

CONSTRUÇÃO 3D DE OBJETOS ATRAVÉS DE FOTOGRAFIAS 2D

GERSON DANTAS ROCHA¹; MARCELO LEMOS ROSSI²

¹Universidade Federal de Pelotas UFPel¹ – dantas_gerson@hotmail.com 1

²Universidade Federal de Pelotas UFPel³ – marcelo.rossi@ufpel.edu.br 2

1. INTRODUÇÃO

A visão computacional é uma área de pesquisa que inclui métodos de aquisição, processamento, análise e reconhecimento de imagens. Os algoritmos de visão computacional são baseados em técnicas de processamento e análise de imagens, as quais permitem extrair informações a partir de imagens (GONZALEZ e WOODS, 2007; TRUCCO E VERRI, 1998; FORSYTH, 2012).

Este trabalho visa desenvolver mecanismos para se obter um modelo 3D de objetos através de fotografias em 2D tiradas em diferentes ângulos. A partir deste trabalho, poderão ser feitas pesquisas aprofundadas nas áreas de visão computacional, objetos e sistemas autônomos, preservação de patrimônio histórico, videogames, realidade aumentada, reconhecimento de espaço e na área da saúde. Encontramos aplicações de visão computacional em diversos tópicos da ciência, desde sistemas de reconhecimento de movimentos ou rostos, sistemas de realidade aumentada (Azuma, 1997), automatização de análise de imagens (útil para grandes bancos de imagens, como em registros de câmera de segurança), até áreas interdisciplinares como a interpretação de imagens médicas e biológicas.

Ao final deste trabalho será acrescido todo o conteúdo e todo mecanismo de processamento de imagem para o reconhecimento da cidade Pelotas.

Para ter o objetivo de se ter uma imagem 3D baseada em fotos 2D foi implementado em laboratório uma bancada constituída de um circuito capaz de reconhecer ordens do computador, um circuito capaz de controlar um motor de passo e um motor de passo onde se é empregada uma base acima do mesmo para se colocar objetos em cima desta base, para assim, apontar uma câmera, a fim de se obter fotos do objeto quando o motor de passo gira conforme o comando do computador.

Com o uso da visão computacional é possível automatizar a tomada de decisões com base em imagens capturadas por uma câmera (SHAPIRO e STOCKMAN, 2001)

Após a extração das imagens em 2D do objeto que está na mesa é feita a parte computacional que consiste em técnicas de segmentação de imagens, calibração da câmera (obtenção das características intrínsecas), obtenção da posição da câmera (características extrínsecas), obtenção de pontos característicos da imagem e triangulação (obtenção da distância do ponto da imagem até a câmera para o reconhecimento da imagem em 3D).

2. METODOLOGIA

A equação da câmera de orifício é mostrada a seguir na equação (1).

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = [p] \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (1)$$

$$[p] = [H] \cdot [E] \quad (2)$$

Onde a primeira matriz da Equação (1) são as coordenadas homogenia do ponto no plano 2D na foto, a matriz $[p]$ é a matriz da câmera que reúne as características intrínsecas (distorções da lente, pixel e foco, $[H]$) e extrínsecas (posição e ângulo da câmera, $[E]$), a última matriz é a coordenada homogenia do ponto no espaço.

A matriz $[p]$ não é invertível e, dessa forma, pode-se dizer que o resultado da Equação 1 é a perda da coordenada de profundidade.

Para recuperar a profundidade do ponto deve-se ter no mínimo duas imagens para, assim, realizar o processo de triangulação. Para fazer o processo de triangulação é necessário conhecer um mesmo ponto nas duas imagens, ou seja, obter um ponto na primeira imagem e encontrar este mesmo ponto na segunda imagem. Essa busca, quando realizada com todos os pontos das imagens, tem um consumo computacional elevado e grandes chances de se obter resultados incoerentes. Assim, para solucionar este problema é necessário encontrar pontos característicos nas imagens, detectar junções, cantos ou regiões específicas e únicas da imagem.

A identificação de junções ou cantos em imagens digitais fornece informação importante para várias aplicações envolvendo análise de imagens, como a localização de pontos homólogos em visão estereoscópica, detecção de movimento, rastreamento de objetos e restauração de imagens (Pedrini, 2008).

Para tal detecção é usado dois detectores que são: Harris e SURF.

O detector de Harris, também conhecido como algoritmo de Plessey calcula derivadas por meio de convolução da imagem com uma máscara de suavização Gaussiana, ou seja, é apropriado para o reconhecimento de quinas ou junções.

O detector SURF (*Speeded – UP Robust Features*) é um algoritmo proposto Bay *et al.* (2006) rápido e robusto para descrição e extração de pontos de característicos em uma imagem (é apropriado para objetos com geometria orgânica).

Para o reconhecimento de um ponto de uma imagem na outra foi utilizado usar o método de Ransac, um algoritmo proposto por *Fischler e Bolles* que tenta reconhecer pontos em um conjunto de amostras através da máxima verossimilhança, ou seja, ele faz um cálculo robusto que permite distinguir pontos de emparelhamento verdadeiros e falsos denominados inliers e outliers (AZEVEDO, TAVARES, VAZ, 2005). Para continuar com o processo de triangulação é necessário obter as características extrínsecas da câmera que é a posição da câmera e o ângulo da mesma que é dado pelo giro do motor de passo controlado pelo computador, esse dado já é conhecido.

