Recreation of Stereo Pairs from 3D Anaglyph Images

Robin Grun
HAW Hamburg | robin.grun@haw-hamburg.de

Digital Version

HAW HAMBURG

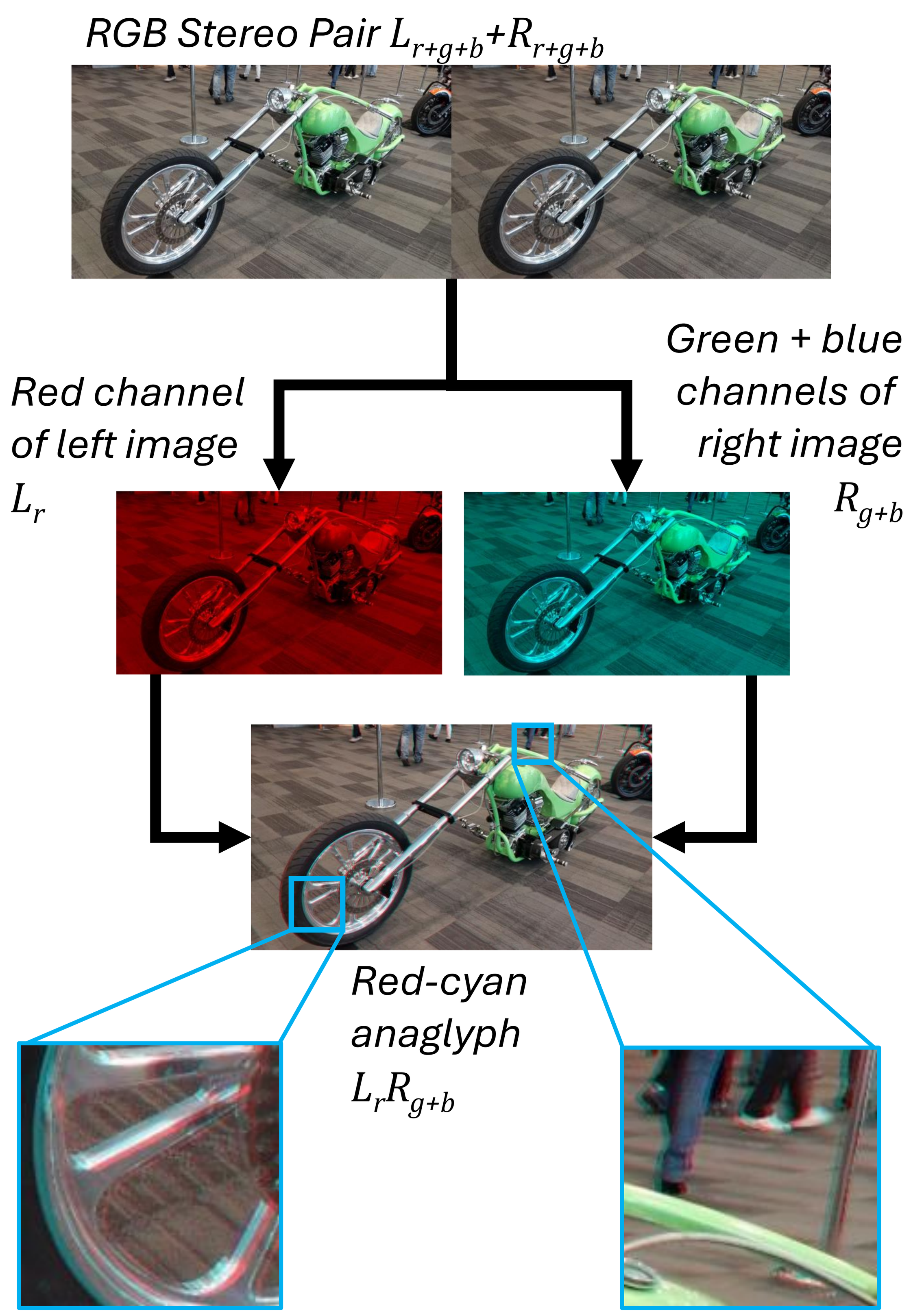
Introduction

- 3D anaglyph images are used for stereo visualization, e. g. movies, video games, macro-visualization of chemical structures, educational purposes
- During creation of anaglyph images, 50% of color information is being lost (green + blue channels of the left image and red channel of the right image)
- Mathematical recreation is not possible, thus missing color channels must be estimated
- Current reconstruction methods need additional information, next to input anaglyph images, e. g. original stereo-pairs, illumination- / depth information [2]
- Solution: recreation via Conditional Inverse Neural Network (cINN) for image colorization
- cINNs provide better training stability, are more comparable and produce higher quality outputs than current used alternatives, e. g. Generative Adversarial Networks (GANs), Variational Auto-encoders (VAEs) [2, 3]

