

Supplementary Material for: Lightweight Optimal-Transport Harmonization on Edge Devices

Anonymous submission

Experiments and Metrics

iHarmony4 Dataset Preprocessing During the metrics calculation, we observed that iHarmony4 composite images exhibit high-frequency artifacts, as illustrated by the difference image in Figure 1. According to the dataset generation procedure, unmasked regions of a composite image should be exactly the same as in the corresponding ground-truth image. However, we found substantial differences that affect the MSE and PSNR metrics. To exclude this noise, we re-assembled the iHarmony4 dataset by replacing the pixels from unmasked regions of composites with pixels taken from their ground-truth images. The resulting artifact-free set is saved in PNG format and named ‘iHarmony4-clean’. We evaluate all baselines on this cleaned dataset to ensure a fair comparison.

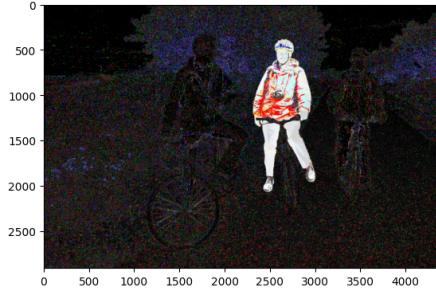


Figure 1: High-frequency artifacts are visible in a difference image between an original iHarmony4 composite and its corresponding real image.

Additional comparison with baselines on the ARCore data is given in Fig. 3, and for comparison with ground-truth, refer to Fig. 4.

Theoretical Analysis

This appendix provides the complete proofs of Lemma 1, Theorem 1 and Lemma 2 stated in Section 6 of the main paper. We adopt the same notation and Assumptions 1–2 introduced there. For convenience we restate each result below together with its proof.

Model Assumptions and Preliminaries Let the color space be the unit cube $\mathcal{X} = [0, 1]^3$. Let the unharmonized

and target color distributions, π_0 and π_1 , be probability measures supported on \mathcal{X} . We impose the following regularity conditions for our analysis.

Assumption 1 (Map Regularity). The true optimal transport map, $T^* : \mathcal{X} \rightarrow \mathcal{X}$, exists and is L -Lipschitz continuous for a constant $L < \infty$. The existence of such a map for a general case is non-trivial and is guaranteed under strong conditions on the measures, such as uniform log-concavity of their densities (?).

Assumption 2 (Distribution Regularity). The source distribution π_0 has mean μ_0 and a non-singular covariance matrix Σ_0 , which means that $\Sigma_0^{1/2}$ needed for the MKL map is well-defined. All expectations $\mathbb{E}[\cdot]$ are taken with respect to $X_0 \sim \pi_0$.

Our algorithm approximates T^* with the MKL map, $T_{\text{MKL}}(x) = \mu_1 + A(x - \mu_0)$, which is the optimal map for transporting between Gaussian surrogates $\mathcal{N}(\mu_0, \Sigma_0)$ and $\mathcal{N}(\mu_1, \Sigma_1)$ (?). Since the MKL map could potentially map part of the density outside of the unit cube, we apply clipping operation $\hat{T}_{\text{MKL}} := \Pi_{\mathcal{X}} \circ T_{\text{MKL}}$, where $\Pi_{\mathcal{X}}$ is the Euclidean projection onto \mathcal{X} , i.e. clipping transformed colors to the valid $[0, 1]^3$ gamut. We proof that the clipping operation coincides with the Euclidean projection in Lemma 1.

Lemma 1 (Clipping equals Euclidean projection). *Let the clipping operator $\text{clip} : \mathbb{R}^d \rightarrow [0, 1]^d$ be defined component-wise by*

$$(\text{clip}(z))_j = \min\{1, \max\{0, z_j\}\}, \quad j = 1, \dots, d. \quad (1)$$

Then for every $z \in \mathbb{R}^d$ the vector $y = \text{clip}(z)$ is the unique Euclidean projection of z onto the cube $\mathcal{X} = [0, 1]^d$; that is, $\text{clip}(z) = \Pi_{\mathcal{X}}(z)$.

