

Representação contextual das atividades no ambiente experimental e didático com robôs móveis

Contextual representation of the activities in the experimental and didactic environment with mobile robots

DOI:10.34117/bjdv7n5-504

Recebimento dos originais: 07/04/2021

Aceitação para publicação: 21/05/2021

Azael de Melo e Sousa

Mestre em Ciência da Computação

Endereço: Av. Eng. Luiz Edmundo C. Coube, 14-01, Bauru - Sp. 17033-360

E-mail: azaelmsousa@gmail.com

Renê Pegoraro

Doutor em Engenharia Elétrica

Departamento de Computação, Faculdade de Ciências, UNESP

Endereço: Av. Eng. Luiz Edmundo C. Coube, 14-01, Bauru - Sp. 17033-360

E-mail: rene.pegoraro@unesp.br

Silas Franco dos Reis Alves

Doutor em Engenharia Elétrica

Software Informatics Research Centre Ontario Tech University

Endereço: 2000 Simcoe Street North, Oshawa, Ontario, Canada L1G 0C5

E-mail: Franco.Alves@ontariotechu.ca

Humberto Ferasoli Filho

Doutor em Engenharia Elétrica

Departamento de Computação, Faculdade de Ciências, UNESP

Endereço: Av. Eng. Luiz Edmundo C. Coube, 14-01, Bauru - Sp. 17033-360

E-mail: ferasoli@fc.unesp.br

RESUMO

No Aedromo (Ambiente Experimental e Didático com Robôs Móveis) robôs móveis e objetos passivos interagem em uma superfície denominada de palco. Esse ambiente tem o objetivo de permitir o controle de robôs de forma simples, na realização de atividades propostas. Para detectar e identificar os robôs e objetos, são usadas etiquetas coloridas e processadas por um sistema de visão global sobre o palco. Logo, o palco não pode ter cores conflitantes com as de identificação dos robôs e objetos. Este trabalho apresenta uma evolução ao sistema de visão de forma que as cores do palco não interfiram na localização e identificação dos robôs e objetos, possibilitando o uso de etiquetas coloridas no palco. O sistema de visão utiliza uma abordagem para a detecção dos robôs e objetos através de relevos e cores, utilizando o sensor Microsoft Kinect. Isso torna o ambiente mais versátil, uma vez que as atividades podem ser melhor ilustradas e compreendidas..

Palavras-chave: Ambiente Educacional, Aedromo, Rastreamento, Visão Computacional.

ABSTRACT

In the Aedromo (Experimental and Didactic Environment with Mobile Robots) mobile robots and passive objects interact on a surface called stage. This environment's goal is to allow a direct control of robots, for the fulfillment of proposed activities. To detect and identify robots and objects, colored labels are used and processed by a global vision system. Thus, the stage cannot have conflicted colors with those of the robots and objects. This work presents an evolution to the vision system in which the colors of the stage do not interfere with the location and identification of the robots and objects, allowing the use of colorful pictures in the stage. The novel vision system uses an approach in the detection of robots and objects through reliefs and colors using the Microsoft Kinect sensor. This makes the environment more versatile once the activities can be better illustrated and understood.

Keywords: Educational Environment, Aedromo, Tracking, Computer Vision.

1 INTRODUÇÃO

O Aedromo (Ambiente Experimental e Didático com Robôs Móveis) é um ambiente que suporta uma ampla gama de experimentos. Em uma área plana e horizontal, semelhante a uma mesa, chamada de palco, robôs e objetos interagem entre si para realizar inúmeras atividades. Os objetivos deste ambiente são tornar mais lúdica e concreta a introdução ao ensino de linguagens de programação e robótica; servir de bancada para validações de pesquisas com robôs móveis ou, meramente, entretenimento.

O Aedromo é um projeto em desenvolvimento contínuo, formado por uma área de trabalho, objetos passivos, dois robôs (ou mais), dois computadores, uma câmera global do tipo webcam e um transmissor (FERASOLI FILHO et al., 2006), conforme ilustra a Figura 1. A flexibilidade de adaptação, uso por várias disciplinas e ciclos de ensino, e o baixo custo são as premissas originais na concepção e desenvolvimento deste ambiente. Com o intuito de facilitar e incentivar o uso deste ambiente, as suas dimensões também foram pontos norteadores desta pesquisa. O robô utilizado no Aedromo pode ser visto na Figura 2.

