

UM SISTEMA DE VISÃO ESTÉREO TRINOCULAR PARALELO

Patrick Magalhães Cardoso

Universidade Federal de Uberlândia
Faculdade de Engenharia Mecânica
patrick@mecanica.ufu.br

José Francisco Ribeiro

jribeiro@mecanica.ufu.br

Resumo: *Este trabalho apresenta um sistema de visão estéreo trinocular. É utilizada uma única web-câmera para adquirir as imagens em três pontos de visão diferentes. A web-câmera é transladada na horizontal mantendo os sistemas de coordenadas paralelos entre as três posições. É apresentado um simples método de calibração onde apenas dois parâmetros são encontrados: a distância focal em pixels horizontais e um parâmetro da distorção radial da lente. É utilizado o método da Soma das Diferenças Absolutas para se encontrar os pontos correspondentes nas imagens adquiridas. São apresentados alguns resultados e o sistema estéreo se mostra atrativo face à sua simplicidade e o seu baixo custo.*

Palavras-chave: *estéreo, visão computacional, trinocular, câmera.*

1. INTRODUÇÃO

Visão estéreo refere-se à habilidade de inferir informações em 3D e distâncias de uma cena, a partir de duas ou mais imagens tomadas de diferentes pontos de visão (Trucco e Verri, 1998). Partindo-se de duas (ou mais) imagens de uma cena, a informação em 3D de um determinado ponto no espaço pode ser obtida a partir da posição que tal ponto ocupa em cada uma das imagens.

Inúmeros pesquisadores têm dedicado seu tempo no desenvolvimento de técnicas para encontrar os pontos correspondentes nas imagens estéreo, de uma maneira rápida e com confiabilidade. Uma boa informação a respeito de câmeras, tratamento de imagens e reconstrução 3D estéreo pode ser obtida em Trucco e Verri, 1998; Forsyth e Ponce, 2003 e; Gonzalez e Woods, 1993.

Este trabalho apresenta uma metodologia simples de reconstrução estéreo. Neste sistema são utilizadas três câmeras dispostas paralelamente e com apenas dois parâmetros a serem estimados: a distância focal expressa em pixels, f_x , e o parâmetro da distorção radial da lente, K_l . O problema de correspondência é resolvido utilizando-se o método da Soma das Diferenças Quadráticas (SSD) ou de modo simplificado e equivalente a Soma das Diferenças Absolutas (SAD). A busca pelos pontos correspondentes é realizada apenas sobre uma linha horizontal, já que as imagens foram tomadas de câmeras que apenas sofreram um deslocamento horizontal entre elas. São apresentados alguns resultados experimentais e o método, apesar de bem simples e de baixo custo mostra-se promissor na obtenção de uma reconstrução densa em 3D.

2. CÂMERA PINHOLE

O modelo geométrico mais comum para as câmeras é o modelo perspectivo ou modelo *pinhole* (Figura 1). O modelo consiste de um plano π , plano da imagem, e um ponto O no espaço, o centro ou foco de projeção. A distância entre o plano da imagem e o centro de projeção é do comprimento

focal f . A linha passando pelo centro de projeção e perpendicular ao plano da imagem é o eixo óptico, e a interseção entre o plano da imagem e o eixo óptico determina o ponto principal ou centro da imagem, o . Como mostrado na Figura 1, p , imagem do ponto P , é o ponto onde a linha reta que passa por P e O intercepta o plano da imagem (Trucco e Verri, 1998).

Considere agora um sistema de referência 3D no qual o ponto O é a origem e o plano da imagem, π , é ortogonal ao eixo Z . O ponto P e p ficam então definidos como $P=[X,Y,Z]^T$ e $p=[x,y,z]^T$. Tem-se então as equações básicas de projeção neste sistema de coordenadas (Trucco e Verri, 1998):

$$x = f \frac{X}{Z} \quad \text{e} \quad y = f \frac{Y}{Z} \quad (1)$$

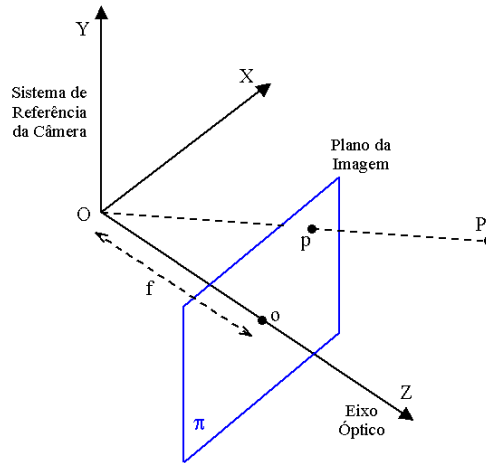


Figura 1: Modelo da câmera perspectiva.

