Optimal Image Transport over Sparse Dictionaries

Junqing Huang[1], Haihui Wang[2], Alexandra Pizzurica[3], Michael Ruzhansky[4,5]

[1] Chair of Computational Mathematics, University of Deusto
[2]Department of Mathematics, Beihang University
[3]Department of Telecommunications and Information Processing, Ghent University
[4]Department of Mathematics: Analysis, Logic and Discrete Mathematics, Ghent University
[5]Department of Mathematics, Queen Mary University of London

Junqing Huang is currently supported by
Grant agreement Nº: 101096251-CoDeFeL
ERC Advanced Grant 2022 - Control for Deep and Federated Learning

June 18, 2026

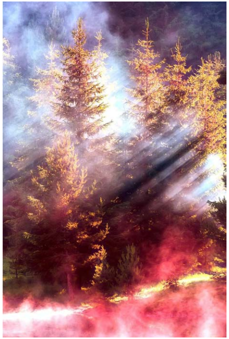
Overview

Q1: Can we change the color of an image to that of a given reference image?

H., Hristina et al.

Input image

Reference image



Results

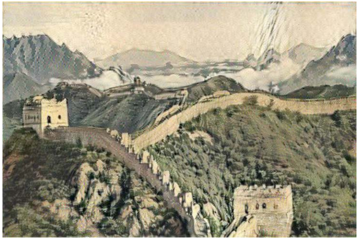
Image Color Transform

Q2: Can we change the style of an image to that of a given reference image?

Y. Jing et al.

Input image

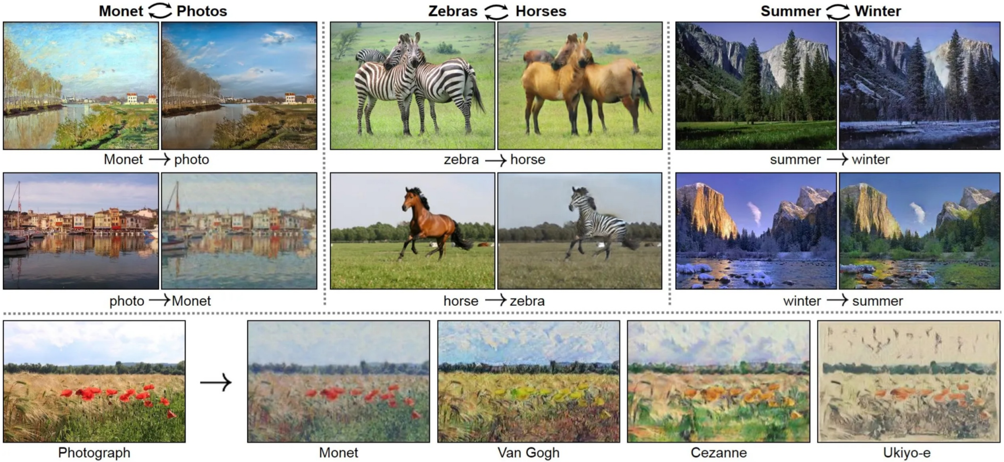
Reference image



Results

Image Style Transfer

Introduction: Image-to-Image Translation



Y. Zhu et al.

Introduction: Color Transform

Existing Solutions:

- Compute color histograms
- Optimal manpping on (sampled) histograms
- Reconstruction by interpolation for high-quality results

Advantages & limitations

- Optimal transport in data space
- Theoretical simplicity
- High computational cost

S., Ferradans et al.

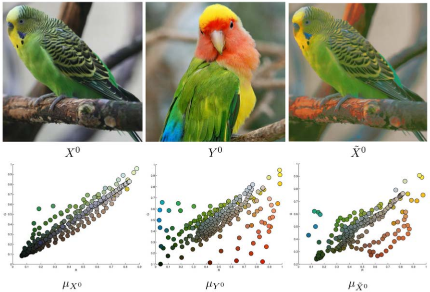


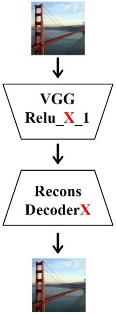
Image Color Transform

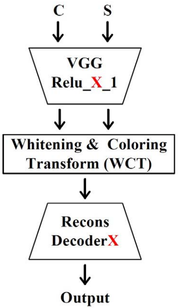
Existing Solutions:

- Deep learning encoding for style representation
- Mapping in (latent) style spaces
- Deep learning decoding for reconstruction

Advantages & limitations

- Optimal transport in latent space
- High cost of data collection
- Unpredicted failure cases





H., Hristina et al.

