

RECONHECIMENTO DE IMAGENS NATURAIS EM REALIDADE AUMENTADA UTILIZANDO REDES NEURAS ARTIFICIAIS

Eduardo Chagas de Oliveira, Keiji Yamanaka, Luciene Chagas de Oliveira, Flávia Gonçalves Fernandes
Universidade Federal de Uberlândia - UFU
Faculdade de Engenharia Elétrica – FEELT
Programa de Pós Graduação em Engenharia Elétrica - PPGEE
Uberlândia – MG, Brasil
edu_ufu@yahoo.com.br, keiji@ufu.br, lchagasoliveira@gmail.com, flavia.fernandes92@gmail.com

Resumo – A tecnologia que permite ao usuário interagir com objetos tridimensionais inseridos no mundo real, também conhecida como Realidade Aumentada, faz muito uso de marcadores como instrumento para posicionar os objetos virtuais no mundo real. A dificuldade maior da utilização de tal tecnologia envolve o reconhecimento em tempo real das imagens dos marcadores, reconhecimento de vários marcadores ao mesmo tempo, uso de marcadores naturais e da manipulação dos marcadores. Esse trabalho realiza um estudo comparativo da forma de reconhecimento tradicional dos marcadores com as diversas formas de reconhecimento com redes neurais artificiais.

Palavras-Chave – marcadores, realidade aumentada, redes neurais artificiais.

RECOGNITION OF NATURAL IMAGES AUGMENTED REALITY USING ARTIFICIAL NEURAL NETWORKS

Abstract - The technology that allows the user to interact with three-dimensional objects inserted into the real world, also known as Augmented Reality, makes much use of markers as a tool to position the virtual objects in the real world. The difficulty of using such technology involves the recognition in real time of the images of the markers, multiple markers recognition while using natural markers and handling of markers. This work makes a comparative study of the traditional form of recognition of markers with the various forms of recognition with artificial neural networks.

Keywords - markers, augmented reality, artificial neural networks.

I. INTRODUÇÃO

Os crescentes avanços tecnológicos fazem com que o reconhecimento de imagens naturais seja um campo de estudos fascinante e ao mesmo tempo desafiador, uma vez que é bastante grande a gama de aplicações, na qual o uso de

mecanismos de Realidade Aumentada (RA) tem o papel de agilizar e facilitar a realização de tarefas cotidianas, buscando extrair informações de imagens detectadas no mundo real, de forma virtual.

O reconhecimento de imagens pode ser utilizado em diversas aplicações, como por exemplo, na Educação, como aprendizado cognitivo através de imagens naturais, associando imagens reais a elementos virtuais.

Além disso, o reconhecimento de imagens naturais pode ser combinado com outras tecnologias, tais como a Realidade Virtual e a Realidade Aumentada.

O reconhecimento de imagens naturais para identificação de objetos reais através de informações virtuais traz, de fato, maior naturalidade para interação do usuário com os sistemas de Realidade Virtual (RV), permitindo o controle de aplicações de visualização que possam estar executando dentro de um ambiente imersivo.

A Realidade Aumentada (RA) é um ambiente real "aumentado" com objetos virtuais que se tornou uma tecnologia promissora para apoiar tarefas em três dimensões (3D) [1]. A maioria dos sistemas de RA utilizam marcadores fiduciais para identificar a posição onde os objetos virtuais serão renderizados. Para desenvolver esta aplicação são necessários vários passos, sendo que um deles é a detecção de marcadores [3].

Estes marcadores contêm um padrão de imagem que deve ser identificada unicamente no ambiente [4]. Os sistemas de marcadores fiduciais consistem em padrões que devem ser unicamente identificados no ambiente por algum algoritmo de detecção. O algoritmo de detecção de um sistema com base em ID, basicamente identifica os marcadores no quadro de imagem e lê o código desenhado dentro deles [4].