Note que a partir destas equações pode-se encontrar apenas a razão entre X e Z ou entre Y e Z das coordenadas em 3D do ponto P . Não é possível, com apenas uma câmera encontrar as 3 coordenadas de um ponto espacial, sem que se conheça pelo menos uma das coordenadas. Tal problema será solucionado mais adiante ao se tratar da visão estéreo.

2.1 Parâmetros Intrínsecos da Câmera Pinhole

Os parâmetros intrínsecos podem ser definidos como um conjunto de parâmetros necessários para caracterizar óptica, geométrica e digitalmente o sistema de visão da câmera. Para uma câmera *pinhole* tem-se um conjunto de três parâmetros intrínsecos (Trucco e Verri, 1998):

- A projeção perspectiva, para a qual o único parâmetro é o comprimento focal f .
- A transformação entre o sistema de referência da câmera e as coordenadas em pixels.
- A distorção geométrica introduzida pelos elementos ópticos.

Para encontrar a transformação entre o sistema da câmera e as coordenadas em pixels é necessário amarrar as coordenadas em pixels de um ponto na imagem (x_{im}, y_{im}) com as coordenadas deste mesmo ponto no sistema de referências da câmera (x, y).

Desprezando as distorções geométricas e assumindo que o sensor CCD (*Charge Coupled Device*, sensor responsável por captar a intensidade de luz proveniente da cena observada) é constituído de uma malha retangular de elementos fotossensíveis, tem-se que a relação entre as coordenadas da câmera e as coordenadas em pixels é (Trucco e Verri, 1998):

$$x = -(x_{im} - o_x) \cdot s_x \quad \text{e} \quad y = -(y_{im} - o_y) \cdot s_y \quad (2)$$

onde: o_x e o_y são as coordenadas em pixels do centro da imagem (ou ponto principal) e s_x e s_y são o tamanho efetivo dos pixels (em milímetros/pixel) na direção horizontal e vertical respectivamente.

Em muitos casos, os sistemas ópticos introduzem distorções que são mais evidentes à medida que se aproxima das bordas da imagem. Tais distorções podem ser bem modeladas como simples distorções radiais utilizando-se a relação (Trucco e Verri, 1998):

$$x = x_d(1 + k_1 r^2 + k_2 r^4) \quad \text{e} \quad y = y_d(1 + k_1 r^2 + k_2 r^4) \quad (3)$$

onde: x_d e y_d são as coordenadas dos pontos distorcidos e $r^2 = x_d^2 + y_d^2$.

A distorção radial é o deslocamento radial dos pontos da imagem. Este deslocamento é nulo no centro da imagem e cresce à medida que se aproxima das bordas. Os valores de k_1 e k_2 fazem parte dos parâmetros intrínsecos da câmera quando se considera a distorção radial. Estes parâmetros são, geralmente, pequenos e a distorção radial pode ser desprezada quando não se necessita de precisão em todas as regiões da imagem ou quando os pontos periféricos são descartados. Como k_2 é bem menor que k_1 é possível considerar apenas o valor de k_1 para modelar a distorção radial, mantendo ainda uma boa precisão. A magnitude da distorção geométrica depende da qualidade da lente utilizada. Com sistemas ópticos de qualidade média e um sensor CCD de tamanho 500x500, a distorção esperada é de poucos pixels (em torno de 5 pixels) nos cantos da imagem. Nestes casos, modelar a distorção radial considerando $k_2=0$ ainda produz resultados precisos (Trucco e Verri, 1998).

2.2 Parâmetros Extrínsecos da Câmera Pinhole

O sistema de referência da câmera foi definido anteriormente com o objetivo de se escrever as equações fundamentais da projeção perspectiva em uma forma simples. No entanto, o sistema de referência da câmera é geralmente desconhecido e um problema comum é determinar a localização e orientação do sistema de referência da câmera em relação a algum sistema de referência conhecido. Os parâmetros extrínsecos são definidos como um conjunto de parâmetros geométricos que relacionam, de modo único, o sistema de referência da câmera desconhecido a um sistema de referência global conhecido (Trucco e Verri, 1998).

Para realizar a transformação entre estes dois sistemas de referência, são utilizados:

- Um vetor de translação 3D, vetor T , que descreve a posição relativa dos dois sistemas de referência.
- Uma matriz de rotação 3x3, R , que relaciona a orientação dos eixos entre os dois sistemas de referência.

Assim, a relação entre as coordenadas de um ponto P no sistema de referência da câmera, P_c , e as coordenadas deste mesmo ponto no sistema de referência global, P_w , fica:

$$P_c = R(P_w - T) \quad (4)$$

3. O SISTEMA ESTÉREO

Em um sistema estéreo tem-se basicamente dois problemas a serem resolvidos: o problema da correspondência e o problema da reconstrução. Os inúmeros trabalhos realizados na área de visão estéreo diferem-se basicamente no modo como tratam estes dois problemas.

