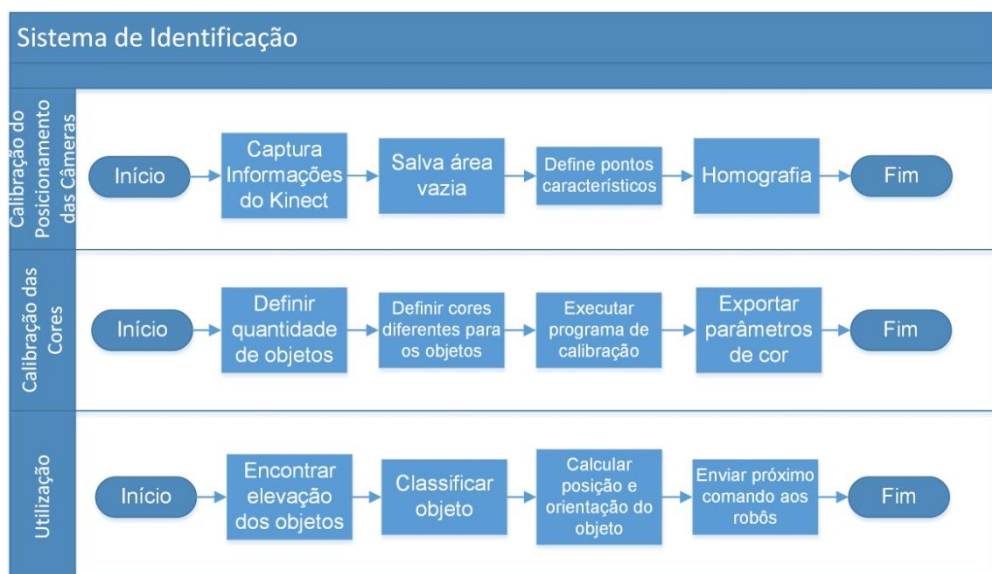


Figura 2.7 - Diagrama de Rota do Sistema de Identificação. Fonte: Autor

O sistema de identificação é dividido em duas fases: calibração e utilização. A fase da calibração é dividida, por sua vez, em calibração do posicionamento das câmeras e das cores. A calibração do posicionamento das câmeras é vista com maiores detalhes no capítulo 4.2. Já a calibração das cores, permanece como é feita na versão anterior do AEDROMO, ou seja, voltada para a classificação de cores e é melhor detalhada no capítulo 4.3. Cada objeto que atua na área de trabalho deve ter cores distintas entre si, para que não haja confusão na classificação do objeto. Desta forma, após a calibração das cores, um arquivo com todos os parâmetros de cores de cada objeto é exportado, evitando recalibrações desnecessárias.

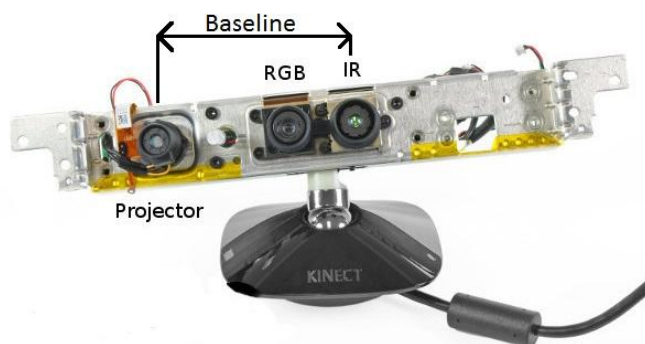
Na fase de utilização, o software encontrará todos os objetos interagindo na área de trabalho do Novo AEDROMO e serão classificados a partir de suas cores. Em seguida, suas posições e orientações serão calculadas por meio de rotinas já implementadas no sistema do AEDROMO. Uma vez que essas informações tenham sido encontradas, o próximo comando será enviado ao robô de tal forma a executar a tarefa inicialmente proposta.

Capítulo 3 Microsoft Kinect

Em 2010, a *Microsoft*, em conjunto com a *PrimeSense*, apresentou o *Kinect* para o *Xbox 360* (Figura 3.1a), um sensor de baixo custo que captura imagens tanto de profundidade como de vídeo em média qualidade e em tempo real, dedicado à detecção de movimentos em um jogo de *videogame* (MACKNOJIA et al., 2013). O dispositivo *Kinect* possui duas câmeras com autofocus, uma câmera IR (sigla em inglês para infravermelho), otimizada para detectar profundidades e a outra é uma câmera padrão de espectro visual RGB (Sigla em inglês para vermelho, verde e azul), um projetor IR, um acelerômetro e quatro microfones. A Figura 3.1b mostra a disposição dessas câmeras no dispositivo. Cada lente esta associada com a câmera ou o projetor. A câmera IR e o projetor IR formam uma *baseline* (distância entre a câmera e o projetor) de 7.5cm. A medição de profundidade consiste num cálculo baseado em triangulação com a câmera IR, o projetor IR e as coordenadas dos *pixels*. (CRUZ; LUCIO; VELHO, 2012).



(a) Kinect



(b) Estrutura do Kinect

Figura 3.1 - Microsoft Kinect (SMISEK; JANCOSK; PAJDLA, 2013).

Este sensor pode ser classificado como uma câmera RGB-D, pois possui uma câmera RGB padrão embutido com um projetor IR e uma câmera IR. Desta forma, o sensor captura imagens RGB junto com informações de profundidade *pixel-a-pixel* (HENRY et al., 2010).

A câmera colorida possui uma resolução de 1280x1024 *pixels* com um campo de visão horizontal de 63 graus e um campo de visão vertical de 50 graus. O *Kinect* é capaz de capturar imagens coloridas numa taxa de 10 *fps* na sua resolução

máxima (1280x1024 *pixels*). Caso a resolução seja reduzida para 640x480 *pixels*, a velocidade de captura aumenta para 30 *fps* (MACKNOJIA, 2013).

A câmera IR opera em uma frequência de 30Hz e captura imagens com resolução de 1200x960 *pixels*. Devido ao baixo rendimento das portas USB 2.0, essas imagens são reduzidas para uma escala de 640x480 *pixels*, com um vetor de profundidade de 11 bits, podendo este variar em 2048 (2^{11}) valores de distância. Como o *Kinect* ocupa, aproximadamente, 70% da transmissão de dados de um único *hub* USB, sobra pouca banda para outros dispositivos (mouse, teclado, etc.). Desta forma, não é possível conectar dois dispositivos *Kinect* simultaneamente no mesmo *hub* (CRUZ; LUCIO; VELHO, 2012; SOLONY; ZEMCÍK, 2011).

O propósito inicial do *Kinect* era para jogar *videogames* em lugares fechados, como um quarto. Assim, o alcance de visão dos sensores varia entre 2,5 pés (800mm) até 13 pés (4000mm). Entretanto, o uso recomendado é de 3 (914mm) a 12 (3657mm) pés (WEBB; ASHLEY, 2012). A Figura 3.2 mostra uma cena capturada pela câmera RGB e sua respectiva cena em profundidade.



(a) Imagem RGB.



(b) Informações de profundidade representadas em escala de cinza.

Figura 3.2 - Imagens tiradas do sensor RGB-D (HENRY et al., 2010).

3.1 Medição de Profundidade

O *Kinect* utiliza o princípio de medição de profundidade por triangulação para encontrar a distância entre um objeto e o sensor (FREEDMAN et al., 2010; CRUZ; LUCIO; VELHO, 2012). O projetor IR emite um único raio que se divide em outros

múltiplos raios por uma grade de difração, criando assim um padrão constante de marcações projetadas no ambiente. Esse padrão é capturado pela câmera IR e comparado com um padrão de referência previamente conhecido e armazenado na memória do *Kinect*. Quando a marcação é projetada num objeto cuja distância até o sensor seja menor ou maior do que aquela no plano de referência, a posição da marcação na imagem IR é deslocada na direção da *baseline* entre o projetor e o centro da câmera IR. Esse deslocamento é medido para todas as marcações por um procedimento de correlação, na qual gera uma disparidade na imagem (KHOSHELHAM, 2011). A figura 3.3 mostra o padrão de marcações.



Figura 3.3 - Padrão de marcações capturado pela câmera IR. Fonte: Autor.

A Figura 3.4 mostra a relação entre a distância de um ponto k de um objeto no plano do objeto (plano no qual o objeto se encontra) até o sensor e a medida de disparidade d . Para cada *pixel*, a distância até o sensor é calculada a partir de sua disparidade. A representação de uma coordenada do R^3 de um objeto é considerado um sistema de coordenadas de profundidade com sua origem no centro da câmera IR. O eixo Z é ortogonal ao plano da imagem (plano no qual a imagem RGB é projetada), ao plano do objeto e o plano de referência (plano no qual se encontra o padrão de marcações). O eixo X é perpendicular ao eixo Z na direção da *baseline* entre o centro da câmera IR e o projetor, e o eixo Y é ortogonal aos eixos X e Z (KHOSHELHAM, 2011).

Assumindo que um objeto o está no plano de referência a uma distância Z_o do sensor, e uma marcação do objeto k é capturado no plano do objeto, se esse

mesmo objeto muda de posição (tanto para longe como para perto do sensor), a localização dessa marcação também muda na direção do eixo X num deslocamento D . Isso é medido como uma disparidade d correspondente a um ponto k do objeto e com essa disparidade, a distância Z_k do objeto pode ser encontrada (KHOSHELHAM, 2011).