Um dos pontos cruciais para se trabalhar com RA com marcadores envolve a identificação dos padrões embutidos nos marcadores utilizados [4]. Para identificá-los é necessário aplicar uma série de algoritmos na imagem de entrada capturada pela câmera.

Este trabalho propõe realizar o reconhecimento de imagens detectadas em marcadores naturais através de Redes Neurais Artificiais para aplicações de RA. Para demonstrar a utilização do reconhecimento de imagens naturais utilizando Redes Neurais Artificiais com Javascript foi desenvolvido um estudo de caso.

O fator motivacional deste trabalho consiste na implementação de uma ferramenta poderosa de reconhecimento, que pode ser utilizada em qualquer dispositivo que utiliza o navegador na web, por exemplo, um dispositivo móvel.

II. TRABALHOS RELACIONADOS

Esta seção apresenta alguns trabalhos relacionados à utilização do reconhecimento de imagens naturais como marcadores para aplicações de RA.



XIV CEEL - ISSN 2178-8308
03 a 07 de Outubro de 2016
Universidade Federal de Uberlândia - UFU
Uberlândia - Minas Gerais - Brasil

O trabalho [10] apresenta o uso de uma técnica de implementação de um algoritmo para processamento de imagens digitais utilizando Redes Neurais Artificiais, mas não envolve o reconhecimento para RA e não faz uso do ambiente WEB.

Em [11], diversas técnicas são apresentadas para análise e reconhecimento de imagens digitais e reais, entre elas, a RNA aplicadas ao reconhecimento de textos que apresentam escritas com diversas fontes e tamanhos.

O trabalho [9], ilustrado na Figura 1, propõe um livro utilizando RA para facilitar o aprendizado de pessoas disléxicas, no qual o sistema repete o som três vezes para o usuário, que deverá pronunciar a palavra.



Fig. 1. Exemplo de aplicação de RA na qual um caminhão é aumentado na palavra fornecida pelo usuário [9].

Com relação ao reconhecimento de marcadores, existem vários sistemas de marcadores já desenvolvidos. Alguns requisitos que devem ser considerados em aplicações de sistemas de marcadores são a precisão da posição de detecção, a detecção rápida e a diversidade de marcadores. Sistemas mais recentes realizam a detecção com oclusão parcial e são robustos para condições de iluminação, o que normalmente afeta a taxa de detecção. A Figura 2 mostra alguns exemplos de sistemas de marcadores [4].

Os marcadores apresentados na Figura 2 possuem um padrão na construção, que são quadrados preto e branco. Por exemplo, os marcadores de ARToolKit [2] são caracterizados por uma moldura preta e grossa em um quadrado contendo tons de cinza padrões. O ARToolKit tem sido amplamente utilizada no desenvolvimento de sistemas de RA [5]. Desde o surgimento do ARToolKit, bibliotecas utilizam marcadores semelhantes a um quadrado de rastreamento, tais como ARTag [6], ARToolKit-Plus [7] e ARLib [8].



Fig. 2. Exemplo de sistemas de marcadores [5].

Diferentemente dos trabalhos apresentados anteriormente, o trabalho proposto neste artigo utiliza o reconhecimento de imagens naturais para realizar a movimentação dos objetos virtuais em Realidade Aumentada, num ambiente web.

III. METODOLOGIA E ARQUITETURA DO SISTEMA

Esta seção descreve os detalhes do projeto e análise da aplicação.

A. Visão Geral do Sistema

A Figura 3 ilustra a ideia geral da proposta deste trabalho, no qual o usuário deve fornecer marcadores de RA para sobreposição dos objetos virtuais no mundo real.

Os marcadores de RA podem ser também objetos reais, como uma capa de revista para identificar dados em tempo real como a editora, ano, assunto, abordagens, etc, sem remover o lacre da revista e folheá-la.

As imagens do mundo real, como as capas de revistas, podem estar armazenadas em uma base de dados ou nas nuvens.