A utilização de sistemas estéreos onde as câmeras possuem diversos parâmetros diferentes, tanto os intrínsecos quanto os extrínsecos, torna o processo de reconstrução 3D mais complexo computacionalmente. Em um sistema estéreo genérico é necessário calibrar o sistema encontrando-se a Matriz Fundamental e a Matriz de Projeção.

Maiores detalhes sobre visão estéreo de um modo geral podem ser encontrados em Trucco e Verri, 1998; Forsyth e Ponce, 2003; Gonzalez e Woods, 1993; Zhang et al, 1997.

3.1 A Configuração Estéreo Utilizada

Optou-se por utilizar uma configuração muito simples para o sistema de visão estéreo. Embora simples, este sistema apresenta uma boa precisão e, o seu principal atrativo é a simplicidade computacional e o baixo custo do sistema físico.

O sistema de visão estéreo é constituído por uma única web-câmera, uma Creative Webcam NX-Pro, e por um carro de deslocamento linear conhecido. A câmera é posicionada sobre o carro e, transladando-se o carro, são tomadas três imagens em pontos de visão diferentes. Os deslocamentos relativos entre as três posições da câmera são medidos com o auxílio de um parafuso micrométrico posicionado no carro. As imagens são transmitidas a um micro-computador utilizando-se uma porta USB. Na Figura 2 tem-se uma foto do sistema estéreo utilizado.

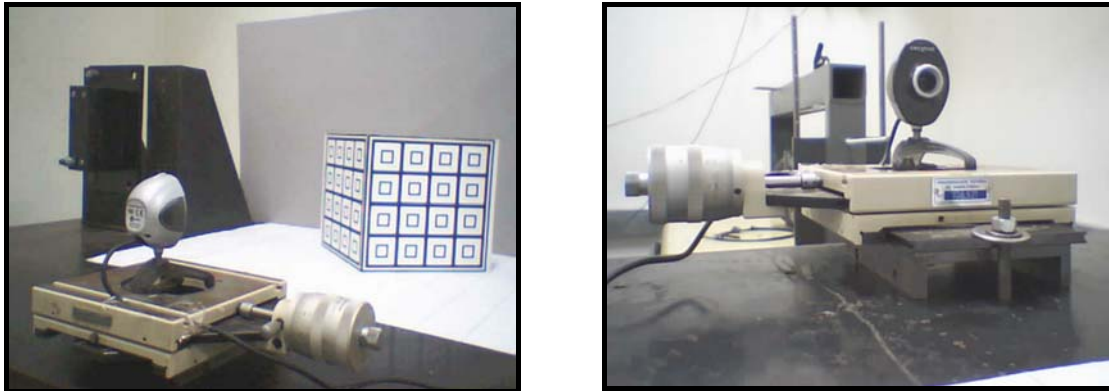


Figura 2: Foto do sistema estéreo utilizado.

Utilizando-se um sistema estéreo com esta configuração, garante-se que as três imagens da cena serão tomadas em posições paralelas, já que é realizada apenas uma translação na horizontal (translação do carro). Com isto elimina-se a necessidade de se determinar a Matriz Fundamental. Isto é possível, pois os pontos correspondentes estão sobre a mesma linha horizontal para as imagens tomadas de pontos de visão paralelos. Elimina-se neste processo também a necessidade de retificar as imagens (veja mais sobre retificação de imagens em Trucco e Verri, 1998).

A câmera é também posicionada de modo que o seu sistema de referência esteja paralelo a um sistema de referência global sobre o qual as coordenadas relativas serão tomadas no processo de calibração.

Para se proceder com a reconstrução 3D é necessário conhecer os parâmetros intrínsecos da câmera e ainda a translação sofrida na horizontal por cada uma das câmeras (linha de base). Neste sistema estéreo são adquiridas três imagens aqui denominadas de imagem esquerda, imagem do centro (base) e imagem direita. A imagem da esquerda é adquirida transladando-se a câmera de uma distância de 10 mm ($B=10$) em relação à imagem de base, em sentido contrário à orientação do eixo X paralelo à linha de base. A imagem direita é adquirida transladando-se a câmera de 10 mm no sentido contrário.

A seguir tem-se passo a passo como é realizada a reconstrução estéreo após o processo de calibração:

- 1) Adquire as três imagens da cena tomadas em pontos de visão diferentes. Deve-se conhecer a translação (linha de base) entre cada uma das posições da câmera.
- 2) Com as três imagens é obtido um novo conjunto de imagens onde a distorção radial é eliminada.
- 3) Com o novo conjunto de imagens, dado um determinado ponto sobre a imagem do centro, utilizando-se o método da Soma das Diferenças Absolutas (SDA) são encontrados os pontos correspondentes na imagem esquerda e direita. Com os pontos correspondentes determinados, pode-se saber a disparidade (diferença em pixels) entre os pontos correspondentes de duas imagens, por exemplo entre os pontos da imagem central e da esquerda.

