

Estereoscopia no cálculo de distância e controle de plataforma robótica

Gustavo Ribeiro Porpino Esteves
POLI - Escola Politécnica de Pernambuco
UPE - Universidade de Pernambuco

Rua Benfica, 455 - Madalena - Recife/PE CEP: 50720-001

Marcílio André Félix Feitosa, Bruno José Torres Fernandes
POLI - Escola Politécnica de Pernambuco
UPE - Universidade de Pernambuco

Rua Benfica, 455 - Madalena - Recife/PE CEP: 50720-001

Resumo--a estereoscopia visual é o processo pelo qual a partir de duas imagens distintas, mas que contemplem a mesma região, pode se ter uma noção da tridimensionalidade dessa região. Nesse artigo ela foi empregada para se calcular distâncias de determinados pontos até seu observador, esse observador é uma plataforma robótica que tem como finalidade se deslocar até determinado objeto mantendo uma distância mínima para que não haja colisão. A fórmula utilizada para o cálculo da distância será provado a partir da triangulação e utilizando representações básicas das câmeras. Será descrito como o algoritmo de reconhecimento deve localizar os pontos de interesse, os cálculos necessários para se obter a distância do objeto até a plataforma e a descrição da plataforma.

Palavras-chave--estereoscopia, robótica, processamento de imagem.

Abstract--a stereoscopic vision is the process by which the from two distinct images, but the contemplates same region, can get a sense of three-dimensionality of this region. In this article it was used to calculate distances determined points to his observer, this observer is a robotic platform whose purpose to travel to a particular object with minimum distance to avoid any collision. The formula used to calculate the distance to be proved and using triangulation from the representations basic camera. Will be described as the algorithm recognition must locate the points of interest, calculations required to obtain the distance to the object to the platform and the description of the platform.

Keywords--stereoscopy, robotics, image processing.

I. INTRODUÇÃO

Dentre os sentidos humanos a visão é a que mais se destaca em relação ao seu posicionamento dentro de um determinado ambiente. Através dela é possível ter a noção de profundidade, altura de objetos, identificação de cores e reconhecimento de figuras. A precisão da visão é relativamente maior para se identificar a distância de determinado objeto até o observador. É possível se posicionar bem próximo de um determinado objeto

bastando apenas que tal objeto esteja dentro do campo de visão.

A visão computacional [1] é dada por um conjunto de algoritmos que tentam captar informações contidas em imagens digitais. Muitas vezes esses algoritmos tentam simular a visão humana em determinados pontos, porém a maioria se limita em extrair apenas algumas determinadas características da imagem para aplicações específicas como o reconhecimento de cores ou de objetos.

Para o cálculo de distâncias a partir de imagens um recurso que pode ser usado é a estereoscopia que consiste na utilização de um par de câmeras capturando imagens de pontos distintos. Essas imagens devem estar direcionadas para um mesmo local, porém terão disparidades significativas. Através dessas disparidades é que se pode calcular a profundidade de determinado objeto.

Uma das áreas onde a visão computacional tem um importante papel é na robótica [2]. A robótica é uma área de pesquisa multidisciplinar e uma de suas vertentes é a robótica móvel que trata de robôs capazes de se deslocar pelos mais diversos tipos de terrenos. As aplicações para a robótica móvel são as mais variadas como: robôs exploradores [3], transporte autônomo [4] e sistema de manufatura flexível [5].

A estereoscopia visual pode ser empregada para solucionar diversos problemas como o mapeamento de ambientes, cálculos de distâncias e reconhecimentos de objetos tridimensionais. Dessa forma ela pode ser usada para resolver problemas na área de robótica como será abordado nesse artigo.

O objetivo é deduzir matematicamente uma equação simplificada capaz de calcular distâncias a partir da estereoscopia e a implementação de um robô capaz de utilizar essa equação para se deslocar até um determinado objeto.

