Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis

SPIE SD&A, San Francisco, January 24, 2011

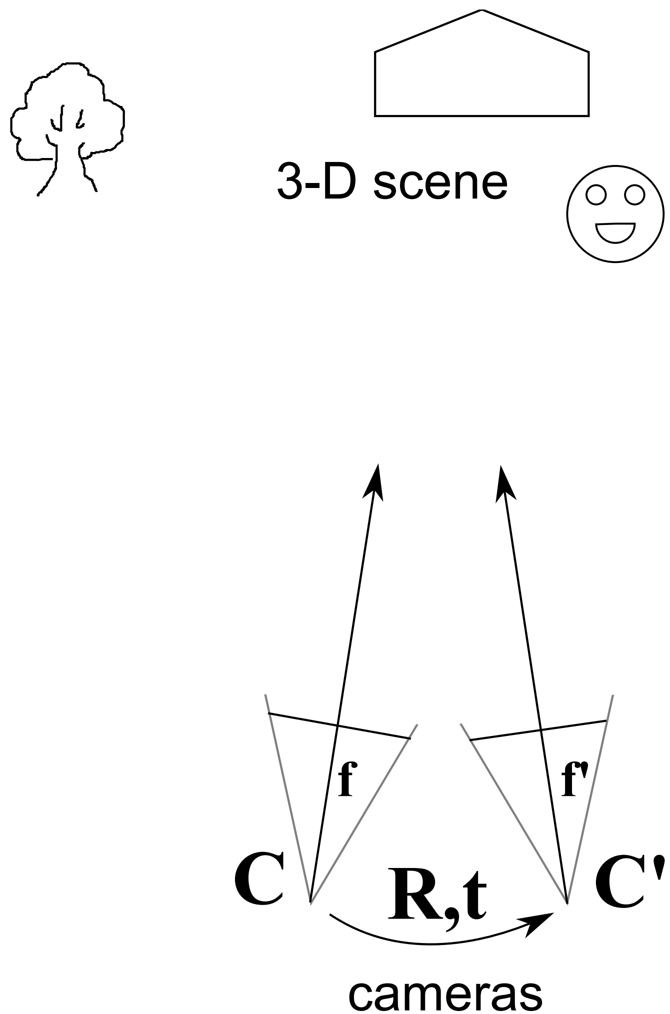
INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
INRIA
centre de recherche GRENOBLE - RHÔNE-ALPES

Frédéric Devernay, Sylvain Duchêne and Adrian Ramos-Peon
INRIA Grenoble - Rhône-Alpes, France

