Automatic 2D-to-3D Video Conversion Techniques for 3DTV

Dr. Lai-Man Po

Email: eelmpo@cityu.edu.hk

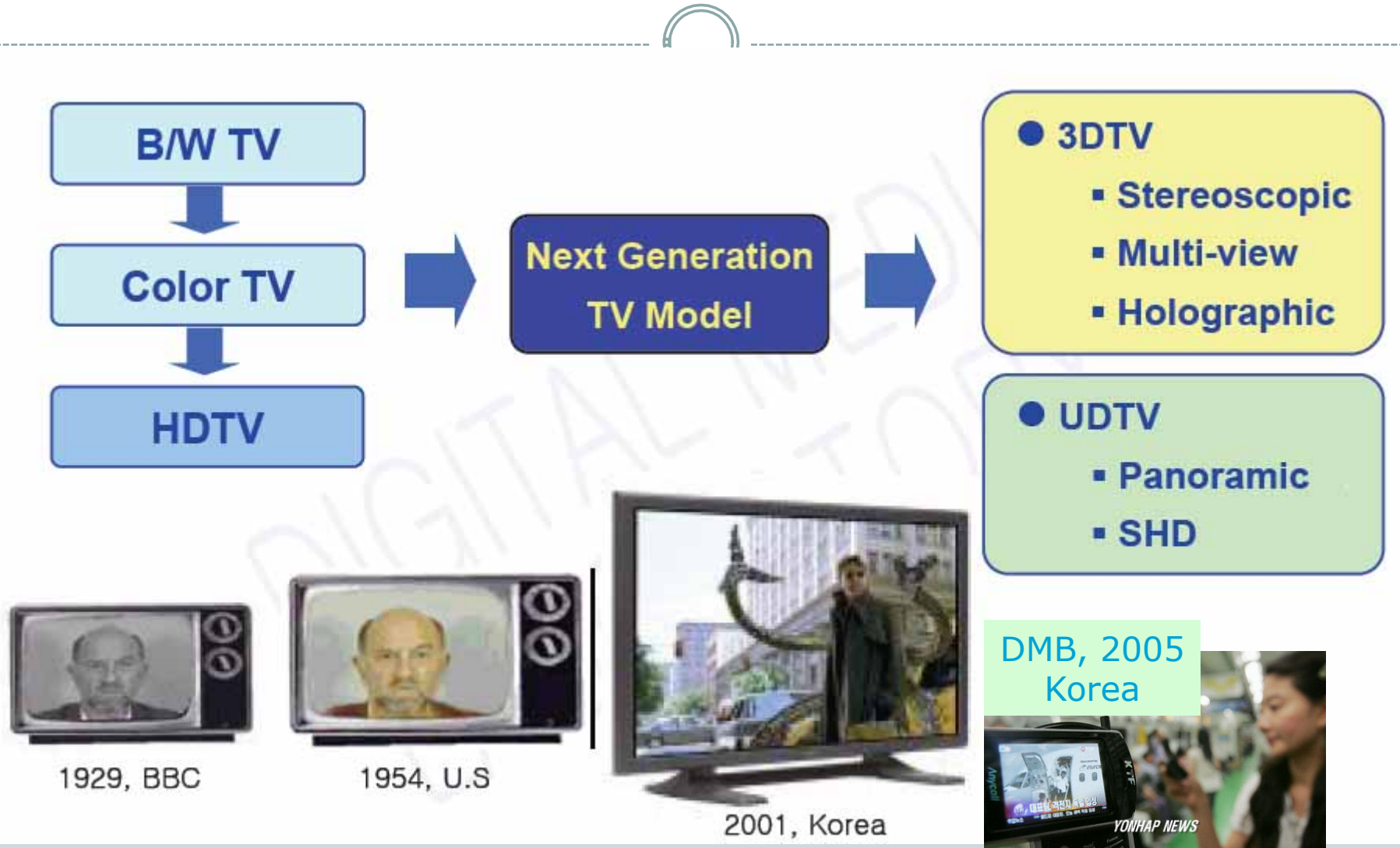
Department of Electronic Engineering
City University of Hong Kong

Date: 13 April 2010

Content

- Why 2D-to-3D Video Conversion?
- Overview of 2D-to-3D Video Conversion System
- Techniques to automatically convert 2D video to 3D:
  - Depth Map Estimation
  - Depth Image Based Rendering (DIBR)
- Demo
- Conclusions

Digital TV Technology Trend Overview

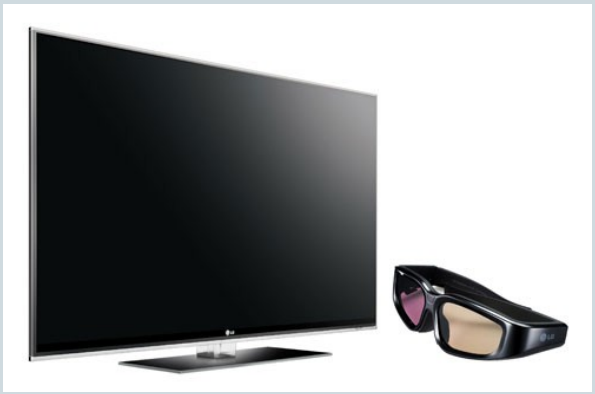


Why 3DTV? – More and more 3D contents will be available

- 3D video is a natural extension to the 2D video but 3D video have never quite moved beyond the gimmick stage – that is until recently.
- In 2005, particularly after Disney's release of "3D Chicken Little" in movie theaters, the industry rediscovered huge business potential of 3D.
- In 2010, "Avatar" raked more than $1.1 billion worldwide. James Cameron is targeting a 3D re-release of "Titanic" in 2012.
  - 3D theatrical releases are now generating more revenue than 2D release.
  - Because tickets to 3D versions of films cost more than their 2D counterpart, people are clearly willing to pay more for 3D.