Proof. Consider the minimization problem

$$\min_{x \in \mathcal{X}} \|z - x\|_2^2. \quad (2)$$

Because \mathcal{X} is the Cartesian product of intervals and the objective decomposes additively,

$$\|z - x\|_2^2 = \sum_{j=1}^d (z_j - x_j)^2, \quad (3)$$

$$\arg \min_{x \in \mathcal{X}} \|z - x\|_2^2 = \left(\arg \min_{x_j \in [0, 1]} (z_j - x_j)^2 \right)_{j=1}^d. \quad (4)$$

For a single coordinate z_j the scalar minimizer is

$$x_j^* = \begin{cases} 0, & z_j < 0, \\ z_j, & 0 \leq z_j \leq 1, \\ 1, & z_j > 1, \end{cases} \quad (5)$$

which is exactly $(\text{clip}(z))_j$. Stacking these x_j^* yields $y = \text{clip}(z)$, and by (4) this y minimizes the full problem. Uniqueness follows from strict convexity of the squared norm, hence $\text{clip}(z) = \Pi_{\mathcal{X}}(z)$. \square

Let us note that any projection operator is 1-Lipschitz since the distance between any two points after projection is never greater than their distance before projection. We seek to bound the expected squared error $\mathcal{E} := \mathbb{E}[\|\hat{T}_{\text{MKL}}(X_0) - T^*(X_0)\|^2]$.

Theorem 1 (Error Bound for L-Lipschitz Color Maps). *Let Assumptions 1 and 2 hold. The total error \mathcal{E} is bounded as:*

$$\mathcal{E} \leq 2\mathcal{E}_{\text{clip}} + 2\mathcal{E}_{\text{lin}}, \quad (6)$$

where the clipping error is $\mathcal{E}_{\text{clip}} := \mathbb{E}[\|T_{\text{MKL}}(X_0) - \hat{T}_{\text{MKL}}(X_0)\|^2]$, and the linearity error, \mathcal{E}_{lin} , is bounded by:

$$\mathcal{E}_{\text{lin}} \leq 2B^2 + 2(\|A\|_{\text{op}} + L)^2 \cdot \text{tr}(\Sigma_0). \quad (7)$$

Here, $B = |\mu_1 - T^*(\mu_0)|$ is a bias term, $\|A\|_{\text{op}}$ is the spectral norm of the MKL matrix, i.e. its largest singular value, which depends only on source and target distribution covariances, and $\text{tr}(\Sigma_0)$ is the trace of the source covariance matrix.

Proof. The decomposition $\mathcal{E} \leq 2\mathcal{E}_{\text{clip}} + 2\mathcal{E}_{\text{lin}}$ follows from the inequality $\|u + v\|^2 \leq 2\|u\|^2 + 2\|v\|^2$. We focus on bounding the linearity error term, $\mathcal{E}_{\text{lin}} = \mathbb{E}[\|T_{\text{MKL}}(X_0) - T^*(X_0)\|^2]$. We introduce and subtract terms centered at the mean of the source distribution, μ_0 :

$$\begin{aligned} & \|T_{\text{MKL}}(x) - T^*(x)\| \\ &= \|(\mu_1 + A(x - \mu_0)) - T^*(x)\| \\ &= \|(\mu_1 - T^*(\mu_0)) + (A(x - \mu_0) - (T^*(x) - T^*(\mu_0)))\|. \end{aligned}$$

Applying the triangle inequality, $\|u + v\| \leq \|u\| + \|v\|$, yields:

$$\begin{aligned} & \|T_{\text{MKL}}(x) - T^*(x)\| \\ &\leq \|\mu_1 - T^*(\mu_0)\| + \|A(x - \mu_0) - (T^*(x) - T^*(\mu_0))\|. \end{aligned}$$

Let $B = \|\mu_1 - T^*(\mu_0)\|$. For the second term, we again use the triangle inequality and the L-Lipschitz assumption on T^* :

$$\begin{aligned} & \|A(x - \mu_0) - (T^*(x) - T^*(\mu_0))\| \\ &\leq \|A(x - \mu_0)\| + \|T^*(x) - T^*(\mu_0)\| \\ &\leq \|A\|_{\text{op}}\|x - \mu_0\| + L\|x - \mu_0\| \\ &= (\|A\|_{\text{op}} + L)\|x - \mu_0\|. \end{aligned}$$