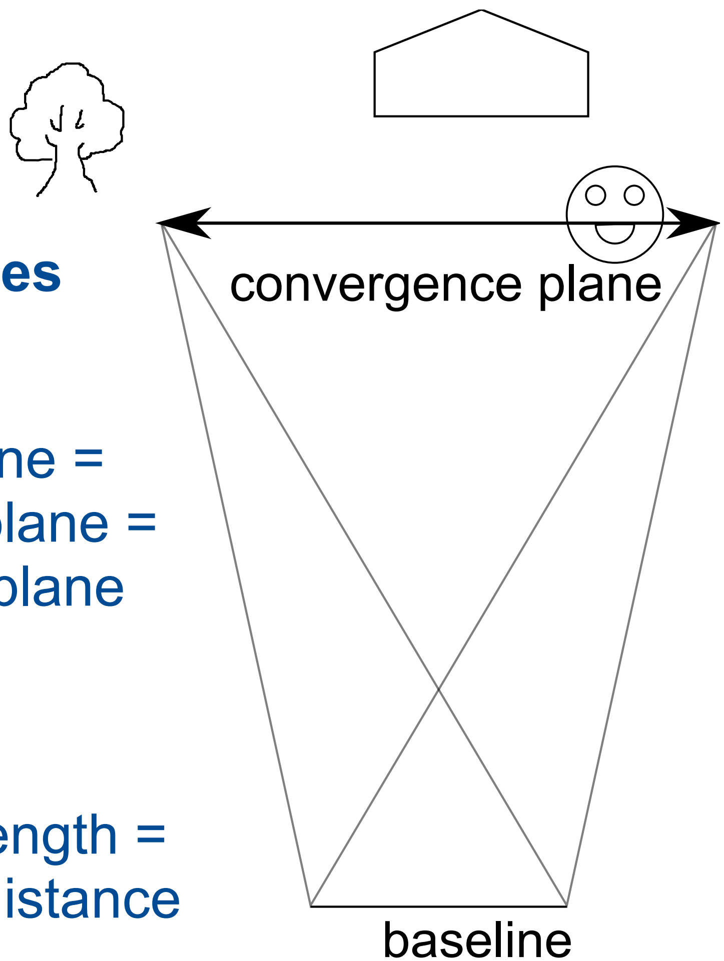
Introduction

- Projecting a stereoscopic movie on **different screen sizes** produces **different perceptions of depth** [Spottiswoode1952]
- Theoretically, a stereoscopic movie **has to be shot for given viewing conditions**, e.g. movie theater or 3DTV
- **Depth distortion**, or even eye divergence, may happen for different viewing geometries: screen size and distance have a strong impact
- **New view synthesis** [Zitnick2004,Rogmans2009] is a possible solution, but what geometric transform should be applied to the images to minimize depth distortions and divergence?
- **Disturbing artifacts** may appear in the synthesized views: we propose to synthesize only one view and to blur out artifacts.

The shooting geometry: classical representation (top view)



The shooting geometry: simplified representation

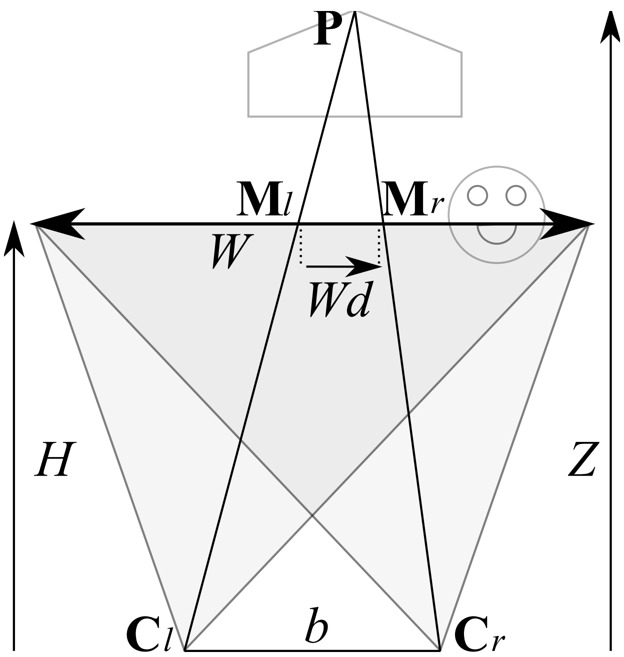


rectified images

rectification plane =
convergence plane =
zero-disparity plane
(// to baseline)

rectified focal length =
convergence distance

Shooting and viewing geometry share the same parameters



Symbol	Camera	Display
b	camera interocular	eye interocular
H	convergence distance	screen distance
W	width of convergence plane	screen size
Z	real depth	perceived depth
d	left-to-right disparity (as a fraction of W)	

Z can be expressed as a function of *d*:

$$Z = \frac{H}{1 - dW/b} \qquad d = \frac{b}{W}\frac{Z-H}{Z}$$

... *same* ***with primes*** *in the* ***viewing geometry***

Viewing the unmodified 3-D movie: depth distortions

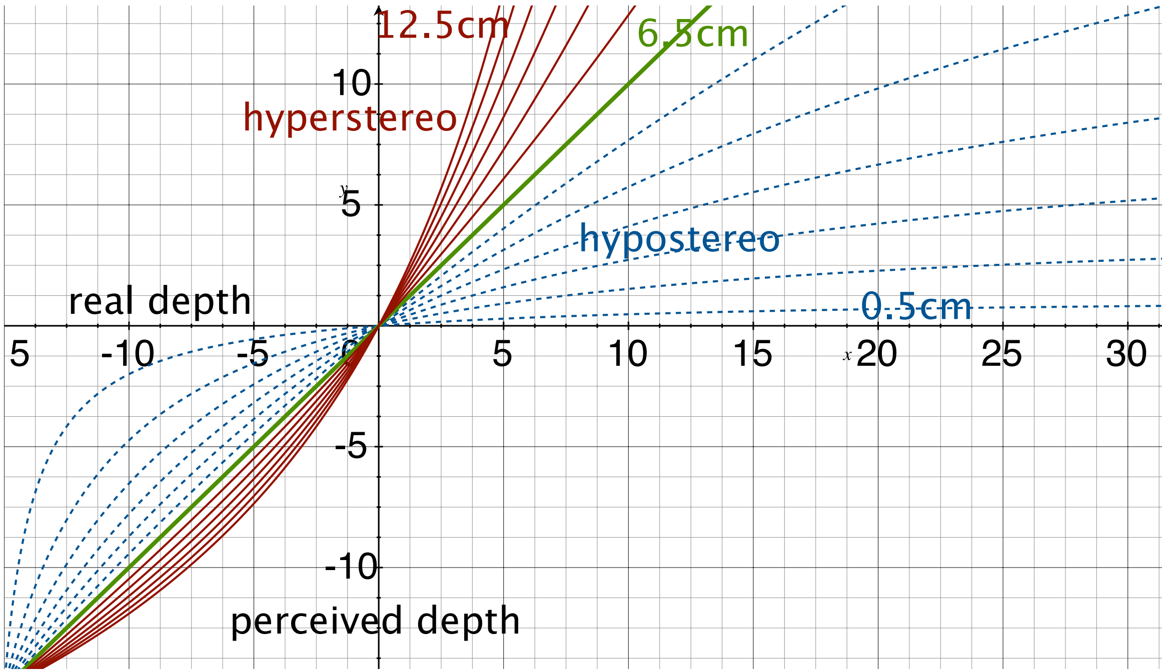
d is the same in both geometries.

substituting *d* gives **perceived depth** *Z'* as a function of real depth.

Highly **nonlinear**!

Eye divergence if:

$Z' < 0$ i.e. $d > b'/W'$



Z' = f(*Z*) for different values of *b* (baseline)

Viewing the unmodified 3-D movie: roundness factor

roundness factor: does a sphere at depth Z appear as flattened ($\rho<1$) or elongated ($\rho>1$)?

on screen: $\rho_{\text{screen}} = \frac{b}{H}\frac{H'}{b'}$ ... *does not depend on screen size*

The ***only*** *shooting geometries that preserve the roundness factor* ***everywhere*** *are* ***scaled (i.e. homothetic) versions of the viewing geometry****!*

Impossible to hold this constraint in practice (sports, wildlife...).

Fixing the roundness factor issue using new view synthesis

Changing the shooting parameters by post-processing the images to fix the roundness factor:

- **Baseline modification** (or view interpolation) corrects on-screen roundness factor, but distorts off-screen depth and image size, and may cause *eye divergence*
- **Viewpoint modification** gives perfect depth and roundness factor, but difficult in practice because of large disoccluded areas
- We propose **hybrid disparity remapping** - fixes the on-screen roundness factor, no depth distortion, not eye divergence

Baseline modification

synthesized baseline b'' computed by setting $\rho_{screen}=1$

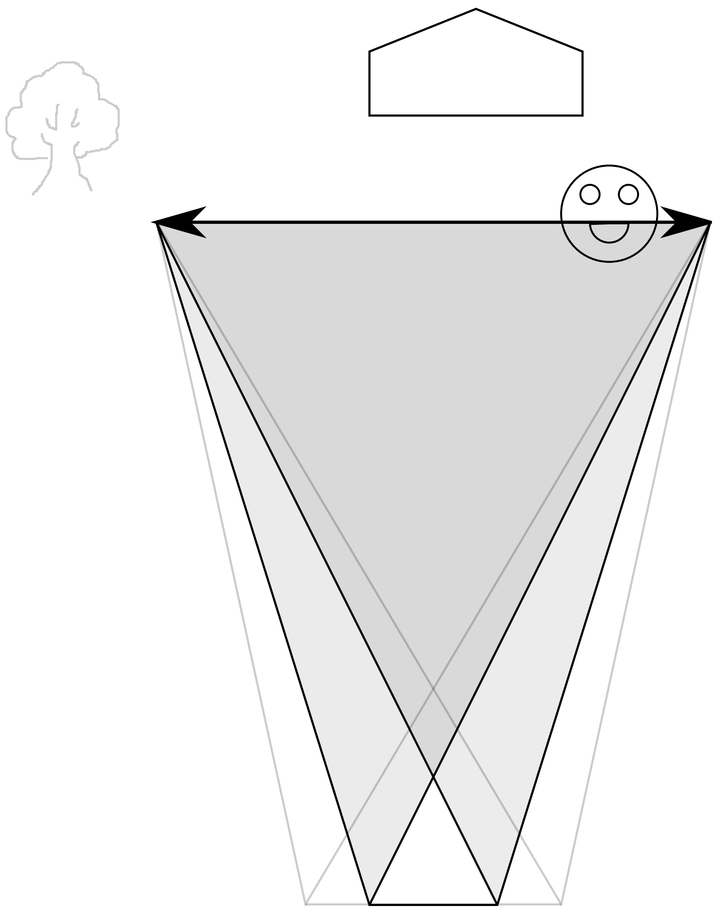
view interpolation ($b'' < b$) or extrapolation ($b'' > b$)

symmetric or asymmetric (one view can be left untouched)