Why 3DTV? – 3D Ready TV will be available in a affordable price.

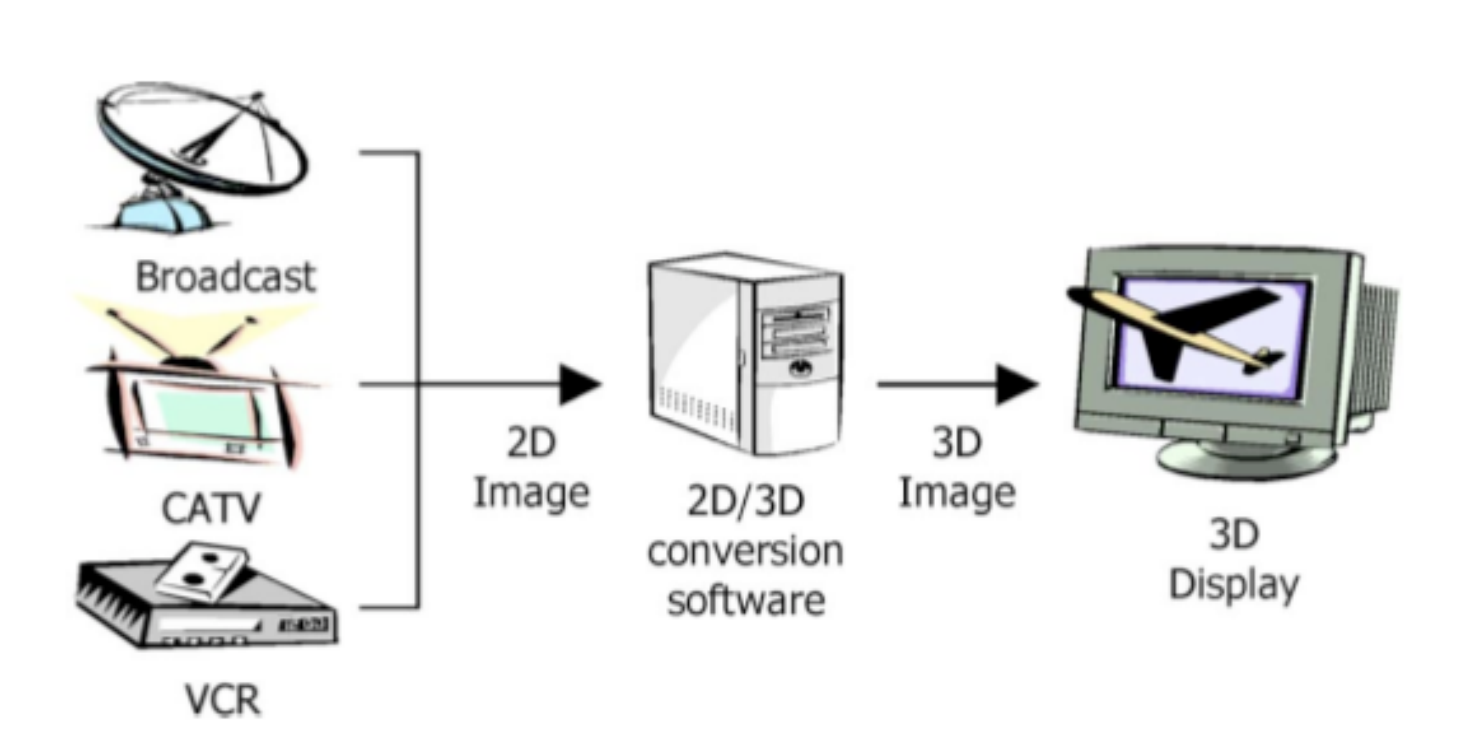
- Many well-known TV brands such as Sony, Panasonic, Samsung, LG will release 3D ready TV in this summer in affordable price.
- Currently, there are three types of 3D display technologies commonly used in 3DTV with glasses:
  - Anaglyph glass
  - Polarization glass
  - Shutter glass



Why we need 3D generation from 2D?

- The 3D stereo contents are not rich enough but there are still large numbers of 2D videos exist in different compressed formats.
- If they are converted to 3D videos, this will offer a more realistic sense of the scene to the viewer.
  - The conversion process can be off-line to generate high-quality 3D version of classic 2D content for redistribution in Blu-ray HDTV format. (Creating revenue from the old content)
  - The conversion can also be implemented in real-time on 3D TV set for stereoscopic displaying of 2D video sources. (Adding 3D feature in the set-top-box, DVD player, or TV set)