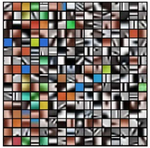
Neural Style Transfer

Motivation

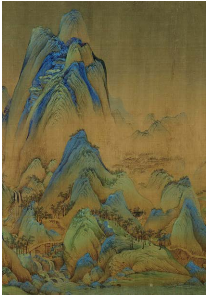
Main Idea: Learning style dictionaries and swapping for specified color/style effects?

Trey Ratcliff

Optimal
$\Longleftrightarrow$
Transfer



H., Hristina et al.



The learned dictionaries of two different images

Motivation

Main Idea: Learning style dictionaries and swapping for specified color/style effects?

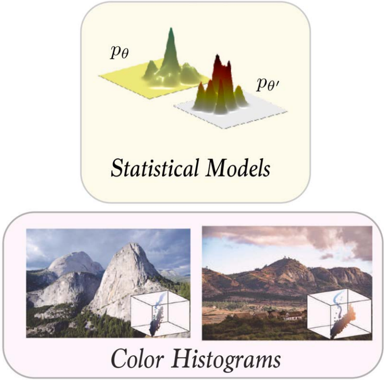
Optimal Transport (OT) Problem:

- Sparse coding for feature (e.g., color, style) representation
- Efficient OT mapping over sparse dictionaries
- Image synthesis via sparse reconstruction

Advantages:

- Theoretical foundation and simplicity
- Computational efficiency
- Easy implementation

Marco Cuturi



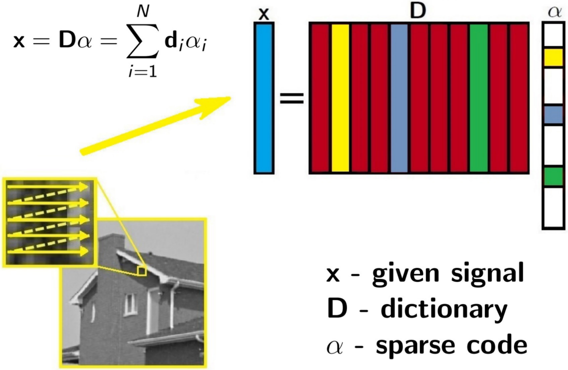
Outline

Sparse Coding and Representation

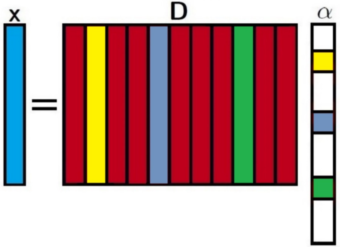
Sparse Coding:

Alexandra Pizzurica et al.

$$\mathbf{x} = \mathbf{D}\alpha = \sum_{i=1}^{N} \mathbf{d}_i \alpha_i$$



Sparse Coding and Representation



$$\hat{\alpha} = \arg\min_{\alpha} \|\mathbf{x} - \mathbf{D}\alpha\|_2^2 \quad \text{s.t.} \quad \|\alpha\|_0 \leq K$$

$$\hat{\alpha} = \arg\min_{\alpha} \|\alpha\|_0 \quad \text{s.t.} \quad \|\mathbf{x} - \mathbf{D}\alpha\|_2 \leq \varepsilon$$

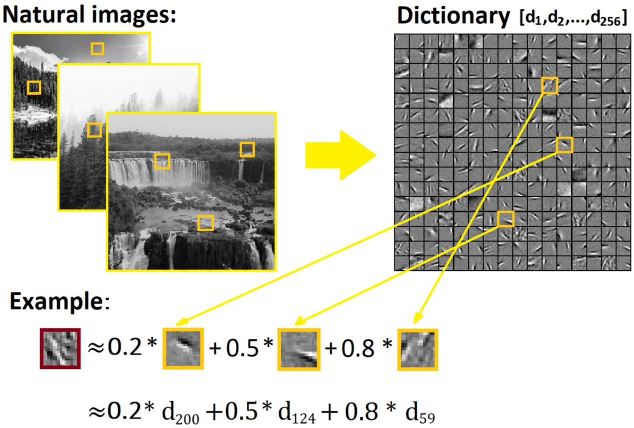
Greedy Algorithms:

- Matching Pursuit (MP) [Mallat and Zhang, '93]
- Orthogonal Matching Pursuit (OMP) [Tropp, '04]

Convex Relaxations:

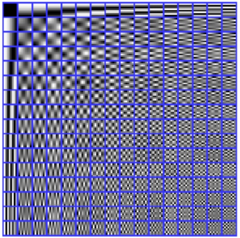
- LASSO [Tibshirani, '93]
- Basis Pursuit Denoising (BPDN) [Chen et al, '01]

Sparse Coding and Representation

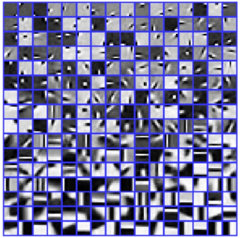


Dictionaries and Dictionary Learning:

- **Pre-defined dictionaries**: DCTs, wavelets, curvelets, sharelets, framelets, ...
- **Learned dictionaries**: trained on the representative datasets



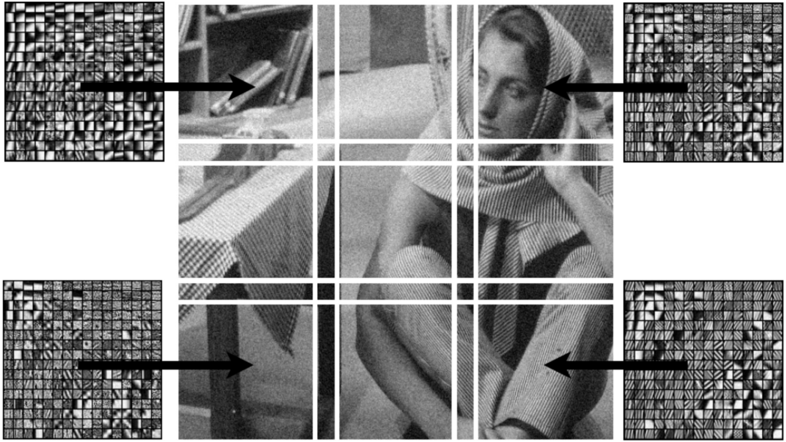
Discrete cosine transform dictionary



Dictionary based on natural images

Sparse Coding and Representation

Learned Local dictionaries:



Dictionaries learned on different parts of the Barbara image

Outline

Optimal Transport (OT)

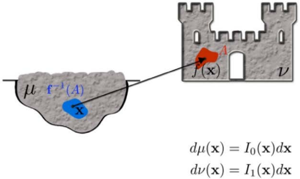
Definition (Monge Problem)

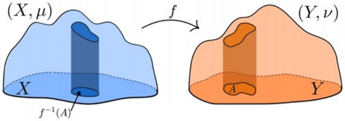
Let μ, ν be two probability measures in spaces $\mathcal{X}$ and $\mathcal{Y}$, a map $T\colon \mathcal{X} \to \mathcal{Y}$ w.r.t. a function $c(x, y)$, which minimizes the transport cost

$$\inf_{T_\sharp \mu = \nu} \int_{\mathcal{X}} c(x, T(x)) d\mu(x), \tag{1}$$

where $\mu(\mathcal{X}) = \int_{\mathcal{X}} d\mu = 1$, $\nu(\mathcal{Y}) = \int_{\mathcal{Y}} d\nu = 1$, and $T_\sharp \mu \overset{\text{def}}{=} \sum_{i:T(x_i)=y_j} \mu_i \delta_{T(x_i)}$ is the *push-forward* operator; $T_\sharp \mu = \nu$ means that T pushes forword the mass of μ to ν.

Gaspard Monge
1746-1818





Optimal Transport

Discrete Case: Let $\boldsymbol{a}$ be a probability vector and $\boldsymbol{a} \in \Sigma_n$ that blongs to the probability simplex:

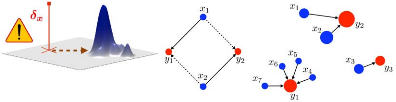
$$\Sigma_n \stackrel{\text{def}}{=} \left\{ \boldsymbol{a} \in \mathcal{R}_+^n : \sum_i^n \boldsymbol{a}_i = 1 \right\}$$

.