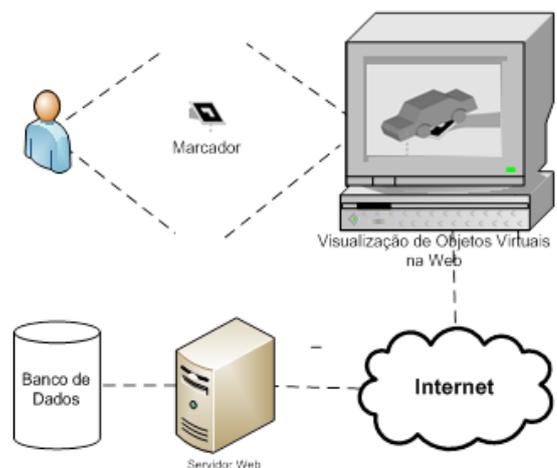


Fig. 3. Reconhecimento de imagens naturais para manipulação de objetos virtuais em RA.

B. Reconhecimento dos Marcadores

Para o reconhecimento de marcadores foi utilizado inicialmente o algoritmo tradicional do Aruco [12]. Além da identificação de marcadores, esse algoritmo já resolve também problemas de rotação e translação de imagens antes da identificação. A arquitetura de código gratuito e aberto Aruco foi mantida, mas evoluída para reconhecimento de imagens através de redes neurais.

IV. FERRAMENTAS AUXILIARES

A finalidade desse projeto consiste em prover uma solução completa para reconhecimento de marcadores, mas também que seja uma solução multiplataforma e também que atenda as regras da realidade aumentada. Dessa forma, todas as ferramentas auxiliares utilizadas para suportar esse objetivo e, portanto, foram construídas neste trabalho para plataforma WEB.

A. Gerador de Marcadores Aleatórios

O gerador automático de marcadores cria aleatoriamente 100 marcadores com 2401 pixels, cada um representado por uma matriz quadrática 7x7 ou por matriz 5x5 sem bordas. O gerador mostra cada imagem que está sendo gerada e pergunta se deseja permitir a geração das 100 imagens para cada marcador. O gerador garante que, apesar da geração aleatória, não existirá a repetição de alguma das imagens geradas no conjunto de imagens.

A ferramenta salva as imagens no disco no caminho default de upload de arquivos do browser, que no caso do Windows geralmente é configurado no caminho `C:\Users\nome_usuario\Downloads`, conforme visto na Figura 4.

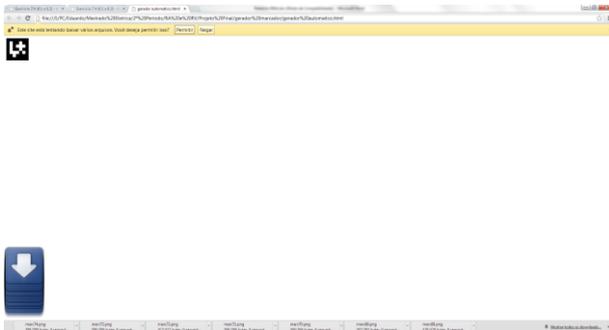


Fig. 4. Geração de marcadores aleatórios.

Quando o usuário permite a geração, a ferramenta gera as imagens em um ciclo de 50 em 50 imagens. Por último, dois arquivos são criados com os nomes de `listOutputVectors.txt` e `listInputVectors.txt`, que contém respectivamente a lista de vetores de imagens para output e input para treinamento das redes neurais artificiais.

B. Identificador de Imagens

O trabalho de reconhecimento de imagens requer constantemente verificar se a imagem gerada pelo vetor de pixels enviado pela câmera se identifica com a imagem esperada. Além de ruídos gerados pela câmera, um erro no algoritmo como deslocamento de um ou mais pixels pode

deformar a imagem totalmente e prejudicar o reconhecimento por completo.

