

Metodologia para Geração de Mapa de Profundidade Térmico para Detecção de Pedestres

Felipe Bueno

Universidade Tecnológica Federal do Paraná
Av. Monteiro Lobato, s/n – Km 04
Ponta Grossa – PR – Brasil
felipebueno@alunos.utfpr.edu.br

Erikson Freitas de Moraes

Universidade Tecnológica Federal do Paraná
Av. Monteiro Lobato, s/n – Km 04
Ponta Grossa – PR – Brasil
emorais@utfpr.edu.br

Max Mauro Dias Santos

Universidade Tecnológica Federal do Paraná
Av. Monteiro Lobato, s/n – Km 04
Ponta Grossa – PR – Brasil
maxsantos@utfpr.edu.br

RESUMO

Estatísticas demonstram a abrangência do número de acidentes envolvendo pedestres. Uma forma de minimizar esses indicadores é através do uso de funções de auxílio ao condutor. Para tal fim uma das possibilidades consiste em utilizar informações de temperatura e coordenadas tridimensionais da rodovia para identificação dos pedestres. No presente trabalho propõe-se uma metodologia capaz de construir um mapa de profundidade utilizando câmeras térmicas organizadas em um par estéreo, o qual indica as distâncias dos pontos observados. Além de outras utilidades, o mapa de profundidade térmico (MPT) representa uma ferramenta importante para sistemas de auxílio ao motorista.

Palavras-chave

Estereoscopia; Termografia; Detecção de Pedestres; Visão Computacional.

ABSTRACT

Statistics evidence the coverage of the number of accidents relating to pedestrians. A way to minimize these indicators is through the use of drive assistance functions. For this purpose one of the possibilities consists in using temperature info and road tridimensional coordinates to pedestrian identification. In the present paper it is proposed a methodology to build a depth map using thermal cameras arranged in a stereo pair, which indicates the distances of the observed points. Besides other applications, the thermal depth map (TDM) represents an important tool for driver assistance systems.

Keywords

Stereoscopy; Thermography; Pedestrian Detection; Computer Vision.

1. INTRODUÇÃO

No Brasil, cerca de 90% dos acidentes de trânsito têm como causa principal falha dos condutores [12]. Nos Estados Unidos, essa estatística chega a aproximadamente 94% [10]. Os fatores variam entre imprudência (desrespeito à legislação, como realização de ultrapassagens em lugares proibidos) e imperícia (falta de experiência do condutor, o que leva o mesmo a não reagir de forma adequada perante situações de risco no trânsito). Segundo a Organização Mundial de Saúde, cerca de 1,25 milhões de pessoas morrem nas estradas todos os anos [4]. Dentro da estatística de acidentes com veículos, nos EUA cerca de 25% deles resultam em morte ou ferimento de pedestres [11].

Uma forma de minimizar essa estatística consiste em utilizar sistemas de auxílio ao motorista capazes de atuar na tomada de decisão do condutor. Esses sistemas são chamados de ADAS (*Advanced Driver Assistance Systems*), e são capazes de gerar alertas em situações de risco ou até mesmo assumir o controle do veículo, como no caso da função de estabilidade de faixa (*LKA – Lane Keep Assist*) [1, 3].

No contexto de acidentes envolvendo pedestres, pode-se encarar como um desafio a criação de sistemas que atuem em sua detecção. Esses tipos de sistemas podem ser usados para alertar o condutor e/ou atuar no veículo conduzido, como foi feito em [2], onde a partir de um sistema que recebe como entrada dados de laser e câmera, é emitido um alerta ao condutor caso seja detectado um pedestre em regiões definidas como críticas. Outras abordagens fazem uso de uma câmera térmica, a partir da fusão de imagens térmicas e convencionais aplicadas em cenas estáticas [5], e detecção e rastreamento de pedestres posicionados a longas distâncias usando uma câmera térmica e análise do histograma do gradiente orientado [7]. Outras ainda utilizam combinações entre dois pares de câmeras convencionais e térmicas [8], promovendo uma fusão de dados multimodal e valendo-se da assinatura de calor característica dos seres humanos, realçada em relação ao ambiente.

Além da detecção do pedestre na cena, é interessante recuperar sua posição no espaço, para que o sistema ADAS consiga tomar alguma ação antes que uma possível colisão ocorra. Uma das formas de se obter a posição do ponto observado é a partir da visão estéreo. Nesse sentido, a pesquisa descrita neste trabalho tem por objetivo a geração de um mapa de profundidade a partir de imagens contendo informações de temperatura e organizadas em um par estéreo calibrado. O Mapa de Profundidade Térmico (MTP) representa uma ferramenta importante para sistemas de auxílio ao motorista, principalmente por apresentar as diferentes temperaturas da cena mapeadas a uma estimativa confiável de distância. O resultado dessa fusão de dados é uma visão privilegiada capaz de identificar corpos quentes mapeados em coordenadas 3D, além de não apresentar sensibilidade a interferências e ruídos tais como fumaça, neblina e variações na iluminação.