Este artigo esta organizado da seguinte forma: na seção 2 será abordada a representação de uma câmera, que dará conceitos básicos importantes para que se possa usar a estereoscopia em imagens digitais, são demonstradas as fórmulas básicas de triangulação que permitem calcular a altura de um triângulo através dos seus ângulos da base, é mostrado como se utilizar as fórmulas da triangulação em um par de imagens estereoscópico, e são explorados os meios de se detectar pontos em comum nas imagens, tal detecção é fundamental para um cálculo preciso de distância; na seção 3 é mostrada uma plataforma utilizada

para mostrar uma das diversas aplicações disponíveis para a estereoscopia visual; e na seção 4 são apresentados os resultados, conclusões e trabalhos futuros.

II. ESTEREOSCOPIA VISUAL ATRAVÉS DE CÂMERAS

Para uma imagem digital ser formada, a câmera projeta o objeto localizado no espaço em um plano imaginário chamado de plano de imagem da câmera [5]. Esse processo de projeção é descrito pela matriz de projeção da câmera que depende do seu tipo e processo de fabricação. Tal matriz pode ser modificada pela alteração do foco da câmera, mas para esse artigo será considerada uma matriz de projeção fixa.

As câmeras escolhidas para serem utilizadas foram as do tipo *pinhole* onde através de um pequeno furo a luz é captada e projetada num sensor, dessa forma sua representação se torna simples. O raciocínio desenvolvido para sua representação pode ser expandido para os diversos tipos de câmeras existentes que usam lentes para focalizar determinadas distâncias. Em ambos os tipos de câmeras, uma característica que deve ser levada em consideração é a distância focal, através dela se obtém o campo de visão que chegará ao plano de imagem da câmera. A distância focal está diretamente ligada ao ângulo de convergência das imagens que chegarão à câmera.

A representação de uma câmera é mostrada na Figura 1, nela é possível visualizar a distância focal representada por f , o ângulo de convergência e o plano da imagem (sensor). Foi feita também uma simplificação na qual o plano da imagem é passado para frente do ângulo de convergência, esse recurso irá facilitar a representação estereoscópica sem perda de informações.

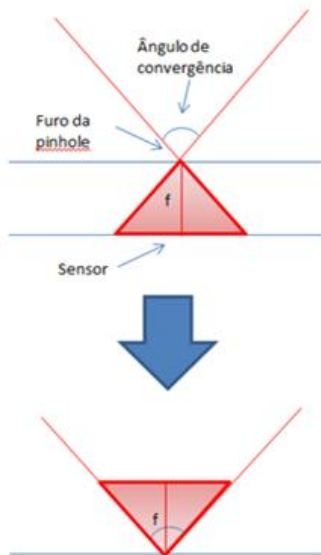


Fig.1: Representação de uma câmera pinhole e sua simplificação

Para se calcular a distância de determinado ponto até seu observador utilizando a estereoscopia visual faz-se necessário o conhecimento dos princípios de triangulação. Dado o triângulo mostrado na Figura 2 é possível elaborar cálculos para saber a distância D do ponto P até a reta formada por A e B utilizando princípios trigonométricos conforme será deduzido.

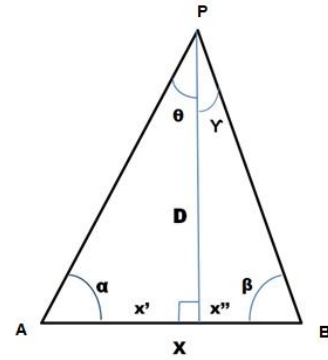


Fig.2: Triângulo representativo para os cálculos

A partir da lei dos senos tem-se:

$$\frac{D}{\text{sen}\beta} = \frac{x''}{\text{sen}\gamma} \therefore x'' = \frac{D\text{sen}\gamma}{\text{sen}\beta}, \quad (1)$$

$$\frac{D}{\text{sen}\alpha} = \frac{x'}{\text{sen}\theta} \therefore x' = \frac{D\text{sen}\theta}{\text{sen}\alpha}, \quad (2)$$

$$X = x' + x'' = \frac{D\text{sen}\gamma}{\text{sen}\beta} + \frac{D\text{sen}\theta}{\text{sen}\alpha}. \quad (3)$$