Overview of the 2D-to-3D Video Conversion System

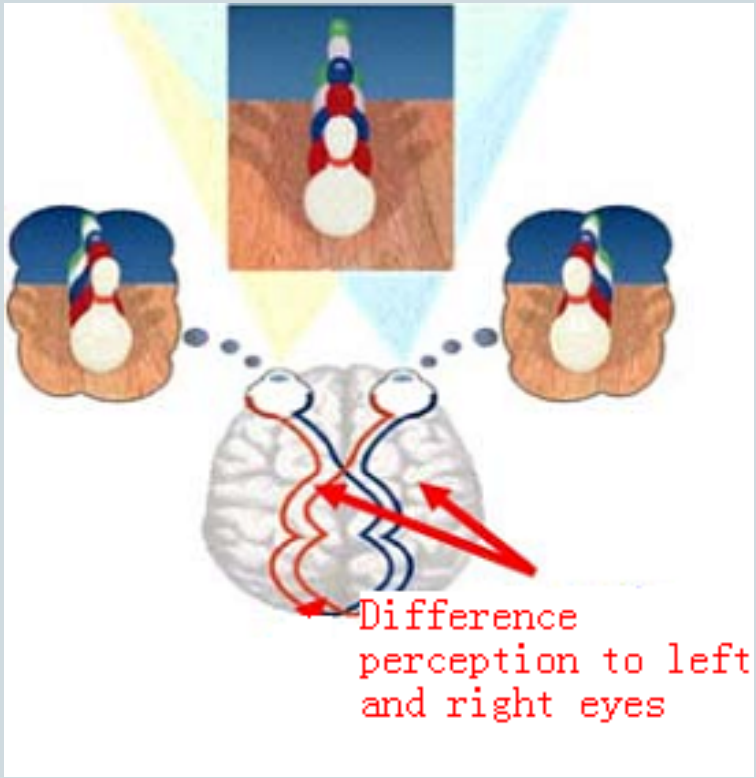


- CyberLink PowerDVD 10 Ultra 3D can convert 2D DVD movies to 3D
- Most of the 3D ready TVs are embedded with real-time automatic 2D-to-3D function for watching 2D video contents with 3D effect.

How to convert 2D Video to 3D Video?

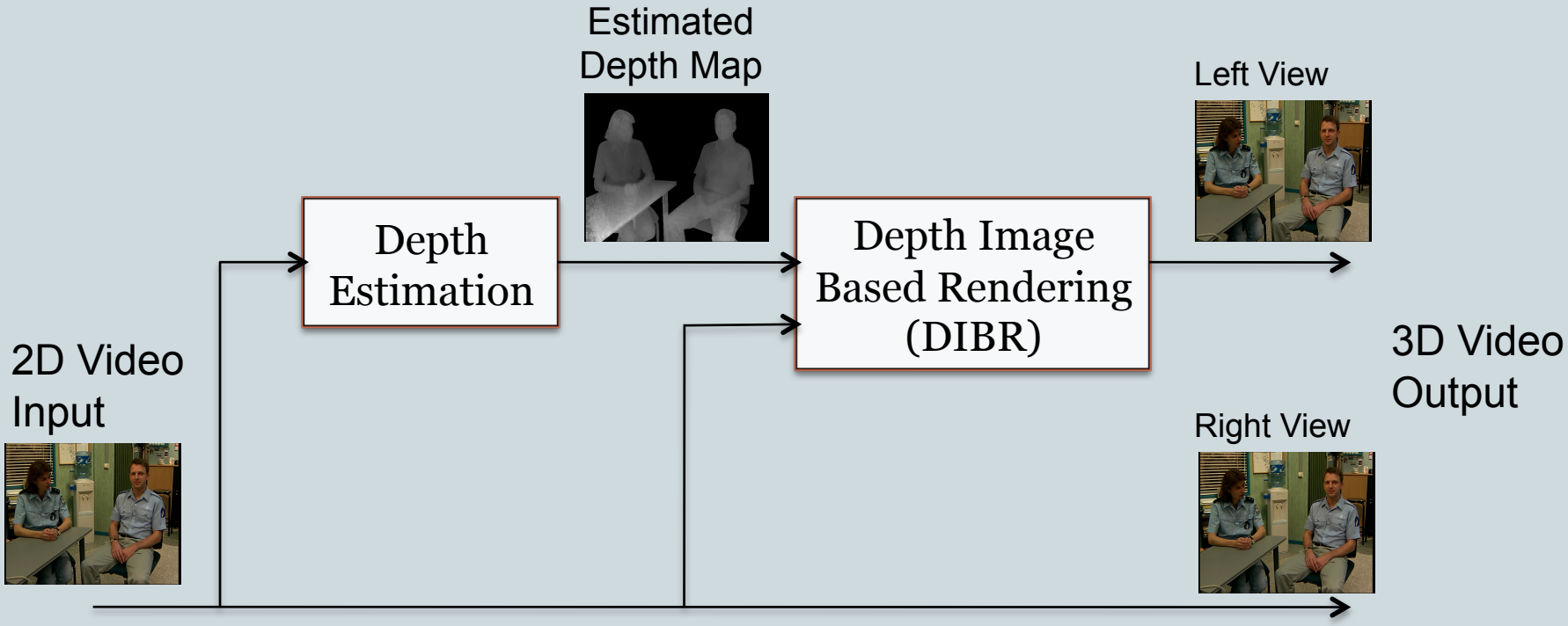
3D Perception of human

- binocular parallax
  - small spatial displacement between the left and right eyes



Block Diagram of 2D-to-3D Video Conversion

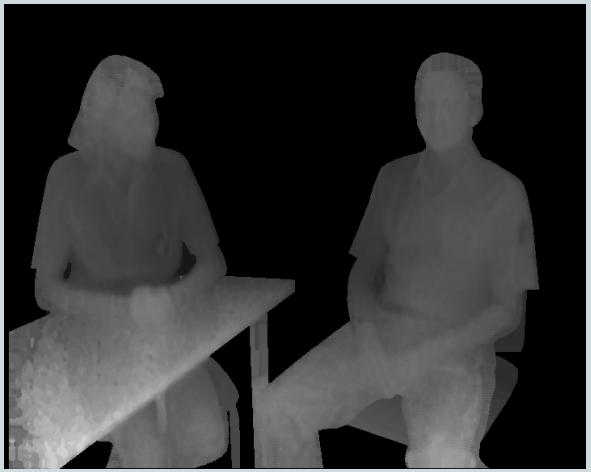
- The main purpose of the 2D-to-3D video conversion is to generate the second view video based on the content of the 2D video, which involve two processes: (1) Depth Estimation and (2) Depth Image Based Rendering (DIBR)