Conforme será discutido na Seção 2, este ambiente permite que sejam realizadas diferentes atividades didáticas onde o robô pode ser programado para mover objetos coloridos no palco de cor neutra. A neutralidade da cor do palco é importante, uma vez que o sistema reconhece a posição do robô e dos objetos por meio de segmentação de cores.

A motivação deste trabalho é melhorar a experiência visual do usuário, permitindo uma melhor compreensão da tarefa a ser realizada pelos robôs, por meio de estampas coloridas apresentadas no palco do Aedromo.

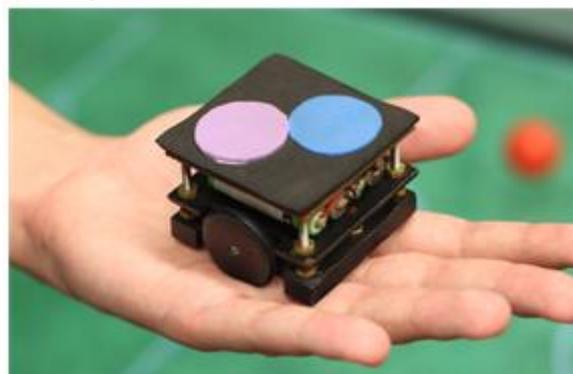
No entanto, a impossibilidade de utilizar cores no palco diminui as capacidades lúdicas do ambiente, pois impede que imagens coloridas (como um mapa) sejam estampadas no palco. Para sobrepor esta dificuldade técnica, o objetivo deste trabalho é o desenvolvimento de um sistema de identificação de elementos (robôs e objetos), usando cores e relevos, permitindo a identificação dos elementos, sem a influência das cores da estampa, sobre o palco do Aedromo.

Portanto, este artigo apresenta, na Seção 2, o ambiente Aedromo e, na Seção 3, o método empregado. O sistema desenvolvido é descrito na Seção 4, enquanto as Seções 5 e 6 apresentam respectivamente a discussão e as conclusões.

Figura 1 - Aedromo com: webcam; robôs, uma bola e palco. Fonte: Autor.



Figura 2 - Robô do Aedromo. Fonte: Autor.

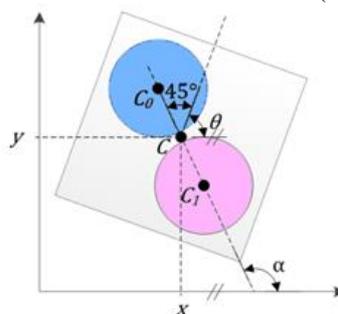


2 AEDROMO

Na arquitetura do Aedromo há dois tipos de computadores: o Computador do Servidor e o Computador do Aluno funcionando como cliente do ambiente. O Computador do Servidor é responsável pela implementação do sistema de visão computacional e pela comunicação com os robôs. Já o Computador do Aluno é onde o aplicativo de controle (Software do Aluno) é desenvolvido. Este aplicativo é um cliente que recebe do servidor as posições cartesianas de todos os objetos presentes no palco, calculadas pelo sistema de visão computacional, e envia ao servidor os comandos de acionamento de cada robô, para que o servidor encaminhe estes comandos ao robô. A comunicação entre o cliente e o servidor é dada através de uma rede Ethernet utilizando o protocolo UDP/IP (*User Datagram Protocol over Internet Protocol*), permitindo que seja escrito um aplicativo de controle para cada robô em qualquer linguagem de programação que suporte “sockets”. Adicionalmente, os aplicativos de controle podem ser executados no mesmo computador ou em diferentes máquinas.

As imagens capturadas pela câmera global são enviadas ao computador servidor. Neste computador, as imagens são adquiridas e processadas. O processamento determina as posições cartesianas (x, y) e o ângulo dos objetos (robôs, bolas, blocos, etc.) inseridos no palco (Alves et al., 2011). Para facilitar o reconhecimento do robô, foram adotados marcadores coloridos que consistem em duas etiquetas de cores diferentes C_0 e C_1 coladas sobre os robôs, como apresentado na Figura 3. Uma vez que o servidor adquire a imagem da câmera, ele é capaz de reconhecer estas cores como regiões e determinar o ponto central de cada uma delas. Com as coordenadas cartesianas de cada etiqueta C_0 e C_1 o servidor calcula a coordenada C , dado pelo ponto médio entre C_0 e C_1 , e o ângulo de orientação do robô T , dado 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).