- 4) Com o valor da disparidade (d_{im}) é possível se determinar as coordenadas 3D do ponto da cena em relação ao sistema de referência da câmera. Isto é realizado conforme o procedimento descrito a seguir.

Substituindo-se a Equação (2) em (1) tem-se:

$$x_{im} = -\frac{f}{s_x} \cdot \frac{X}{Z} + o_x \quad \text{e} \quad y_{im} = -\frac{f}{s_y} \cdot \frac{Y}{Z} + o_y \quad (5)$$

No sistema estéreo utilizado, verificou-se que o centro da imagem está no ponto (0,0), ou seja, o_x e o_y são iguais a zero. Considerando que para o sistema físico em questão tem-se que a razão f/s_x é igual à razão f/s_y e igual a f_x (comprimento focal em pixels horizontais), tem-se para o sistema estéreo:

$$x_{im} = -f_x \cdot \frac{X}{Z} \quad \text{e} \quad y_{im} = -f_x \cdot \frac{Y}{Z} \quad (6)$$

Na Figura 3 tem-se um simples sistema estéreo onde as câmeras estão paralelas. A partir da relação de triângulos é possível encontrar:

$$\frac{B-d}{Z-f} = \frac{B}{Z} \quad (7)$$

onde: d é a disparidade ou $x_r - x_l$ e; B é a linha de base.

Resolvendo a equação acima para Z e transformando o valor da disparidade de mm para pixels ($d_{im} = d/s_x$), tem-se:

$$Z = f_x \frac{B}{d_{im}} \quad (8)$$

onde d_{im} é o valor da disparidade em pixels.

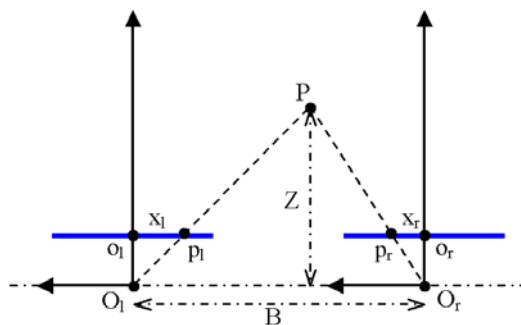


Figura 3: Um simples sistema estéreo com duas câmeras.

De posse do valor da disparidade em pixels, pode-se calcular o valor da profundidade Z do ponto no espaço através da Equação (8) e, com este valor de profundidade, pode-se encontrar os valores de X e Y fazendo-se uso da Equação (6) modificada para se computar X e Y .

3.2 Calibração do Sistema Estéreo

Como a câmera utilizada trata-se de uma web-câmera, não sendo portanto uma câmera com um conjunto óptico preciso, é necessário encontrar a distorção radial provocada pela lente da câmera. Outro parâmetro a ser encontrado é o comprimento focal expresso em pixels horizontais (f_x).

O valor da linha de base (B) de uma câmera em relação à outra é determinado através de um parafuso micrométrico que mede o deslocamento horizontal do carro sobre o qual está fixada a câmera (Figura 2).

3.2.1 Distorção Radial

Conforme a Equação (3), a distorção radial pode ser modelada pelos parâmetros k_1 e k_2 . Para se determinar o valor destes parâmetros colocou-se uma grade uniforme em um plano paralelo à câmera (Figura 4). O centro desta grade também coincide com o centro da imagem (o_x e o_y são nulos).

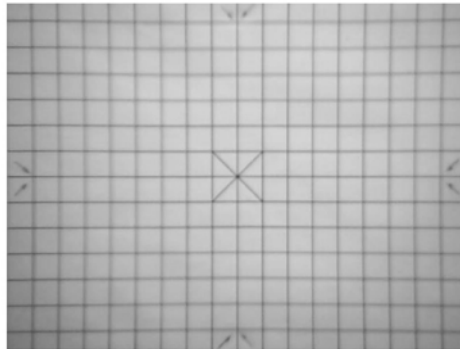


Figura 4: Grade utilizada para estimar a distorção radial da lente da câmera.

Como a região central da imagem praticamente não sofre distorção radial, considerou-se que o grande quadrado central (marcado por um X, diagonais do quadrado) tinha suas dimensões perfeitas (sem distorções) e considerou-se a metade do lado deste quadrado como sendo a medida em pixels das divisões da grade. Com isto, é possível se saber qual a posição que cada ponto de interseção das linhas da grade deveria ocupar para o caso de uma lente sem distorção radial.

