



PROCESSAMENTO DE NUVEM DE PONTOS 3D
CGEO- 7028

**Determinação de coordenadas 3D com
Luz estruturada**

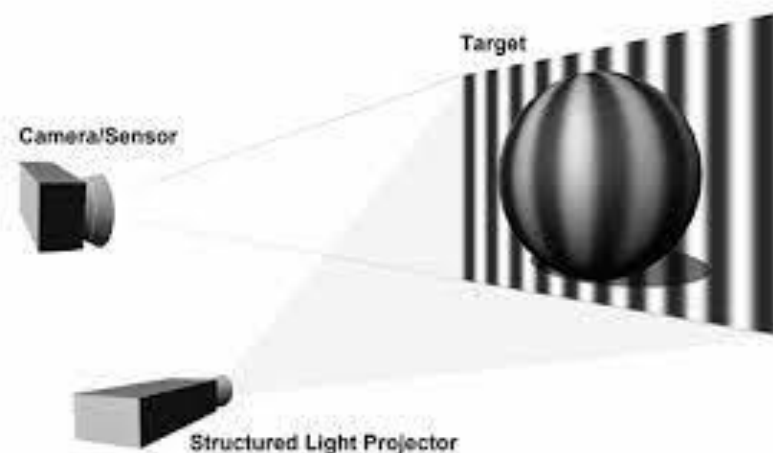
UFPR – Departamento de Geomática
Dr. Ing. Jorge A.S. Centeno
2023
copyright@ centenet

Luz estruturada

Um sistema de luz estruturada se baseia no reconhecimento de padrões conhecidos que são projetados na superfície dos objetos. Geralmente usa-se um projetor de luz e uma câmara.

- O projetor projeta uma imagem conhecida (padrões) sobre a superfície do objeto
- A câmara captura a imagem destes padrões

Os padrões aparecem distorcidos na imagem por causa da variação de profundidade na superfície dos objetos.



Luz estruturada

Vantagem:

- Vários alvos/padrões podem ser medidos simultaneamente
- É possível medir superfícies com textura uniforme.

Problemas:

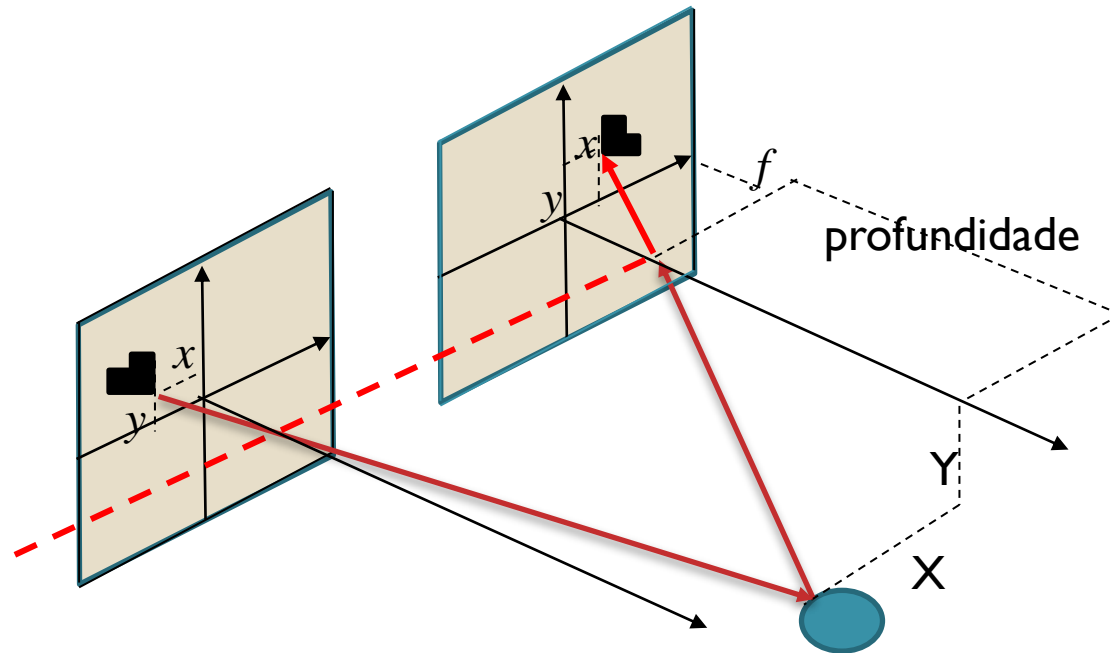
- Contraste,
- diferenças de escala,
- deformação do padrão.



Triangulação

A geometria envolvida relembra o cálculo de profundidade por fotogrametria (intersecção de raios de luz).