Em um triângulo a soma de seus ângulos internos é de 180° logo,

$$A + \theta + 90^\circ = 180^\circ \therefore \theta = 90^\circ - \alpha, \quad (4)$$

$$\beta + \gamma + 90^\circ = 180^\circ \therefore \gamma = 90^\circ - \beta. \quad (5)$$

A partir das equações (3), (4) e (5) obtém-se a expressão:

$$X = \frac{D\text{sen}(90^\circ - \alpha)}{\text{sen}\alpha} + \frac{D\text{sen}(90^\circ - \beta)}{\text{sen}\beta}, \quad (6)$$

Como,

$$\text{sen}(90^\circ - \alpha) = \text{cos}\alpha, \quad (7)$$

então,

$$X = \frac{D\text{cos}\alpha}{\text{sen}\alpha} + \frac{D\text{cos}\beta}{\text{sen}\beta}. \quad (8)$$

A seguinte relação também é conhecida:

$$\frac{\text{cos}\alpha}{\text{sen}\alpha} = \frac{1}{\text{tg}\alpha}. \quad (9)$$

Utilizando a relação (8) na equação (8) chega-se a:

$$X = \frac{D}{\text{tg}\alpha} + \frac{D}{\text{tg}\beta} = D \left(\frac{1}{\text{tg}\alpha} + \frac{1}{\text{tg}\beta} \right). \quad (10)$$

Passando D para o primeiro termo e X para o segundo:

$$D = \frac{X}{\left(\frac{1}{\operatorname{tg}\alpha} + \frac{1}{\operatorname{tg}\beta}\right)}. \quad (11)$$

Pela equação (11) fica evidente que é possível se calcular a distância de um determinado ponto P até a base de um triângulo X bastando apenas ter os ângulos α e β entre esse ponto e a base desse triângulo.

Para a utilização da estereoscopia visual serão posicionadas duas câmeras cada uma em um vértice da base do triângulo da Figura 2. As câmeras vão estar alinhadas de acordo com a Figura 3.

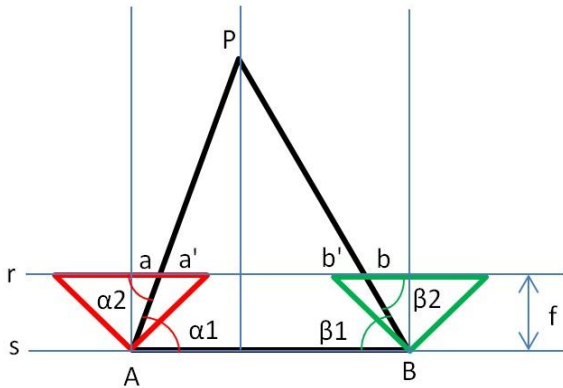


Fig.3: Representação do posicionamento das câmeras e do ponto de interesse

Para a equação (11) poder ser usada se faz necessário o conhecimento dos ângulos $\alpha 1$ e $\beta 1$, porém tais ângulos não são obtidos de forma direta. Para se adaptar a equação para seu uso direto, algumas adaptações devem ser feitas.

A reta r e a reta s são paralelas entre si, dessa forma os ângulos $\alpha 1$ e $\alpha 2$ são iguais por serem alternos internos, da mesma forma que os ângulos $\beta 1$ e $\beta 2$ também são equivalentes. Dessa forma é possível se achar as tangentes de α e β pelas seguintes equações:

$$\operatorname{tg}\alpha = \frac{f}{a}, \quad (12)$$

$$\operatorname{tg}\beta = \frac{f}{b}. \quad (13)$$

Utilizando as equações (12) e (13) na equação (11) obtém-se:

$$D = \frac{X}{\left(\frac{1}{f/a} + \frac{1}{f/b}\right)} = \frac{X}{\left(\frac{a}{f} + \frac{b}{f}\right)} = \frac{X}{\frac{a+b}{f}} = \frac{fX}{a+b}. \quad (14)$$

Os valores de **a** e **b** são as distâncias entre o centro da imagem e a localização onde o ponto de interesse está projetado no plano da imagem. Esses valores são absolutos

e independem do ponto estar projetado a direita ou a esquerda do centro da imagem.