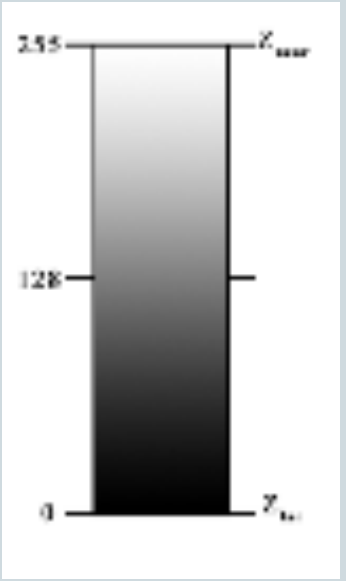
What is Depth Map or Depth Image?

- Each depth image stores depth information as 8-bit grey values with the grey level 0 indicating the furthest value and the grey level 255 specifying the closest value.

2D Image



Depth Image (Map)

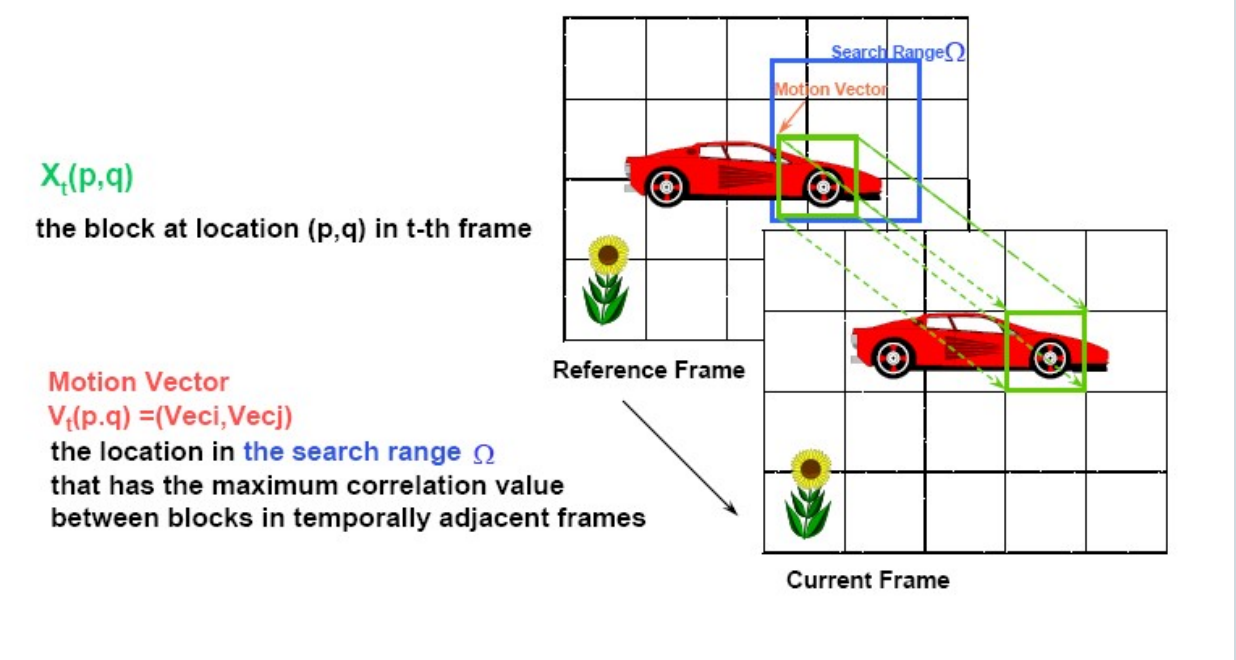


Depth Map Estimation Methods

- There are three commonly used depth estimation methods from 2D-to-3D conversion applications:
- **Depth from blur** : The basic idea is to estimate the depth information based on the amount of blur of the object.
- **Vanishing Point based Depth Estimation** : The main idea is to find out the vanishing point that is the farthest point of the whole image.
- **Depth from Motion Parallax** : This is based on the fact that objects with different motions usually have different depths. For example, near objects move faster than far objects, and so relative motion can be used to estimate the depth map. This method is widely used for the depth estimation in 2D-to-3D video conversion.

Block-Matching based Depth from Motion

- The motion information can be easily obtained by block-matching algorithm between two consecutive frames.



- The relative depth information is calculated by

$$D(i,j) = c\sqrt{MV(i,j)_x^2 + MV(i,j)_y^2}$$

Why block-matching based depth from motion?

- The block-matching motion estimation is very easy to implement in hardware and very suitable for real-time applications.
- In addition, the inputs of the block matching based depth estimation are motion vectors which are easily extracted from the compressed video bit stream such as MPEG-2 (DVD), and H.264/AVC (Digital TV broadcasting).