O projetor substitui a segunda câmara com a vantagem de que se conhece a posição exata dos pontos nesta imagem.



Câmera estéreo

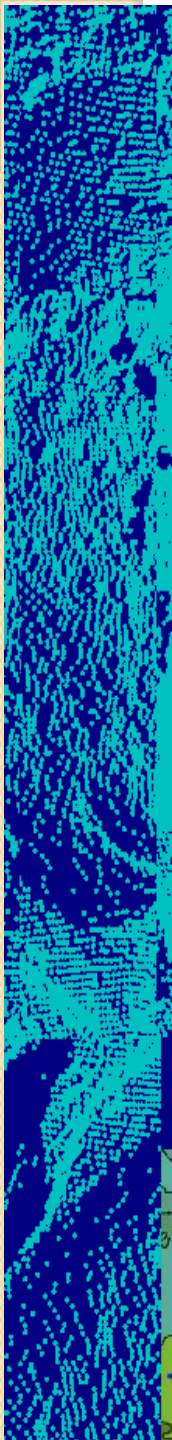
Como a câmera e o projetor estão alinhados, pode-se considerar o caso de duas câmeras (estéreo) paralelas alinhadas na horizontal. Com isto a profundidade é dada por

$$z = b * f / d$$

- z é a profundidade (em metros),
- b é a linha de base horizontal entre as câmeras (em metros),
- f é a distância focal das câmeras (em pixels) e
- d é a disparidade (em pixels).

Se a disparidade for zero, considera-se que os raios de cada câmera são paralelos, e a profundidade é infinita. Valores maiores para a disparidade significam distâncias menores.





Método I – Template matching

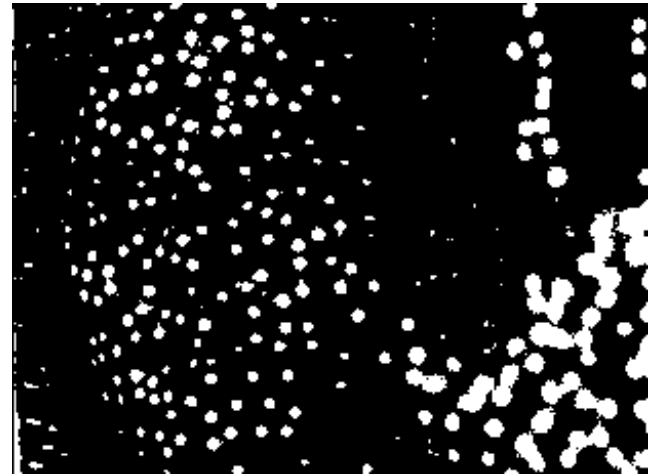
Etapas:

- a) Obtenção da imagem
- b) Segmentação dos padrões
- c) Cálculo da profundidade

Eventualmente, a imagem pode ser afetada por variações de iluminação ou pela cor da superfície. Na fase de pré-processamento pode-se ajustar o contraste para facilitar a detecção dos padrões.

Segmentação

O objetivo desta etapa é delimitar regiões uniformes na imagem que podem corresponder aos padrões projetados.



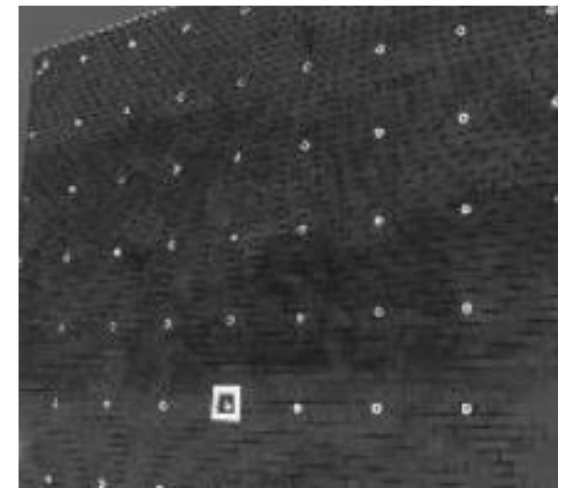
Limiarização

consiste em definir um valor digital (limiar) que permita separar elementos claros (o padrão) de escuros.

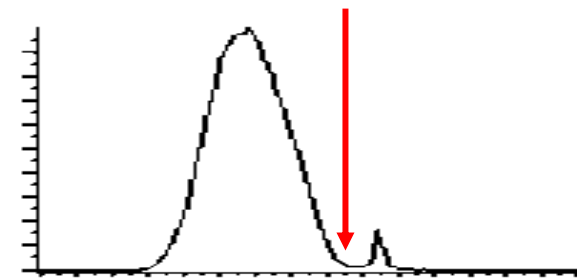
O resultado é uma imagem binária, com valor 0 para os pixels do fundo (background) e 1 para os pixels de interesse (foreground)

A limiarização pode ser global, usando um único limiar para toda a imagem, ou local, variando o limiar para se ajustar às variações de contraste.