Existem fatores que podem deformar a imagem antes de enviar para o reconhecimento, como por exemplo, a rotação e inclinação da imagem. Outro fator consiste no reconhecimento realizado com marcadores neste trabalho que possui um algoritmo que remove as bordas laterais das imagens. Um erro neste algoritmo pode deslocar a imagem ou mesmo deformá-la.

Neste trabalho, outra ferramenta auxiliar foi desenvolvida, capaz de identificar marcadores gerados pela câmera, com e sem bordas laterais, conforme visto na Figura 5.

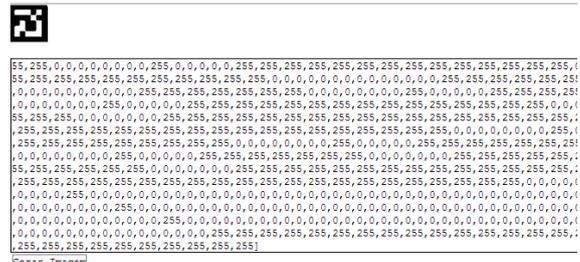


Fig. 5. Identificação de marcadores.

Essa ferramenta de identificação de imagens também é muito útil para se trabalhar com marcadores naturais, para verificar qual o formato que vem da câmera quando se deseja reconhecer imagens com várias tonalidades de cor.

C. Rotação de Matrizes e Imagens

Na realidade aumentada, o reconhecimento de imagens através de câmera requer um algoritmo que suporte a inclinação e rotação da imagem. Sendo assim, antes de cada treinamento das redes neurais artificiais, é necessário que o vetor de pixels que identifica a imagem seja convertido em uma matriz de pixels, para que ela seja rotacionada para direita, para esquerda e para baixo. Logo, a rotação das matrizes deve ser realizada com a matriz sem as bordas dos marcadores. A Figura 6 mostra uma ferramenta capaz de rotacionar as matrizes de imagens.

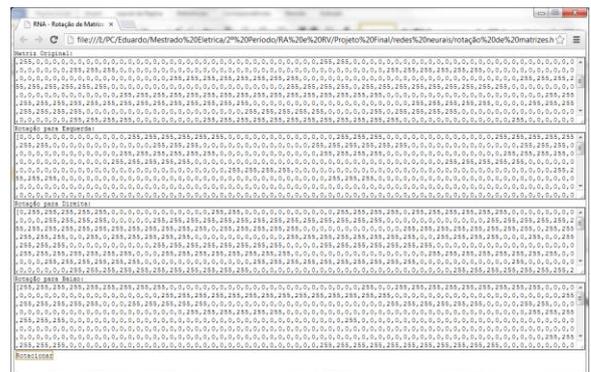


Fig. 6. Rotação de matrizes de imagens.

V. ANÁLISE

A análise de reconhecimento de marcadores envolve o estudo do treinamento das redes neurais desenvolvidas neste trabalho, Perceptron, Adaline e Multi Layer Perceptron (MLP) individualmente, para que, como resultado desta análise, seja utilizado os melhores parâmetros de treinamento

das redes e para o uso de marcadores na realidade aumentada.

A análise realizada nestes próximos tópicos utiliza-se do cálculo da somatória dos erros quadráticos em cada iteração para compor o gráfico da curva de aprendizagem.

A. Perceptron

A análise da rede neural artificial Perceptron com o reconhecimento de 100 imagens de marcadores envolve fixar o número de iterações e também os pesos iniciais da rede e variar a taxa de aprendizagem, buscando a melhor taxa de aprendizagem para treinamento da rede.

Assim, uma ferramenta de treinamento e análise da curva de aprendizagem é necessário para se obter a análise cada rede neural artificial. Também é necessário o armazenamento dos pesos finais para serem utilizados no projeto de realidade aumentada, para que não seja preciso aguardar o treinamento durante o uso do programa, pelo usuário final.

Observa-se que com o aumento da taxa de aprendizagem de 0,001 para 0,07 é possível diminuir o erro quadrático gerados pelos pesos na última iteração. Veja, na Figura 7 com o comparativo dos treinamentos.