Figura 3 - Posicionamento do robô reconhecido (ALVES et al., 2011)



Outros objetos passivos (bolas e blocos), que compõem o ambiente, também podem ser identificados no palco e suas posições calculadas através de suas cores ou por etiquetas colocadas sobre eles.

Estas informações de posição processadas pelo servidor são disponibilizadas ao aplicativo de controle no computador cliente através de uma API (*Application Programming Interface*) para desenvolvimento de clientes do Aedromo, que contém os métodos necessários para estabelecer a comunicação com o servidor. Estas informações são utilizadas pelo módulo de estratégia do aplicativo de controle de acordo com as regras da aplicação em execução. Então, comandos são encaminhados ao servidor para serem enviados aos robôs, que conseqüentemente agem no ambiente segundo estas regras. A atuação dos robôs é vista pela câmera e o processo se repete numa taxa equivalente à velocidade de captura da câmera, que em geral varia de 10 a 20 *fps* (*frames per seconds*).

Alves et al. (2011) apresenta diversas opções de aplicações para o Aedromo. A capacidade de reconhecer dois robôs e vários objetos coloridos através da câmera possibilita atividades que envolvam sistemas colaborativos e multiagentes.

O objetivo das atividades propostas com o Aedromo é experimentar aspectos que vão desde os mais simples como reconhecimento do ambiente, passando pelo controle dos robôs até os mais complexos da robótica e da computação.

Incontáveis atividades podem ser desenvolvidas neste ambiente, apenas a termo de ilustração, citamos as atividades: Colecionando Coisas, Caça e Caçador e Resolvendo o Labirinto.

A atividade, intitulada Colecionando Coisas, trabalha os aspectos básicos do ambiente. No palco são colocadas várias bolas de cores diferentes. A tarefa é recolher determinadas bolas de mesma cor (Coisa) para uma área reservada indicada no palco. O tempo para recolher as bolas é o objetivo deste trabalho. Existem duas situações para esta aplicação. A primeira se caracteriza por um robô por vez trabalhando no ambiente e a segunda, com dois robôs. Principalmente, a segunda situação é rica, devido ao dinamismo imposto pelo ambiente e as possibilidades de variantes.

Na atividade Caça e Caçador, os dois robôs são colocados no ambiente e com objetivos diferentes. O primeiro foge do segundo enquanto o segundo tenta pegar o primeiro. O ambiente pode ser enriquecido com objetos para facilitar a fuga ou a caça.

Uma atividade que pode envolver disciplinas de Teoria dos Grafos ou Inteligência Artificial é Resolvendo o Labirinto. No palco são colocados obstáculos (reais ou virtuais), propiciando caminhos alternativos até uma determinada área. O objetivo está em

encontrar esta área e retornar para a origem no menor tempo ou pelo menor caminho. Neste caso, somente um robô é utilizado.

Em cada uma destas aplicações, apesar das dimensões serem mantidas entre elas, existe a necessidade do palco ser decorado com estampas coloridas específicas para cada uma, permitindo que os usuários tenham uma representação mais concreta da tarefa e entenda melhor as finalidades das atividades. Porém, as cores destas estampas não podem conflitar com as dos robôs e objetos.

Caso haja similaridade entre as cores das estampas com as dos robôs, o sistema de visão e identificação poderá detectar as etiquetas dos robôs em lugares inesperados, resultando em informações incorretas sobre os estados dos robôs e causando anomalias no controle e na realização das atividades.

3 MÉTODOS

A aplicação de uma estampa colorida para cada atividade no Aedromo poderia ser feita manualmente com a substituição da superfície do palco. Porém, esta aplicação manual gera dificuldades, pois exige a impressão e fixação, e limita a troca em atividades onde sequências de decorações podem melhorar o entendimento da atividade. Portanto, a projeção de imagens coloridas e decorativas no palco pode facilitar a utilização do ambiente.

Para que as decorações possam ser aplicadas de forma automatizada, uma TV LCD passa a ser o palco, assim, nesta TV são projetadas as estampas dedicadas para cada atividade. 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 identificação dos elementos sobre o palco baseou-se na abordagem proposta por Wilson (2010). Nela é proposta a utilização de um sensor de profundidade *Kinect* da *Microsoft* para obter informações da aproximação de mãos humanas sobre uma 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 o palco. 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 gerencia os dados do *Kinect* e

funciona como o servidor do Aedromo. Após o tratamento destes dados, como a identificação dos elementos no palco, as coordenadas destes elementos são disponibilizadas à uma máquina cliente, que definirá os comandos a serem executados pelos robôs.