2D Video Sequence

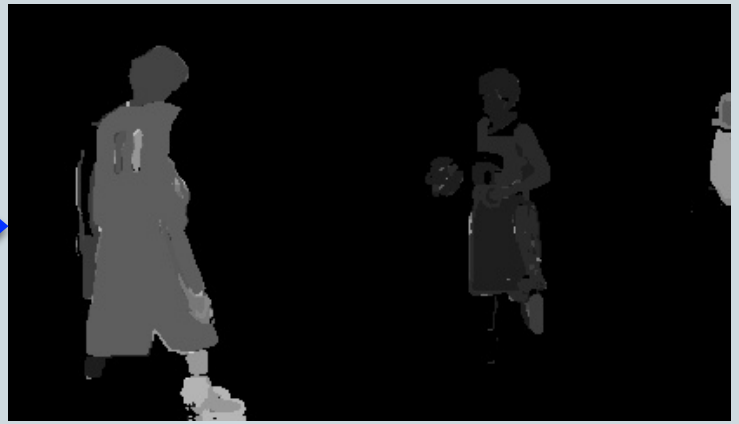
Estimated Depth Map by Block-Matching Motion Estimation

Depth Map Enhancement by Color Segmentation

Depth Map Estimated by Block-Matching Motion Estimation

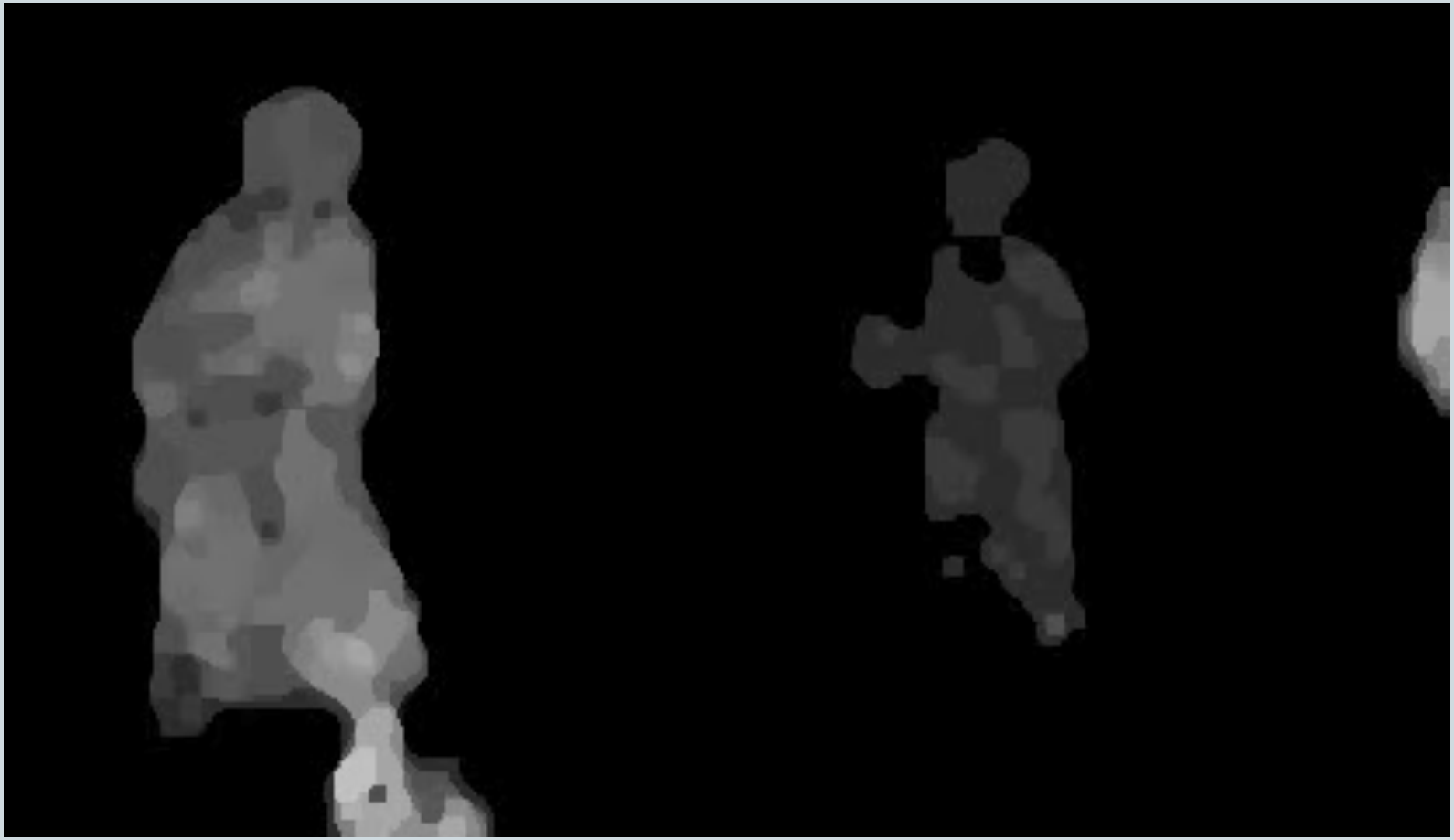
Color Segmented Frame

Depth Fusion



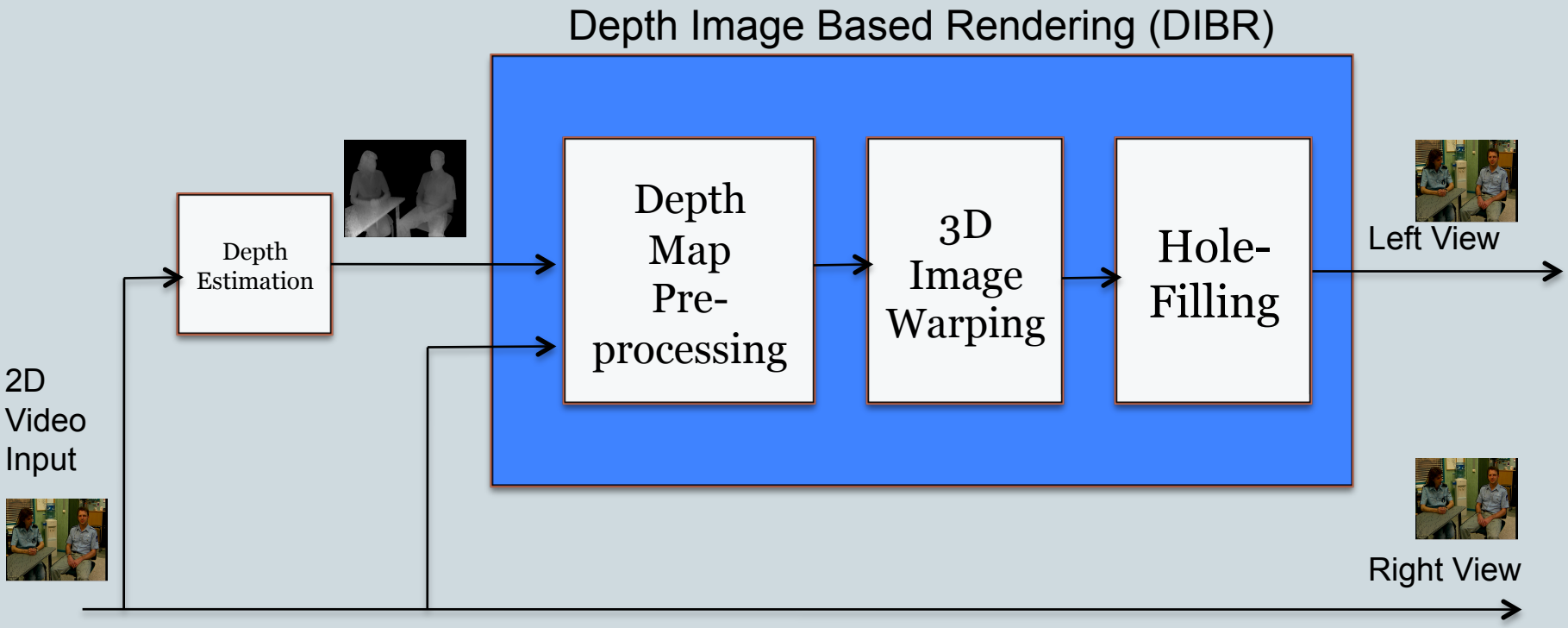