O trabalho está dividido da seguinte maneira: Na seção 2 são apresentados conceitos necessários para o desenvolvimento da proposta. A seção 3 apresenta a metodologia proposta para cálculo do mapa de profundidade, e na seção 4 é descrita a conclusão e os passos futuros.

2. REFERENCIAL TEÓRICO

Nesta seção serão introduzidos conceitos necessários para melhor entendimento deste trabalho:

A termografia refere-se ao registro da informação térmica. No contexto de processamento de imagens, o produto da termografia é o termograma, o qual é uma imagem que indica a temperatura de cada ponto do ambiente registrado a partir de uma determinada escala (tons de cinza ou paleta de cores).

A Radiação Infravermelha, por sua vez, caracteriza a transferência de energia térmica (calor) através de radiação. Todo objeto com temperatura superior ao zero absoluto emite essa radiação. Quando esta radiação incide sobre um corpo, uma parte é absorvida, outra é refletida e outra passa através do mesmo [20]. Quando toda a radiação é absorvida tem-se o chamado corpo negro.

Emissividade é a capacidade da superfície do material de emitir radiação. É a razão entre a quantidade de energia emitida por um determinado corpo, e seu corpo negro, respeitando-se as mesmas condições de temperatura e comprimento de onda [9]. A emissividade geralmente é um parâmetro de configuração das câmeras térmicas, referente à sensibilidade do sensor.

A Visão Estéreo, ou estereoscopia, permite a obtenção de informações tridimensionais e distância do ponto observador até determinado objeto, a partir de sua observação de dois ou mais pontos de vista [17].

A geometria da visão estereo é chamada de Geometria Epipolar. Tem como base o modelo de câmera *pinhole*, descrito em [6] e ilustrado na Figura 1, que em linhas gerais define que em uma câmara escura com um orifício com diâmetro tendendo a zero (ponto focal), a convergência dos raios de luz captados através deste orifício faz com que seja projetada no fundo da câmara a imagem observada. Esse modelo é importante para entender como uma imagem nas câmeras atuais é formada, em função do que é captado através da lente da câmera. É interessante perceber que a imagem projetada no fundo também pode ser obtida antes de ser invertida pela projeção através do orifício, como pode ser observado na Figura 1, no plano da imagem e plano virtual da imagem. A distância focal, por sua vez, é a distância entre o ponto de convergência da luz (ponto focal) e a projeção da imagem no plano da imagem. Em câmeras convencionais essa distância é determinada a partir do modelo da lente utilizada.

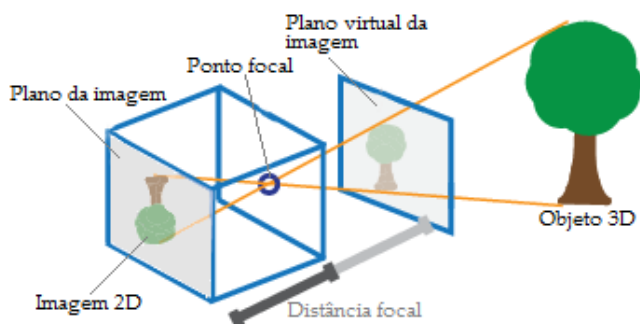


Figura 1: Modelo de câmera *pinhole*. Adaptado de [19].

Com base no entendimento da projeção de uma imagem a partir de uma câmera, pode-se adotar o mesmo modelo, porém para projeções a partir de n câmeras, e então traçar relações geométricas

que nos permitem descobrir informações acerca da posição de um objeto no espaço e suas dimensões. A Figura 2 apresenta um modelo básico da geometria epipolar a partir de dois pontos de vista distintos.

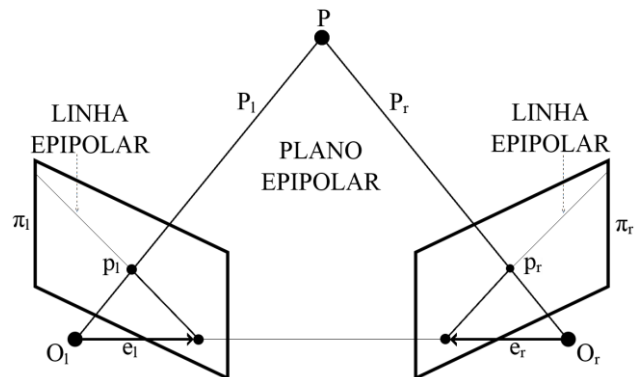


Figura 2: Geometria epipolar. Adaptado de [17].