Na tentativa de se reduzir os erros de todos os processos, feitos até então, é feita a segmentação de imagem. A segmentação é um processo que consiste em dividir as imagens em segmentos homogêneos (GONZALEZ; WOODS, 2000). A classificação pode ser definida de forma genérica como um processo que consiste em agrupar pixels em classes pré-estabelecidas pelo usuário (JENSEN, 2005). A técnica de segmentação escolhida é baseada em limiar de cores utilizando o sistema de cores cieLab, que consiste em um plano de cores em que as componentes separadas assumem L, a e b. A componente L é o brilho da imagem e as componentes a e b são os eixos do diagrama de cores da matriz. desta forma as cores são diferenciadas e é possível isolar apenas uma cor que no caso deste trabalho será o verde limão.

3. RESULTADOS E DISCUSSÃO

Na figura.1 vemos o objeto em cima da base onde o motor será ordenado a girar computacionalmente, na mesma figura a cor do fundo e a cor da base são a mesma, verde-limão, para facilitar a remoção deste chroma key. Na figura 2 é mostrado o objeto já segmentado com a remoção do chroma key, note que o cieLab conseguiu remover todo o verde limão, mostrando que esta técnica é útil e de baixo consumo computacional.

Ainda na Figura.2 é mostrado o resultado da triangulação com o algoritmo para a obtenção de uma nuvem de pontos em 3D (3D esparso). Note que em comparação com a imagem segmentada (fundo preto) é possível detectar visualmente o 3D pelas cores dos pontos, desta forma, é afirmado que as técnicas empregadas conseguiram êxito o resultado desejado.



Figura.1. Objeto em cima da base que é girada pelo motor de passo.

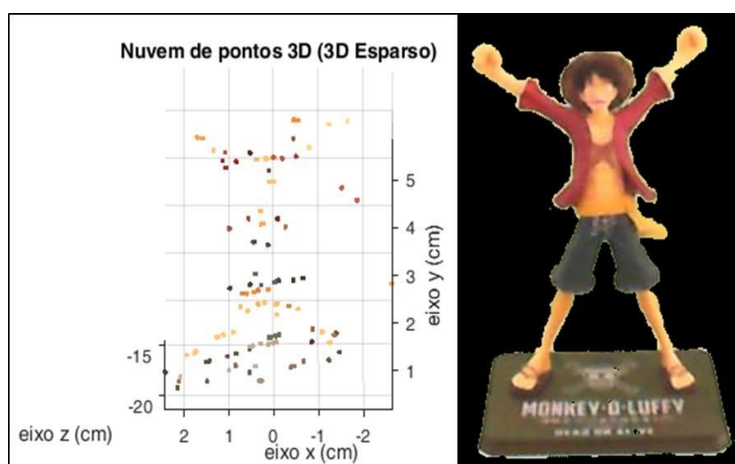


Figura.2 Imagem segmentada (fundo preto) e a nuvem de pontos 3D.

4. CONCLUSÕES

Seguindo o pensamento de que dentro da área de processamento de imagens, em relação a segmentação foram abordados procedimentos que se baseiam na afinidade do objeto em questão ou nos padrões da imagem. Todas essas técnicas, apesar de serem aplicadas em campo geral, podem ser aplicadas em casos particulares específicos, já que, o uso das mesmas tem resultados muito satisfatórios.

Ao longe deste trabalho é proposto um aproveitamento computacional, com interface gráfica adequada para, assim, admitir a empregabilidade e a comparação das cátedras consideradas.

Seguindo todas essas técnicas é possível criar a nuvens de pontos e permite que o computador visualize em 3D e aplique em técnicas possíveis, em trabalhos futuros pode ser feito o refinamento da nuvem de pontos e o 3D denso.

5. REFERÊNCIAS BIBLIOGRÁFICAS

GONZALEZ, R. C.; WOODS, E. R. **Digital Image Processing**. 3 ed. Prentice Hall, 2007.

TRUCCO, E.; VERRI, A. **Introductory Techniques for 3-D Computer Vision**. Prentice Hall, 1998.

FORSYTH, D. A.; PONCE, J. **Computer Vision: A Modern Approach**. 2 ed. Prentice Hall, 2012.

SHAPIRO, L.; STOCKMAN, G. **Computer vision**. New Jersey: Prentice Hall, 2001.

BAY, H.; TUYTELAARS, T.; VAN GOOL, L. J. SURF: **Speeded up robust features**. In **European Conference on Computer Vision**, pp. 404-417, 2006.

Pedrini, H. e Schwartz, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações**. 1ª. ed. [S.I.]: PIONEIRA THOMSON, 2008.

AZEVEDO, T. C. S; TAVARES, J. M. R. S.; VAZ, M. A. P. Obtenção da forma 3D de objetos usando uma metodologia de reconstrução de estruturas a partir do movimento In: **CONGRESO DE MÉTODOS NUMÉRICOS EN INGENIERÍA** ; Granada, 2005, SEMNI Departamento de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia da Universidade do Porto , 2005. v. 1. p.6.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de Imagens Digitais. (Traduzido do Inglês, Título original: Digital Image Processing)**: Editora Edgard Blücher. 509 p., 2000.

JENSEN, J. R. **Introductory Digital Image Processing**. 3rd. ed. [S.I.]: Pearson Prentice Hall, New Jersey. 526 p., 2005.

AZUMA, R. T. **A Survey of Augmented Reality. Teleoperators and Virtual Environments** v. 6, n. 4, p. 355-385, 1997.