Enhanced Depth Map

Enhanced Depth Map by Color Segmentation



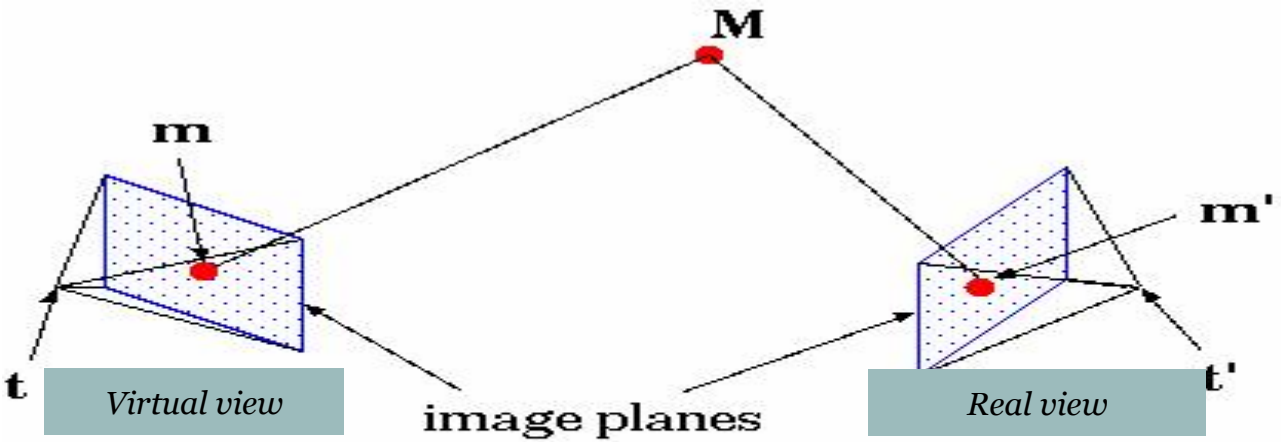
Depth Image Based Rendering (DIBR)

- To generate the 3D video, DIBR is used to syntheses the second view video based on the estimated depth map and the 2D video input.
- DIBR consists of three processes.