Combining these yields the pointwise bound: $\|T_{\text{MKL}}(x) - T^*(x)\| \leq B + (\|A\|_{\text{op}} + L)\|x - \mu_0\|$. To obtain the bound

on \mathcal{E}_{lin} , we square this expression and take the expectation. Using $(u + v)^2 \leq 2u^2 + 2v^2$:

$$\begin{aligned} \mathcal{E}_{\text{lin}} &= \mathbb{E}[\|T_{\text{MKL}}(X_0) - T^*(X_0)\|^2] \\ &\leq \mathbb{E}[2B^2 + 2(\|A\|_{\text{op}} + L)^2\|X_0 - \mu_0\|^2]. \end{aligned}$$

Since $\mathbb{E}[\|X_0 - \mu_0\|^2] = \text{tr}(\Sigma_0)$ by definition of covariance, the result follows. \square

This theorem bounds the approximation error, linking it to the Lipschitz constant L of the true optimal transport map and the tail probability $\mathcal{E}_{\text{clip}} \leq d \cdot \mathbb{P}[T_{\text{MKL}}(X_0) \notin \mathcal{X}]$. The latter bound holds due to the following Lemma 2.

Lemma 2 (Tail-probability bound for the clipping error). *Let $\Pi_{\mathcal{X}} = \text{clip}(\cdot)$ be the Euclidean projection onto $\mathcal{X} = [0, 1]^d$, defined as*

$$(\text{clip}(z))_j := \min\{1, \max\{0, z_j\}\}, \quad (8)$$

for $j = 1, \dots, d$, $z \in \mathbb{R}^d$ and define

$$\mathcal{E}_{\text{clip}} := \mathbb{E}[\|Z - \Pi_{\mathcal{X}} Z\|^2], \quad Z \in \mathbb{R}^d. \quad (9)$$

Then

$$\mathcal{E}_{\text{clip}} \leq d \mathbb{P}[Z \notin \mathcal{X}], \quad (10)$$

and for our application with $Z = T_{\text{MKL}}(X_0)$ and $d = 3$,

$$\mathcal{E}_{\text{clip}} \leq 3 \mathbb{P}[T_{\text{MKL}}(X_0) \notin \mathcal{X}]. \quad (10')$$

Proof. Because clip acts coordinate-wise,

$$|z_j - (\text{clip } z)_j| \leq 1 \quad \text{for each } j.$$

Hence

$$\|z - \text{clip } z\|^2 = \sum_{j=1}^d (z_j - (\text{clip } z)_j)^2 \leq d \mathbf{1}_{\{z \notin \mathcal{X}\}},$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. Taking expectations with $z = \bar{Z}$ yields (10). \square

3D Model Credits



Figure 2: The 3D objects used in our experiments.

3D objects used in our experiments are collected from open source and converted to the Wavefront .obj format and single “baked” texture file. Some of them are shown in Fig. 2. Models’ authors and sources are listed below

- **AR Core pawn**

github.com/google-ar/arcore-android-sdk

- **Chair Author:** River Yang,
<https://www.blenderkit.com/asset-gallery-detail/703fe964-84bb-4468-84bd-f6aa3fb337b4/>
- **Chicken**
- **Beach ball** Author: Mohammed Hossain,
<https://www.blenderkit.com/asset-gallery-detail/cb647b58-d1c8-4a34-a6b8-263cd002bf30/>
- **Green sofa** Author: Joris LM,
<https://www.blenderkit.com/asset-gallery-detail/573bf228-019b-4072-8086-bc45a6a2b2fa/>
- **IKEA armchair** Author: Branislav Kubečka,
<https://www.blenderkit.com/asset-gallery-detail/82020230-9451-4683-bb6b-1e01c7c4fa36/>
- **Burger** Author: Raymond Gabriel,
<https://www.blenderkit.com/asset-gallery-detail/5fe9d2e3-31ac-44d1-b392-cdb00ae1d490/>
- **Skeleton** Author: Aidan Sanderson,
<https://www.blenderkit.com/asset-gallery-detail/130ec931-f4df-45b8-9a81-2660ccceb581/>
- **Knight's armor** Author: AnomalyFound,
<https://www.blenderkit.com/asset-gallery-detail/4d71b492-be49-4748-b159-6ec337aefa50/>
- **Car** Published by: WyattP,
https://open3dmodel.com/3d-models/lexus-free-3d-model-car_12127.html
- **Tangerines** Author: Yahku le Roux,
<https://www.blenderkit.com/asset-gallery-detail/fc4e728b-c25b-4059-bdc9-d211b4fadcf8/>
- **Manhole** Author: RaunoX, Poly Heaven
<https://www.blenderkit.com/asset-gallery-detail/847ef18f-3f72-4893-b5c2-389a1898c21b/>
- **Giant fern** Author: Amanpreet Bajwa,
<https://www.blenderkit.com/asset-gallery-detail/25d6a477-7a01-46b1-b5dd-d162fe2ab2fc/>
- **Penguin** Author: nabarun1011,
<https://free3d.com/3d-model/emperor-penguin-601811.html>, <https://sketchfab.com/3d-models/penguin-f65e799bf9534c66a12724a93bb72c39>
- **Coffee mug** Generated,
<https://www.meshy.ai/>
- **Anime girl** Generated,
<https://www.meshy.ai/>
- **Guitar** Author: nabarun1011,
<https://sketchfab.com/3d-models/electric-guitar-d51983607d404791acb42f460259f23a>
- **Beagle dog**
<https://www.cadnav.com/3d-models/model-55200.html>
- **Deer** Author: TheCaitasaurus,
https://open3dmodel.com/3d-models/little-deer-animal_352076.html
- **Pigeon** Published by: Zoe_A,
https://open3dmodel.com/3d-models/3d-model-rock-pigeon_119874.html