Para simplificar os cálculos sem a necessidade de se utilizar o valor médio da imagem, é possível se substituir os valores de **a** e **b** por Xe e Xd respectivamente, onde Xd e Xe seriam as medidas da lateral esquerda da imagem até o ponto de interesse projetado no plano da imagem.

Dada que a largura total de ambos os planos de imagem seja **l**, o valor de **a** pode ser substituído por $l/2 - a'$, bem como o valor de **b** pode ser substituído por $l/2 - b'$ usando essas substituições na equação (14) resulta em:

$$D = \frac{fX}{(l/2 - a' + l/2 - b')} = \frac{fX}{(l - a' - b')}. \quad (15)$$

O valor de $l - a'$ é exatamente o valor medido da lateral esquerda da imagem esquerda até o ponto projetado no plano de imagem esquerdo, logo pode ser substituído por Xe que é definido da mesma forma. O valor **b'** é o valor medido da lateral esquerda da imagem direita até o ponto de interesse projetado no plano de imagem direita, logo pode ser substituído por Xd. Dessa forma a equação (15) resulta em:

$$D = \frac{fX}{(Xe - Xd)}. \quad (16)$$

Conforme um objeto se afasta a diferença $Xe - Xd$ diminui, pois de acordo com a Figura 2 os ângulos α e β aumentariam, logo os valores de **a** e de **b** da Figura 3 diminuiriam, resultando numa diminuição de Xe e aumento de Xd. Os valores de Xe e Xd são discretos e medidos em pixels, dessa forma conforme um objeto se afasta do par estereoscópico pior se tornará a precisão assim como na visão humana.

Se os valores de f e X forem mantidos constantes, é possível traçar um gráfico relacionando a diferença entre o ponto de interesse nos planos das imagens Xe-Xd e o valor equivalente a essa distância D. Para o caso de um valor de fX igual a 100 unidades o gráfico resultante entre a variação da diferença $Xe - Xd$ e a distância D é mostrado na Figura 4.

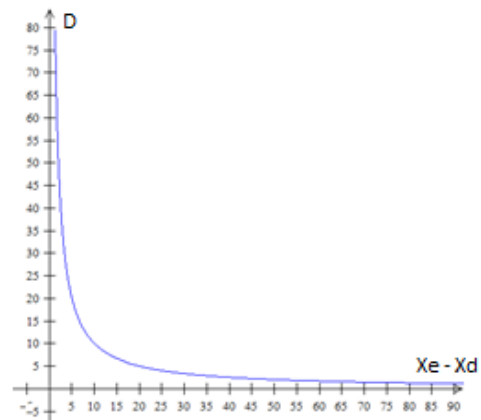


Fig.4: Gráfico da distância do ponto de interesse D em função da disparidade $Xe - Xd$

Um dos principais problemas para a utilização da estereoscopia visual é a localização e correlação correta dos pontos de interesse em ambas as imagens [6]. Dado um par

estereoscópico, primeiramente deve-se achar diversos pontos que serão candidatos ao emparelhamento em ambas as imagens, e após isso esses diversos pontos devem ser correlacionados de forma adequada para que se possa efetuar os cálculos de distância par a par.

Para se encontrar os candidatos ao emparelhamento, diversas técnicas podem ser utilizadas como a identificação por cor ou por máxima variância. Após achados os pontos, duas abordagens podem ser utilizadas para correlacioná-los. A primeira abordagem consiste em achar pontos de interesse apenas na primeira imagem e procurar seus correspondentes na segunda imagem. A segunda abordagem consiste em se encontrar pontos em ambas as imagens e depois tentar correlacioná-los.

A primeira abordagem pode se tornar demorada devido ao grande espaço combinacional existente, mas em alguns casos pode ser o único meio de se achar um determinado emparelhamento. A segunda abordagem tem uma complexidade reduzida, mas alguns dos pontos encontrados podem gerar ambiguidades ou não ter um par correto.

O método utilizado para correlacionar os pontos neste artigo foi uma forma direta já que apenas um único cálculo de distância era necessário. Dessa forma os valores de X_d e X_e devem ser escolhidos adequadamente sendo eles a lateral esquerda do objeto em relação a borda esquerda da imagem.