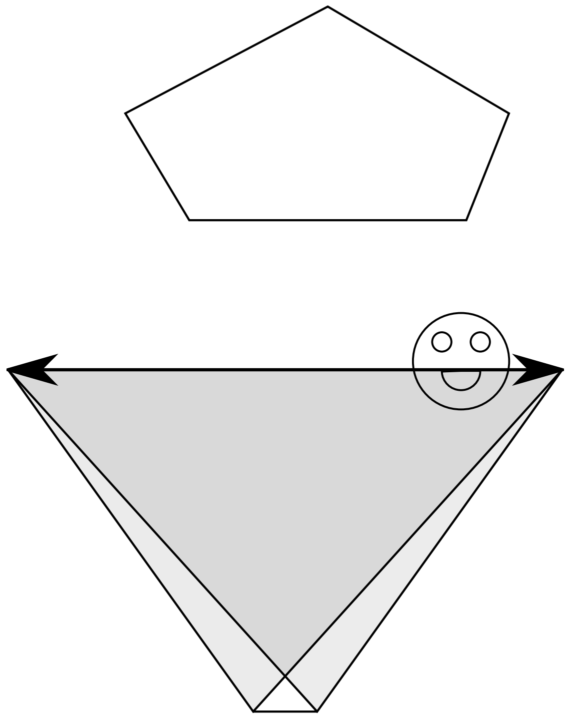
Baseline modification

3-D geometry is distorted

eye divergence may happen



shooting geometry



viewing geometry

Viewpoint modification

Synthesized geometry is homothetic to the viewing geometry.

Both views must be synthesized (**symmetric**)

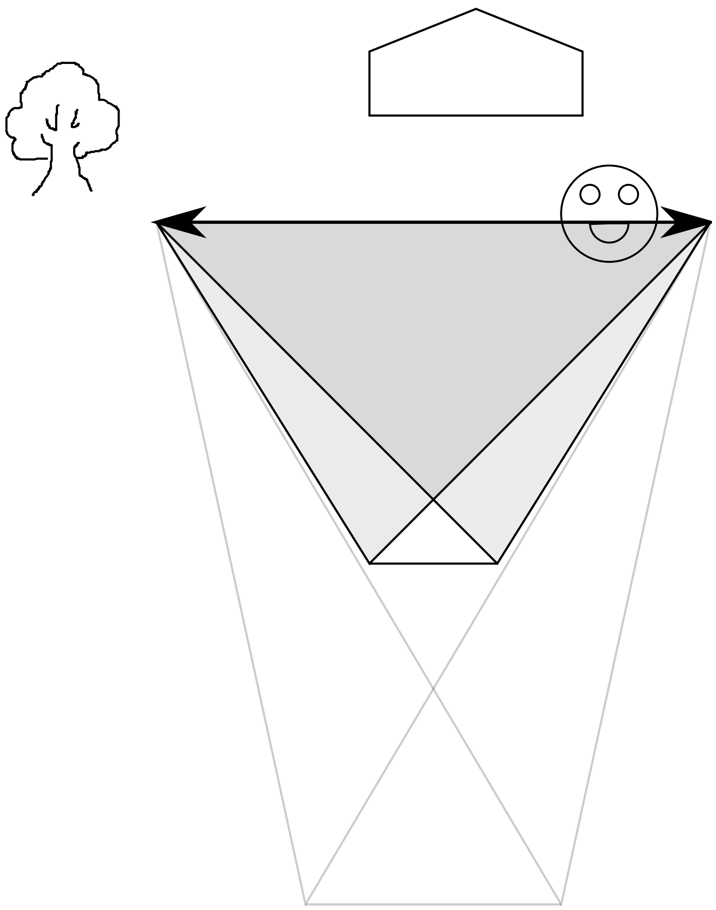
Large scene parts that are not visible in the original views may become **disoccluded**

⇒ Produces many holes and image artifacts...