O método de OTSU é uma alternativa viável, especialmente eficiente quando o histograma da imagem é bimodal.



Histogramm



Componentes conexos

Quando se usam padrões com formas e tamanhos diferentes, pode-se identificar cada um deles analisando sua forma. Para isto, é necessário, inicialmente, identificar e separar cada elemento. Isto é feito usando o método dos componentes conexos.

Objetivo: identificar e numerar na imagem binária os agrupamentos de pixels conexos.

```
BW = [0 0 0 0 0 0 0 0 0;
      0 1 1 0 0 1 1 1;
      0 1 1 0 0 0 1 1;
      0 1 1 0 0 0 0 0;
      0 0 0 1 1 0 0 0;
      0 0 0 1 1 0 0 0;
      0 0 0 1 1 0 0 0;
      0 0 0 0 0 0 0 0];
```

Connected Components

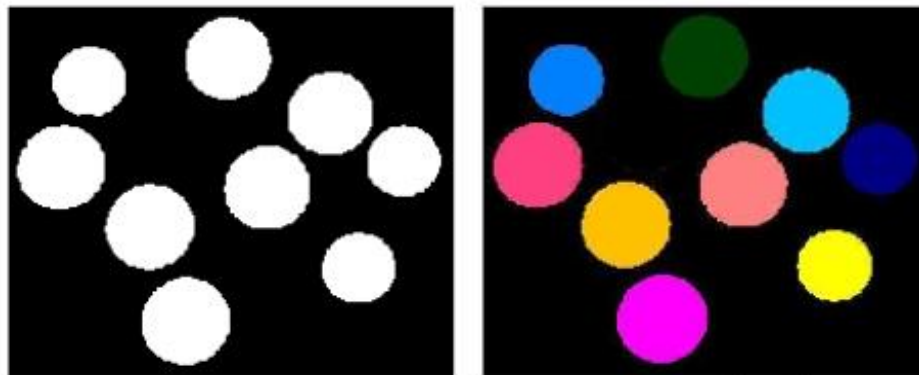
```
BW = [0 0 0 0 0 0 0 0 0;
      0 1 1 0 0 0 3 3 3;
      0 1 1 0 0 0 0 3 3;
      0 1 1 0 0 0 0 0 0;
      0 0 0 0 2 2 0 0 0;
      0 0 0 0 2 2 0 0 0;
      0 0 0 0 2 2 0 0 0;
      0 0 0 0 0 0 0 0 0];
```

Labeled Connected Components

Rotular pixels conexos

Problema: dada uma imagem com pixels rotulados com o mesmo valor e o fundo:

- Encontrar quantos grupos existem e
- Rotular todos os pixels de um mesmo grupo com o mesmo valor



Cada “mancha” é identificada como um grupo uniforme e seus pixels recebem o mesmo identificador.



identificação

A identificação consiste em classificar os segmentos nas classes de padrões previamente definidas.

Para isto, os segmentos podem ser descritos por parâmetros espaciais (tamanho, forma, orientação...) ou cor (quando se usam imagens coloridas) e este vetor de estado pode ser usado para classificar o segmento na classe mais similar.

Porém, o método mais usado é o de comparação do segmento com modelos dos padrões esperados (Template matching)

Também podem ser aplicados métodos de inteligência artificial.

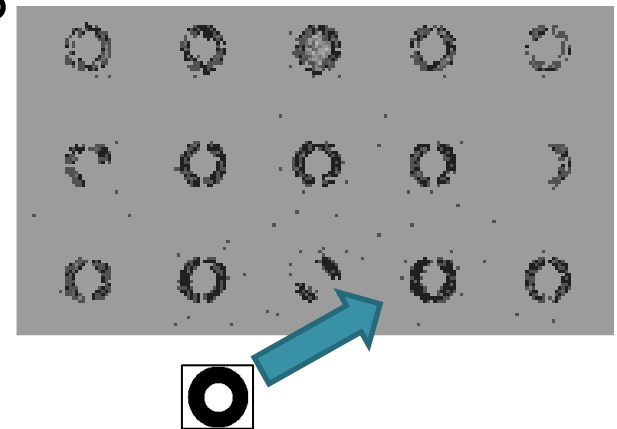
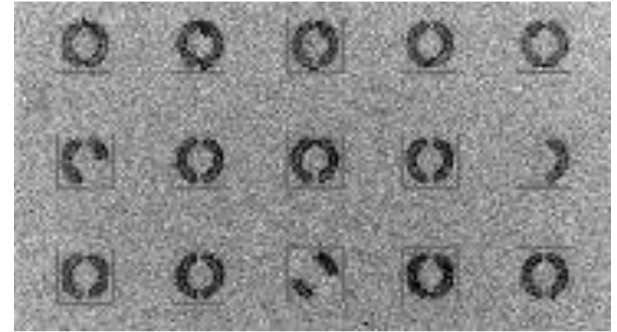
Template matching

