

# Sistemas de Visão e Percepção Industrial

## 8-Outras Formas de Percepção - Parte 3

Outros Sensores 3D

- 1 Visão Estéreo
- 2 Princípios de medição - TOF
- 3 Princípios de medição – Luz estruturada
- 4 KINECT

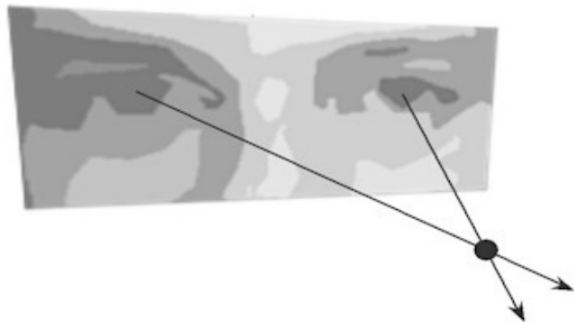
# Imagem 3D e técnicas de obtenção

- Há múltiplas formas de representar 3D. Entre as principais estão:
  - Núvens de pontos (coordenadas  $x,y,z$ )
  - Imagens de distância (range images) – distâncias a um ponto (câmara)
  - Mapas de profundidade (depth maps) – distâncias a um plano (da imagem ou focal)
  - Imagens RGB-D
  - Mapas de disparidade
- Há uma grande diversidade de técnicas e princípios para obter imagens 3D:
  - **Visão Estéreo**,
  - Varrimento Laser,
  - Triangulação Laser,
  - **Triangulação com luz estruturada**,
  - Fotogrametria,
  - **Time-Of-Flight (Tempo de Voo)**,
  - Interferometria,
  - Holografia,
  - Moire fringe range contour,
  - Shape from texture (forma a partir da textura),
  - Shape from shadows (forma a partir da sombra),
  - Shape from focus (forma a partir da focalização),

# Visão Estéreo

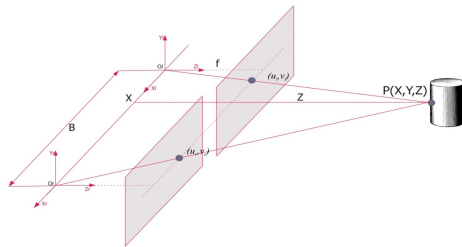
# Visão Estéreo – generalidades

- Inspirado no princípio biológico da visão humana,
- Visão binocular, cria um mapa de disparidade entre duas imagens,
- Como criar a noção de profundidade com apenas um ponto de vista?
  - Sombras,
  - Textura dos objetos,
  - Movimento do olhar



# Visão estéreo com duas câmaras

- Necessita de duas câmaras calibradas,
- Elevados requisitos de processamento,
- Influência da iluminação,
- Influência da textura da superfície,



- A profundidade ( $Z$ ) depende da distância focal ( $f$ ) e da disparidade ( $D$ ) entre os pixels de cada imagem

- Por congruência de triângulos:

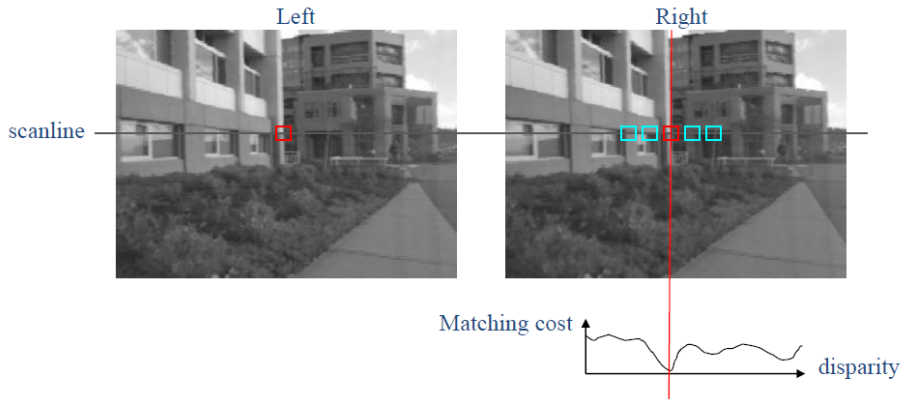
- $$\frac{B}{Z} = \frac{B + x_R - x_L}{Z - f}$$