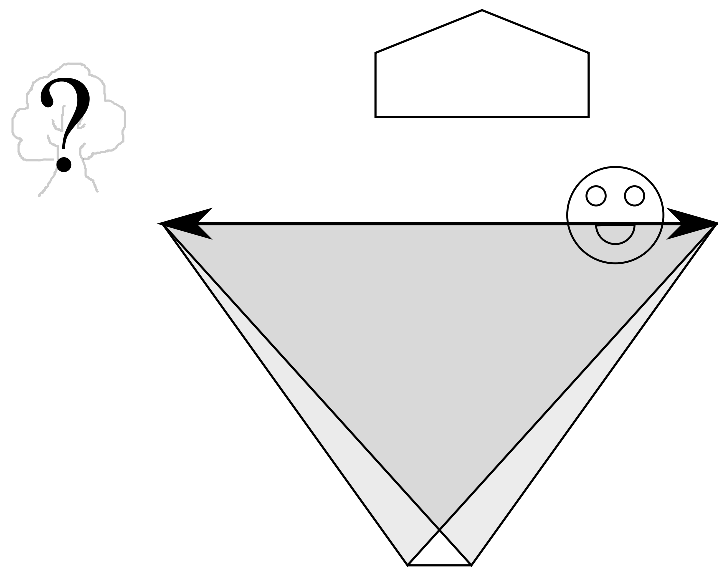
Viewpoint modification

3-D geometry is preserved

unseen objects become disoccluded



shooting geometry



viewing geometry

Hybrid disparity remapping

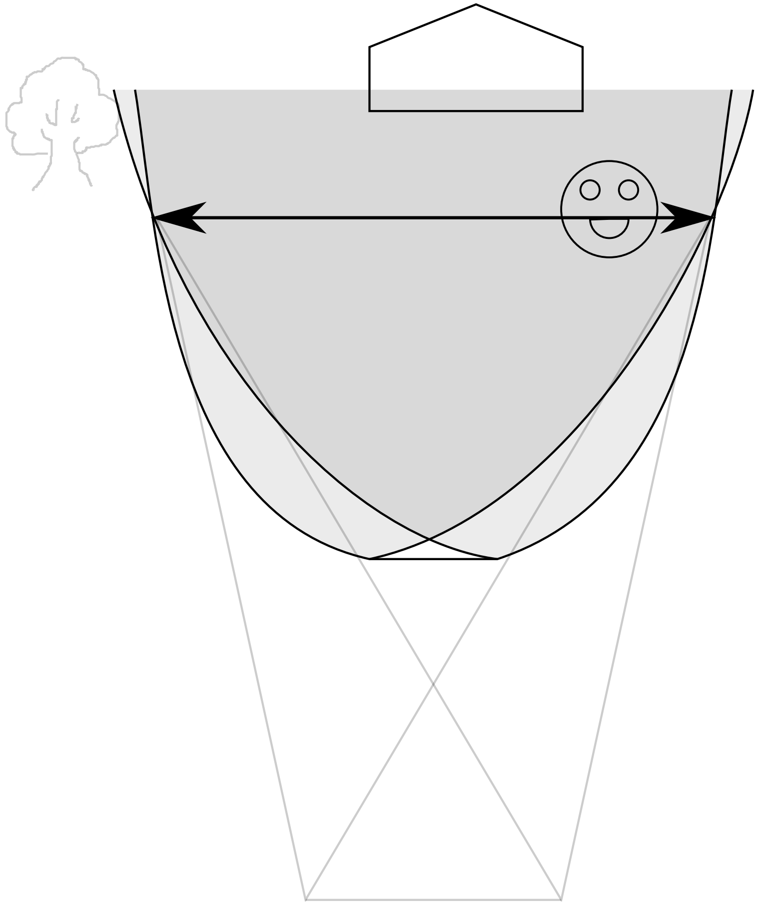
Compute a disparity remapping function $d''(d)$ so that

$$\rho_{\text{screen}} = 1 \text{ and } Z' = \alpha Z$$

⇒ *same disparity as viewpoint modification, but no depth-dependent image scaling.*