Neste caso, a matriz de referência/template (contendo o padrão) é deslocada ao longo de toda a imagem e em cada possível posição é avaliada a similaridade entre o *template* e a região correspondente na imagem é medida.

Se as duas regiões são parecidas, é admitido que o padrão foi encontrado neste local da imagem.

Para medir a similaridade das regiões (match) pode-se usar:

- A soma das diferenças ao quadrado entre pixels
- A correlação
- A soma de coincidências (hits)

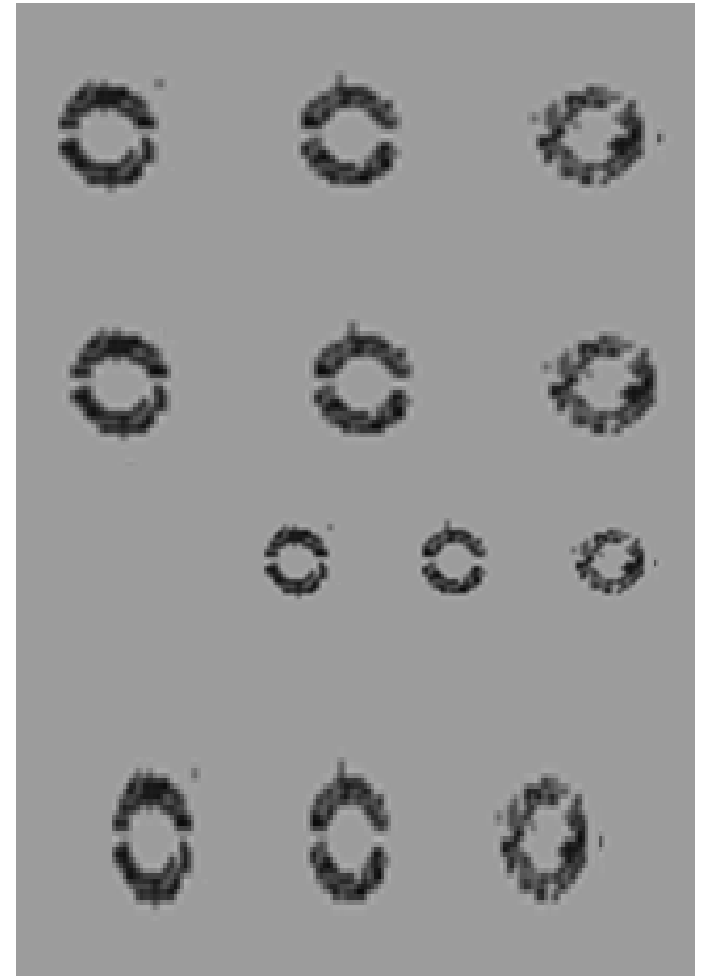


Desvantagem

Este método é fortemente afetado pelas mudanças de escala que podem ser causadas pelas diferenças de profundidade.

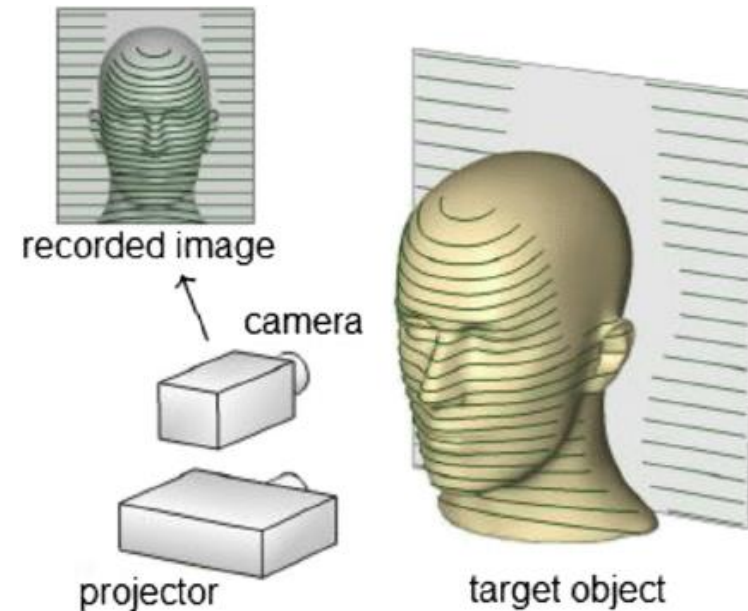
Superfícies inclinadas podem causar deformações de escala diferentes nas direções linhas/colunas e com isso deformar significativamente o segmento.