Research Questions

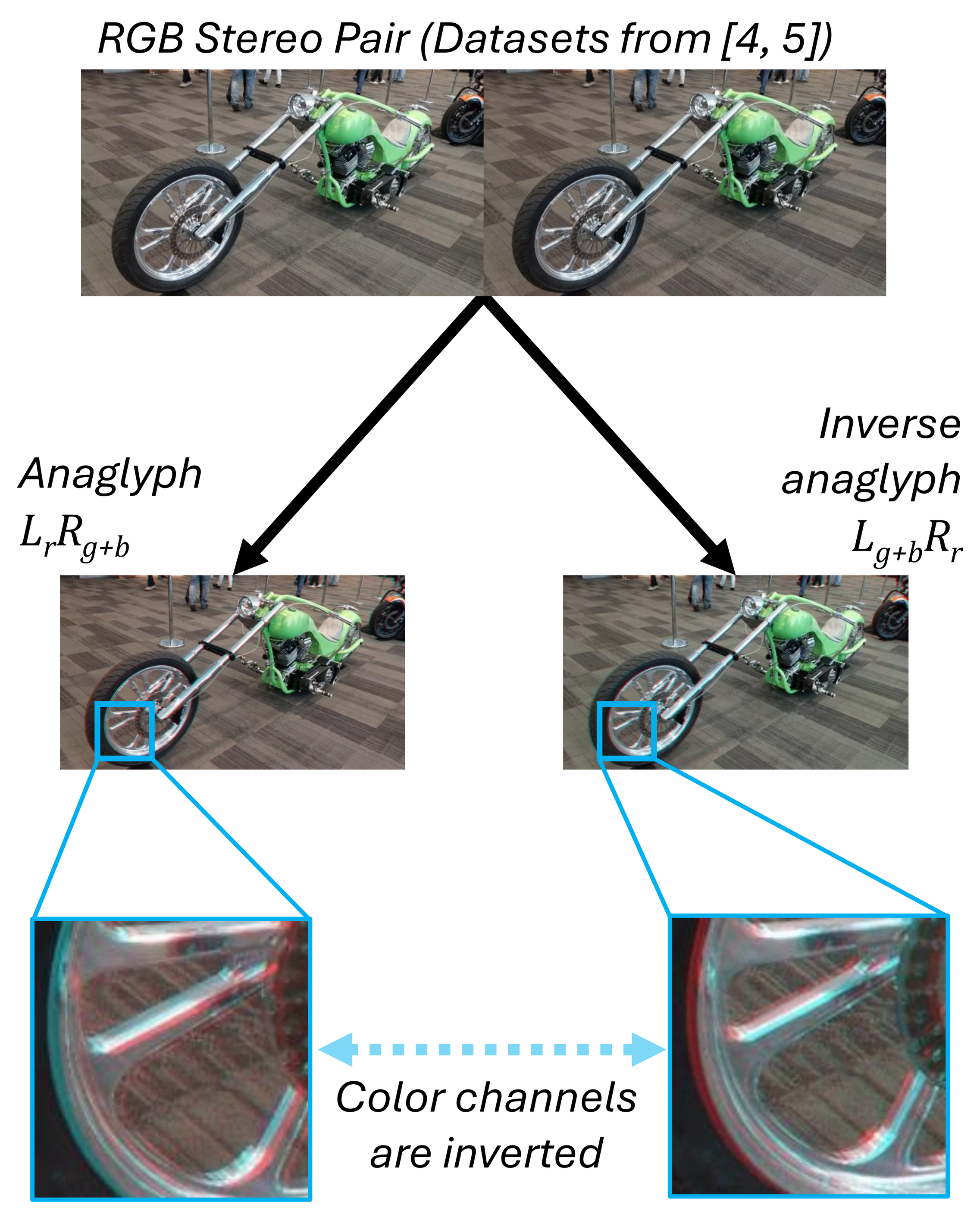
- How precise are estimation methods to recreate original stereo pairs without additional input information?
- How do cINN estimation methods perform in comparison to currently used alternatives (GANs/VAEs)?