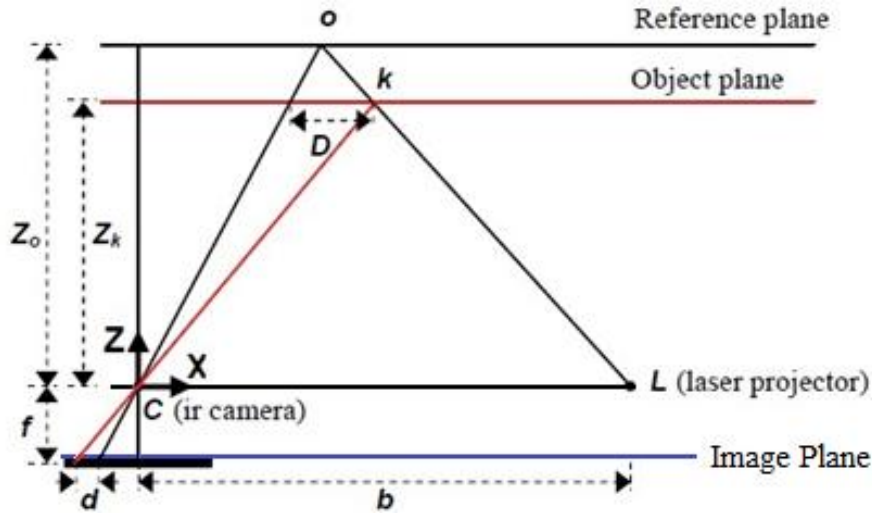


Figura 3.4 - Esquema representando a disparidade de profundidade (KHOSHELHAM, 2011).

A partir de semelhança de triângulos, tem-se:

$$\frac{D}{b} = \frac{Z_o - Z_k}{Z_o} \quad (3.1.1)$$

e:

$$\frac{d}{f} = \frac{D}{Z_k} \quad (3.1.2)$$

Onde, D é o deslocamento do ponto k no plano do objeto, b é a *baseline*, já Z_k representa a distância (profundidade) do ponto k no plano do objeto, d é a disparidade observada no plano da imagem e f é a distância focal da câmera IR. Substituindo D da equação (3.1.1) pelo da equação (3.1.2) e isolando a variável Z_k tem-se:

$$Z_k = \frac{Z_o}{1 + \frac{Z_o}{fb}d} \quad (3.1.3)$$

A equação (3.1.3) é o modelo matemático básico para a profundidade a partir de uma divergência observada dado os parâmetros constantes Z_o , f e b (KHOSHELHAM, 2011).

Capítulo 4 Calibração

De acordo com o dicionário Aurélio da língua portuguesa, a palavra "Calibração" significa: "Operação em que estabelece uma correspondência entre as leituras de um instrumento e valores de uma grandeza física que é medida ou não, diretamente, pelo instrumento".

Na área da visão computacional, Faugeras, Luong e Maybank (1992) descrevem que a calibração de câmeras é uma tarefa importante na visão computacional, cujo propósito é estabelecer a projeção do sistema de coordenadas do mundo para o sistema de coordenadas do plano da imagem. Uma vez que essa projeção é conhecida, informações do mundo podem ser obtidas a partir de informações da imagem, e vice versa. Zhang (2000) reforça essa ideia dizendo que a calibração de câmeras é um passo necessário na visão computacional 3D, a fim de extrair informações métricas de imagens 2D.

Câmeras RGB-D, tais como o *Kinect*, podem ser usadas em uma extensa gama de aplicações como: robótica, reconhecimento facial, reconhecimento de movimentos, monitoramento de corpos, entre outros. Todas essas aplicações exigem a transformação de pontos entre o plano da imagem de vídeo, o plano de profundidade e o mundo (STARANOWICZ; MARIOTTINI, 2012). Seguindo o conceito de Faugeras, Luong e Maybank (1992), o objetivo da calibração de câmeras é determinar a matriz de transformação responsável por realizar tais transposições.

No âmbito deste trabalho, existem dois tipos de calibração: a calibração do posicionamento das câmeras e a calibração das cores. A primeira consiste em obter um modelo matemático que permita a transição das coordenadas obtidas nos planos das imagens, RGB e profundidade, entre si e com as coordenadas cartesianas do plano da arena. A partir dessa calibração é possível relacionar um objeto nestes três planos. A segunda calibração está voltada à definição das cores de cada robô e objeto, que são utilizados posteriormente no ambiente. Ambas as calibrações são necessárias. Com a calibração do posicionamento das câmeras é possível sobrepor as imagens de profundidade e RGB, e com a calibração das cores, as cores das regiões de elevação, agora vistas na imagem RGB, podem ser classificadas.

4.1 Noções Matemáticas

A calibração está baseada num modelo que relacione um ponto na imagem de uma câmera com seu correspondente no mundo real. Assim, um ponto 2D em uma imagem obtida por uma câmera pode ser relacionado matematicamente ao respectivo ponto 3D no mundo. Um ponto 2D é representado como $m = [u, v]^T$, e um ponto 3D será representado por $M = [X, Y, Z]^T$. Os pontos M' e m' representam as coordenadas homogêneas dos pontos M e m respectivamente. Assim: $m' = [u, v, 1]^T$ e $M' = [X, Y, Z, 1]^T$. A imagem de um ponto 3D M , denotado por m é formado a partir de um raio ótico de M através do centro ótico C e intersectando o plano da imagem. Esses três pontos são colineares. Na figura 4.1 o plano da imagem está posicionado entre o ponto da cena e o centro ótico, na qual é matematicamente equivalente à configuração física sobre a qual o plano da imagem está localizado (ZHANG; 2000, 2004, p. 8).

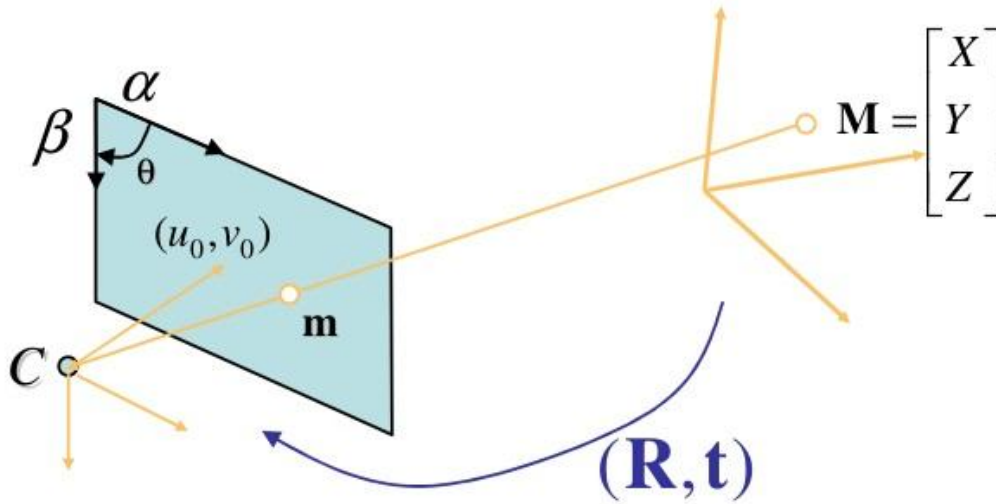


Figura 4.1 - Projeção dos pontos C , M , m (ZHANG, 2004, p. 8).

A relação entre o ponto 3D M e a projeção no plano da imagem m é dado por:

$$sm' = A[R \ T]M' = PM' \quad (4.1.1)$$

$$\text{com } A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.1.2)$$

$$\text{e } P = A[R \ T] \quad (4.1.3)$$

Onde s é um fator de escala arbitrário, a matriz $[R \ T]$ é a matriz de rotação e translação na qual relaciona o sistema de coordenadas do mundo com a câmera, A

é chamada de matriz da câmera, sendo u_0 e v_0 as coordenadas do centro da câmera, α e β são as distâncias focais no eixo das abscissas e das ordenadas respectivamente, e γ é o parâmetro que descreve a distorção dos eixos das duas imagens. A matriz 3×4 P é chamada de matriz de projeção da câmera, que mistura ambas matrizes A e $[R \ T]$. (ZHANG; 2000, 2004, p. 8).

4.2 Calibração do Posicionamento das Câmeras

Em geral, técnicas de calibração envolvem a análise de imagens da projeção de uma série de pontos característicos. Para conseguir isso, uma sequência de imagens, na qual o modelo de calibração foi capturado, é disponibilizado. Dentro desse modelo, uma série de pontos são localizados, que providenciam informações vitais no aspecto de distância entre os pontos a serem definidos. Uma vez que esses pontos e suas coordenadas bidimensionais tenham sido extraídas da sequência de imagens, a relação entre cada ponto do modelo é encontrada para a série em toda sequência. Essas informações são usadas como dados de entrada no algoritmo de calibração (ESCALERA; ARMINGOL, 2010).

Existem diversos métodos para a calibração de câmeras tais como os métodos de Zhang, Herrera e DCCT (STARANOWICZ; MARIOTTINI, 2012). Estes métodos de calibração encontram os parâmetros indicados no item 4.1. Porém, a homografia é um método mais simples de calibração que se aplica ao AEDROMO, uma vez que o AEDROMO tem seu funcionamento restrito ao plano da arena.