Tomou-se então 26 pontos sobre as duas diagonais principais da grade (formando um X). Para estes pontos são conhecidas as coordenadas em pixels dos pontos distorcidos e as coordenadas em pixels para os pontos sem distorção (valor desejado encontrado utilizando-se o valor estimado para as divisões da grade).

Utilizando-se o método de otimização BFGS encontrou-se o valor de k_1 e k_2 . No entanto, verificou-se que o valor de k_2 tende à zero durante o processo de otimização. Assim, seguindo as características de um sistema óptico de média qualidade (citado no item 2.1) considerou-se apenas o parâmetro k_1 para estimar a distorção radial.

O valor do parâmetro k_1 da distorção radial é de $-1,9 \cdot 10^{-7}$ para a câmera utilizada.

3.2.2 Comprimento Focal f_x

No sistema estéreo mostrado neste trabalho, utilizou-se a mesma câmera para adquirir as três imagens em pontos de observação diferentes. Sendo assim, é necessário encontrar apenas uma vez o valor de f_x , já que não se altera o foco da câmera ao se mudar de ponto de visão.

Para se determinar o valor de f_x utilizou-se um plano padrão cujas coordenadas relativas de alguns pontos são conhecidas. Este plano é paralelo ao plano XY , que é o plano da câmera. Não é necessário, para o procedimento de calibração adotado, conhecer as coordenadas absolutas dos pontos sobre o plano padrão. Foram tomados 20 pontos sobre este plano sendo conhecidas as coordenadas relativas entre estes pontos. As imagens adquiridas, e que serão utilizadas neste procedimento de estimativa de f_x devem ser previamente corrigidas quanto à distorção radial, conforme mostrado no item anterior.

Tomando-se 2 pontos (pontos 1 e 2) quaisquer sobre o plano padrão e fazendo uso da Equação (6), tem-se que:

$$\begin{aligned} X_1 &= \frac{-x_1}{f_x} \cdot Z_1 & Y_1 &= \frac{-y_1}{f_x} \cdot Z_1 \\ X_2 &= \frac{-x_2}{f_x} \cdot Z_1 & Y_2 &= \frac{-y_2}{f_x} \cdot Z_1 \end{aligned} \quad (9)$$

onde: x e y representam as coordenadas do ponto na imagem expressos em pixels (foi omitido o sub-índice *im* que aparecia na Equação (6)).

Nota-se que como os pontos 1 e 2 pertencem ao mesmo plano eles possuem a mesma coordenada Z_1 e também o mesmo parâmetro f_x . Conforme mencionado anteriormente, são conhecidas as coordenadas relativas entre os pontos deste plano. Assim, tomando-se as coordenadas relativas entre os pontos 1 e 2, tem-se:

$$\Delta X = (x_2 - x_1) \cdot \frac{Z_1}{f_x} \quad \text{e} \quad \Delta Y = (y_2 - y_1) \cdot \frac{Z_1}{f_x} \quad (10)$$

Tomando-se os demais pares de pontos sobre este plano padrão é possível encontrar o valor médio para a razão $\frac{Z_1}{f_x}$ para o plano padrão 1. Esta razão será denominada R_1 .

Tomando-se o mesmo plano padrão, porém agora deslocado sobre o eixo Z (deslocamento paralelo ao plano XY da câmera) tem-se o plano Z_2 . Realizando-se o mesmo procedimento exposto acima, pode-se encontrar uma nova razão média $\frac{Z_2}{f_x}$, denominada R_2 . Note que como a câmera permanece a mesma, o parâmetro f_x também é o mesmo. São tomados um total de 6 planos, em seis posições diferentes, fornecendo seis valores para a razão média R . Estes planos padrão utilizados na estimativa de f_x podem ser vistos na Figura 5 (nesta figura as imagens ainda não foram corrigidas quanto à distorção radial).

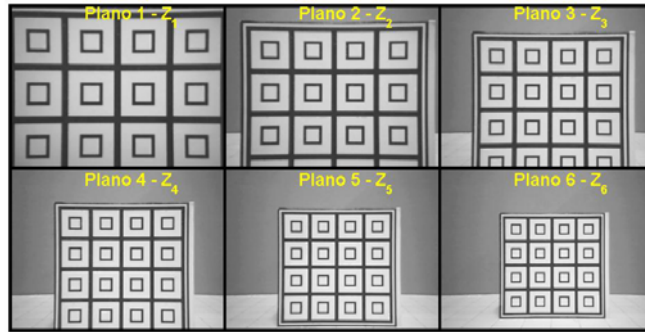


Figura 5: Planos padrão utilizados na estimativa do parâmetro f_x .