- $$(Z - f)B = ZB + Z(x_R - x_L) \Leftrightarrow -fB = Z(x_R - x_L)$$

- $$Z = \frac{fB}{x_R - x_L} = \frac{fB}{D}$$

# Visão estéreo com duas câmaras

- Correspondência entre pontos nas imagens das duas câmaras. Como fazer?
  - Projeção de luz estruturada (exige fonte de luz adicional)
  - Utilização de arestas em cada linha,
  - Correspondência entre pontos/regiões (e.g. por correlação normalizada)



# Imagem de disparidade (estéreo)

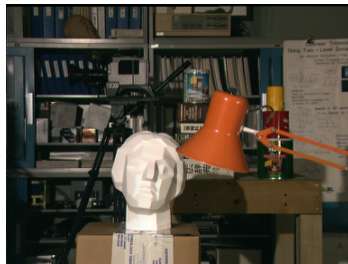


Imagem Esquerda



Imagem direita

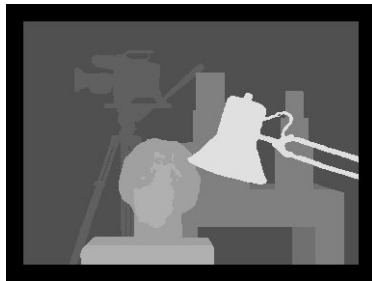


Imagem de  
disparidade

Cada pixel tem como intensidade o valor da diferença entre as coordenadas  $x$  dos pixels correspondentes nas duas câmaras



# Exemplos de câmaras estéreo

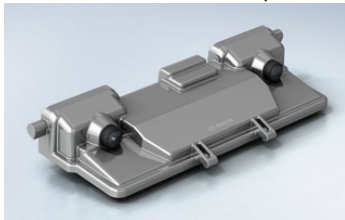
Elphel NC353LSTEREO



Pointgrey Xb3



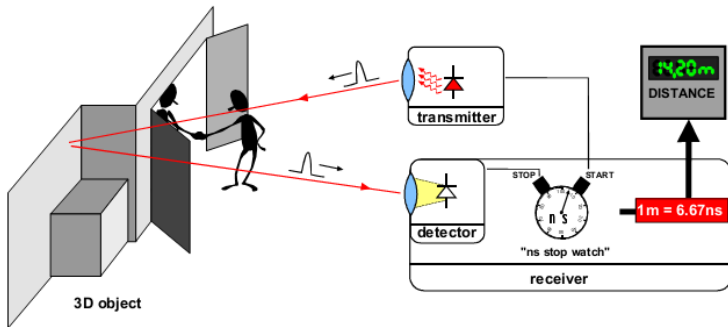
Bosch stereo camera (ADAS)



# Princípios de medição - TOF

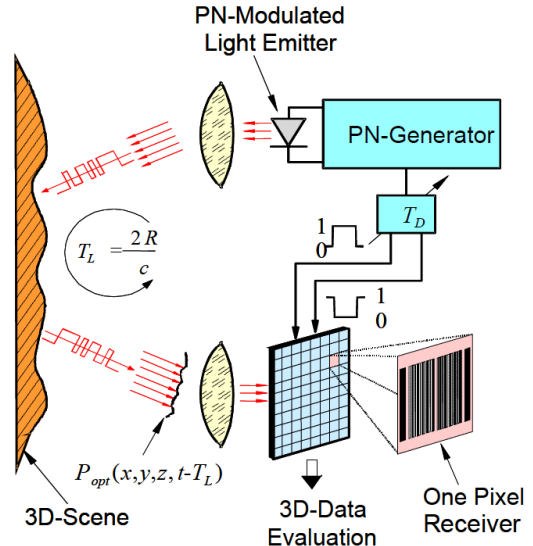
# Princípios de Medição ToF – Pulsed Modulation

- Determina a profundidade dos objetos pela medição do tempo ou fase entre o envio e a recepção de um sinal de luz,
- Necessita apenas de uma câmara;
- Aquisição 3D em tempo real;
- Redução da dependência da iluminação do ambiente;
- Pouca dependência da textura dos objetos.