4.2.1 Homografia

Homografia é um mapeamento entre duas imagens planas. Outros termos para essa transformação incluem projetividade, colinearização e transformação projetiva plana. Hartley e Zisserman (2003) providenciaram uma definição mais específica: homografia é um mapeamento de um plano projetivo para si mesmo tal que três pontos estão na mesma reta se e somente se os pontos mapeados são também colineares. O teorema algébrico dita que o mapeamento de $P^2 \rightarrow P^2$, sendo P^2 uma imagem plana 2D, é possível se e somente se existe uma matriz H

não singular 3×3 , tal que para qualquer ponto em P^2 , seu ponto mapeado é igual à Hx (DUBROFSKY, 2009; MALIS et al, 2010; HARTLEY; ZISSERMAN, 2003). A figura 4.2 apresenta a aplicação da matriz de homografia H para transformar um ponto 2D no plano do mundo para outro ponto 2D no plano da imagem.

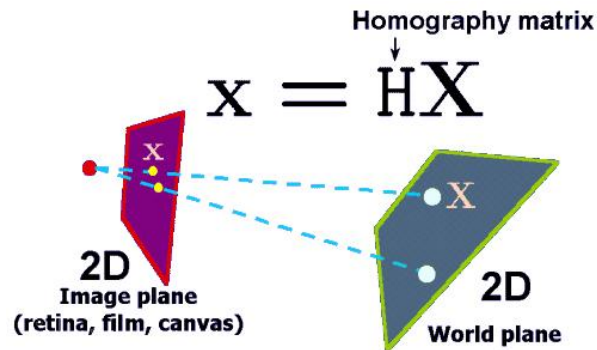


Figura 4.2 - Aplicação da matriz de Homografia H . Fonte: <http://plus.maths.org/issue23/features/criminisi/homography.gif>, acesso em: 24/11/2014.

A figura 4.3 mostra o resultado de uma homografia aplicada numa imagem.

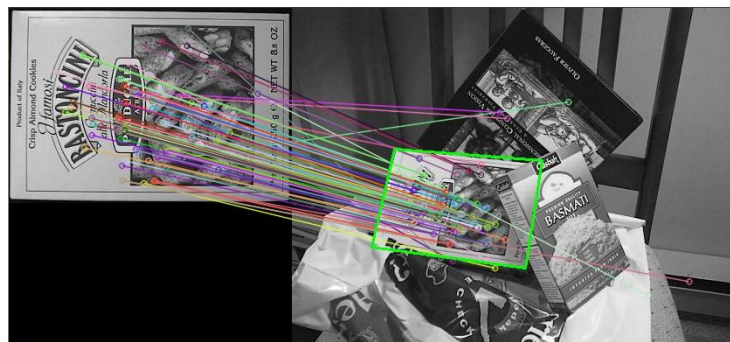


Figura 4.3 - Inserção de um quadro em uma imagem através da homografia. Fonte: docs.opencv.org/_images/Feature_Homography_Result.jpg, acesso em: 24/11/2014.

Uma vez definida a matriz de homografia H , aplica-se uma transformação perspectiva em todos os elementos do plano origem, tratando cada elemento como um vetor 2D ou 3D. As equações (4.2.1.1) e (4.2.1.2) apresenta a fórmula da transformação.

$$(x, y, z) \rightarrow \left(\frac{x'}{w}, \frac{y'}{w}, \frac{z'}{w} \right) \quad (4.2.1.1)$$

Onde,

$$(x', y', z', w') = H [x \ y \ z \ 1] \quad (4.2.1.2)$$

Sendo, H a matriz de homografia, o ponto (x', y', z', w') a transformação do vetor $[x \ y \ z \ 1]$ representado na forma de coordenadas homogêneas.

A partir da matriz de homografia H , os pontos das regiões de elevação dos robôs e objetos, que se encontram na imagem de profundidade, podem ser transformados em pontos na imagem colorida do *Kinect*, onde estão as identificações coloridas dos elementos, o que possibilita sua análise. Além disso, esses pontos podem ser transformados em pontos no mundo possibilitando o rastreo das posições métricas reais dos robôs e objetos na arena.

4.2.2 Aplicação da Metodologia de Calibração

A calibração do posicionamento das câmeras para o Novo AEDROMO, precisará de três matrizes de homografia. Será necessária a transformação do plano de profundidade para as coordenadas do mundo, destas, por sua vez, para o plano da imagem e deste de volta para o plano de profundidade. Para fazer o caminho inverso, basta transformar os pontos com as matrizes de homografia inversas. No momento da calibração, o usuário deverá selecionar os pontos característicos em cada plano. Os pontos no mundo já são conhecidos. A figura 4.4 apresenta um diagrama de atividade mostrando o funcionamento desse procedimento.

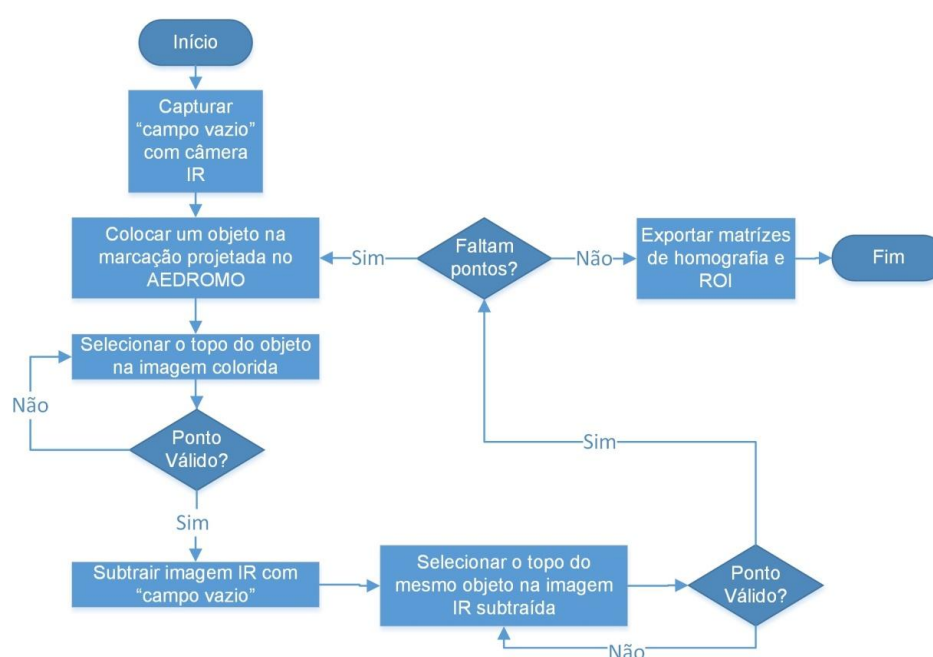


Figura 4.4 - Diagrama de Atividade da Calibração do Posicionamento das Câmeras. Fonte: Autor.

O primeiro passo na calibração do posicionamento das câmeras é salvar as informações de profundidade do Novo AEDROMO sem nenhum robô ou objeto sobre a TV que será chamada de área vazia. Ela servirá como uma máscara, para que qualquer imagem projetada na arena abaixo desta profundidade seja desconsiderada. Em seguida, com um domínio de pontos conhecidos sobre a arena, são gerados os conjuntos imagens da imagem colorida e da imagem de profundidade. Com essas informações, as matrizes de homografia são geradas. As câmeras do *Kinect* capturam mais do que apenas a região da arena, nelas aparecem o mundo ao redor, que não é desejado para o AEDROMO. Cria-se então, nas imagens, uma *region of interest* (ROI) da arena.

4.3 Calibração das Cores

A calibração das cores está voltada à definição das cores de cada robô e objeto, que serão utilizados posteriormente no ambiente. Após definido o número total de robôs e objetos, cada um terá uma cor, ou par de cores, específicas. Desta forma não haverá confusão na detecção e classificação dos objetos durante a utilização do ambiente. O sistema, ou espaço, de cores usado na calibração é o coordenadas esféricas (SCT), que originalmente, em inglês, significa *Sphere Coordinate Transformation*. A figura 4.5 exibe um diagrama do processo de calibração das cores.



Figura 4.5 - Diagrama do funcionamento da Calibração das Cores. Fonte: Autor.

4.3.1 O SCT

O sistema de cores fornecido pela biblioteca² que opera com a câmera do *Kinect* é o RGB. Entretanto, esse sistema apresenta algumas limitações, tais como, a presença da parte negativa dos espectros visíveis, os quais não possibilita a representação de certas cores por meio da sobreposição desses três espectros, a dificuldade de determinar as características das cores, como a presença ou ausência de uma dada cor, e a inabilidade da distância Euclidiana capturar corretamente as diferenças de cores no sistema RGB (SARIFUDDIN; MISSAOUI; 2005).

Para facilitar a quantificação das cores, uma transformação em coordenadas esféricas é necessária. Essa transformação divide o sistema de cores em um espaço bidimensional representado por dois ângulos, α e β , e em um espaço unidimensional de intensidade (luminosidade), representado por um vetor de magnitude L . A figura 4.6 apresenta o sistema de cores RGB transformado (UMBAUGH et al., 1989).