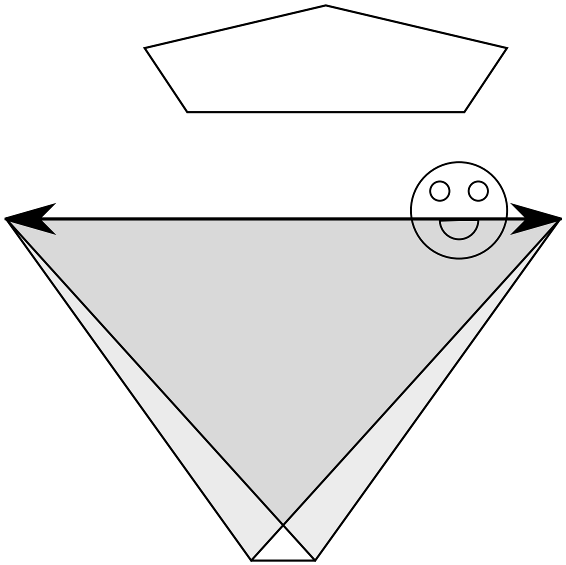
Depth is preserved, but image scale is not respected for off-screen objects - *Just like when zooming with a 2-D camera.*

Hybrid disparity remapping



depth is preserved

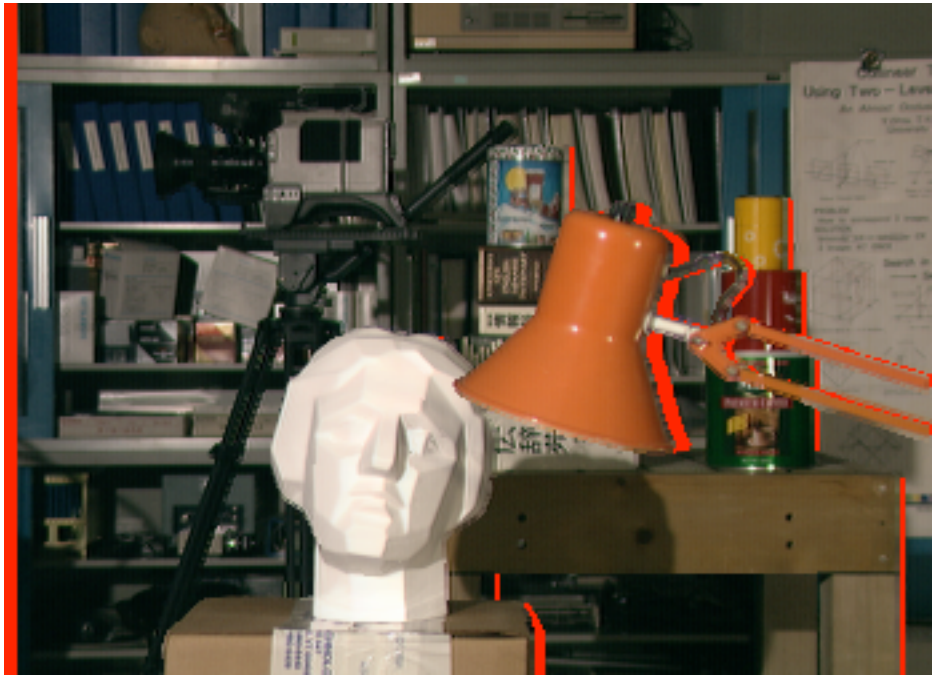
X-Y scale is distorted



shooting geometry

viewing geometry

Example showing disoccluded areas



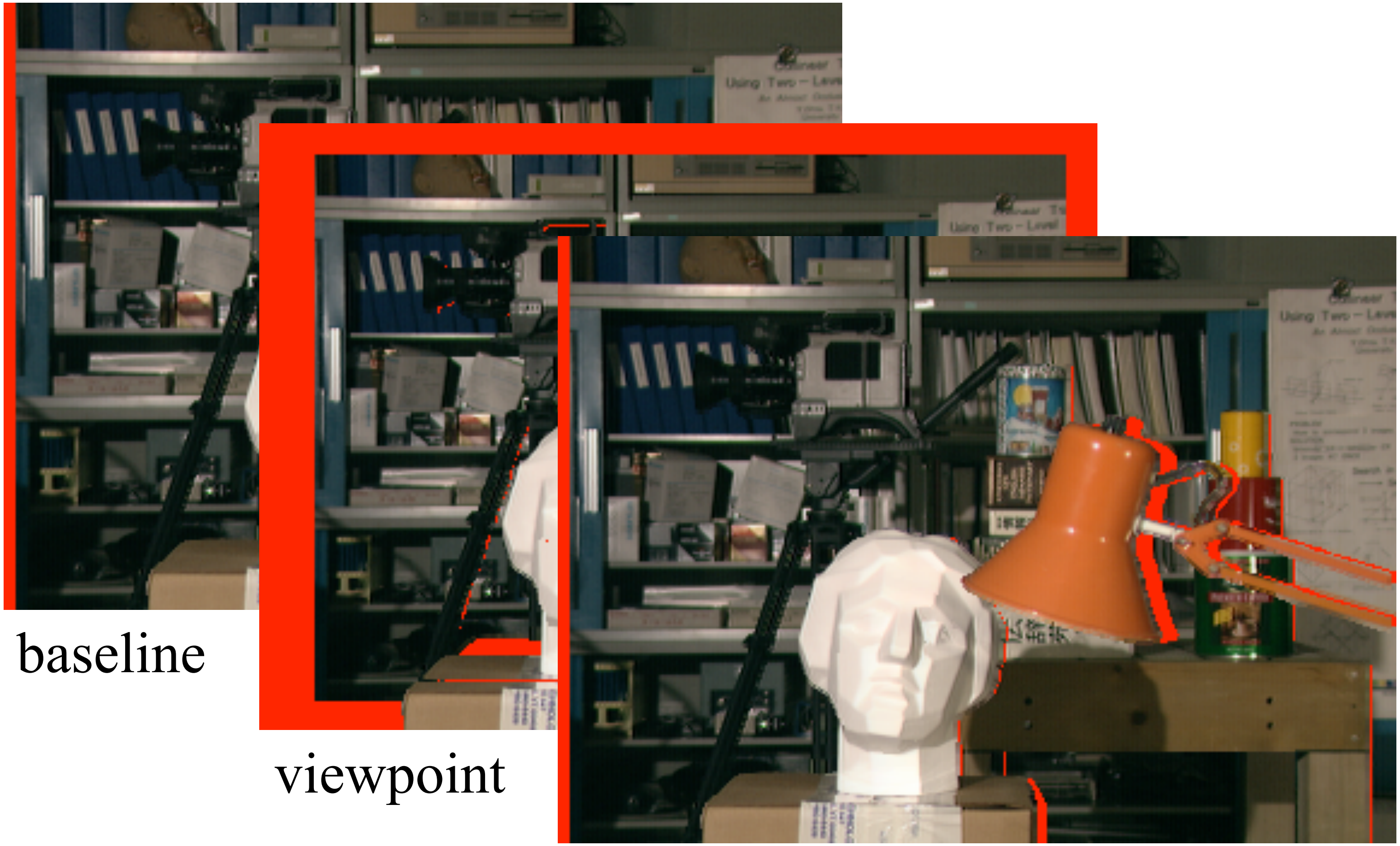
baseline

Example showing disoccluded areas

baseline

viewpoint

Example showing disoccluded areas



baseline

viewpoint

hybrid disparity remapping

Dealing with the vergence-accomodation conflict

Human depth of field for a screen at 3m is from 1.9m to 7.5m.

Corresponds to disparities from -3.8cm to 2.6cm.

In-focus objects should not be displayed out of this range!