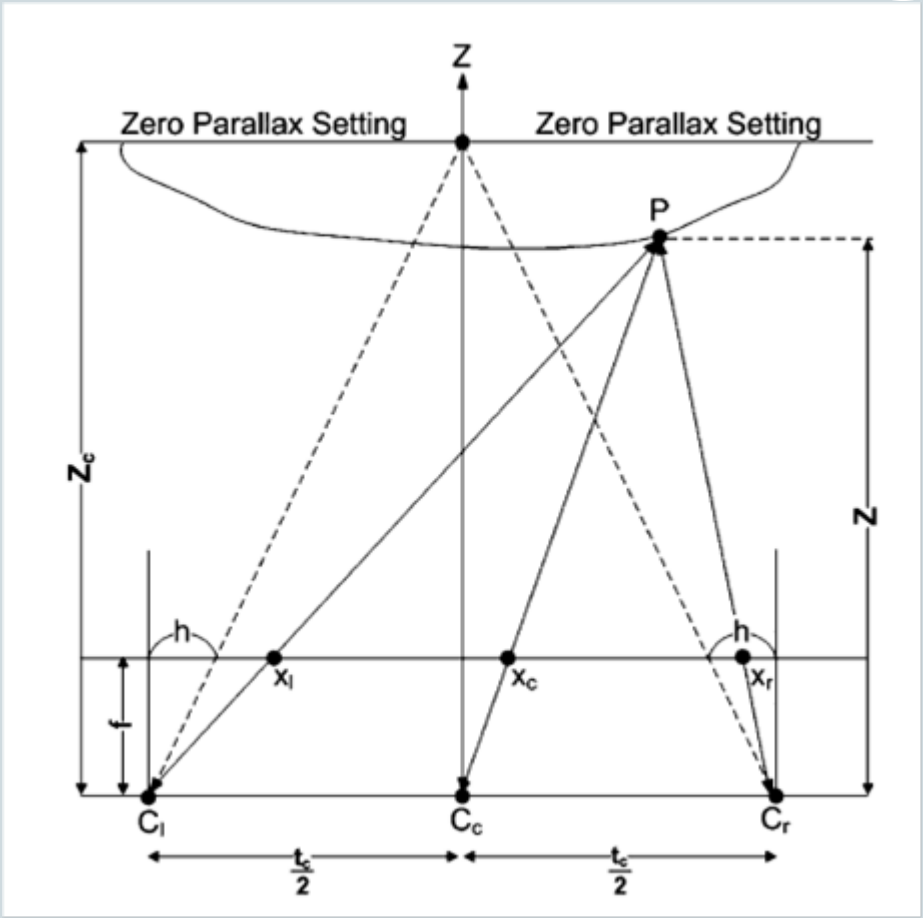
3D Image Warping

- The process includes two steps:
- Original image points (e.g. m'(x',y')) from the real view are re-projected into the 3D world
- The 3D space points (e.g. M(X,Y,Z)) are projected into the image plane of the "virtual" view (e.g. m(x,y)).



M: point on 3D world coordinate
m, m': projections of **M** on the image planes
t, t': centers of cameras

3D Image Warping



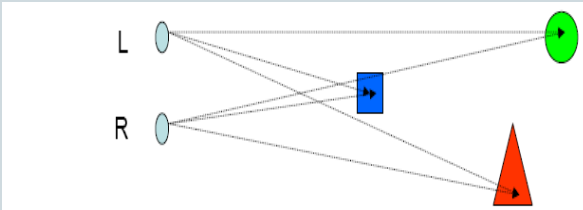
$$x_l = x_c + \frac{t_c f}{2Z(x_c, y)} + h$$

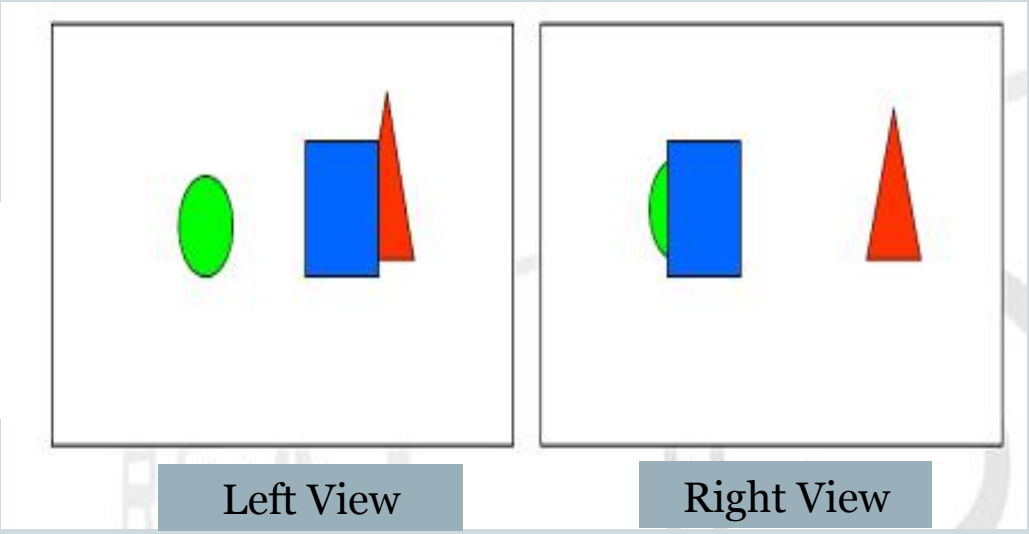
$$x_r = x_c - \frac{t_c f}{2Z(x_c, y)} - h$$

$$h = -\frac{t_c f}{2Z_c}$$

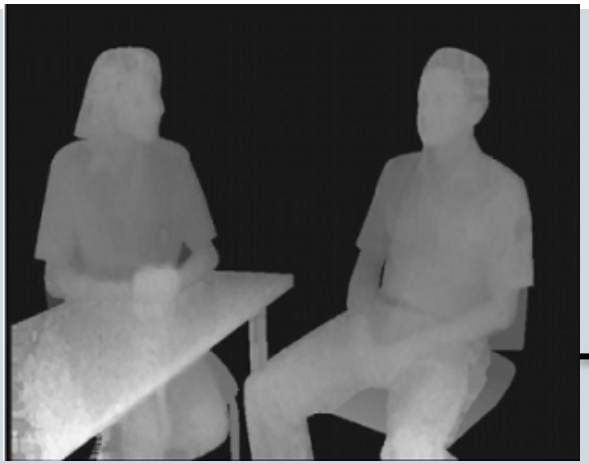
Major Challenges in DIBR

- **Occlusion:** Two different points in the image plane at the real view can be warped to the same location in the virtual view. To resolve this, the point with position appear closer to the camera in the virtual view will be used.
- **Disocclusion:** Occluded area in the real view may become visible in the virtual view. Disocclusion can be resolved by (1) **Hole-filling** and (2) **Depth Map Pre-processing**





Holes Created in 3D Image Warping



Depth Map

Right View Image

3D Image Warping

Left View Image created by 3D Image Warping with holes due to disoculsion.