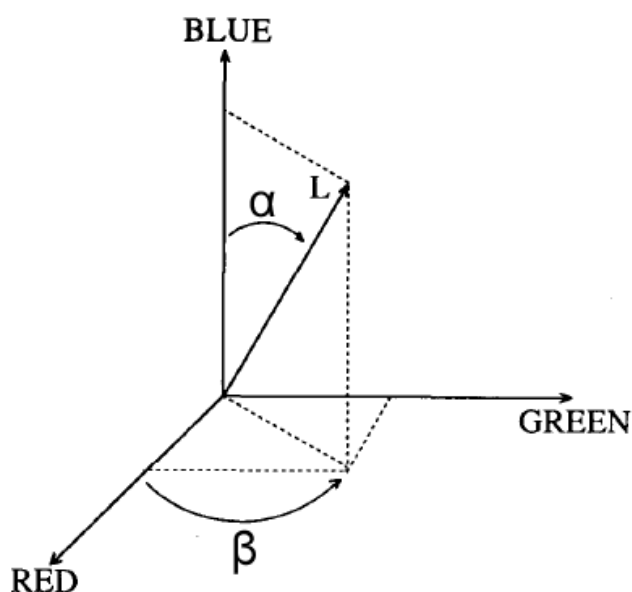


Figura 4.6 - Espaço RGB com a transformação SCT (UMBAUGH et al., 1989).

A quantificação de cores reduz a quantidade de dados necessários a serem processados. Os dois ângulos são usados para definir o espaço que é então quantificado com a ideia de reduzir esse espaço angular bidimensional para um

² Biblioteca Freenect, da comunidade OpenKinect. Fonte: openkinect.org/wiki/Main_Page. acesso: julho/2012.

espaço unidimensional de intensidade (UMBAUGH et al., 1989). A formula da transformação é como se segue:

$$L = \sqrt[2]{R^2 + G^2 + B^2} \quad (4.3.1)$$

$$\alpha = \cos^{-1} \left[\frac{B}{L} \right] \quad (4.3.2)$$

$$\beta = \cos^{-1} \left[\frac{R}{L \times (\sin \alpha)} \right] \quad (4.3.3)$$

Onde (R, G, B) são os valores correspondentes aos espectros de cor vermelho, verde e azul, respectivamente, no espaço de cores RGB e (α , β , L) são os novo valores SCT (UMBAUGH et al. 1989).

Capítulo 5 Resultados e Discussões

Com o objetivo de encontrar os robôs e os objetos passivos numa área de trabalho dinâmica, alguns softwares foram desenvolvidos para cada uma das fases do processo de calibração e utilização, visto no diagrama da figura 2.6.

5.1 Software de Calibração do Posicionamento das Câmeras

A metodologia de calibração do posicionamento das câmeras adotada nesse trabalho foi da homografia. Desta forma, para se obter as matrizes de transformação de homografia é necessário encontrar pontos característicos na imagem de vídeo colorida (*pixels*), na imagem de profundidade (*pixels*) e no mundo do Novo AEDROMO (cm). Os pontos no mundo (cm) são conhecidos e usados como base para encontrar os outros. A figura 5.1 mostra o funcionamento do *software*.

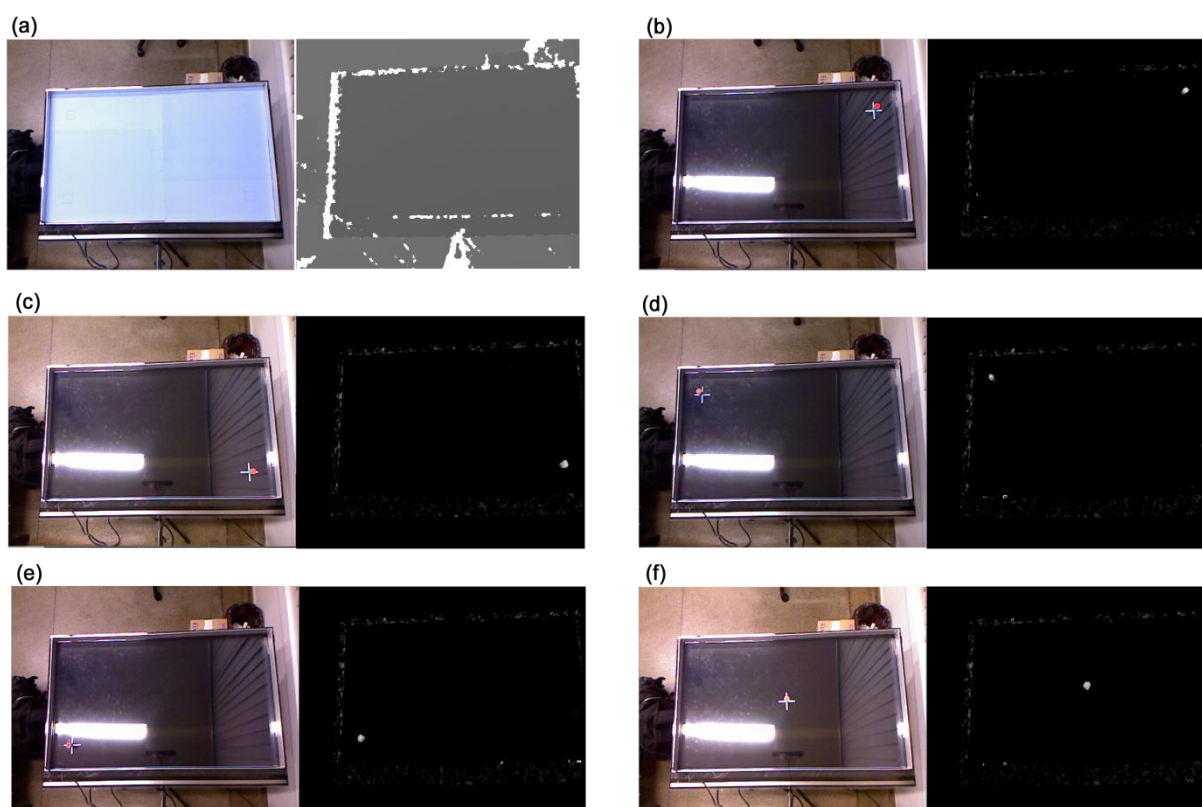


Figura 5.1 - Software de Calibração do Posicionamento das Câmeras. Área vazia (a), Captura dos pontos característicos (b) até (f). Fonte: Autor.

O *software* segue o diagrama da figura 4.4. Inicialmente há a necessidade de capturar a arena sem nenhum robô ou objeto e com uma proteção sobre ela. Isso

será usado numa subtração *pixel-a-pixel* nas imagens de profundidade que virão a seguir. A proteção é necessária pois os raios infravermelhos disparados pelo *Kinect* sofrem um efeito de refração ao atravessar o vidro instalado sobre a TV. A figura 5.2 mostra esse efeito. Mesmo assim, devido ao ângulo de inclinação dos raios, não são todos os pontos que se perdem.

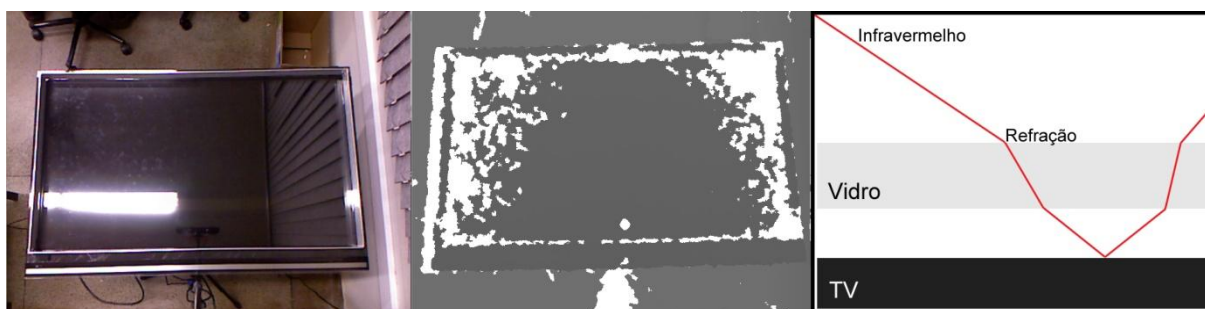


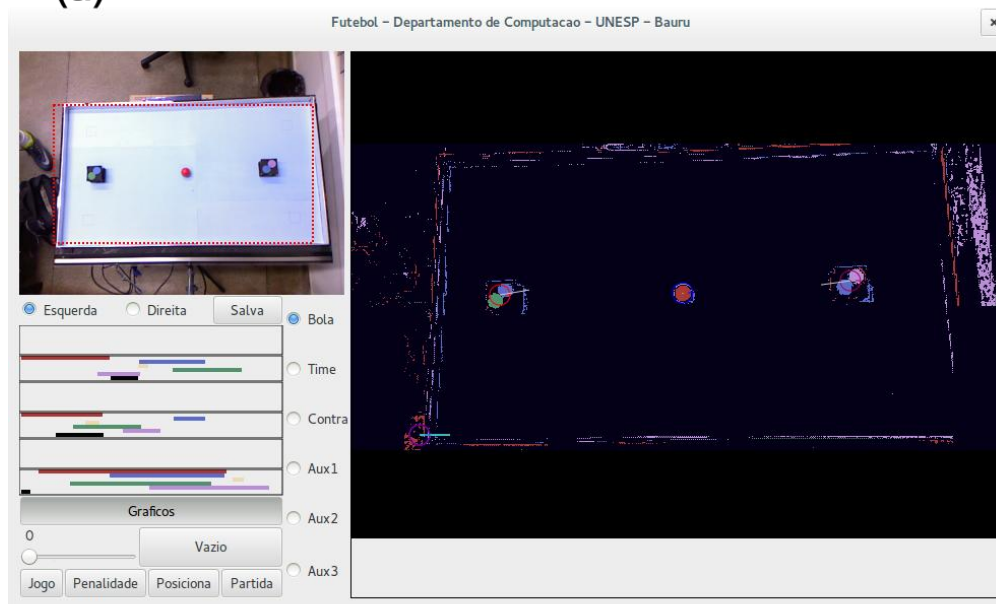
Figura 5.2 - À esquerda a área sem a proteção do vidro. No meio a imagem de profundidade. À direita o efeito físico da refração no AEDROMO. Fonte: Autor.