Para se utilizar os cálculos foram medidos os pontos de interesse nas imagens manualmente, mas para que um robô se guie através da estereoscopia é necessário que o computador saiba localizar esses pontos e aplicar o cálculo de profundidade corretamente. Para se identificar o objeto como se tratava de um cubo verde utilizou-se um algoritmo que fazia operações entre os níveis de verde, vermelho e azul da imagem e após as operações o resultado é uma imagem em tons de cinza é binarizada utilizando um limiar calculado pelo método de Otsu.

Por se tratar de um único ponto de interesse o único valor que precisa ser achado é a distância até a lateral esquerda em ambas as imagens. Para isso dada a imagem da Figura 5a, a imagem resultante está binarizada conforme a figura 5b, o próximo passo é varrer as imagens na vertical a partir da lateral esquerda até achar uma coluna na qual a soma de pontos brancos passe de um limiar. Esse recurso se localizar um grupo de pontos brancos ao invés de um único pixel se deve ao fato de que podem haver ruídos que resultem em pixel brancos por toda a imagem, atrapalhando na localização real do objeto.

Após se encontrar a coluna onde a soma de pixels passou do limiar, esse valor na coluna é armazenado. O processo é feito em ambas as imagens. Esses valores das colunas encontradas são usados na equação 16 para se achar a distância do objeto.

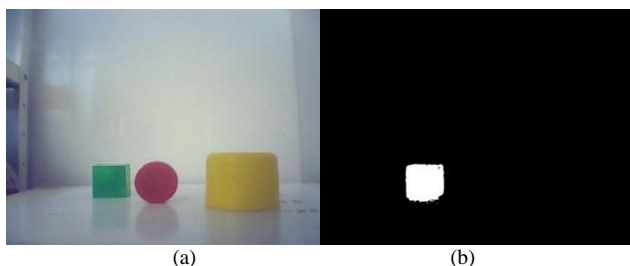


Fig.5: (a) Imagem original, (b) imagem binarizada.

III. PLATAFORMA ROBÓTICA E USO DA ESTEREOSCOPIA

A plataforma robótica utilizada nesse trabalho consiste em uma estrutura na qual se localizam as câmeras que enviam as imagens para o computador. Também estão localizados os motores [7] que permitirão a plataforma se locomover bem como a eletrônica responsável por receber os comandos do computador e controlar os motores. Um esquemático da plataforma é mostrado na Figura 6.

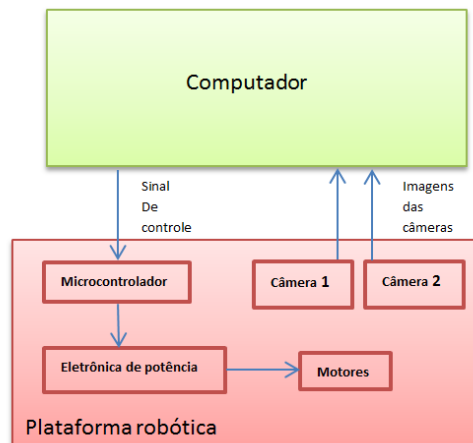


Fig.6: Esquemático das ligações da plataforma robótica

A plataforma tem um local onde as câmeras possam ser colocadas de forma que não haja nenhuma peça que atrapalhe a captação das imagens. O deslocamento do robô é feito de forma suave para que as câmeras não sofram impactos conforme o robô avança. Para limitar a velocidade foi usado um controle por modulação por largura de pulso (PWM - *Pulse-width modulation*), esse tipo de controle pôde ser utilizado de forma adequada já que os motores usados foram de corrente contínua e ímãs permanentes. Esse tipo de motor pode ter sua velocidade controlada de acordo com a tensão aplicada em seus terminais e o PWM controla a tensão eficaz aplicada no motor.