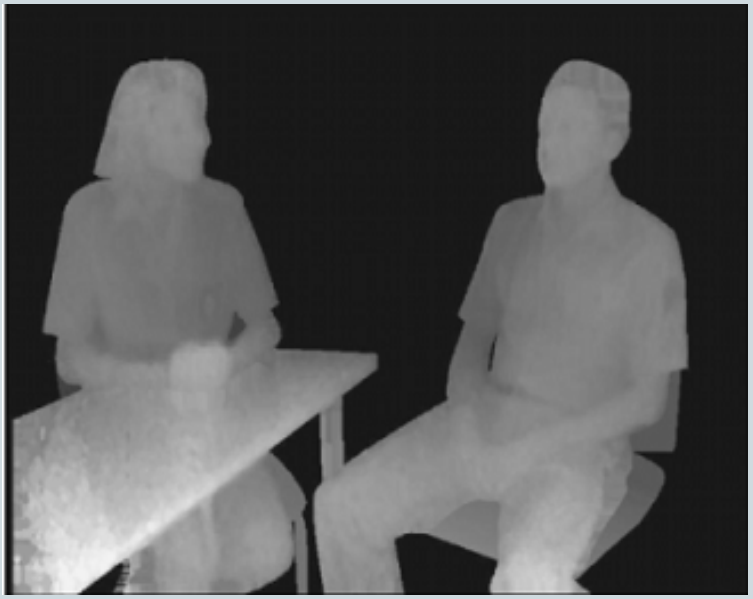
Hole-Filling

- Detect holes
- Fill holes by averaging textures from neighborhood pixels
- Linear interpolation technology

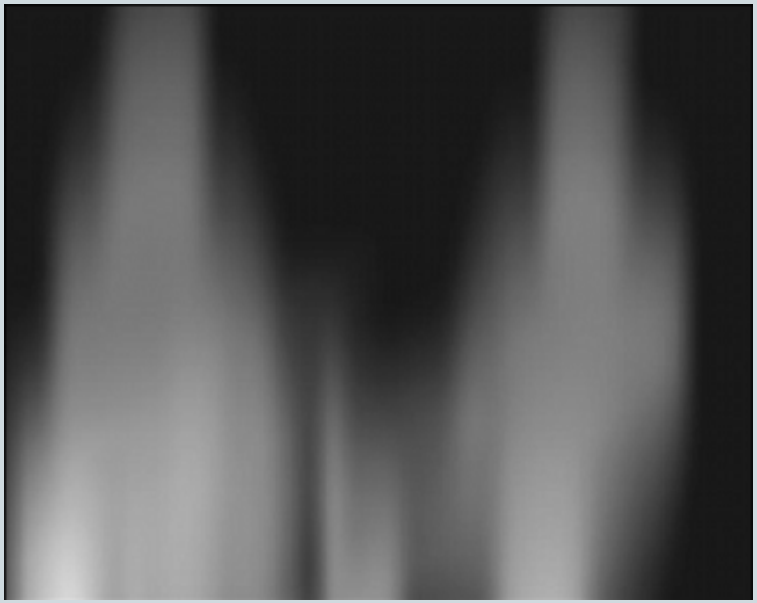
Hole-Filling by Interpolation

Pre-Processing of Depth Map by Smoothing Filter

- Reduce disocclusion (holes) in the virtual views
- Less significant texture artifacts

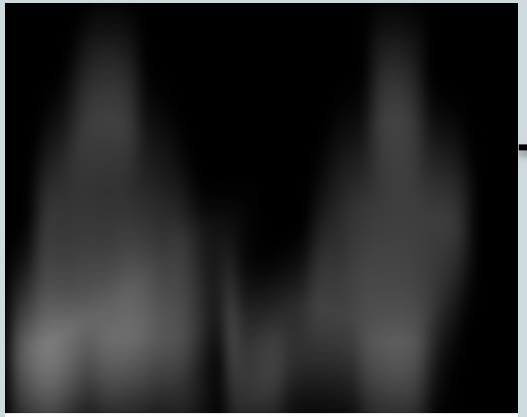


Original Depth Map



Depth Map after Smoothing Filter

Pre-Processing of Depth Map



Asymmetric Smoothed Depth Map

Right View Image

Left View Image created by 3D Image Warping using the smoothed depth map with much fewer holes.

Left View Image after hole-filling

Enlarged Left View Image after hole-filling

Demo: 2D Video Sample

Demo: Converted 3D Video Sample

Conclusion

- It is possible to convert 2D video to 3D video automatically for some video with good 3D perception using depth from motion estimation and DIBR techniques.
- When you buy your 3D Ready TV, the quality of the 2D-to-3D conversion function should be one of your consideration as different brands use different technologies for this conversion.

Thank You for your attention!!