A Figura 4 apresenta a aparência da nova estrutura do Aedromo.

3.1 MICROSOFT KINECT

Em 2010, a *Microsoft*, em conjunto com a *PrimeSense*, apresentou o *Kinect* para o *Xbox 360*, um sensor de baixo custo que captura imagens tanto de profundidade como colorida com uma taxa de 30 *fps*, dedicado à aplicação em jogos de videogame.

Figura 4 - Aedromo com sensor *Kinect* e TV. Fonte: Autor.



As câmeras de profundidade e de imagem colorida do *Kinect* não são alinhadas de fábrica o que exige um ajuste de sobreposição entre as imagens. Para que esta sobreposição funcione com precisão este ajuste deve ser realizado para cada *Kinect*.

Sensores, tais como o *Kinect*, podem ser usados 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 a imagem colorida, a imagem de profundidade e as coordenadas do mundo (Staranowicz; Mariottini, 2012).

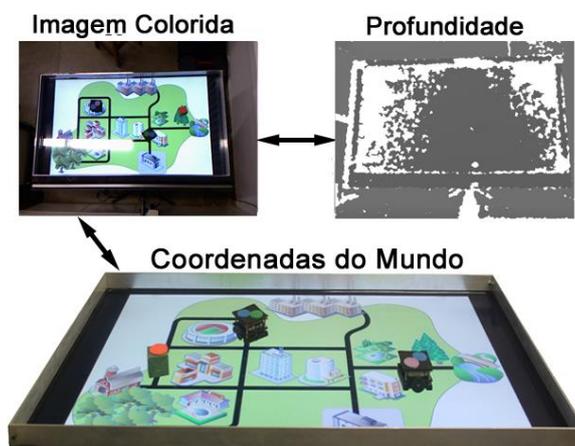
O sensor de profundidade do *Kinect* é utilizado no Aedromo para encontrar objetos no palco por meio de elevações, ou seja, regiões acima do plano do palco. Uma vez identificadas as regiões de elevação, somente estas regiões tem importância na imagem colorida do *Kinect*. Caso, na imagem colorida, ao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

A transformação de coordenadas das regiões de elevação na imagem de profundidade para as mesmas na imagem colorida e destas para as do mundo é feita através de homografia que realiza um ajuste de perspectiva. Para que este ajuste funcione coerentemente, entre as imagens do sensor e o mundo, é preciso realizar uma calibração.

3.2 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 x em P^2 , seu ponto mapeado é igual à Hx (Dubrofsky, 2009; Malis et al, 2010; Hartley; Zisserman, 2003). A Figura 5 mostra a aplicação da homografia no Aedromo.

Figura 5 - homografia aplicada no Aedromo. Fonte: Autor.



As setas da Figura 5 representam a aplicação da homografia, ou seja, é a transformação das coordenadas de interesse entre as imagens e o mundo.

3.3 SISTEMA DE IDENTIFICAÇÃO

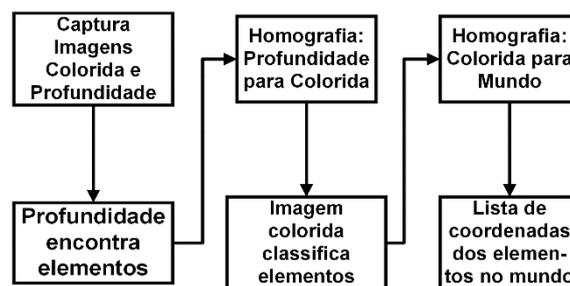
O novo sistema de identificação captura as imagens de profundidade e por meio das elevações no palco, encontra os objetos. A partir da matriz de homografia, as coordenadas em *pixels* desses objetos são transformados em pontos na imagem colorida do *Kinect*. Na imagem colorida estão as identificações coloridas dos elementos, o que

possibilita sua análise. Assim, é possível classificar os elementos entre robôs e objetos. Por fim, as coordenadas na imagem colorida são transformadas para as coordenadas do mundo, possibilitando o rastreamento das posições métricas reais dos elementos no palco. A Figura 6 mostra um diagrama sobre o novo sistema de identificação.

Assim, o Aedromo precisa de um total de duas matrizes de homografia. Uma é necessária para a transformação da imagem de profundidade para as coordenadas de imagem colorida e outra para a transformação das coordenadas desta imagem para as posições métricas do palco. O caminho inverso, pode ser realizado pela aplicação das inversas das mesmas matrizes de homografia.