São conhecidos os deslocamentos relativos (D) entre cada um dos planos padrão. O deslocamento relativo entre o plano 1 e 2, por exemplo, será denominado de D_{1-2} , e assim segue-se para os demais deslocamentos relativos.

De posse do deslocamento relativo entre os planos padrão 1 e 2, D_{1-2} , e das razões médias R_1 e R_2 , pode-se fazer:

$$\frac{Z_2}{f_x} - \frac{Z_1}{f_x} = R_2 - R_1 \quad \text{ou} \quad \frac{Z_2 - Z_1}{f_x} = R_2 - R_1 \quad (11)$$

Como $Z_2 - Z_1$ é o valor do deslocamento relativo D_{1-2} , pode-se encontrar o valor de f_x considerando-se os planos 1 e 2:

$$f_x^{1-2} = \frac{D_{1-2}}{R_2 - R_1} \quad (12)$$

Encontrando-se os valores de f_x para os demais pares de planos padrão, pode-se obter o valor de f_x como sendo o valor médio dentre os valores encontrados para todos os pares de planos.

Após realizados os procedimentos descritos acima, obteve-se o valor de $f_x = 731.72$ pixels para a câmera utilizada no sistema estéreo.

3.2.3 Busca pelos Pontos Correspondentes

Neste trabalho utilizou-se o método da Soma das Diferenças Absolutas (SDA) que é derivado do método da Soma das Diferenças Quadráticas (SDQ) (Okutomi e Kanade, 1993; Kanade et al, 1995; Kanade et al, 1996). Utilizando-se o primeiro tem-se a vantagem de diminuir a quantidade de cálculos matemáticos, sem se perder a generalidade do método.

Da Equação (8) tem-se que:

$$\frac{d_{im}}{B} = f_x \cdot \frac{1}{Z} = \zeta \quad (13)$$

onde: a disparidade em pixels (d_{im}) dividida pelo comprimento da linha de base (B) é chamada de profundidade inversa (ζ).

A Equação (13) acima, indicada que para um ponto particular na imagem o valor de ζ é constante, uma vez que há apenas uma distância Z para este ponto. Tendo-se uma certa medida de correspondência para um determinado ponto representada em função de ζ , esta medida deve mostrar uma boa consistência para um único valor correto de ζ , que é independente da linha de base (Kanade et al, 1995). A representação em termos de ζ é muito útil no caso de se utilizar um sistema estéreo com mais de duas câmeras, e portanto mais de um comprimento de linha de base. Para este caso, ao se representar em termos de ζ a função avaliada na correspondência dos pontos, tem-se uma independência em relação à linha de base, possibilitando comparar os valores da função de correspondência entre os diversos pares de imagens adquiridos. Isto permite também que ao se adicionar estas medidas de correspondência, considerando cada valor de ζ , se possa ter um único ponto de solução, que deverá ser mais evidente que a utilização das funções de correspondência SDA em separado. Esta melhor evidência do ponto de solução pode ser visualizada na Figura 6, onde se tem um típico gráfico de SDAs e SSDA para um determinado ponto na imagem de base. O ponto de solução é aquele onde a função SSDA assume o menor valor. A utilização de apenas um par de câmeras, apenas uma curva de SDA, pode gerar ambigüidades ao se determinar o mínimo desta função.

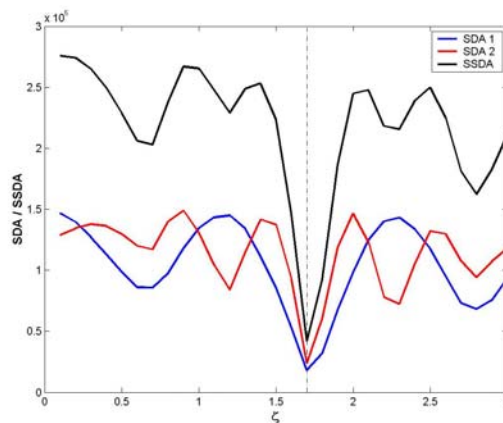


Figura 6: Gráficos típicos de SDA e SSDA em função de ζ .

A SDA (ou equivalentemente a SDQ) sobre uma pequena janela é uma das mais simples e efetivas medidas de emparelhamento (correspondência) de pontos na imagem (Kanade et al, 1995). Para um determinado ponto particular em uma imagem de base, é tomada uma pequena janela em torno deste ponto. Tomando-se este mesmo tamanho de janela nas outras imagens, e fazendo esta janela se transladar ao longo da linha sobre a qual estão os pontos correspondentes, os valores de SDA são computados para cada valor de disparidade. No caso deste trabalho e no caso de se utilizar imagens retificadas translada-se ao longo da linha horizontal, mas no caso de imagens não retificadas, translada-se ao longo da Linha Epipolar (Trucco e Verri, 1998). O cálculo de SDA para cada janela é realizado somando-se as diferenças absolutas entre o valor de intensidade de cada pixel na janela de base e o seu correspondente pixel na janela que foi translada na outra imagem, de modo que se tenha um valor de SDA para cada translação da janela, ou melhor para cada valor de disparidade em pixels.