O objetivo desse projeto foi de criar um robô que utilizando apenas um par de câmeras fosse capaz de se locomover até uma distância mínima de um determinado objeto. Para isso as câmeras localizadas na frente da plataforma capturam um par de imagens estereoscópicas que são enviadas até o computador. No computador essas imagens são usadas para se achar os pontos de interesse e calcular a distância que o objeto está do robô. Caso o objeto ainda esteja muito longe, o computador envia um sinal para que os motores sejam acionados e para a plataforma se locomover por um determinado tempo. Após esse tempo novamente a imagem é capturada e enviada ao computador. Esse ciclo se repete até que o objeto esteja suficientemente perto do robô, e nesse momento não será mais enviado o comando para que ele se desloque em direção ao objeto.

Para o controle do robô foi projetada uma placa de controle e potência, mostrada na Figura 7, capaz de receber sinais enviados por um computador e controlar a potência exigida pelos motores. Um microcontrolador da família PIC [8] é responsável por receber os sinais e controlar a parte de potência. Apesar do robô só precisar se deslocar para frente, a eletrônica projetada também contempla a possibilidade do robô fazer curvas e se locomover para trás.

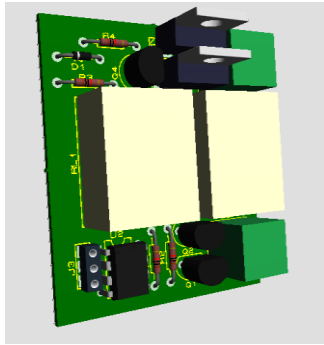


Fig.7: Circuito de controle dos motores da plataforma

A plataforma após montada é mostrada na Figura 8. Os motores ficaram posicionados na mesma linha de forma a facilitar as curvas. As câmeras foram devidamente colocadas na parte frontal do robô como foi proposto, para que nenhuma parte da plataforma atrapalhasse a captação das imagens.



Fig.8: Plataforma robótica construída

IV. RESULTADOS

O fator fX da equação 16 depende de duas características que se mantem fixas, mas nem sempre podem ser obtidas de forma direta. Pare se solucionar esse problema pode-se recorrer a um método empírico para encontrar esse fator. Para se achar esse valor é necessário se ter um par de imagens estereoscópicas obtido pelo par de câmeras das quais se quer achar o fator bem como saber a distância real de um determinado objeto até as câmeras.

A Figura 9 representa as distâncias entre as câmeras componentes do par estereoscópico e o objeto que será usado como referência. As Figuras 10a e 10b são as imagens capturadas. A partir dos dados obtidos é possível se calcular o fator fX da seguinte forma: a distância D do objeto ao par de câmeras é conhecido tem o valor de 30cm, as distâncias X_e e X_d também são conhecidas medindo 784px e 282px. A partir desses valores e da equação 16 é possível calcular fX como se segue:

$$30 = \frac{fX}{(784 - 282)}, \quad (17)$$

$$fX = 30 * 502, \quad (18)$$

$$fX = 15060, \quad (19)$$

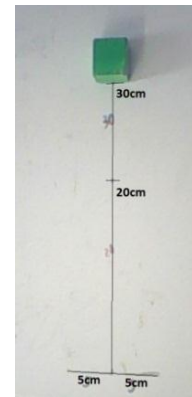


Fig.9: Vista superior do experimento



(a)



(b)

Fig.10: Par de imagens estereoscópicas, (a) imagem captada pela câmera esquerda, (b) imagem inferior captada pela câmera direita.

Após encontrado o fator fX ele pode ser utilizado para calcular quaisquer outras distâncias. Esses cálculos podem ser empregados para calcular diversos pontos na imagem bastando apenas que os pontos sejam devidamente correlacionados. Para comprovar a possibilidade de utilização dos cálculos, com o mesmo par estereoscópico foi tirado outro par de imagens mostrados na Figura 11a e 11b e se efetuaram os cálculos conforme segue:

$$D = \frac{15000}{(428 - 124)} = 49,3\text{cm},$$

$$D = \frac{15000}{(1044 - 546)} = 30,1\text{cm}.$$

De acordo com a Figura 12 esses valores estão bem próximos aos valores reais de distâncias dos objetos. Algumas imprecisões podem ocorrer devido ao posicionamento das câmeras, mas sem afetarem a utilização para o experimento em questão.

Dessa forma todo o ambiente onde a plataforma estiver pode ser mapeado gerando um mapa tridimensional. Tal mapa pode ser utilizado pela plataforma para traçar a melhor trajetória a se seguir.