Quando muitos padrões diferentes são usados, é necessário varrer a imagem com todos eles, o que eleva o custo computacional.



linha

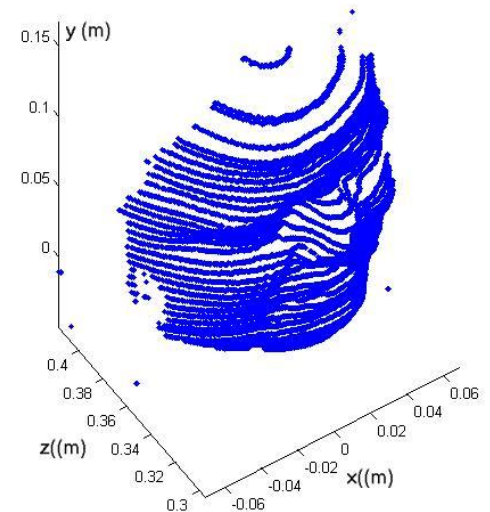
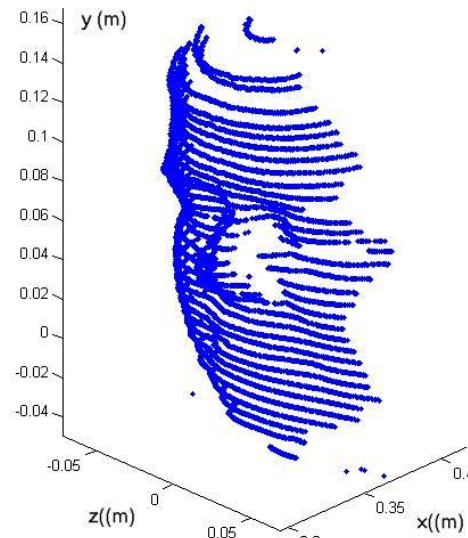
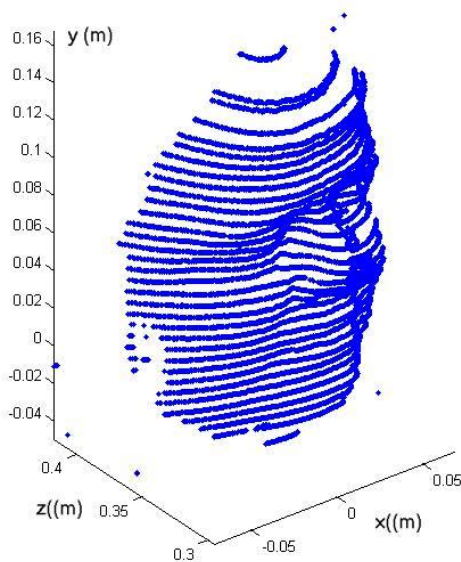
Projeta-se uma linha em diferentes alturas. Quando a linha é projetada sobre objetos a diferentes profundidades, ocorre um deslocamento vertical na linha em função da profundidade.



Passos:

- Segmentar linha/fundo (pode ser por limiarização)
- Calcular o centro da linha na imagem, pois ela pode aparecer grossa)
- Calcular o deslocamento (disparidade) causado pela variação da profundidade (em pixels)
- Calcular a profundidade