Anaglyph Creation



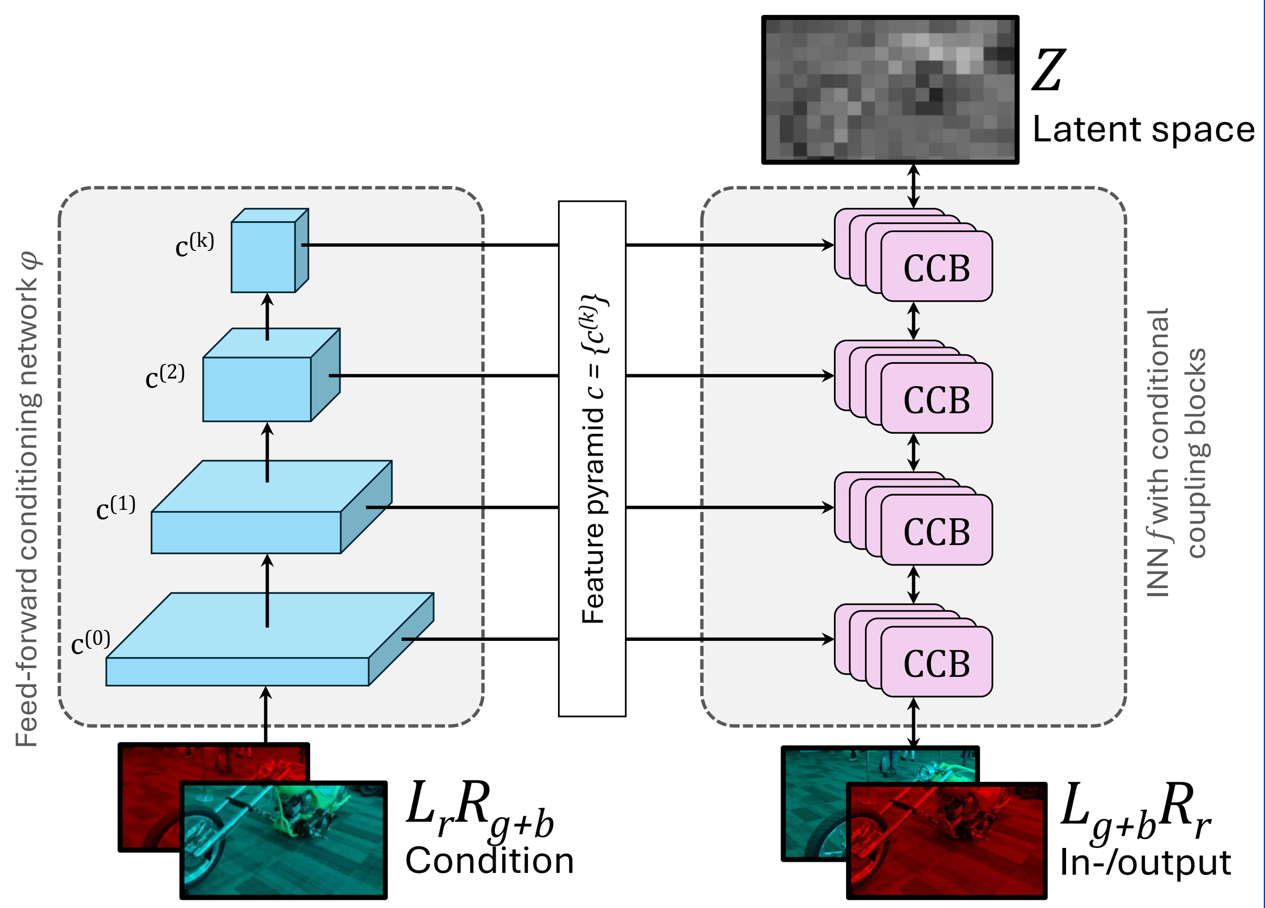
Training Data Generation

- Creation of 3D anaglyph images and inverse anaglyphs from stereo pairs



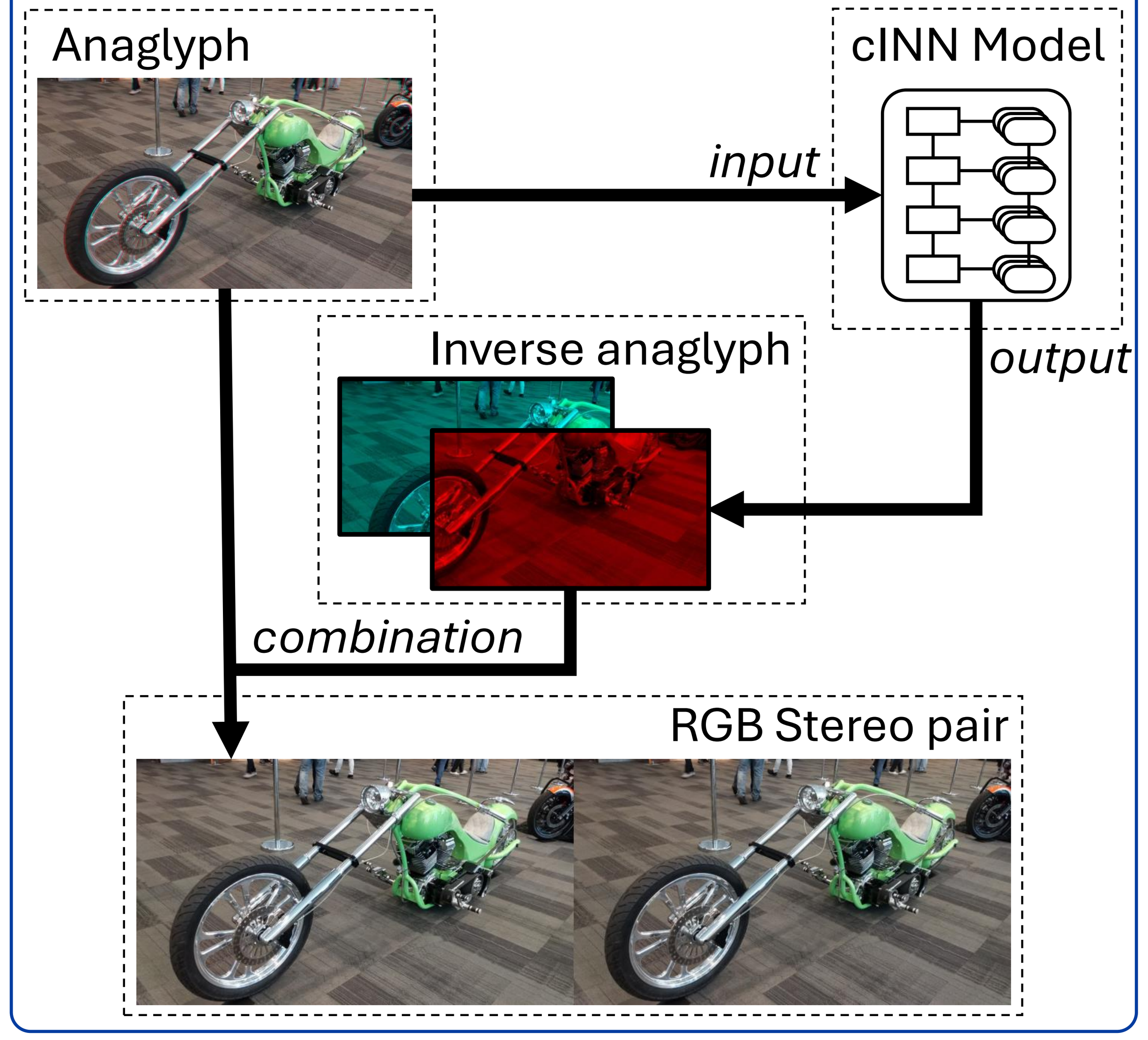
cINN Model [1]

- Anaglyphs are conditional input L_rR_{g+b}
- Training data are inverse anaglyphs as input $L_{g+b}R_r$
- Training input is transformed into latent space Z
- Model applies invertible wavelet downsampling between Conditional Coupling Blocks (CCBs)
- Estimated output by model (after training) are inverse anaglyphs $L_{g+b}R_r$



Conditional Invertible Neural Network (cINN) Architecture [1]

Model Usage



Future Steps

- Conduct experiments with cINN architecture models
- Comparison to current methods
- Transfer to anaglyph images with different color channel variations, e. g. anachrome, triscopic, ColorCode-3D

References