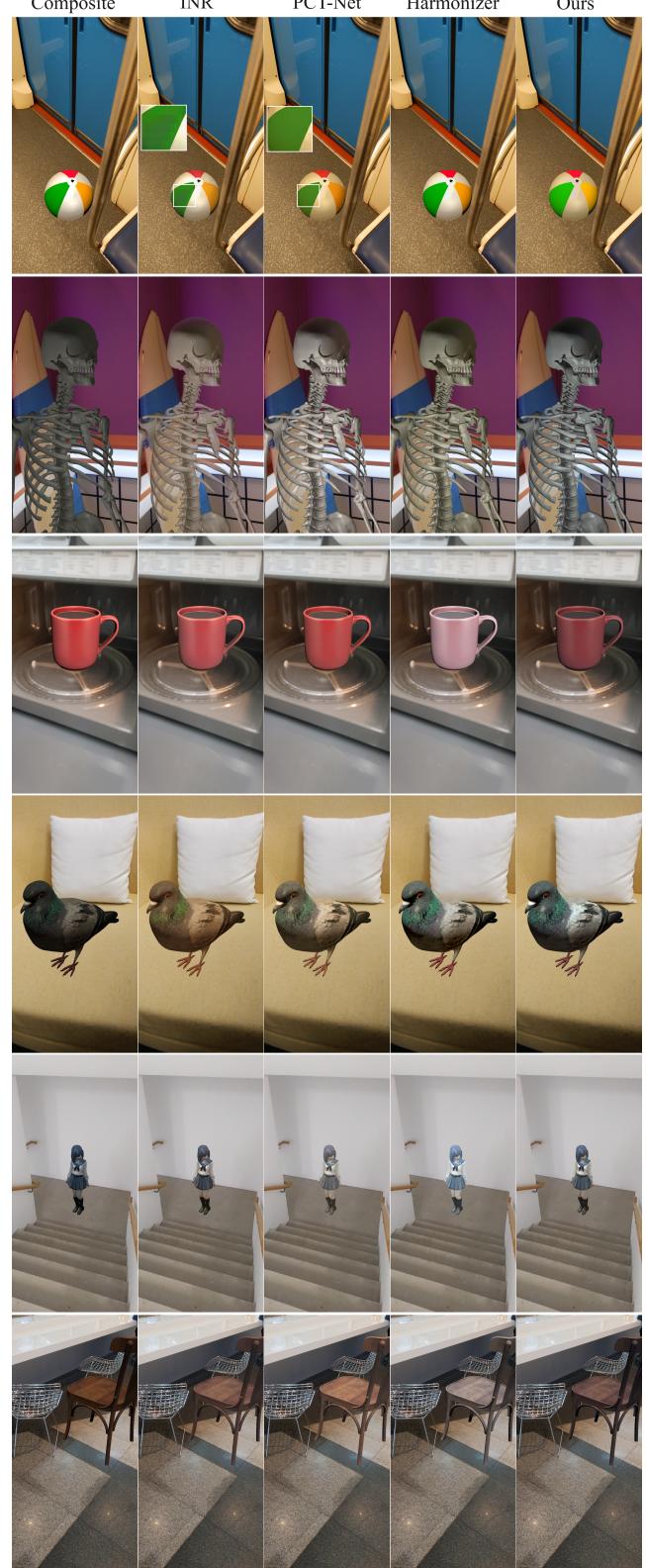


Figure 3: Qualitative comparison with baselines.