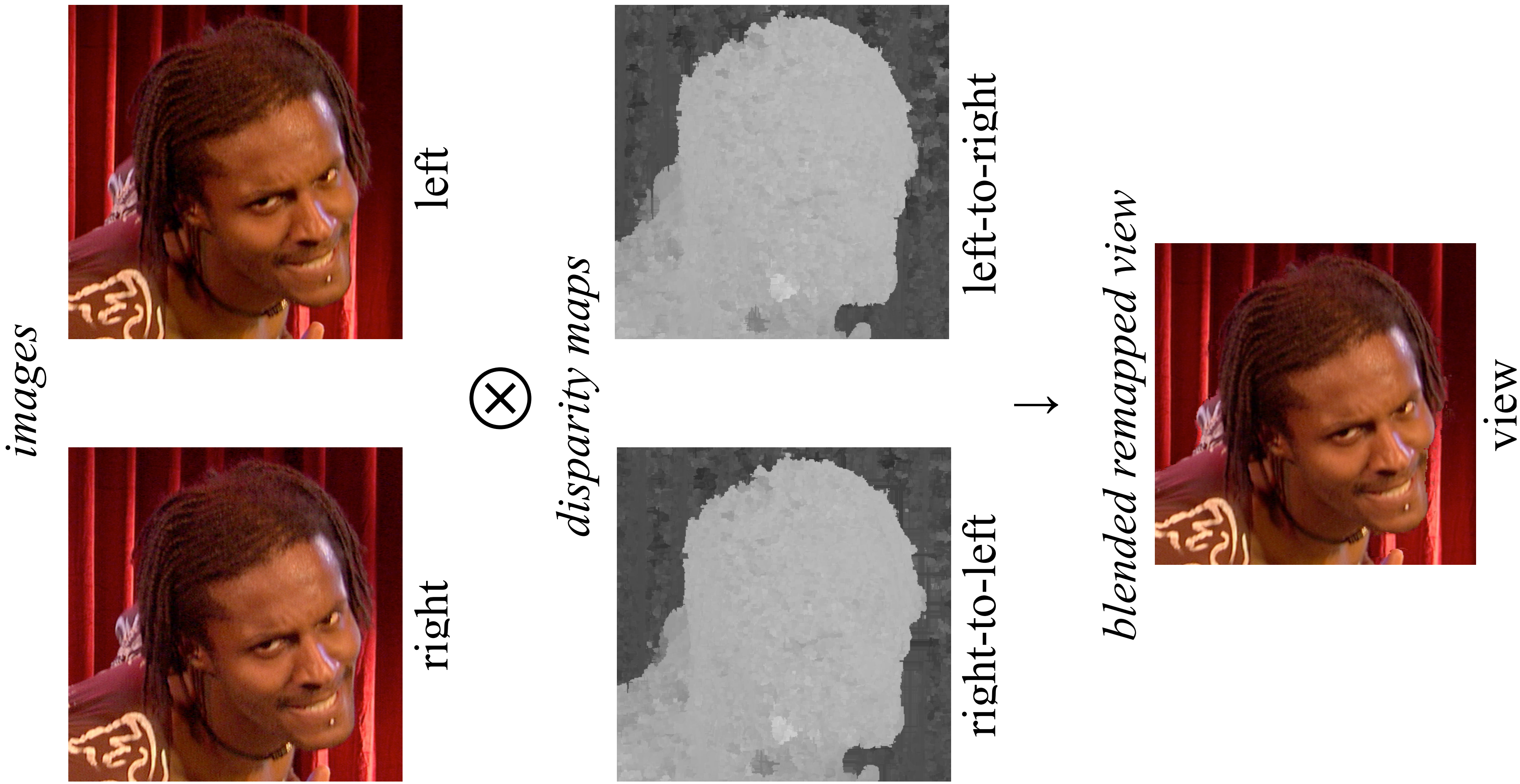
Hybrid disparity remapping can be used to adapt movies so that:

- The on-screen roundness factor is 1
- The disparity at infinity is no more than 2.6cm

Just synthesize views for a screen at the same distance, but 2.5 times wider! (6.5/2.6=2.5)

Similarly, the on-screen roundness factor can be adjusted by changing the screen distance.

New View Synthesis from Stereo



Artifacts!

Artifacts detection and removal

Our approach:

• Use **asymmetric** synthesis, so that one view keeps the highest possible image quality

• **Detect** artifacts in the synthesized view

• **blur out** the artifacts by anisotropic filtering

Why it should work:

• This locally reduces the high frequency content on artifacts

• The visual system will use *other 3-D cues* from the other (original) view to perceive 3-D in these areas [Stelmach 2000,Seuntiens 2006]

• Temporal consistency *should not be critical* because of low spatial frequency (to be validated)

Detecting and removing artifacts

Comparison of interpolated image with the original images:

- **colors** should be **similar**
- **Laplacian** should be **similar** too: *an edge can not appear!*

We compute a **confidence map** combining both, and use it as the conduction in the Perona-Malik anisotropic diffusion/blur equation:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(x,y,t)\nabla I) = c(x,y,t)\Delta I + \nabla c \cdot \nabla I$$

conduction $c \in [0,1]$

Laplacian

gradients

Interpolated frame

Interpolated frame, artifacts removed

Interpolated frame

Interpolated frame, artifacts removed

Conclusions

Hybrid disparity remapping of stereoscopic content solves most issues caused by classical novel view synthesis methods.

Asymmetric synthesis helps preserving perceived quality.

Artifact removal is performed by detecting and blurring out artifacts in the synthesized view

Demo movie

Work done within the 3DLive project: http://3dlive-project.com