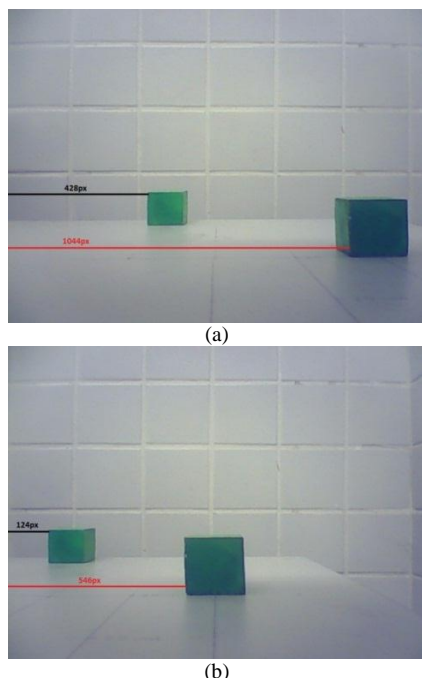


Fig.11: Par de imagens estereoscópicas, (a) imagem captada pela câmera esquerda, (b) imagem captada pela câmera direita.

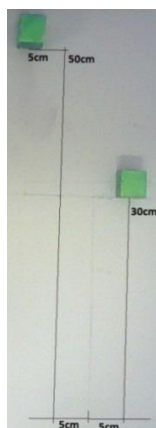


Fig.12: Vista superior do experimento que gerou as imagens da Figura 11

V. CONSIDERAÇÕES FINAIS

A plataforma pode ser melhorada de diversas formas, tanto no algoritmo de detecção que pode calcular diversos pontos e mapear todo o ambiente, quanto no algoritmo de controle que pode controlar o robô em diversas trajetórias para chegar no objeto em questão.

Este trabalho apresentou o desenvolvimento de equações que podem ser usadas para se calcular distâncias de determinados pontos até um observador utilizando a estereoscopia. As equações desenvolvidas se aplicam ao caso de câmeras alinhadas e com mesma resolução simplificando a fórmula resultante.

Um dos problemas apresentados foi a detecção dos pontos de interesse que para o caso deste artigo foi um cubo de cor verde. Outro problema apresentado foi o de correlacionar diversos pontos.

Para a utilização dessas equações foi usada uma plataforma robótica que permitiu uma utilização prática para o algoritmo desenvolvido. Na plataforma o par estereoscópico ficou posicionado de forma frontal. As imagens captadas eram enviadas para um computador onde os cálculos de distância eram feitos e após isso um sinal de controle era enviado a um microcontrolador para que a plataforma pudesse se aproximar mais do cubo.

REFERÊNCIAS

- [1] J. L. Crowley and H. I. Christensen. *Vision as Process*. [S.l.]: Springer-Verlag, 1995
- [2] G. BORGES. Um sistema óptico de reconhecimento de trajetórias para veículos automáticos. Master's thesis, Universidade Federal da Paraíba, Brasil, 1998.
- [3] K. P. VALAVANIS and G. N. SARIDIS *Intelligent Robotic Systems. Theory. Design and Applications*. Kluwer Academic Publishers, 1992.
- [4] M. BERTOZZI and A. B. GOLD: a parallel real-time stereo vision system For generic obstacle and lane detection. *IEEE Transactions on Image Processing*, 1997. URL <http://citeseer.nj.nec.com/bertozzi98gold.html>
- [5] J. A. FRANÇA. Desenvolvimento de Algoritmos de Visão Estereoscópica para Aplicações em Robótica Móvel. UFSC, Florianópolis - SC, Brasil
- [6] M. GALO e C. L. TOZZI, "Inclusão de Injunções Epipolares na Solução do Problema de Correspondência", Anais do X SIBGRAPI, (1997)
- [7] A.E. FRITGERALD, C. KINGSLEY e S. D. UMANS. *Máquinas Elétricas*, 6ª edição, editora Bookman.
- [8] F. PEREIRA. *Microcontroladores PIC - Técnicas Avançadas*, editora ERICA.