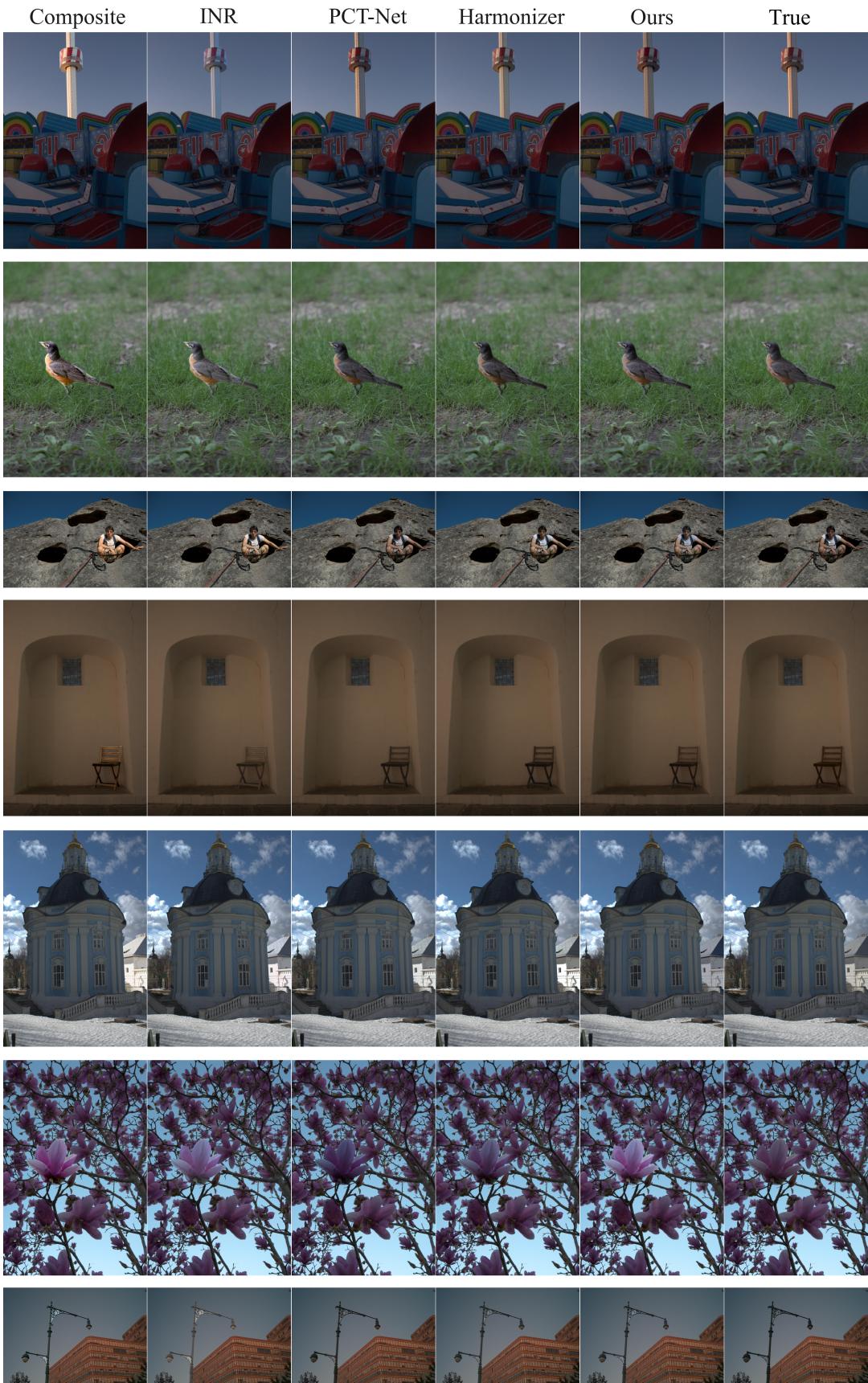


Figure 4: Qualitative comparison with baselines and true images on HAdobe5k test set.