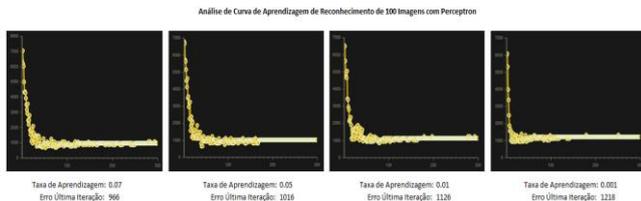


Fig. 7. Análise de curva de aprendizagem de reconhecimento de 100 imagens com Perceptron.

Entretanto, a rede neural artificial Perceptron não possui um decréscimo acentuado do erro quadrático da última iteração, assim como as redes Adaline e MLP, que serão analisados nos próximos tópicos.

B. Adeline

Em comparação à rede neural artificial Perceptron, analisado no tópico anterior, a rede Adaline apresenta uma evolução mais acentuada de diminuição do erro quadrático na última iteração quando se aumenta a taxa de aprendizagem.

Observa-se que com a Adaline é possível diminuir o erro quadrático em 70% com a variação da taxa de aprendizagem de 0,00001 para 0,01, enquanto que a rede Perceptron diminuiu somente 30% do erro quadrático com a variação da taxa de aprendizagem de 0,001 para 0,07. Isto pode ser observado através da evolução comparativa dos gráficos da Figura 8.

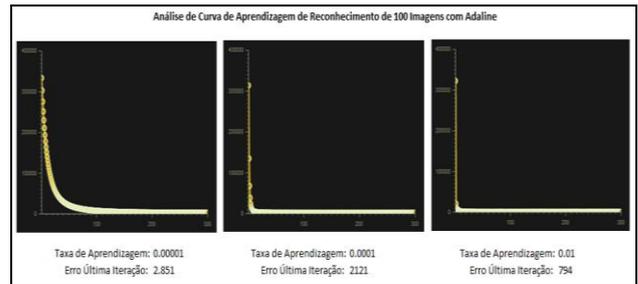


Fig. 8. Análise de curva de aprendizagem de reconhecimento de 100 imagens com Adaline.

C. Multi Layer Perceptron (MLP)

A rede neural MLP além de possibilitar a aprendizagem através da variação da taxa de aprendizagem, ainda possibilita a aprendizagem através de outros artifícios como momento e topologia. A análise no reconhecimento de imagens para marcadores utiliza-se principalmente da variação da topologia da MLP, cuja representação é: (tamanho do vetor entrada, [neurônios da camada escondida], [neurônios da camada de saída]).

Variando a topologia de (1225, [5, 10]) para (1225, [15, 10]), e mantendo as demais variáveis da rede, é possível diminuir o erro quadrático da última iteração em quase 90%. É possível verificar esta evolução realizando a comparação dos gráficos da Figura 9.

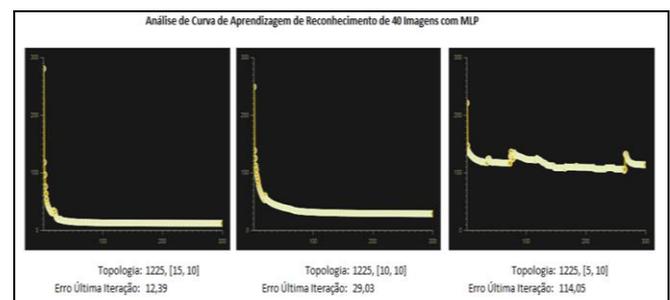


Fig. 9. Análise de curva de aprendizagem de Reconhecimento de 100 imagens com MLP.