Para cada par de câmeras estéreo tem-se um valor da linha de base (B). Ao se dividir o valor da disparidade (d_{im}) pela linha de base, pode-se ter a curva de SDA expressa em valores de ζ , que será conveniente para realizar a comparação entre os diferentes pares de imagens.

A curva de SSDA (soma de SDA) é obtida somando-se todos os valores de SDA para um mesmo valor de ζ . Observando-se a curva SSDA, nota-se que esta curva possui um mínimo mais efetivo e claro, o que melhora a eficiência do método de busca dos pontos correspondentes.

Encontrando-se o mínimo da função SSDA, pode-se determinar o valor da profundidade Z através da Equação (8), já que o valor de ζ é o inverso de B/d_{im} . Utilizando-se a Equação (6) pode-se encontrar o valor das coordenadas X e Y em relação ao sistema de referência da câmera de base (câmera do centro), partindo-se dos valores de x_{im} e y_{im} na imagem de base.

4. RESULTADOS OBTIDOS

Na Figura 7 tem-se a imagem de base (imagem do centro) de um objeto e o mapa de disparidade obtido pelo método descrito no item anterior. Nota-se que se obteve um bom mapa de disparidade, embora se tenha, neste mapa, apenas 15 níveis diferentes de disparidade para este objeto, o que não permite distinguir, em alguns pontos, os picos e vales do objeto.

Têm-se apenas 15 níveis de disparidade devido ao sistema estéreo utilizado. Para se ter uma melhor discretização na disparidade é necessário que os valores de linha de base dos pares de câmera estéreos sejam maiores, aumentando assim o número de disparidades possíveis entre o valor máximo e mínimo. Isto pode ser verificado através da Equação (8), sendo que, por exemplo, para um mesmo intervalo de profundidade Z , ao se dobrar o valor do comprimento da linha de base, se dobra o número de discretizações do intervalo entre a mínima disparidade e a máxima disparidade observadas para a máxima profundidade e para a mínima profundidade respectivamente na cena.

Já na Figura 8 tem-se um objeto cujas coordenadas de alguns pontos são conhecidas, objeto calibrado. Utilizou-se este objeto para medir a distância em 3D entre alguns pontos, validando assim o processo de reconstrução estéreo. Os pontos tomados sobre este objeto também são mostrados na Figura 8.

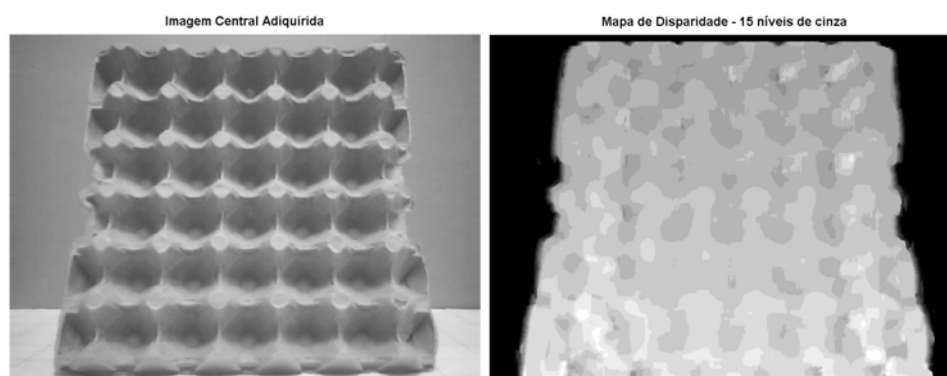


Figura 7: Imagem de base (à esquerda) e mapa de disparidade (à direita).

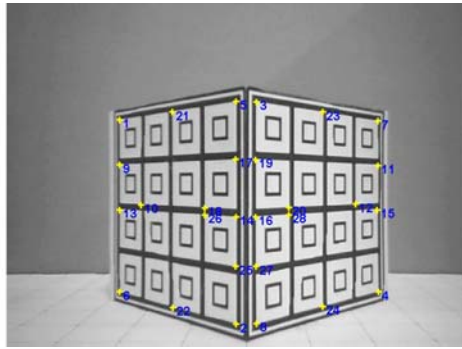


Figura 8: Objeto calibrado e pontos de medição utilizados para validar a reconstrução 3D.