Após a captura da área vazia, o usuário deve colocar um bloco, de altura entre 4 e 8 cm, sobre o marcador projetado na TV. As posições onde estes blocos devem ser colocados são indicadas pelo sistema na imagem projetada, totalizando cinco pontos característicos. Para cada uma das posições do bloco, o usuário deve clicar no topo do mesmo, em ambas imagens de vídeo e de profundidade para capturar suas coordenadas em *pixels*. A figura 5.1 de (b) até (f) ilustra esse processo. Após todos os pontos característicos terem sido capturados, o *software* exportará para o disco todas as três matrizes de homografia no formato *XML* e a ROI da arena. Essas informações serão usadas na fase de utilização.

5.2 Software de Calibração das Cores

Este *software* de calibração está voltado à definição das cores de cada elemento, podendo ser robôs ou objetos passivos, e segue o diagrama da figura 4.5. Inicialmente, o usuário deverá colocar todos os elementos que serão calibrados na arena. Em seguida, a cor específica de cada elemento é identificada. Conforme as cores forem sendo identificadas, a imagem capturada pelo *Kinect* é gradativamente indexada conforme uma tabela de cores criada em tempo de execução. A figura 5.3 apresenta esse *software*.

(a)



(b)

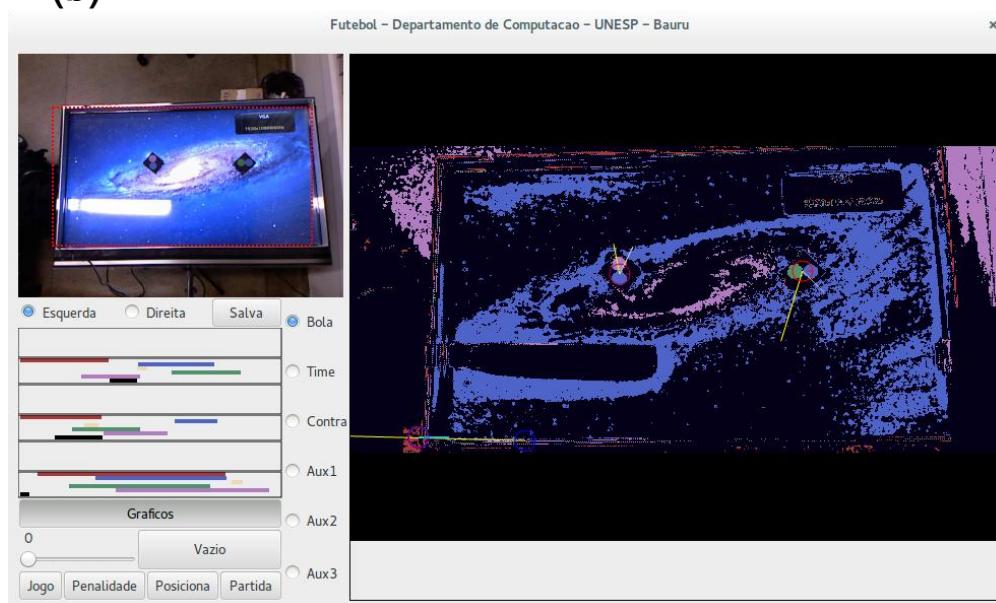


Figura 5.3 - Software de Calibração das Cores. (a) Com proteção, (b) Sem proteção e imagem sendo projetada na TV.
Fonte: Autor.

Na figura 5.3a, foram calibrados dois robôs e um objeto passivo. Na tela menor à esquerda, é exibida as imagens capturadas pela câmera de vídeo do *Kinect*. À direita, encontra-se a imagem correspondente à de vídeo com as cores identificadas na calibração e indexada. As cores calibráveis são divididas em: cor do primeiro robô, cor do segundo robô e cores de três objetos passivos. Todos os robôs possuem uma cor de etiqueta comum e uma segunda cor que de fato irá defini-lo.

Para calibrar um elemento, primeiro seleciona-se seu tipo, ou seja, se é um dos robôs ou um objeto passivo. Em seguida, clica-se repetitivamente sobre a

etiqueta colorida ou a cor desse mesmo elemento na imagem de vídeo. A cada clique o sistema de cores RGB é transformado em SCT. Como podem existir tonalidades diferentes para a mesma cor na imagem de vídeo, um gráfico com as faixas de valores, no formato SCT, é criado e exibido ao usuário. A figura 5.4 apresenta esse gráfico de cores. A partir das informações das faixas de valores, a imagem à direita é indexada. Possibilitando ao usuário verificar se sua calibração foi bem sucedida ou não. Caso a cor calibrada esteja com ruído ou não apareça na imagem indexada, a recalibração daquela cor é necessária.



Figura 5.4 - Gráfico da faixa de valores. Ângulos α e β e luminosidade do objeto de cor laranja. Fonte: Autor.

Uma vez que todos os elementos tenham suas próprias faixas de valor, o *software* exportará esses dados num arquivo de texto para ser usado no momento da utilização do ambiente.

5.3 Utilização do Novo AEDROMO

Um terceiro *software* completa o sistema do lado do servidor do AEDROMO, este *software* quando executado carrega as informações de calibração e inicia o *Kinect*, sendo então capaz de identificar e disponibilizar as posições dos elementos na arena e receber os comandos a serem enviados aos robôs.

O Novo AEDROMO requer que alguma atividade seja elaborada. Uma vez que a atividade tenha sido escolhida e as matrizes de homografia e os parâmetros de cor dos elementos tenham sido definidos, o ambiente estará pronto para uso. Dois *softwares* são necessários: o *software* servidor e o *software* cliente. O servidor enviará para o cliente, a cada nova imagem fornecida pelo *Kinect*, informações

referentes ao estado da arena naquele momento, tais como, as coordenadas em centímetros de cada elemento na arena, seu ângulo de direção e sua cor. A partir da cor, o usuário saberá se aqueles dados recebidos são referentes ao robô ou a algum objeto passivo.

O usuário poderá escrever um programa cliente, que a partir dos dados fornecidos pelo servidor, envia os comandos a serem executados pelos robôs, com o objetivo de alcançar a meta proposta na atividade.

O programa servidor procura na imagem de profundidade capturada pelo *Kinect* todas as áreas de elevação acima da superfície da arena. As áreas de elevação detectadas são elementos. Uma vez que toda a arena tenha sido processada, as regiões de elevação serão transformadas, com o uso das matrizes de homografia, para a imagem colorida de vídeo. Então, na imagem de vídeo, somente os *pixels* dos elementos serão analisados. Todo o resto, ambiente externo e a arena, é desconsiderado. Desta forma, não haverá problema se algo colorido for projetado na arena. A figura 5.5 ilustra esse resultado. Nessa imagem, foram postos dois robôs e uma bola de cor laranja sobre a arena. Desta sequência de operações, com os parâmetros de cor previamente obtidos na calibração, a imagem de vídeo é analisada e uma nova imagem indexada é gerada seguindo esses mesmos parâmetros. A figura 5.6 mostra a imagem final indexada.

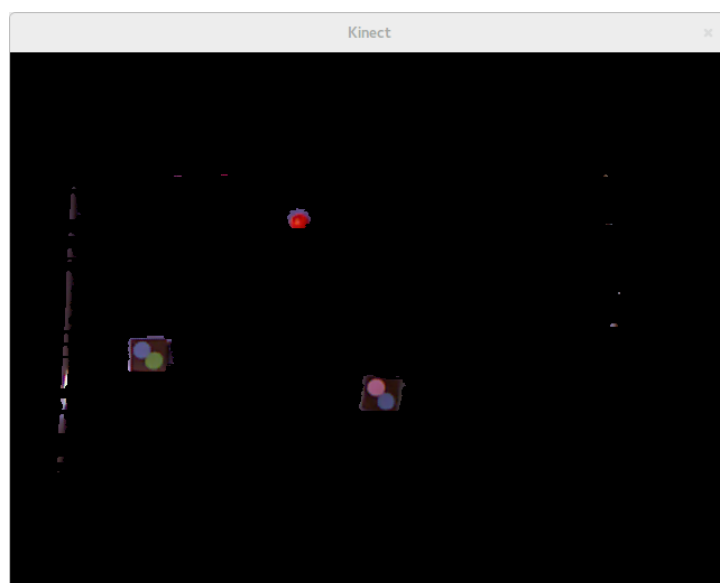


Figura 5.5 - Imagem colorida analisada. Processo de detecção de elevações e aplicação da homografia. Fonte: Autor.

O processo de classificação dos elementos é feito após a indexação da imagem. Pode-se observar que a imagem possui algum ruído. Esses ruídos são tratados por meio de um algoritmo de otimização que verifica as dimensões ideais das etiquetas, as cores próximas e a relação das posições atual e anterior dos objetos. Esse tratamento foi incorporado no AEDROMO antigo e, portanto, não está no escopo do trabalho.