Baseado na Figura 2, pode-se definir terminologias que serão importantes para a compreensão deste trabalho. Os símbolos O representam o ponto focal, para as câmeras da esquerda e direita. Os símbolos π representam os planos da imagem. Os caracteres e são chamados de epipolos, pois indicam os pontos onde ocorre a interseção entre as projeções dos pontos focais. As linhas representadas por P indicam o mesmo ponto em coordenadas do mundo (tridimensionais), e as linhas indicadas por p indicam as projeções desse ponto nos planos da imagem. As linhas epipolares indicadas na Figura 2 indicam a localização do plano epipolar. Ambas fazem parte da definição de restrição epipolar, a qual significa que para um determinado ponto no plano da imagem de uma das câmeras, o ponto correspondente na outra deve estar ao longo da linha epipolar [17]. A identificação do exato ponto correspondente ao longo da linha epipolar caracteriza o chamado problema de correspondência.

A Calibração de Câmera refere-se ao processo de extração (identificação) dos parâmetros intrínsecos e extrínsecos da câmera. Os parâmetros intrínsecos são distância focal (em pixels), coordenadas do ponto principal (centro do plano da imagem), e tamanho do pixel (vertical e horizontal). Os parâmetros extrínsecos são os valores de rotação e translação que permitem retificar as imagens para um mesmo plano [17].

Matriz Fundamental é a matriz capaz de correlacionar um ponto na imagem A do par estereo com a linha epipolar da imagem B (transformação de espaço), dessa forma otimizando a análise da imagem para resolver o problema de correspondência [6].

O Mapa de Disparidade pode ser caracterizado como uma imagem que denota a diferença da localização de um mesmo ponto tridimensional projetado a partir da perspectiva de pontos de vista distintos [14]. Se um ponto A em uma imagem 1 está localizado na coordenada (100, 20), e B (135, 20) é seu correspondente na imagem 2, a disparidade será de 35 pixels. Dessa forma, no mapa de disparidade, pode ser aplicada uma normalização nos valores de disparidade para que estes assumam um valor no intervalo de 0 a 255 permitindo sua visualização em tons de cinza.

O Mapa de Profundidade, por sua vez, é uma imagem que representa as distâncias dos objetos observados em relação ao

ponto observador, que pode ser considerado como sendo o ponto médio entre as câmeras. Nesse caso a profundidade é o valor da distância de um ponto da cena ao ponto observador e o mapa é uma matriz de profundidades referentes à cena.

3. METODOLOGIA PROPOSTA

Para a obtenção do mapa de profundidade, é necessário completar uma série de etapas, detalhadas na Figura 3.

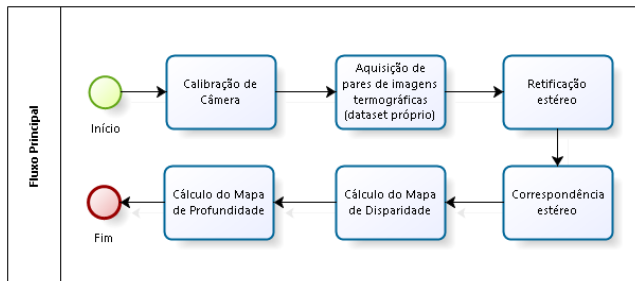


Figura 3: Fluxo da metodologia proposta. Autoria própria.

A primeira etapa apresentada consiste na calibração das câmeras utilizadas. Essa calibração pode ser dividida em geométrica e radiométrica, além da definição dos parâmetros de emissividade, relativos ao sensor térmico da câmera. A calibração radiométrica não faz parte do escopo deste trabalho, portanto não será considerada. A calibração será realizada individualmente em cada câmera que compõe o par estéreo.

O processo pode ser realizado a partir de rotinas que analisam um alvo com um padrão de calibração de geometria conhecida e que possua características visuais que possam ser localizadas facilmente. Normalmente usa-se um objeto planar com dimensões conhecidas tal como um tabuleiro de xadrez, como ilustrado na Figura 4 [16]. Esse padrão também é importante pois é utilizado para a caracterização das distorções radiais presentes na câmera. A análise e extração dessas distorções também não fazem parte do escopo deste trabalho e não serão consideradas.



Figura 4: Alvo de calibração. Autoria de [13].