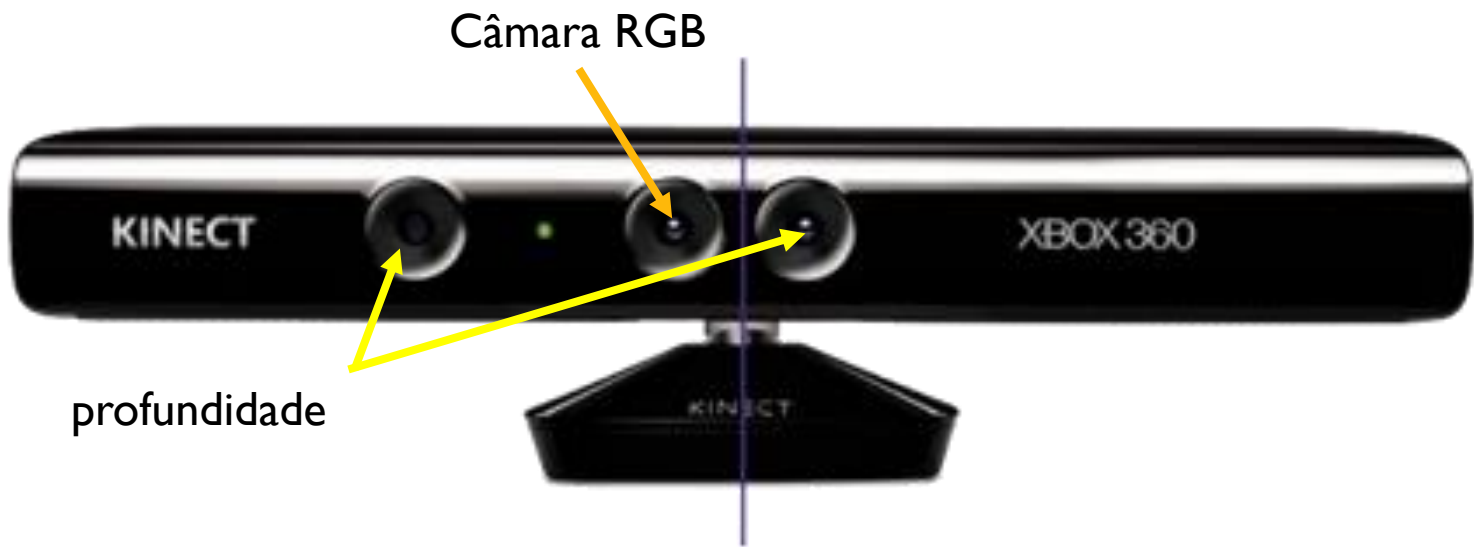
Projeção de linha



O sistema deve ser previamente calibrado para considerar as deformações causadas pelo sistema de lentes e eventuais erros na medição exata da distância base entre o projetor e a câmera.

Kinect

- Originalmente desenvolvido para jogos, pode ser usado para medir profundidades.
- Este sistema captura imagens RGB (coloridas) e usa um sistema projetor-câmera (em infravermelho) para estimar a profundidade ao longo do campo de visão.

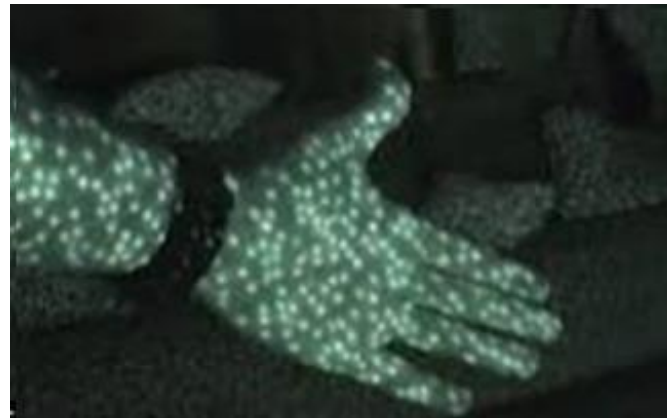
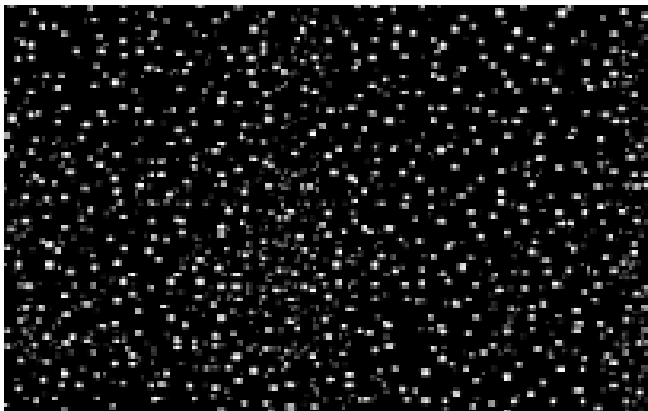