# Tecnologia TOF (Time-Of-Flight)

- Baseia-se na medição do tempo entre a emissão e a recepção de um feixe de luz modulada;
- O tempo de voo é proporcional à profundidade da cena;
- A aquisição da profundidade da imagem é realizada em simultâneo para todos os pixels;
- A qualidade da medição de profundidade é influenciada negativamente pela inclinação dos objetos e pela proximidade da câmara, tornando os objetos distorcidos ou ambíguos.

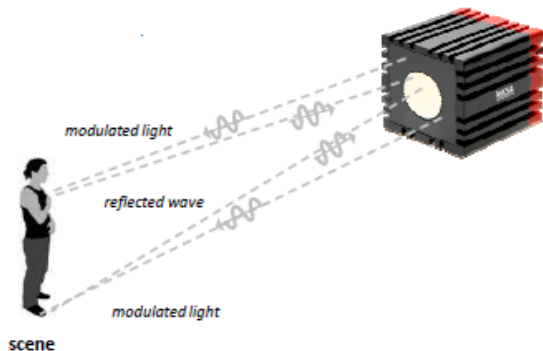


# Tecnologia TOF (Time-Of-Flight)

- Devido à utilização de luz modulada, não deve existir mais do que uma câmara no mesmo local.
- Nesse caso, utilizar uma frequência modulante diferente;

- Aplicações:

- Medição de volumes,
- Robótica,
- Seguimento de objetos,
- Controlo de acessos,
- Aplicações agrícolas;



# Tecnologia TOF (Time-Of-Flight)

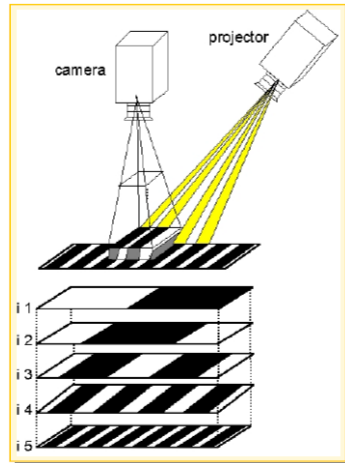
- Exemplo de sensor industrial – Swiss Range, Mesa.

Pixel Array Size	176(h)×144(v)
Field fo view [degree]	43,6(h)×34,6(v)
Pixel pitch[μm]	40
Illumination wavelength [nm]	850
Working range [m]	0.3 – 5.0
Maximum frame rate [fps]	54
Dimensions [mm]	65×65×68
Weight [g]	470



# Princípios de medição – Luz estruturada

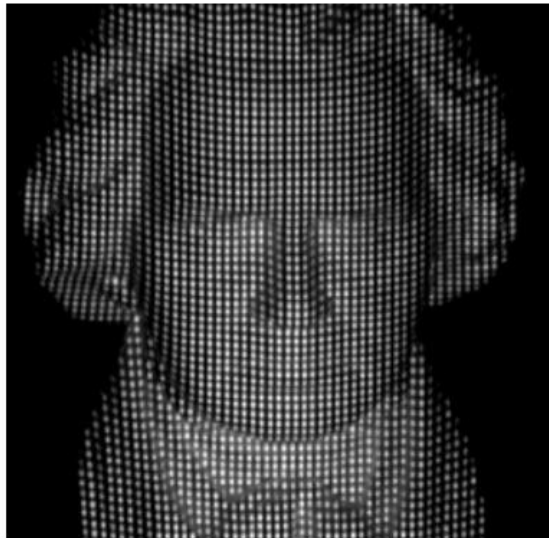
- Projeções de padrões de luz sobre os objetos:
  - Grelhas
  - Listras,
  - Padrões elípticos e aleatórios
- As formas das superfícies são deduzidas das distorções dos padrões produzidas nas superfícies,
- O conhecimento da geometria entre o projetor e a câmera, permite calcular a profundidade através de triangulação,
- O desafio da triangulação ótica é a obtenção de correspondência entre os pontos projetados e a imagem adquirida.