Cada elemento operando no palco do Aedromo possui uma etiqueta colorida, com uma ou duas cores, específicas, que também serão definidas durante a calibração. Desta forma não há confusão na detecção e classificação dos elementos durante a utilização do ambiente.

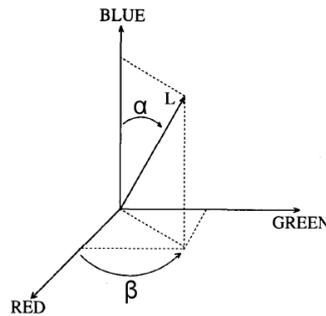
Figura 6 - Diagrama do sistema de identificação. Fonte: Autor.



As imagens coloridas capturadas pelo Kinect são fornecidas no sistema RGB de cores. Como este formato sofre grande influência da intensidade de luz sobre as etiquetas, o que dificulta a identificação de cores em condições pouco controladas de iluminação do ambiente, a identificação das etiquetas passa por uma transformação para o sistema SCT (*Sphere Coordinate Transformation*).

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 7 apresenta o sistema de cores RGB transformado em SCT (Umbugh et al., 1989).

Figura 7 - Sistemas RGB e SCT (Umbaugh et al., 1989).



O espaço bidimensional gerado pelos ângulos é utilizado para identificação das cores, com pouca influência da intensidade da iluminação.

4 DESCRIÇÃO DO SISTEMA

Com o objetivo de encontrar os robôs e os objetos no palco decorado, foram desenvolvidos três softwares. Dois deles são de calibração, sendo um responsável pela calibração prévia das cores usadas nos elementos e outro pelo posicionamento das câmeras, ou seja, cálculo das matrizes de homografia. O terceiro é o servidor do Aedromo, onde é realizada a análise das imagens, as classificações de cada elemento, o cálculo de suas coordenadas no mundo e o suporte ao controle do robô.

4.1 SOFTWARES DE CALIBRAÇÃO

A metodologia de calibração do posicionamento das câmeras adotada nesse trabalho gera as matrizes de homografia. Desta forma, para se obter estas matrizes é necessário encontrar pontos correspondentes na imagem de vídeo colorida (pixels), na imagem de profundidade (pixels) e no mundo do Aedromo (cm). Pontos específicos no mundo (cm) são conhecidos e usados como base para encontrar os pontos correspondentes nas outras imagens

O software de calibração das câmeras utiliza as imagens colorida e de profundidade. Inicialmente é capturada uma imagem de profundidade do palco vazio, sem elementos. Então o usuário deve colocar cinco blocos, com os topos coloridos em laranja e de 6 cm de altura, sobre cinco marcações projetadas na TV em posições conhecidas que são observados nas duas imagens. Para cada uma das cinco posições dos blocos, o sistema identifica na imagem colorida os topos laranja dos mesmos e na imagem de profundidade os pontos mais elevados dos cubos, capturando, assim, suas coordenadas em pixels em ambas imagens. Com os cinco pares de pontos capturados, juntos aos seus

correspondentes projetados, o software calcula todas matrizes de homografia e as exporta junto com a imagem de profundidade vazia.

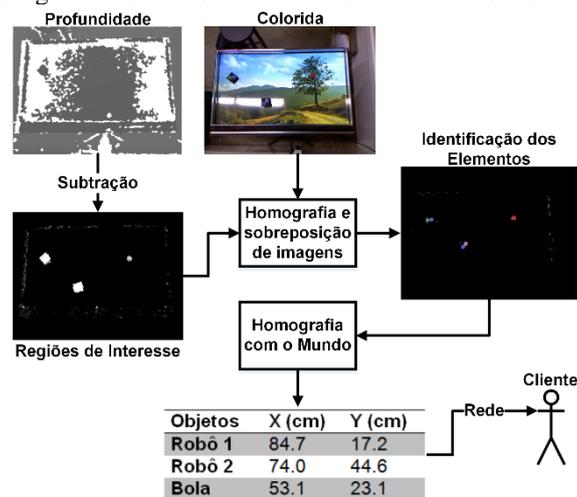
O software de calibração das cores está voltado à definição das cores de cada elemento. O usuário deve selecionar um elemento e relacioná-lo com faixas das componentes α e β do sistema SCT da sua cor correspondente. No final, estas faixas são exportadas.