VI. RESULTADOS

Este tópico apresenta como finalidade mostrar os resultados visuais obtidos com o uso das redes neurais artificiais no reconhecimento de imagens de vários marcadores provenientes de uma mesma câmera, mas com o processamento independente um do outro. A realidade aumentada reconhece o marcador e insere um objeto tridimensional para o marcador. Como o objetivo desse trabalho envolve principalmente o reconhecimento dos marcadores em tempo real, uma imagem bidimensional em vermelho com o texto de cor azul para o reconhecimento da imagem é suficiente para analisar a identificação do marcador.

Os resultados são a comparação das soluções de redes neurais com o reconhecimento tradicional. Um campo texto com a matriz da última imagem reconhecida é mostrado logo abaixo da câmera de cada representação.

A. Reconhecimento de Marcadores

A solução possibilita identificar vários marcadores ao mesmo tempo. O primeiro teste é treinando as redes neurais com cem imagens diferentes de marcadores para as redes neurais e uma combinação bem próxima para a solução tradicional. O teste consiste em oito marcadores com a mesma imagem. Este reconhecimento pode ser observado através da Figura 10.

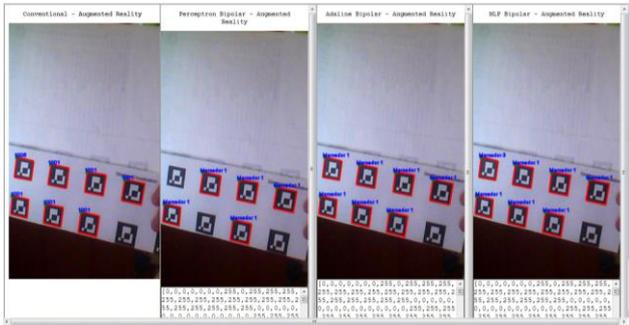


Fig. 10. Reconhecimento de marcadores.

B. Reconhecimento de Marcadores Especiais

Ruídos proveniente da câmera, da impressão da imagem do marcador, da envergadura da superfície da imagem, da iluminação do ambiente, distorções da rotação e inclinação da imagem, dentre outros fatores, podem interferir no reconhecimento. Dessa maneira, é necessário realizar testes utilizando um conjunto de imagens que representam tais problemas.

Um conjunto de dezenove marcadores com efeitos, como, por exemplo, reflexo de flash, efeito de molhado e efeito de iluminação foi utilizado para teste. Isto pode ser observado através da Figura 11.



Fig. 11. Reconhecimento de marcadores especiais.

Os marcadores com cores claras, com efeito de distorção de lente, reflexo de flash e efeito de molhado apresentam maiores problemas no reconhecimento, independente da solução aplicada.

C. Reconhecimento de Marcadores Naturais

Outro tipo de reconhecimento de marcadores também utilizado na realidade aumentada consiste nos marcadores naturais. Consistem em imagens do mundo real, como por exemplo, as imagens de amostra do Windows ou de propaganda de marcas famosas. A solução realizada neste trabalho, de reconhecimento de marcadores naturais, com redes neurais artificiais treinadas com vinte e cinco imagens

diferentes pode ser observado nas três últimas imagens da Figura 12.



Fig. 12. Reconhecimento de marcadores naturais.

Outras soluções tradicionais que utilizam javascript e são executadas no navegador, geralmente não reconhecem imagens de marcadores naturais ou possuem dificuldades durante o reconhecimento, como pode ser visto na primeira imagem da Figura 12. As soluções de Perceptron e Adaline podem reconhecer esse tipo de marcadores, mas também tem dificuldades principalmente quando as imagens são próximas e contém pouca quantidade de áreas claras ou possuem pouca quantidade de áreas escuras. Em contra partida, a solução de MLP pode reconhecer com taxas de acerto maiores que 80% das imagens, utilizando um conjunto de vinte e cinco imagens de marcadores.

A imagem natural proveniente da câmera é trabalhada em uma representação preto e branco antes de ser enviada para o processamento de reconhecimento. Ela pode vir também rotacionada. Segue, logo abaixo, alguns exemplos de imagens provenientes da câmera para reconhecimento, conforme visto na Figura 13.