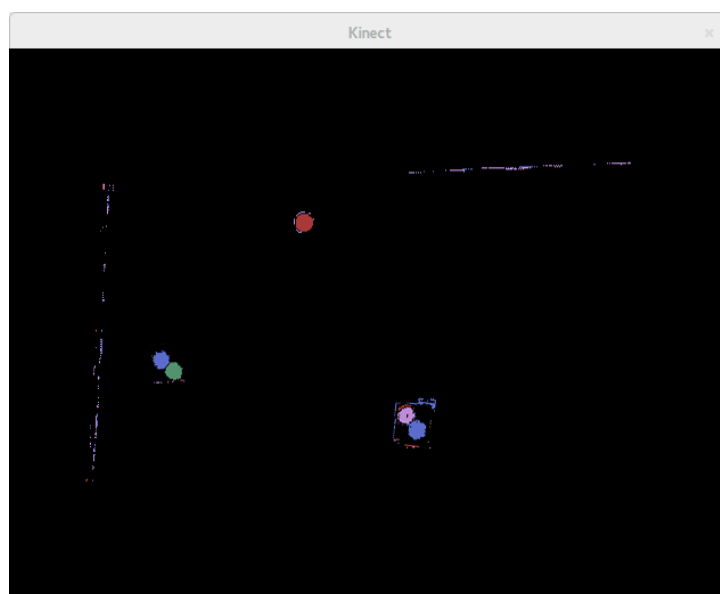


Figura 5.6 - Imagem colorida indexada. Fonte: Autor.

Nesta imagem, o *software* servidor procura todas as etiquetas com as cores calibradas e classifica os elementos. Obtendo assim o estado da arena.

O usuário do AEDROMO deve desenvolver um programa cliente para controlar um robô. O programa do cliente permanecerá com as mesmas características que são usadas no AEDROMO. Os dados das posições dos elementos na arena são enviados via protocolo UDP/IP do programa servidor para o programa cliente. O programa cliente, por sua vez, calcula, a partir das posições dos elementos, o(s) próximo(s) comandos do robô sendo controlado e os envia de volta para o servidor. O servidor então, envia os comandos aos robôs de acionamento via rádio.

Para concluir, os *softwares* de calibração do posicionamento das câmeras, calibração das cores, e utilização do Novo AEDROMO, juntos, formam o novo sistema de identificação dos objetos. As cores da área de trabalho e qualquer outra cor fora do ambiente são desconsideradas.

Capítulo 6 Conclusão

O desenvolvimento do presente trabalho proporcionou a investigação de processos de identificação de objetos em superfícies, ignorando possíveis cores e formas conflitantes no ambiente externo. Essa investigação permitiu o estudo minucioso de métodos existentes para a calibração de câmeras RGB-D (Apêndice B). A viabilidade deste projeto foi comprovada por meio de sua implementação e testes experimentais.

Por meio dos experimentos realizados, observou-se que o dispositivo *Microsoft Kinect*, apesar de ser um sensor de baixo custo, mede o nível de profundidade com boa precisão. Contudo, os raios infravermelhos emitidos pelo *Kinect* na geração de padrões de marcações no ambiente podem sofrer o efeito físico da refração resultando em medidas incorretas e irregulares.

Verificou-se que o método de calibração do *Microsoft Kinect* aplicado a esse trabalho, apresenta resultados satisfatórios para uma distância de aproximadamente um metro entre a superfície e o sensor. Entretanto, esse método de calibração deve ser repetido toda vez que o sensor ou a superfícies mudarem de lugar ou posição. Já a calibração de cores só deve ser refeita caso existam novos objetos interagindo com a tarefa proposta.

Este trabalho de conclusão de curso deu origem a três trabalhos científicos, sendo dois deles apresentados no Congresso de Iniciação Científica da UNESP (CIC 2013 e 2014): "Área de trabalho dinâmica para o ambiente experimental e didático com robôs móveis" (disponível no Apêndice A) e "Avaliação dos Parâmetros Intrínsecos de um Sensor RGB-D" (disponível no Apêndice B). O terceiro trabalho "*Educating Recycling with Robotics*" foi publicado no *2015 International Workshop on Serious Gaming = Serious Business* (disponível no Apêndice C).

Referências Bibliográficas

- ALVES, S. F. R., FERASOLI FILHO, H., PEGORARO, R., CALDEIRA, M. A. C., YONEZAWA, W., ROSÁRIO, J., Ambiente Educacional de Robótica Direcionado a Aplicações em Engenharia. **Anais** do SBAI - Simpósio Brasileiro de Automação Inteligente, 2011.
- AVANZATO, R., Mobile robotics for freshman design, research, and high school outreach. In 2000 IEEE International **Conference on Systems, Man, and Cybernetics**, 2000.
- BARKER, B.S., ANSORGE, J., Robotics as means to increase achievement scores in an informal learning environment. **Journal of Research on Technology in Education**, 2007, p.229-243.
- CRUZ, L., LUCIO, D., VELHO, L., Kinect and RGBD Images: Challenges and Applications. In: **Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2012 25th SIBGRAPI Conference on**. IEEE, 2012. p. 36-49.
- DUBROFSKY, E., **Homography Estimation**. Master's Essay as requirements for the degree of Master of Science. University of British Columbia, Canada. 2009
- ESCALERA, A. D. L., ARMINGOL, J. M., Automatic Chessboard Detection for Intrinsic and Extrinsic Camera Parameter Calibration. **Sensors**, v. 10, n. 3, p. 2027-2044, 2010.
- FAUGERAS, O. D.; LUONG, Q.-T.; MAYBANK, S. J. Camera self-calibration: Theory and experiments. In: **Computer Vision—ECCV'92**. Springer Berlin Heidelberg, 1992. p. 321-334.
- FERASOLI FILHO, H., PEGORARO, R., CALDEIRA, M.A.C., ROSÁRIO, J., AEDROMO- An Experimental and Didactic Environment with Mobile Robots. In: **Proceedings of The 3rd International Conference on Autonomous Robots and Agents**, 2006.
- FREEDMAN, B., SHPUNT, A., MACHLINE, M., ARIELI, Y., **Depth Mapping Using Projected Patterns**. U.S. Patent 2010/0118123, 13 maio 2010.
- HARTLEY, R., ZISSERMAN, A., **Multiple View Geometry in Computer Vision**. Cambridge University Press, second edition, 2003.
- HENRY, P., KRAININ, M., HERBST, E., REN, X., FOX, D., RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments. In: **Proceedings of the 12th International Symposium on Experimental Robotics**, 2010.
- KHOSHELHAM, K., Accuracy Analysis of Kinect Depth Data. In: **ISPRS workshop laser scanning**. 2011. p. W12.

- MACKNOJIA, R., **Design and Calibration of a Network of RGB-D Sensors for Robotic Applications over Large Workspaces**. Thesis submitted to the Faculty of Graduate and Postdoctoral Studies, University of Ottawa, Canada, pg. 25-32. 2013
- MACKNOJIA, R., ARAGÓN, A. C., PAYEUR, P., LAGANIÈRE, R., Calibration of a Network of Kinect Sensors for a Robotic Inspection over a Large Workspace. In: **Robot Vision (WORV), 2013 IEEE Workshop on**. IEEE, 2013. p. 184-190.
- MALIS, E., HAMEL, T., MAHONY, R., MORIN, P., Estimation of Homography dynamics on the special linear group. in **Visual Servoing via Advanced Numerical Methods**, Springer. Chapter 8, pages 139–158, 2010.
- PAPERT, S., **The Children's Machine: Rethinking School in the Age of the Computer**, ISBN 0-465-01830-0, Harper-Collins Publishers, 1993.
- SARIFUDDIN, M.; MISSAOUI, Rokia. A New Perceptually Uniform Color Space with Associated Color Similarity Measure for Content-Based Image and Video Retrieval. In: **Proc. of ACM SIGIR 2005 Workshop on Multimedia Information Retrieval (MMIR 2005)**. 2005. p. 1-8.
- SMISEK, J., JANCOSSEK, M., PAJDLA, T., 3D with Kinect. In: **Consumer Depth Cameras for Computer Vision**. Springer London, 2013. p. 3-25.
- SOLONY, M., ZEMCÍK, P., Scene Reconstruction from Kinect Motion. In: **Conference and Competition Student EEICT** (Electrical Engineering, Information Technology and Communication) [S.l.: s.n.]. 2011.
- STARANOWICZ, A., MARIOTTINI, G. L., A Comparative Study of Calibration Methods for Kinect-style cameras. In: **Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments**. ACM, Crete, Greece, 2012.
- UMBAUGH, S. E.; MOSS, R. H.; STOECKER, W. V. Automatic color segmentation of images with application to detection of variegated coloring in skin tumors. **Engineering in Medicine and Biology Magazine, IEEE**, v. 8, n. 4, p. 43-50, 1989.
- WEBB, J., ASHELEY, J., **Beginning Kinect Programming With the Microsoft Kinect Sdk**. [S.l.]: Apress,. 2012. 51 p.
- WILSON, A. D., Using a depth camera as a touch screen. In **ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)**. ACM, New York, NY, USA, 69-72. 2010.
- ZHANG, Z., A Flexible New Technique for Camera Calibration. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 22, n. 11, p. 1330-1334, 2000.

ZHANG, Z., Camera Calibration. In: MEDIONI, G.; KANG, S. B. **Emerging Topics in Computer Vision**. Prentice Hall Professional Technical Reference, 2004. Chapter 2, p. 5-43.