[1] L. Ardizzone, J. Kruse, C. Lüth, N. Bracher, C. Rother, and U. Köthe, "Conditional Invertible Neural Networks for Diverse Image-to-Image Translation," in *Pattern Recognition (Lecture Notes in Computer Science)*, Z. Akata, A. Geiger, and T. Sattler, Eds., Cham: Springer International Publishing, 2021, pp. 373–387.

[2] L. Ardizzone, C. Lüth, J. Kruse, C. Rother, and U. Köthe, "Guided Image Generation with Conditional Invertible Neural Networks," Jul. 2019. [Online]. Available: http://arxiv.org/pdf/1907.02392

[3] H. Hoyez, C. Schockaert, J. Rambach, B. Mirbach, and D. Stricker, "Unsupervised Image-to-Image Translation: A Review," *Sensors* (Basel, Switzerland), early access. doi: 10.3390/s22218540.

[4] Y. Hua et al., "Holopix50k: A Large-Scale In-the-wild Stereo Image Dataset," Mar. 2020. [Online]. Available: http://arxiv.org/pdf/2003.11172

[5] X. Liu, S. Iwase, and K. M. Kitani, "StereOBJ-1M: Large-scale Stereo Image Dataset for 6D Object Pose Estimation," Sep. 2021. [Online]. Available: http://arxiv.org/pdf/2109.10115