# Projeção de padrões

- Projeção de um padrão em vez de um ponto,
- Necessita apenas de uma imagem, para reconstruir o objeto,
- A correspondência entre as arestas observadas e a imagem projetada não é diretamente mensurável,
- A correspondência entre pontos não é única.

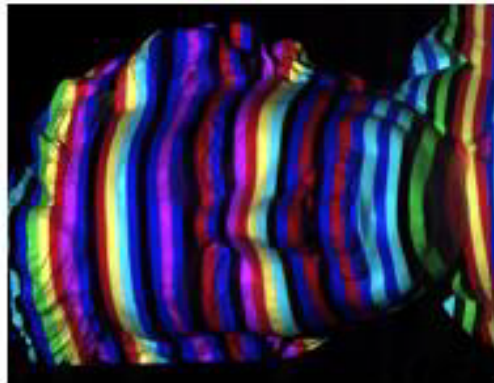
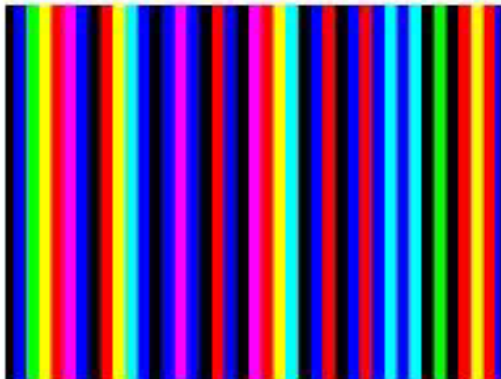


- A projecção de padrões de pontos ou de listras não permite uma comparação única,
  - A que ponto corresponde?
- Como solução pode-se utilizar padrões de luz especiais, para simplificar a comparação e de forma a cobrir toda a imagem,
- A luz estruturada é um método robusto e eficaz, utilizado ao nível industrial e noutras situações.



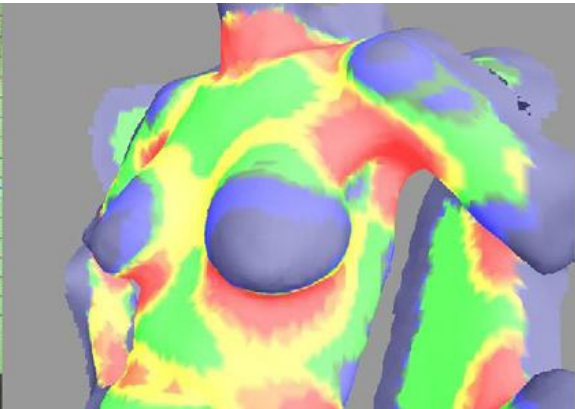
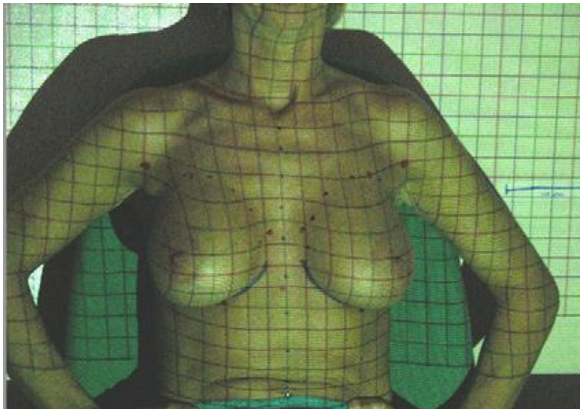
## Listras coloridas (Coloured stripes)

- Projeção de listras únicas,
- Não aconselhável utilizar em superfícies coloridas,