4.2 SOFTWARE DE CONTROLE DO AEDROMO

Este software é a aplicação do sistema de identificação. Quando executado, carrega as informações de calibração e inicia o *Kinect*, sendo então capaz de identificar e disponibilizar as posições dos elementos no palco e receber os comandos a serem enviados aos robôs.

O software de controle do Aedromo captura a imagem de profundidade do *Kinect* e a imagem colorida. Com essas imagens, é feita uma subtração entre a imagem vazia e a de profundidade gerando regiões de interesse nas posições dos elementos. Em seguida, com a homografia, as regiões de interesse são sobrepostas na imagem colorida e cada elemento é identificado pela sua cor no sistema SCT. Por fim, é feita a homografia com o mundo e uma lista de elementos em coordenadas métricas no palco é gerada e disponibilizada ao software cliente. A Figura 8 apresenta um diagrama do funcionamento do software.

Figura 8 - Diagrama do funcionamento do software de controle do Aedromo.

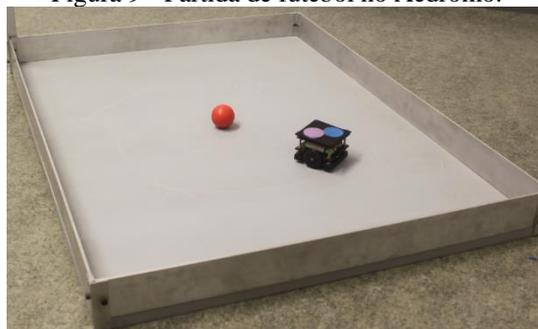


Fonte: Autor.

5 DISCUSSÕES

No Aedromo, sem a possibilidade de estampas, o palco não pode ser colorido, o que dificulta o entendimento da tarefa a ser realizada. A Figura 9 mostra a foto de uma aplicação em jogo de futebol nesse ambiente. Observa-se que é difícil inferir o tipo de atividade em execução.

Figura 9 - Partida de futebol no Aedromo.



Fonte: Autor.

Após as modificações estruturais e digitais, o Aedromo melhora a compreensão da atividade em curso o que permite a introdução de conceitos de diferentes complexidades aos alunos de uma forma cativante e divertida. A Figura 10 apresenta a foto de uma partida de futebol nesse novo ambiente.

Figura 10 - Aedromo com sistema de identificação pronto para um jogo de futebol.



Fonte: Autor.

A atividade em ambos ambientes é a mesma. Porém, na Figura 9 o propósito da tarefa não está claro!

Não foram realizados testes quantitativos referentes aos aspectos de precisão e desempenho do ambiente obtidos com o método apresentado, porém, estes se mostraram semelhantes ao do Aedromo original, na realização das mesmas atividades.

6 CONCLUSÃO

Este trabalho contribui com uma modificação do Aedromo possibilitando a representação da atividade de forma visual, com a decoração do palco, facilitando assim a compreensão da tarefa a ser realizada pelos robôs.

A identificação de elementos coloridos através de relevos sobre o palco do Aedromo, mesmo quando decorado com uma estampa colorida, foi possível pela integração, através da homografia, das imagens fornecidas pelas câmeras de um sensor *Kinect* da *Microsoft*.

Com o método de calibração empregado, os pontos nas duas câmeras são relacionados ao ambiente real, resultando em informações das posições métricas dos elementos sobre o palco.

Devido à flexibilidade adicionada ao Aedromo, outras atividades que exijam uma estampa colorida do palco podem ser elaboradas.

REFERÊNCIAS

Alves, S. F. R., Ferasoli Filho, H., Pegoraro, R., Caldeira, M. A. C., Yonezawa, W. and Rosário, J., Ambiente Educacional de Robótica Direcionado a Aplicações em Engenharia. Anais do SBAI - Simpósio Brasileiro de Automação Inteligente, 2011.

Dubrofsky, E., Homography Estimation. Master's Essay as requirements for the degree of Master of Science. University of British Columbia, Canada. 2009

Ferasoli Filho, H., Pegoraro, R., Caldeira, M.A.C. and Rosário, J., AEDROMO- An Experimental and Didactic Environment with Mobile Robots. In: Proceedings of The 3rd International Conference on Autonomous Robots and Agents, 2006.

Hartley, R., Zisserman, A., Multiple View Geometry in Computer Vision. Cambridge University Press, second edition, 2003.

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.

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