Considerando que a entrada será composta por imagens que representam um mapa de calor, será necessário utilizar um alvo de calibração apropriado onde seja possível destacar as diferenças de temperatura, como foi feito em [13], onde as regiões de maior calor equivalem aos quadrados pretos da Figura 4.

Com a câmera calibrada, é possível dar início à criação de um *dataset* próprio, que será composto de imagens de pessoas, captadas a partir de pontos fixos. Essas imagens conterão o termograma, que é uma representação 2D em uma escala de cores referente à temperatura medida pela câmera em cada ponto,

também chamadas de imagens térmicas. O resultado dessa etapa consiste em pares de imagens térmicas captadas a partir de duas posições distintas. Após a captura, o próximo passo apresentado na Figura 3 é a retificação estéreo. Esse processo consiste em alinhar as imagens em um mesmo plano de modo que os pares das linhas epipolares conjugadas tornem-se colineares e paralelos a um dos eixos [17]. Com isso, é possível promover um alinhamento 2D, de modo a restringir a busca de pontos de correspondência entre as imagens 1 e 2 ao longo da linha epipolar.

Após a geração do *dataset* com as imagens retificadas, será realizado o processo de correspondência estéreo. Esse processo, consiste em identificar os pixels correspondentes entre as imagens. Existem diversos algoritmos que tratam do problema de correspondência. Contudo, será realizado um levantamento bibliográfico de forma a selecionar uma técnica ou um conjunto de técnicas, de modo a promover uma correspondência em imagens térmicas adequada ao propósito do trabalho.

De posse das posições correspondentes de cada ponto das imagens retificadas, a etapa a seguir refere-se ao cálculo do mapa de disparidade, o qual pode ser representado por uma matriz com as distâncias entre os pontos homólogos nas imagens do par estéreo, ou seja, é uma imagem contendo as distâncias entre os pontos correspondentes identificados na etapa anterior.

O mapa de profundidade, por fim, pode ser obtido a partir do mapa de disparidade, distância focal e distância entre os centros de projeção das câmeras (*baseline*), pois quanto maior a disparidade, mais próximo o objeto está das câmeras, portanto menor sua profundidade. A equação (1) ilustra o cálculo da profundidade em imagens retificadas para cada pixel,

$$z = bf \frac{1}{d} \quad (1)$$

onde d corresponde à disparidade, b à distância *baseline*, f à distância focal e z à profundidade. A Figura 5 ilustra a imagem gerada a partir do mapa de profundidade a partir de pontos de vista distintos. A escala de cinza indica a distância em relação ao ponto observador, onde quanto mais claro mais próximo.

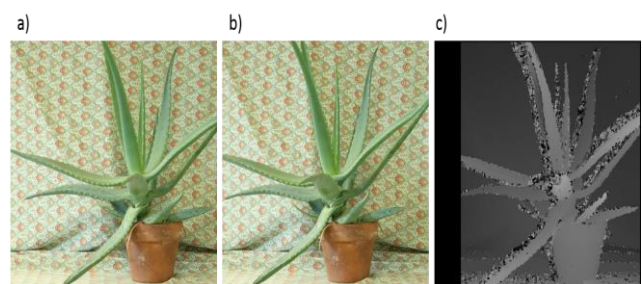


Figura 5: Geração do mapa de disparidade. a) Imagem captada pela câmera da esquerda. b) Imagem captada pela câmera da direita. c) Mapa de disparidade. Autoria própria.

A detecção do pedestre na imagem pode ser feita a partir da aplicação de uma técnica de segmentação que se valha da assinatura térmica característica de seres vivos na imagem. Após isso, a partir do mapa de profundidade calculado e utilizando-se um pixel referente ao pedestre segmentado, obtém-se a distância entre este e o ponto observador (veículo). De posse dessas informações, um sistema ADAS pode alertar o condutor a partir da ativação de um

sinal sonoro, vibrando o volante do veículo ou acionando seu sistema de freios, caso este esteja em rota de colisão, evitando um possível acidente.

4. CONCLUSÃO

O mapa de profundidade representa uma etapa importante em qualquer sistema baseado em geometria estéreo. A partir da metodologia apresentada é possível estabelecer um mapa de profundidade, identificando distâncias, associadas à informação de temperatura em uma cena. No contexto de identificação de pedestres, a aquisição dessas informações pode representar um fator crucial para evitar colisões, diminuindo assim os indicadores referentes a ferimentos ou mortes em acidentes com veículos.

Além da possibilidade de utilização como entrada de dados para uma função ADAS, a informação térmica e de profundidade de uma cena possui aplicações em diversos cenários, como sistemas de vigilância, sistemas de monitoramento de equipamentos [18], auxílio no diagnóstico médico [15], entre outras.