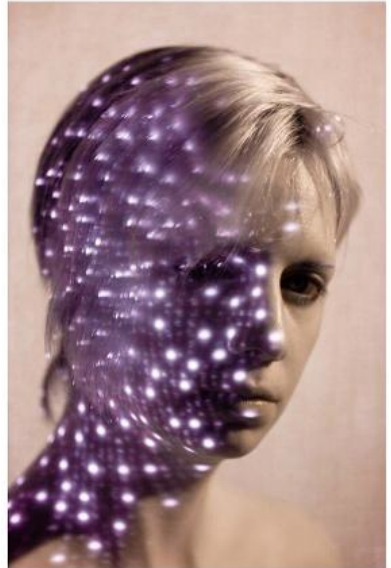
## Coloured stripes

- Análise da forma da mama para avaliação da evolução dos resultados cirúrgicos do tumor da mama (2005),
- INT: Istituto Nazionale per lo studio e la cura dei Tumori Milano



# Pseudo random pattern

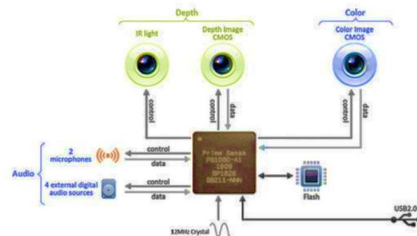
- Projeção de uma codificação de padrões única, e perfeitamente identificável pela restante vizinhança,
- A decodificação nem sempre é possível, devido à oclusão e a sombras,
- Método utilizado pela Kinect.



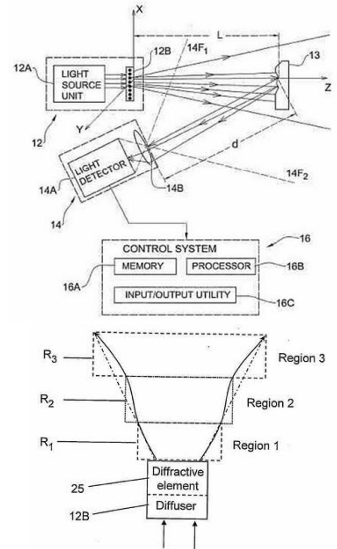
KINECT

# Kinect

- Constituído por:
  - Projetor de Infravermelhos,
  - Câmara RGB,
  - Sensor de Infravermelhos,
  - (microfone e sistema de posicionamento);
- Captação de vídeo 3D, independente das condições de iluminação;
- No software do fabricante (jogos) pode detetar e fazer o seguimento de até 6 pessoas.
- Video frame rate: 30Hz,
- RGB video: 8-bit (640x480 pixels),
- Codificação Bayer,
- Profundidade: utiliza um sensor monocromático (640x480 pixels) com 11-bit,
- Campo de trabalho: 40 cm a 3.5 m

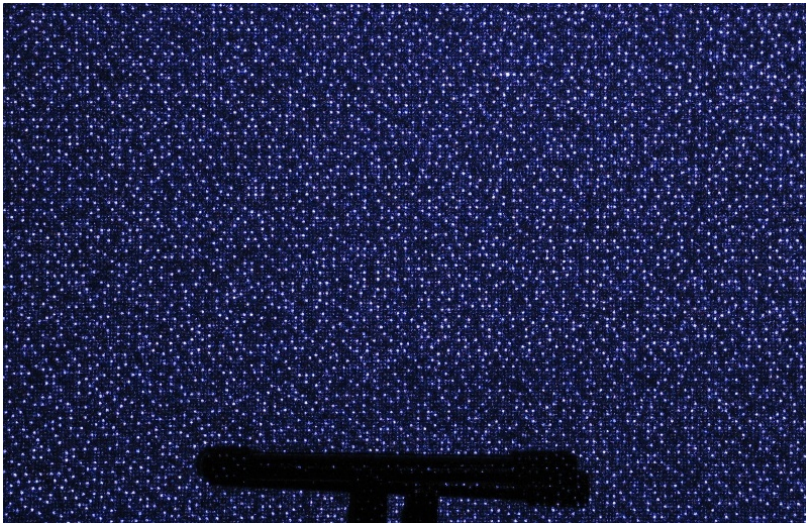


- Projeção de um padrão conhecido em infravermelho
- Aquisição da imagem com um sensor CMOS IR
- Conhecida a geometria entre o projetor de IR e o sensor CMOS
- Triangulação de cada ponto entre a imagem virtual (padrão) e o padrão observado
- Utilização de 3 diferentes padrões em função da distância, para melhoria da precisão de aquisição da imagem