Profundidade

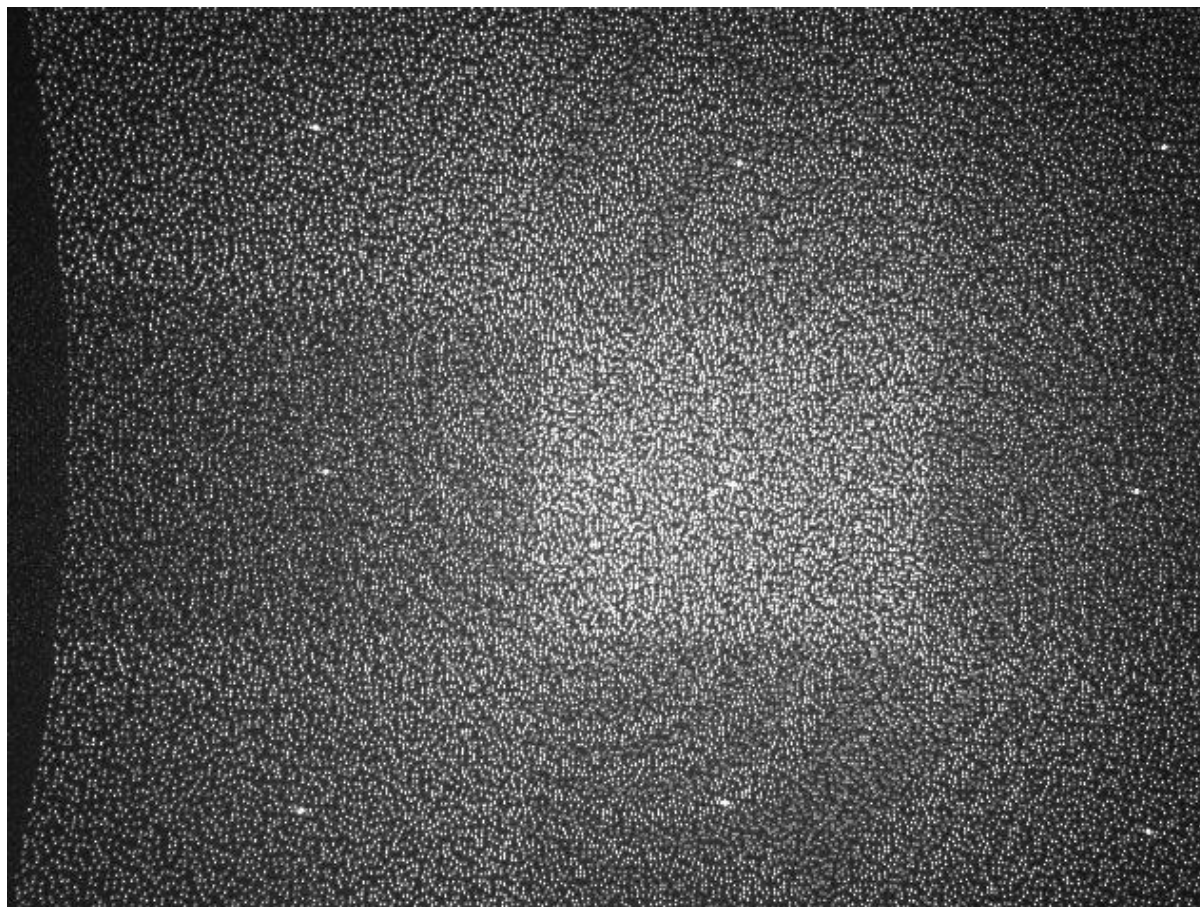
A profundidade é estimada criando-se um mapa de profundidades (luz estruturada)

O mapa de profundidades é obtido projetando padrões conhecidos (círculos) claros na superfície. Para isto, usa-se radiação no infravermelho.

O projetor projeta um padrão preto e branco (em infravermelho) que é composto basicamente por muitas circunferências brancas em um fundo preto.



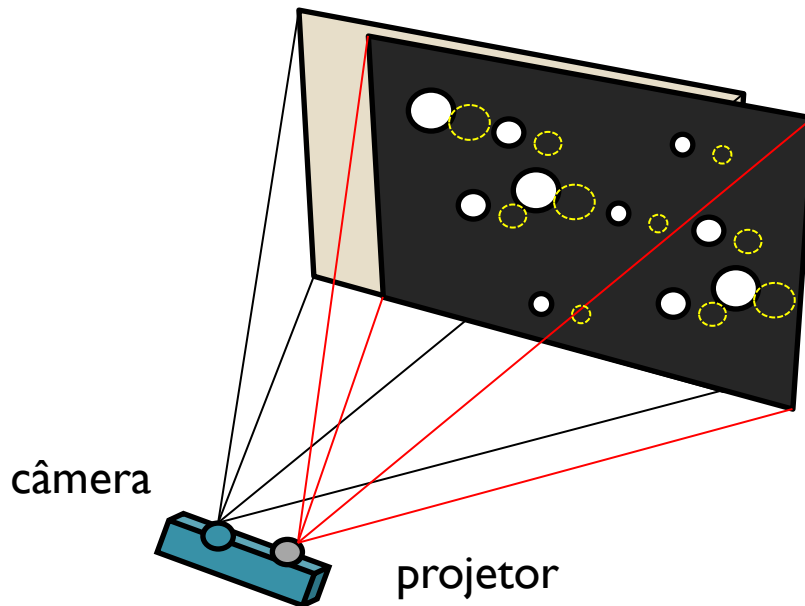
Padrão Kinect



KINECT

A câmera IR e o projetor IR estão alinhados na horizontal, separados aproximadamente 7,5 cm.