The Monge transport problem w.r.t. discreted probability measures $\boldsymbol{\mu} := \sum_{i=1}^n \boldsymbol{a}_i \delta_{x_i}$ and $\boldsymbol{\nu} := \sum_{j=1}^m \boldsymbol{b}_j \delta_{y_j}$ is,

$$\min_T \left\{ \sum_i c(x_i, T(x_i)) : T_\sharp \boldsymbol{\mu} = \boldsymbol{\nu} \right\}. \tag{2}$$

- Optimal solution T^* is not always available
- T^* may be unstable and not always unique
- High computational cost (e.g., L.P. $\mathcal{O}(n^3)$)

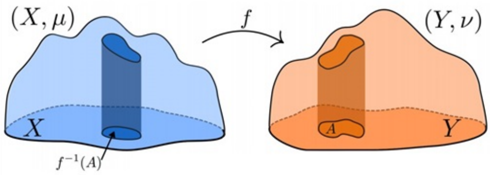


Definition (Kantorovich Relaxation)

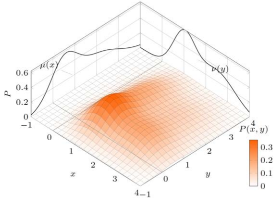
Given two probability measures $\mu, \nu \in \mathcal{P}(\Omega)$ and a cost function $c(x, y)$ on $\Omega \times \Omega$, Kantorovich problem minimizes the transport cost with probabilistic map $T \in \mathcal{P}(\Omega \times \Omega)$, that is,

$$\inf_{T \in \Pi(\mu,\nu)} \iint c(x, y) T(dx, dy), \tag{3}$$

where $\Pi(\mu, \nu) \overset{\text{def}}{=} \{T \in \mathcal{P}(\Omega \times \Omega) | \forall A, B, T(A \times \Omega) = \mu(A), T(B \times \Omega) = \nu(B)\}$



$\Longleftrightarrow$

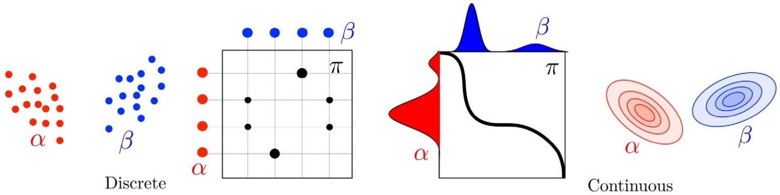


Kantorovich Relaxation

Discrete Case: Consider $\boldsymbol{\mu} := \sum_{i=1}^{n} \boldsymbol{a}_i \delta_{x_i}$ and $\boldsymbol{\nu} := \sum_{j=1}^{m} \boldsymbol{b}_j \delta_{y_j}$, the Kantorovich relaxation is then generalized as,

$$\boldsymbol{L}_{\mathbf{C}}(\boldsymbol{a}, \boldsymbol{b}) \stackrel{\text{def}}{=} \min_{\mathbf{T} \in U(\boldsymbol{a}, \boldsymbol{b})} \langle \mathbf{C}, \mathbf{T} \rangle \stackrel{\text{def}}{=} \min_{\mathbf{T} \in U(\boldsymbol{a}, \boldsymbol{b})} \sum_{i,j} \mathbf{C}_{i,j} \mathbf{T}_{i,j}, \tag{4}$$

where $U(\boldsymbol{a}, \boldsymbol{b}) \stackrel{\text{def}}{=} \{\mathbf{T} \in R_{+}^{n \times m} | \mathbf{T}\mathbf{1}_m = \boldsymbol{a}, \mathbf{T}^\top \mathbf{1}_n = \boldsymbol{b}\}$ and $\mathbf{C}_{\boldsymbol{i},\boldsymbol{j}} \stackrel{\text{def}}{=} \mathbf{C}(x_i, y_j)$ is the cost to move a single unit resource from location i to j.