**Apêndice A - Artigo Apresentado no Congresso de
Iniciação Científica da UNESP 2013**

ÁREA DE TRABALHO DINÂMICA PARA O AMBIENTE EXPERIMENTAL E DIDÁTICO COM ROBÔS MÓVEIS

Azael de Melo e Sousa, René Pegoraro, Campus de Bauru, Faculdade de Ciências, Bacharelado em Ciência da Computação, azael_gel@hotmail.com, Bolsa PROPe – Projeto Renove

Palavras Chave: AEDROMO, RENOVE, KINECT

Introdução

O AEDROMO é um Ambiente Experimental Didático com Robôs Móveis, formado por robôs, objetos passivos, dois computadores, um dispositivo de captura dos objetos presentes em uma área de trabalho (arena) e um transmissor que envia comando aos robôs. Desde criação deste ambiente em 2006, os objetos são identificados por suas cores, assim a arena não pode ser colorida livremente.

Explorou-se a integração dos dados dos sensores de profundidade e imagem visível do Kinect da Microsoft para detectar objetos por meio da distância e cores, permitindo desenhos coloridos na arena.

Objetivos

Desenvolver um novo sistema de identificação dos objetos sem a influência das cores da estampa na arena.

Material e Métodos

O projeto contou com o AEDROMO como base e um sensor de profundidade Kinect para obter informações da arena e assim identificar os objetos sobre a superfície. Este sensor é instalado a uma altura aproximada de um metro sobre o plano da arena e captura as informações de distância do sensor até ela. Objetos entre o sensor e este plano são identificados por variações nas distâncias medidas. Estas medidas, após transformações geométricas tridimensionais, são convertidas em valores métricos das coordenadas sobre a área de trabalho e junto com suas cores identificam os objetos (Figura 1). Antes do uso, é necessário fazer uma calibração para identificar os parâmetros de configuração do ambiente e sensor. Na calibração, identifica-se o plano e os limites da arena. Com estes dados, calcula-se uma máscara que será utilizada como base durante o uso.

Resultados e Discussão

O ambiente apresenta agora um mecanismo que permite a troca dos desenhos, mesmo coloridos, na área de trabalho, possibilitando maior versatilidade ao ambiente. A aplicação do futebol no AEDROMO, utiliza agora um fundo decorado com a imagem de um gramado, que é distinguido das etiquetas sobre

os robôs pela variação de distância ao sensor Kinect. Uma nova aplicação está sendo desenvolvida sobre uma imagem do planisfério, onde os robôs se deslocam entre regiões, com o objetivo de explorar o emprego da robótica no ensino de geografia. Novamente, as cores do planisfério não interferem no reconhecimento dos robôs.

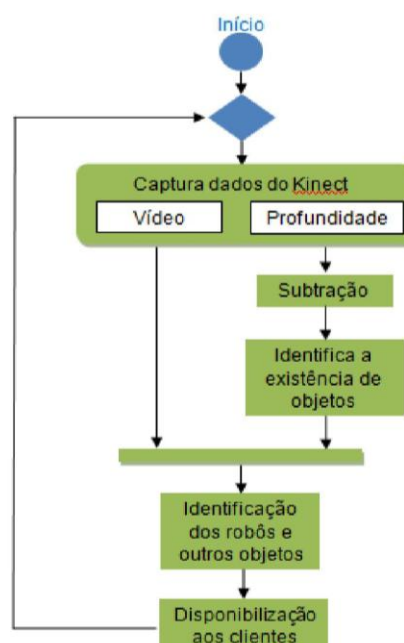


Figura 1. Primeiramente captura os dados do Kinect. A imagem de profundidade sofre uma subtração pixel-a-pixel, com uma máscara previamente gerada. Com isso é possível encontrar as variações nas medidas e consequentemente todos objetos. Comparando com a imagem do vídeo, identifica-se os robôs.

Conclusões

O ambiente apresenta agora um mecanismo que permite a troca dos desenhos na arena, possibilitando maior versatilidade ao ambiente. Com este desenvolvimento, novas atividades são possíveis no AEDROMO, pois a identificação dos objetos independe das cores utilizadas sobre a arena.

Agradecimentos

Pró-Reitoria de Pesquisa da UNESP – Projeto Renove e Empresa MSTech pelo empréstimo do sensor Kinect.

**Apêndice B - Artigo Apresentado no Congresso de
Iniciação Científica da UNESP 2014**

Avaliação dos parâmetros intrínsecos de um sensor RGB-D.

Azael de Melo e Sousa, Renê Pegoraro, Campus de Bauru, Faculdade de Ciências, Bacharelado em Ciência da Computação, azael_gel@hotmail.com

Palavras Chave: Kinect, Calibração, RGB-D, Parâmetros, Intrínsecos

Introdução

O Ambiente Experimental e Didático com Robôs Móveis (AEDROMO)¹ usa uma arena plana horizontal de 80cmx60cm onde robôs interagem, nele os robôs são identificados através de visão global. Continuando o desenvolvimento deste ambiente pretende-se utilizar um dispositivo RGB-D para identificar os robôs e permitir a interação dos usuários. O Kinect da Microsoft foi o dispositivo RGB-D escolhido neste trabalho. Ele é composto por duas câmeras, uma colorida e uma de profundidade². Nele os parâmetros intrínsecos, usados na conversão do sistema de coordenadas da câmera para o da arena, precisam ser calibrados. Um método de calibração para o Kinect² frequentemente usado é o de Zhang³ implementado no openCV. A utilização dos parâmetros intrínsecos obtidos por este método demanda cuidados pois estes parâmetros dificilmente se repetem. Macknoja (2013) coloca como solução a escolha de 10 experimentos realizados com 30 imagens escolhidas aleatoriamente em um conjunto de 100 imagens.

Objetivos

Verificação dos parâmetros de um câmera RGB-D utilizada para identificação global no AEDROMO.

Material e Métodos

Neste trabalho, usou-se a arena do AEDROMO e o sensor de profundidade do Kinect posicionado-o a uma altura de 1m sobre ela.

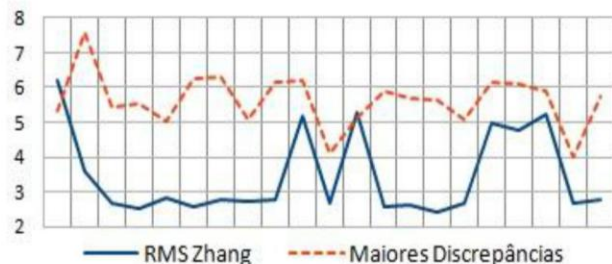
Inicialmente, capturou-se 100 imagens da área de trabalho do AEDROMO, então, seguindo a proposta de Macknoja (2013), calibrou-se 100 vezes, com 30 imagens escolhidas aleatoriamente de cada vez, obtendo 100 conjuntos de parâmetros intrínsecos que foram avaliados quanto aos seus erros. Sobre a arena, alguns blocos com 4cm de altura foram postos um em cada canto, e outro no centro com altura de 25cm, totalizando 5(cinco) blocos em diferentes posições. O topo dos blocos são identificados através do Kinect, obtendo suas coordenadas no plano da câmera e a profundidade em cm de cada bloco. As distâncias entre os topos de dois em dois blocos são usadas para calcular os erros utilizando a equação de projeção da câmera,

uma vez que se conhece as posições dos blocos tanto nas imagens quanto no ambiente real. Foram calculados os erros de duas formas. A primeira RMS (Root Mean Square) com o método do Zhang (openCV). A segunda usando a discrepância entre as distâncias de dois em dois blocos com o maior erro de cada conjunto de parâmetro intrínseco.

Resultados e Discussão

O **Gráfico 1** mostra a comparação dos resultados obtidos pelo método RMS de Zhang e o desenvolvimento nesse trabalho.

Gráfico 1. Erros RMS do Zhang e maiores discrepâncias.



Observa-se no gráfico que os menores erros obtidos pelo método do Zhang, para os parâmetros intrínsecos, não representam os parâmetros nos quais se obtêm as menores discrepâncias entre o medido manualmente no ambiente e o calculado pela equação de projeção.

Conclusões

Com os resultados obtidos, especificamente para o AEDROMO, a identificação dos parâmetros intrínsecos pelas menores discrepâncias, pode representar uma melhor forma de escolha. Ou seja, os melhores parâmetros intrínsecos, para essa determinada situação são aqueles cuja diferença entre o erro calculado pelo método do Zhang e as discrepâncias foram menores.

¹ Ferasoli Filho, H., Pegoraro, R., Caldeira, M.A.C., Rosário, J., AEDROMO - AN Experimental and Didactic Environment with Mobile Robots. **2006**.

² Macknoja, R., Aragón, A. C., Payeur, P., Laganière, R. Calibration of a Network of Kinect Sensors for a Robotic Inspection over a Large Workspace. **2013**, Canada.

³ Zhang Z., Camera Calibration. In: Medioni, G., Kang, S. B., Emerging Topics in Computer Vision. **2004**, Chapter 2, p. 4-43.

**Apêndice C - Artigo Apresentado no *Workshop
Serious Games = Serious Business* 2015**

Educating Recycling with Robotics

Azael de Melo e Sousa
Computer Department.
Universidade Estadual
Paulista
Bauru, São Paulo, Brazil
azaelsousa@gmail.com

Silas F. R. Alves
Electrical and Computer Eng.
Dep.
São Paulo University
São Carlos, São Paulo, Brazil
salves@usp.br,

Celso de Oliveira Lisboa
Computer Department
Universidade Estadual
Paulista
Bauru, São Paulo, Brazil
celso.lisboa@yahoo.com.br

Renê Pegoraro
Computer Department
Universidade Estadual
Paulista
Bauru, São Paulo, Brazil
pegoraro@fc.unesp.br

Abstract— Just like computers, robotics is a technological resource used in the educational process which can contribute to the student's cognitive development and specific intellectual skills. It is offered as an interesting pedagogic tool [1]. Therefore, the use of robotics in education is, mainly, motivated by the active learning which it provides and happens in two aspects. Firstly, teaching particular and general concepts of subjects linked with robotics itself. And secondly, to teach and examine concepts of related fields, apart from robotics [2]. Considering the educational factors of robotics and the importance of correctly recycling nowadays due to the lack of world's naturally resources, this project presents the development of an interactive workspace where children may control mobile robots in order to collect, separate and drag object-like garbage to their specific rubbish bins located in the waste disposal complex.