Como a imagem projetada é conhecida, se conhece a posição exata de cada circunferência, para uma determinada profundidade.



REFERÊNCIA

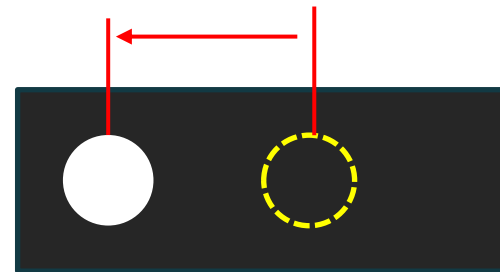
O padrão é projetado sobre um plano a uma distância conhecida para se obter o deslocamento (disparidade) de referência

KINECT

Ao projetar o padrão sobre uma nova superfície, as circunferências sofrem pequenos deslocamentos. Por isso, é feita uma busca (por correlação) nas proximidades da posição esperada (usando uma janela de busca de 9x9 pixels) variando a posição ao longo da linha (uns 64 pixels).

O local onde a correlação é máxima permite determinar o deslocamento causado pela variação da profundidade (em pixels) : isso é chamado de disparidade. Kinect executa ainda uma interpolação adicional para obter precisão de subpixel (em até 1/8 de pixel).

Conhecida a disparidade e a profundidade de referência, a profundidade pode ser estimada.



Relação disparidade / profundidade

Como a câmera e o projetor estão alinhados, pode-se considerar o caso de duas câmeras (estéreo) paralelas alinhadas na horizontal. Com isto a profundidade é dada por

$$z = b * f / d$$

- z é a profundidade (em metros),
- b é a linha de base horizontal entre as câmeras (em metros),
- f é a distância focal das câmeras (em pixels) e
- d é a disparidade (em pixels).

Se a disparidade for zero, considera-se que os raios de cada câmera são paralelos, e a profundidade é infinita. Valores maiores para a disparidade significam distâncias menores.



Disparidade Kinect

Kinect mede uma disparidade bruta, não é normalizada. Por isso, quando sua disparidade é zero isto não corresponde à distância infinita.

Assim, a disparidade deve ser normalizada

$$d = 1/8 * (d_{off} - d_k),$$

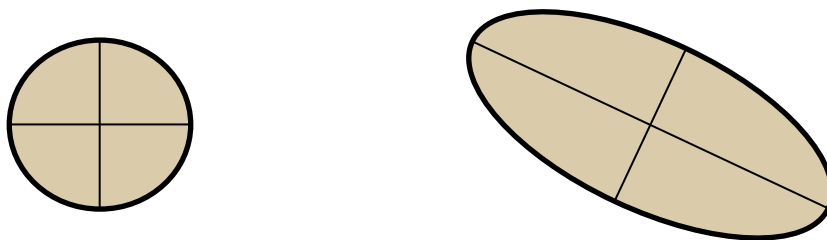
- d é a disparidade normalizada,
- d_k é a disparidade medida pelo Kinect e
- d_{off} é um valor de deslocamento específico, parâmetro de cada equipamento Kinect.
- O fator $1/8$ aparece porque os valores de d_k estão em unidades de $1/8$ pixel, após o refinamento da disparidade

A lente com aberração (astigmatismo) faz com que o padrão que é projetado no objeto seja deformado com uma direção de alongamento que varia com a distância a objeto.

Pode-se estimar a distância sensor-objeto com base na direção da deformação (alongamento da elipse)

Podem ser projetadas várias circunferências, o que permite cobrir todo o campo de visão

Pra calcular a distância se compara o tamanho do eixo maior e o eixo menor da elipse.



Consulte...

- Classificação automática de padrões projetados por um sistema de luz estruturada. Kokubum, C.N.C.; Antonio Maria Garcia Tommaselli, A.M.G. ; Reiss, M.L.L. Bol. Ciênc. Geod., sec. Artigos, Curitiba, v. 11, no 1, p.89-116.
- How Does The Kinect Work? John Maccormick .
<https://pages.cs.wisc.edu/>
- Freedman Et Al. United States. Patent Application Publication
- Konolige, K., Mihelich, P., 2010. Technical description of Kinect calibration