Na Tabela 1 têm-se os valores das distâncias medidas, os valores reais e os erros obtidos na reconstrução. Verifica-se na Tabela 1 que se obteve na maioria dos pontos um valor de erro razoável, no entanto em alguns pontos o erro foi superior a 10%, que é um valor elevado. Este valor elevado pode ser consequência dos poucos valores de disparidade que se tem para este sistema estéreo, um menor erro pode ser conseguido aumentando-se o valor dos comprimentos das linhas de base. Isto não foi realizado neste momento por se ter uma limitação física na translação da câmera, limitação esta imposta pelo parafuso micrométrico utilizado.

Tabela 1: Valores medidos e valores reais das distâncias em 3D entre os pontos da Figura 8.

Pontos	Distância Medida [mm]	Distância Real [mm]	Erro [%]
1-2	261	252	3,6
3-4	242	252	-4,0
5-6	236	252	-6,3
7-8	267	252	6,0
9-10	64	57	12,3
11-12	66	57	15,8
13-14	190	178	6,7
15-16	198	178	11,2
17-18	56	57	-1,8
19-20	57	57	0,0
21-22	182	178	2,2
23-24	181	178	1,7
25-26	55	57	-3,5
27-28	56	57	-1,8
5-2	182	178	2,2
3-8	182	178	2,2
1-6	181	178	1,7
7-4	186	178	4,5

5. CONCLUSÃO

Neste trabalho apresentou-se um sistema estéreo de baixo custo e que utiliza três imagens da cena adquiridas em pontos de visão diferentes. Embora se tenha utilizado uma web-câmera, que não possui uma boa qualidade da imagem e nem um bom conjunto óptico, os resultados se mostraram atraentes e promissores.

O processo de calibração do sistema estéreo é simples, não exige vários ensaios experimentais e são apenas dois os parâmetros a serem encontrados, f_x e k_t , além dos comprimentos das linhas de base (B) que devem ser conhecidos.

O método de busca pelos pontos correspondentes também é simples e se mostra eficaz com um custo computacional relativamente baixo. Quanto à reconstrução 3D, melhores resultados podem ser obtidos com o aumento do comprimento da linha de base entre os pares de câmeras estéreo.

Em trabalhos futuros será desenvolvido um método de posicionamento e calibração de um sistema estéreo com a mesma configuração utilizada neste trabalho, porém utilizando realmente três câmeras. As três câmeras deverão ter as mesmas características de posicionamento que aquelas presentes no sistema estéreo apresentado neste trabalho. Esta calibração será realizada utilizando-se um objeto padrão, capaz de permitir o posicionamento e o ajuste corretos da câmera.

6. AGRADECIMENTOS

Os autores agradecem o apoio financeiro proporcionado a este trabalho pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

4. REFERÊNCIAS

- Forsyth, D. A. e Ponce, J., 2003, “Computer Vision – A Modern Approach”, Prentice Hall, ISBN-0-13-085198-1.
- Gonzalez, R. C. e Woods, R. E., 1993, “Digital Image processing”, Addison-Wesley Pub. Comp. Inc., ISBN-0-201-60078-1.
- Kanade, T., Kano, H., Kimura, S. Yoshida, A. e Oda, K., 1995, “Development of a Video-Rate Stereo Machine”, Proc. of International Robotics and Systems Conference, Pittsburgh.
- Kanade, T., Yoshida, A., Oda, K., Kano, H. e Tanaka, M., 1996, “A Stereo Machine for Video-Rate Dense Depth Mapping and Its New Applications”, Proc. Of 15th Computer Vision and Pattern Recognition Conference, São Francisco.
- Okutomi, M. e Kanade, T., 1993, “A Multiple-Baseline Stereo”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, N° 4, pp. 353-363.
- Trucco, E. e Verri, A., 1998, “Introductory Techniques for 3-D Computer Vision”, Prentice Hall, ISBN-0-13-261108-2.
- Zhang, Z., Faugeras, O. e Deriche, R., 1997, “An Effective Technique for Calibrating a Binocular Stereo Through Projective Reconstruction Using Both a Calibration Object and the Environment”, Videre: A Journal of Computer Vision Research (MIT Press), Vol.1, pp. 58-68.

A PARALELL TRINOCULAR STEREO VISION SYSTEM

Patrick Magalhães Cardoso

Federal University of Uberlândia
School of Mechanical Engineering
patrick@mecanica.ufu.br

José Francisco Ribeiro

jribeiro@mecanica.ufu.br

Abstract: *This paper presents a trinocular stereovision system. A web cam is used to acquire the images in three different points. The web cam has a translation in horizontal axis keeping the frame system parallel. It is present a simple calibration method where are only two parameters for to estimate: the focal length in horizontal pixels and a radial distortion parameter. It is used the Sum of Absolute Differences (SAD) to search the matches points in the images. Some results are presents and the stereovision system is attractive for the simplicity and the low cost of it.*

Keywords: *stereo, computer vision, trinocular, camera.*