Como ponto de melhoria propõe-se a definição de estratégias que atenuem o impacto da possível baixa eficácia na detecção de pedestres em cenários onde o plano de fundo captado está com temperatura elevada, próxima à do pedestre.

Como passos futuros propõe-se a geração de um modelo 3D a partir dos dados coletados e estabelecimento de métricas para validação de testes e verificação de eficácia. Além disso, propõe-se a geração tridimensional *on the fly* em um sistema móvel, além da adaptação do modelo para utilização de câmeras convencionais em conjunto com as térmicas (sistema multimodal).

5. REFERÊNCIAS

- [1] Bishop, R. 2005. Lane Keeping Assist Systems. *Intelligent vehicle technology and trends*. Artech House.
- [2] Broggi, A., Cerri, P., Ghidoni, S., Grisleri, P. and Jung, H.G. 2009. A New Approach to Urban Pedestrian Detection for Automatic Braking. *IEEE Transactions on Intelligent Transportation Systems*. 10, 4 (Dezembro 2009), 594–605. DOI:<https://doi.org/10.1109/TITS.2009.2032770>.
- [3] Geronimo, D., Lopez, A.M., Sappa, A.D. and Graf, T. 2010. Survey of Pedestrian Detection for Advanced Driver Assistance Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 32, 7 (Jul. 2010), 1239–1258. DOI:<https://doi.org/10.1109/TPAMI.2009.122>.
- [4] Global status report on road safety 2015: 2015. http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/. Accessed: 2017-07-23.
- [5] Goubet, E., Katz, J. and Porikli, F. 2006. Pedestrian tracking using thermal infrared imaging. (2006), 62062C.
- [6] Hartley, R. and Zisserman, A. 2004. Camera Models. *Multiple view geometry in computer vision*. Cambridge University Press.
- [7] Kim, D.E. and Kwon, D.S. 2015. Pedestrian detection and tracking in thermal images using shape features. *2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (Outubro 2015), 22–25.
- [8] Krotosky, S.J. and Trivedi, M.M. 2007. On Color-, Infrared-, and Multimodal-Stereo Approaches to Pedestrian Detection. *IEEE Transactions on Intelligent Transportation Systems*. 8, 4 (Dezembro 2007), 619–629. DOI:<https://doi.org/10.1109/TITS.2007.908722>.
- [9] Minkina, W. and Dudzik, S. 2009. Measurements in Infrared Thermography. *Infrared Thermography: Errors and Uncertainties*. John Wiley & Sons, Ltd.
- [10] NHTSA 2015. *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*. NHTSA - National Highway Traffic Safety Administration.
- [11] NHTSA 2015. *TRAFFIC SAFETY FACTS 2015*. NHTSA - National Highway Traffic Safety Administration.
- [12] OBSERVATÓRIO | 90% dos acidentes são causados por falhas humanas, alerta OBSERVATÓRIO: 2015. <https://www.onsv.org.br/90-dos-acidentes-sao-causados-por-falhas-humanas-alerta-observatorio/>. Accessed: 2017-05-01.
- [13] Saponaro, P., Sorensen, S., Rhein, S. and Kambhamettu, C. 2015. Improving calibration of thermal stereo cameras using heated calibration board. *2015 IEEE International Conference on Image Processing (ICIP)* (Sep. 2015), 4718–4722.
- [14] Shapiro, L. and Stockman, G.C. 2001. Perceiving 3D from 2D Images. *Computer Vision*. Prentice Hall.
- [15] Souza, M.A. de, Sanches, I.J. and Gamba, H.R. 2012. A New Method for Generating 3D Thermography Models. (2012).
- [16] Trucco, E. and Verri, A. 1998. Camera Calibration. *Introductory techniques for 3-D computer vision*. Prentice Hall.
- [17] Trucco, E. and Verri, A. 1998. Stereopsis. *Introductory techniques for 3-D computer vision*. Prentice Hall.
- [18] Vidas, S., Moghadam, P. and Bosse, M. 2013. 3D thermal mapping of building interiors using an RGB-D and thermal camera. *2013 IEEE International Conference on Robotics and Automation* (Maio 2013), 2311–2318.
- [19] What is Camera Calibration? 2017. <https://www.mathworks.com/help/vision/ug/camera-calibration.html>. Accessed: 2017-08-23.
- [20] Živčák, J., Hudák, R., Madarász, L. and Rudas, I.J. 2013. Physical Nature of Thermography. *Methodology, models and algorithms in thermographic diagnostics*. Springer Science & Business Media. 31--54.