Keywords—recycling; robotics; serious games.

I. INTRODUCTION

Solid waste is essentially garbage. Waste is produced in people's homes, businesses and some industrial sources. Solid waste production is growing in volume and toxicity. More and more products contain toxic chemicals, such as mercury or flame retardant chemicals (PBDEs) and many others which impact public health and the environment. In order to change this reality we live in, people have to be educated accordingly, but sometimes children might find this lessons boring, tiresome and not take them seriously. For this reason, this paper proposes the use of a workspace full of mobile robots and passive objects as a interactive tool for learning to recycle.

There are many literature describing the use of robotics in education, such as [3], [4] and [5]. In each one, an objective is established, focusing in collaborative work, math and science enhancement, improvement in team work which can be expanded beyond the classroom to include much larger projects and distributed teams. At university level computer science and engineering concepts are reinforced due to its intrinsically multidisciplinary [8]. Robots have also been used to teach non-majors computer science, for instance, introducing programming to teachers pursuing a degree in Educational Technology [6].

The successful use of robotics in plenty applications related to education motivated this work. This paper presents the continuous development of an ongoing project of [7] but with a slightly difference in the computer vision and image projection.

II. DEVELOPMENT

The goal of this work is to develop an interactive environment where robots can collaborate with each other and with passive objects in the workspace, which projects decorative images that changes according to the given task.

To create this environment, the Robotelo mobile robot (Fig. 1) was adopted. It has a two colored labels on the top, which are used in its detection, classification and localization. In order to separate these labels from the projected images, the Microsoft Kinect's depth sensor was used.

The Kinect sensor was installed about 1 meter high from the workspace. According to Wilson [9], the Kinect can scan the whole workspace, with its depth sensor, and encounter every object taller than the field. Once all objects have been detected, they are properly classified by their colors, and their coordinates are sent back to the user's application. Thus, it is the user's application that guides the robot in order to accomplish the goal. Since the colors of the workspace aren't considered, images can be projected at will by a TV located right under the workspace. Fig 2 shows this environment.

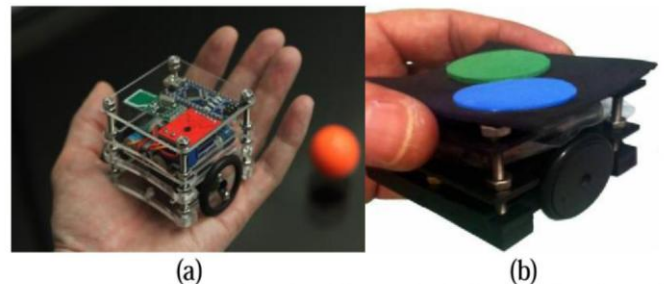


Fig. 1. Images of Robotelo robot. (a) only the skeleton of the robot and (b), attached with a label.

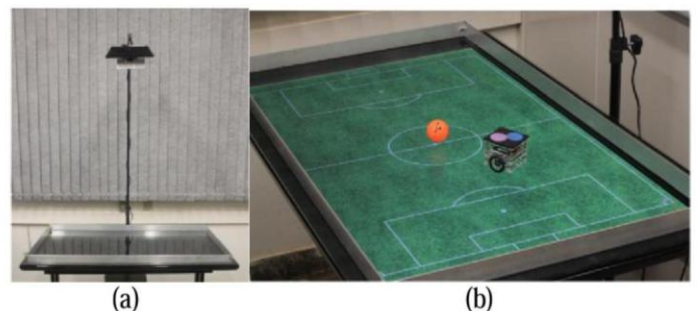


Fig. 2. (a) Workspace with the TV and the Microsoft Kinect. (b) Workspace ready for a football match.

Using an homography method of transformation [10], it is possible to map both depth and video images from Kinect to each other and to the world coordinates of the Workspace. As a result, all pixels, captured through depth image, belonging to objects will be classified based on the colors of those same pixels transformed to the colored image. Fig 3 displays an diagram of how the identification system works.



Fig. 3. Diagram of the identification system operation.

After the first scan, everything that is not an object is disregarded, though there is still noise. This noise is reduced by a post-optimization algorithm.

III. GAME DESIGN

Playing, especially during childhood, performs important roles in psychological, social and intellectual development; is a voluntary activity that is intrinsically motivating; involves some level of activity and often possesses ludic qualities. Such attributes are similar to those contained in modern educational theories where learning should be a self-motivated and rewarding activity [11]. In this work, the objective of the game is to control Robotelo to collect, separate and drag all objects to the rubbish bins located in the solid waste disposal complex represented by the gray building on Fig.4. The object-like trash will be scattered around the workspace. Also, there will be places like forests and parks in the scenario. If the student takes or leaves garbage in one of those areas, he will get less points in the final score. On the other hand, since the complex don't have height, it will not be considered a real object, as a consequence, both real and virtual objects can be used during the game. Every player will have to perform the task within a time limit. The final score is measured by the right actions, time taken for the robot to perform the activity and the amount of trash thrown away. The most difficult aspect of the game is to elaborate the proper path that will guide the robot.

The user controls Robotelo through a client application which receives the coordinates of all objects in the area from the server. Through simple command lines written by the player, like "goto(x,y)", where x and y are the next coordinates the robot has to go, the client application sends back the robot activation command to the server and then, the server controls the robots. The communication between the client and the server is made through an Ethernet network and UDP/IP protocol, allowing a software to control each robot individually. Fig. 5 shows this Control Architecture.



Fig. 4. Example of the game scenario projected in the workspace.

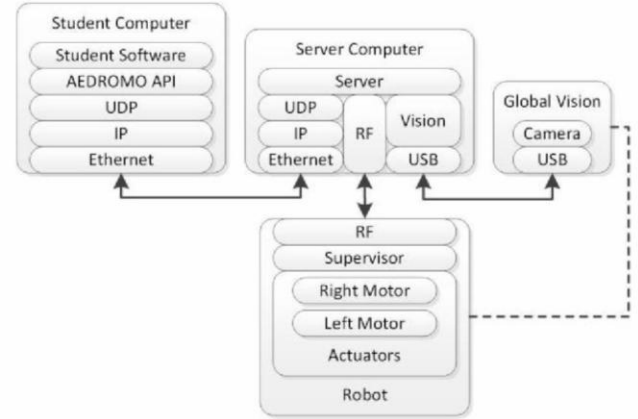


Fig. 5. Client/Server Communication Diagram

IV. FINAL DISCUSSION

The presented environment allow the creation of several games. A recycling game was proposed to teach children the importance of preserving our world's natural resources.

Due to the flexibility and amount of sensors in the workspace, a variety of activities can be elaborated, not only social related like recycling. Tasks based on basic education, such as history (project the world map and pretend the robots are nations or ships) and logic (find your way through a maze), can be executed without the addition of extra sensors.

REFERENCES

- [1] S. Paper, "The Children's Machine: Rethinking School in the Age of the Computer.", ISBN 0-465-01830-0, Harper-Collings Publishers, 1993.
- [2] B. S. Barker, J. Ansorge, "Robotics as Means to Increase Achievement Scores in an Informal Learning Environment". in *Journal of Research on Technology in Education* , 2007, pp. 229-243.
- [3] D. Brigitte, S. Hubert. "Collaborative learning in an educational robotics environment." *Computers in Human Behavior* 17.5 (2001): pp. 465-480.
- [4] G. Rachel, A. Eguchi, and E. Sklar. "Using educational robotics to engage inner-city students with technology." *Proceedings of the 6th international conference on Learning sciences*. International Society of the Learning Sciences, 2004.
- [5] R. D. Beer, H. J. Chiel, and R. F. Drushel. "Using autonomous robotics to teach science and engineering." *Communications of the ACM* 42.6 (1999): 85-92.
- [6] K. J. O'Hara, J. S. Kay. "Investigating open source software and educational robotics." *Journal of Computing Sciences in Colleges* 18.3 (2003): 8-16.
- [7] H. F. Filho, R. Pegoraro, M. A. C. Caldeira, and J. M. Rosario, "AEDROMO: An Experimental and Didatic Environment with Mobile Robots." in *ICARA*, 2006, pp. 403-408
- [8] R. Avanzato, "Mobile robotics for freshman design, research, and high school outreach". In *2000 IEEE International Conference on Systems, Man, and Cybernetics*.
- [9] A. D. Wilson, "Using a Depth camera as a Touch Screen". In *AMC International Conference on Interactive Tabletops and Surfaces (ITS'10)*. ACM, New York, NY, USA, 2010, pp 69-72.
- [10] E. Dubrofsky, "Homography estimation." *Diss. Univerity of British Columbia (Vancouver)*, 2009.
- [11] A. Amory, et al. "The use of computer games as an educational tool: identification of appropriate game types and game elements." in *British Journal of Educational Technology*, 30.4 (1999): 311-321.