Fig. 13. Reconhecimento de marcadores naturais.

Observe que a imagem é modificada quando há mudanças de luminosidade, sendo preciso um tratamento mais especializado e utilização de tonalidades intermediárias ao preto e branco.

Este trabalho, realiza a captura dos *pixels* através da arquitetura já existente do Aruco.js, modifica e os vetores de *pixels* para depois passar para o reconhecimento de imagens usando RNA.

Como trabalho futuro, a evolução arquitetural construída neste trabalho será novamente reutilizada, mas outra ferramenta auxiliar será desenvolvida para a melhor captura

das imagens utilizadas como treinamento da RNA, uma vez que este processo foi feito de forma manual, neste trabalho.

VII. CONCLUSÕES

É possível criar uma solução completa para reconhecimento de marcadores na realidade aumentada que seja também multiplataforma, através de Javascript. O uso de redes neurais artificiais é eficaz no reconhecimento de imagens para marcadores, principalmente quando se tratam de marcadores naturais.

Das soluções de redes neurais para reconhecimento de imagens de marcadores, a MLP consegue um melhor reconhecimento das imagens, variando tanto a taxa de aprendizagem quanto a topologia da rede. Quanto maior o tamanho da camada escondida, melhor é o reconhecimento, entretanto, requer maior tempo de treinamento. O reconhecimento da MLP é superior ao reconhecimento da solução tradicional.

Esta técnica de Inteligência Artificial possibilitou o reconhecimento com qualidade de imagens reais e, em paralelo às técnicas de Realidade Aumentada possibilitou a imersão do usuário com acompanhamento em tempo real de dados virtuais ao mundo real.

REFERÊNCIAS

- [1] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier and B. MacIntyre (2001): “Recent Advances in Augmented Reality”, IEEE Computer Graphics and Applications, 21(6):34–47.
- [2] ARToolKit. “ARToolKit”. Disponível em: <http://www.hitl.washington.edu/artoolkit/>. Acesso em 29 de agosto de 2016.
- [3] R. Azuma, “A survey of augmented reality,” 1995. [Online]. Available: citeseer.ist.psu.edu/azuma95survey.html
- [4] B. F. Reis, J. M. X. N. Teixeira, V. Teichrieb, and J. Kelner, “Perspective correction implementation for embedded (marker-based) augmented reality,” in Proceedings of Workshop de Realidade Virtual e Aumentada, 2008.
- [5] Ma. Xiaohu, Shi Gang and Tian Hong. “Adaptive Threshold Algorithm for Multi-marker Augmented Reality System”. 2010. In VRCAI '10 Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry, ACM Press, 71-74.
- [6] M. Fiala. “Artag, a fiducial marker system using digital techniques”. 2005. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 590–596.
- [7] W. Daniel, and S. Dieter, S. “Artoolkitplus for pose tracking on mobile devices”. 2007. In Proceedings of 12th Computer Vision Winter Workshop (CVWW'07).
- [8] D. Diggins. “ARLib: A C++ Augmented Reality Software Development Kit”. 2005. Master’s thesis, MSc Computer Animation N.C.C.A Bournemouth University.
- [9] Microsoft Speech. “Microsoft Speech Platform”. Disponível em: <http://www.microsoft.com/en-us/download/details.aspx?id=10208>. Acesso em 29 de agosto de 2016.
- [10] R. R. Silva. “Reconhecimento de imagens digitais utilizando Redes Neurais Artificiais”. Lavras, Minas Gerais, 2005.
- [11] S. Mariani. Introduction to Document Analysis and Recognition, Studies in Computational Intelligence (SCI) 90, p. 1-20, 2008.
- [12] ArUco. Mini biblioteca para aplicações de Realidade Aumentada baseada em OpenCv. Disponível em: <http://code.google.com/p/js-aruco/>. Acesso em agosto de 2013.