Perceptually-Optimized Content Remapping for Automultiscopic Displays

Belen Masia[1,2] Gordon Wetzstein[2] Carlos Aliaga[1] Ramesh Raskar[2] Diego Gutierrez[1]
[1]Universidad de Zaragoza [2]MIT Media Lab

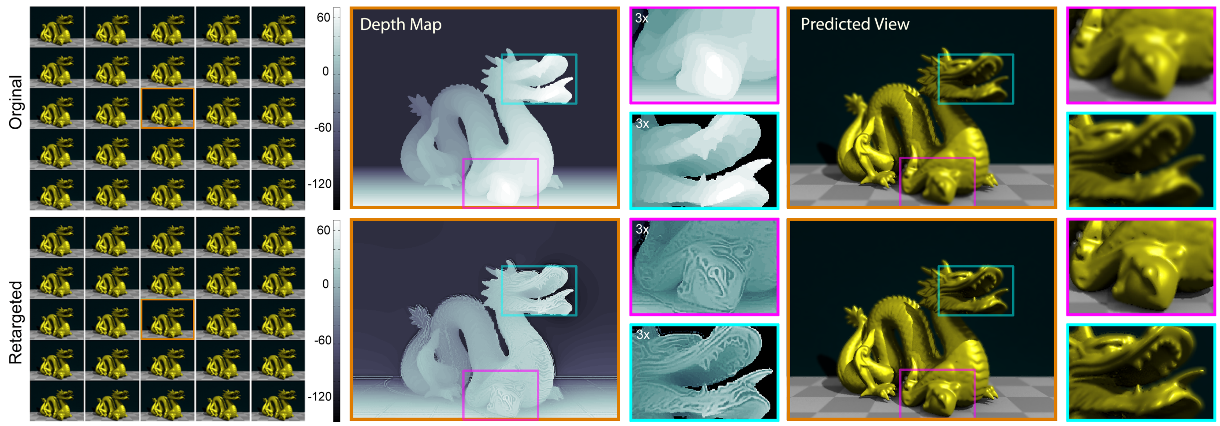

Figure 1: *Light field retargeting for automultiscopic 3D displays. Just like cameras, all automultiscopic 3D displays exhibit a limited depth of field that blurs virtual objects extruding from the physical display enclosure (top right). Using a perceptually-driven optimization framework, we retarget an input light field (left column) compressing the depth of the scene (center column) in such a way that the perceived 3D appearance is preserved but we also retain sharp details of the observed images (bottom right).*

Introduction 3D content and display technology is now widely available. However, available displays range from large-screen cinematic projection systems to hand-held devices and from screens supporting glasses-free 3D modes to glasses-bound systems. Counterintuitively, the produced content is usually only generated for a single display configuration, making labor-intense, manual post-processing of the data necessary. Recently, several content remapping techniques for stereo content have been proposed (see e.g. [Lang et al. 2010]). We present a perceptually-driven optimization framework for automatic light field remapping, specially designed for *automultiscopic* displays. It takes into account both the limitations of the target display as well as those of a human observer, and poses the problem as a non-linear least squares optimization. Our model includes depth of field limitations and the contrast sensitivity function, as well as sensitivities to binocular disparity and motion parallax in the perception of depth. Additionally, the model can be adapted for more common stereo remapping.

Disparity Remapping Glasses-free automultiscopic displays share the vergence-accommodation mismatch of stereoscopic displays. Their main limitation, however, is the reduced depth of field (similar to cameras with large apertures) [Wetzstein et al. 2012]; hence, 3D objects extruding from the physical display enclosure appear blurred (see Fig. 1, top right). To preserve the 3D appearance of the scene but also sharp image details (Fig. 1, bottom right), our objective function includes a term that accounts for the perceived blur using a model for contrast sensitivity [Mantiuk et al. 2011].

The most important depth cues provided by automultiscopic displays not supported by conventional 2D displays are binocular disparity and motion parallax. Our function includes two terms which aim at preserving the original depthmap. These terms are based on published perceptual findings, and take into account our ability to discriminate depth from these two cues. Sensitivity to both cues depends on the spatial frequency of the signal, and thus a muti-scale pyramid decomposition is needed. For binocular disparity our function incorporates threshold detection and discrimination data from Didyk et al. [2011]. For motion parallax, detection thresholds come from experiments by Bradshaw and colleagues [2006].

Our objective function has the following form, which we optimize for per-pixel depth d:

$$\arg\min_{d} \left(\mu_{DOF} \left\|\omega_{CSF} \left(\rho_S \left(L_{orig}\right) - \rho_S \left(\phi_b \left(L_{orig}, d\right)\right)\right)\right\|_2^2 \right.$$
$$+ \mu_{BD} \left\|\omega_{BD} \left(\rho_L \left(\phi_v \left(D_{orig}\right)\right) - \rho_L \left(\phi_v \left(d\right)\right)\right)\right\|_2^2$$
$$\left. + \mu_{MP} \left\|\omega_{MP} \left(\rho_L \left(\phi_v \left(D_{orig}\right)\right) - \rho_L \left(\phi_v \left(d\right)\right)\right)\right\|_2^2\right) \quad (1)$$

where μ_{DOF}, μ_{BD} and μ_{MP} are the weights given to each of the terms. $\phi_b\left(L, d\right)$ models the blurring of the original luminance image L_{orig} in a depth-dependent manner, and D_{orig} represents the original depthmap, converted into vergence values with operator $\phi_v\left(\cdot\right)$. The multi-scale decompositions into frequency levels are given by $\rho_S\left(\cdot\right)$ and $\rho_L\left(\cdot\right)$. Finally, ω_{CSF}, ω_{BD} and ω_{MP} are the weights accounting for sensitivity to the different aspects involved, explained above. For further details on our objective function and explanatory figures, please refer to the supplementary material.

Discussion As opposed to previous disparity remapping approaches, which focused on stereo displays, our framework automatically retargets an input light field for a given *automultiscopic* display. Our model is the first to take into account both the specific limitations of such displays, as well as the characteristics of the human visual system. As shown in the supplementary material, it can easily be adapted for stereo content as well.

References

BRADSHAW, M. F., HIBBARD, P. B., PARTON, A. D., ROSE, D., AND LANGLEY, K. 2006. Surface orientation, modulation frequency and the detection and perception of depth defined by binocular disparity and motion parallax. *Vision Research 46* (September), 2636–2644.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2011. A perceptual model for disparity. *ACM Trans. Graph. 30*, 96:1–96:10.

LANG, M., HORNUNG, A., WANG, O., POULAKOS, S., SMOLIC, A., AND GROSS, M. 2010. Nonlinear disparity mapping for stereoscopic 3d. *ACM Trans. Graph. 29* (July), 75:1–75:10.

MANTIUK, R., KIM, K. J., REMPEL, A. G., AND HEIDRICH, W. 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Trans. Graph. 30*, 40:1–40:13.

WETZSTEIN, G., LANMAN, D., HIRSCH, M., AND RASKAR, R. 2012. Tensor Displays: Compressive Light Field Display using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. 31*, 1–11.

SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008