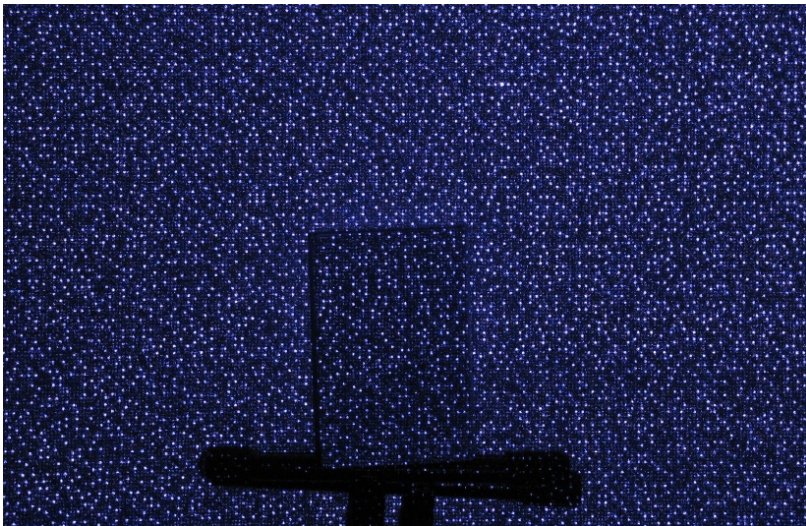




- Projeção sem objecto



- Projeção com objecto



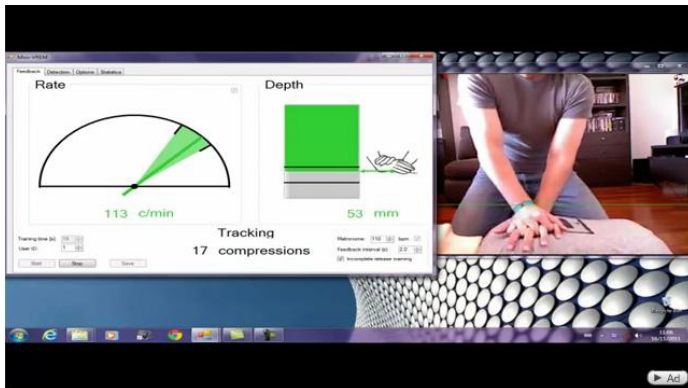
# Kinect – Aplicações e Características

- Reconhecimento dos movimentos do corpo
- Reconhecimento de gestos
- Reconhecimento facial



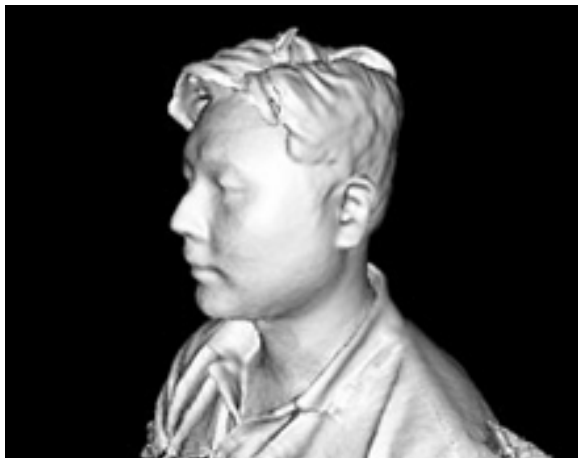
# Kinect – Aplicações e Características

- Treino interativo
- Mini-VREM usou o Kinect para criar um sistema que analisa a posição das mãos, velocidade, pressão e força das compressões para ensiná-los a executar corretamente as massagens



# Kinect – Aplicações e Características

- Geração de modelos tridimensionais em tempo real
- Possibilidade de realizar o scan em 3D de objetos, pessoas, cenários completos



# Kinect – Aplicações e Características

- Revisão de imagens médicas, permitindo acesso a outras imagens sem que haja contaminação