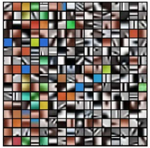
Kantorovich optimal transport with discrete and continuous cases

Outline

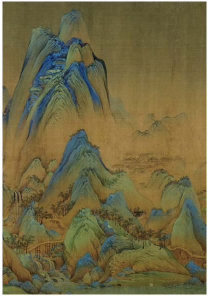
Main Idea: Learning and Swapping Dictionaries

Trey Ratcliff

Optimal
⟺
Transfer



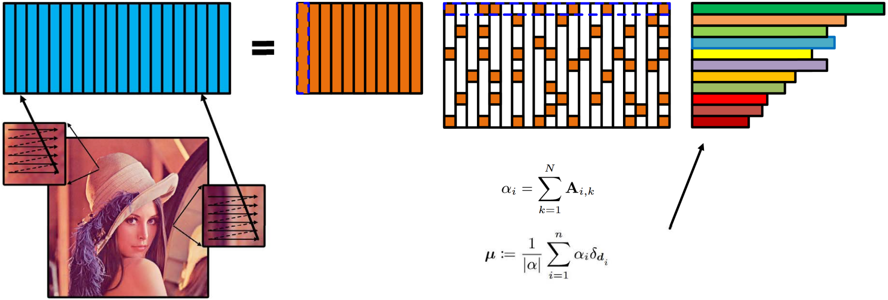
H., Hristina et al.



The learned dictionaries of two different images

Motivation

- **Sparse Representation:**



$$\alpha_i = \sum_{k=1}^{N} \mathbf{A}_{i,k}$$

$$\boldsymbol{\mu} := \frac{1}{|\alpha|} \sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{d}_i}$$

$$\mathbf{X} = \mathbf{DA}, \qquad s.t. \quad \forall k : \quad ||\boldsymbol{a}_k||_0 \leq K.$$

Problem Formulation

- **Sparse Representation:**

$$\mathbf{X}_{d\times N} = \mathbf{D}^x_{d\times n}\mathbf{A}_{n\times N} \qquad \mathbf{Y}_{d\times M} = \mathbf{D}^y_{d\times m}\mathbf{B}_{m\times M}$$

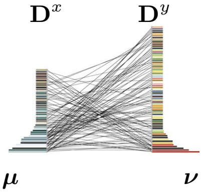
$$s.t. \quad ||\boldsymbol{a}_i||_0 \leq K_1,\, \mathbf{D}^x \in \mathcal{D}_1, \qquad s.t. \quad ||\boldsymbol{b}_j||_0 \leq K_2,\, \mathbf{D}^y \in \mathcal{D}_2,$$

$$\mathcal{D}_1 \equiv \left\{\mathbf{D}^x \in \mathbb{R}^{d\times n} : \|\boldsymbol{d}^x_i\|_2 \leq 1, \forall i = 1\cdots n\right\}. \qquad \mathcal{D}_2 \equiv \left\{\mathbf{D}^y \in \mathbb{R}^{d\times m} : \left\|\boldsymbol{d}^y_j\right\|_2 \leq 1, \forall j = 1\cdots m\right\}.$$

- **Discrete distributions of dictionaries:**

$$\alpha_i = \sum_{k=1}^{N} \mathbf{A}_{i,k}$$

$$\boldsymbol{\mu} := \frac{1}{|\alpha|}\sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{d}^x_i}$$



$$\beta_j = \sum_{k=1}^{M} \mathbf{B}_{j,k}$$

$$\boldsymbol{\nu} := \frac{1}{|\beta|}\sum_{j=1}^{m} \beta_j \delta_{\boldsymbol{d}^y_j}$$

Problem Formulation

- **Optimal Transport over Sparse Dictionaries:**

Let $\boldsymbol{\mu} := \frac{1}{|\alpha|}\sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{d}_i^x}$ and $\boldsymbol{\nu} := \frac{1}{|\beta|}\sum_{j=1}^{m} \beta_j \delta_{\boldsymbol{d}_j^y}$ be probability distributions of dictionaries. Then, the Kantorovich OT holds,

$$\boldsymbol{L}_{\mathbf{C}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \stackrel{\text{def}}{=} \min_{\mathbf{T} \in U(\boldsymbol{\mu}, \boldsymbol{\nu})} \langle \mathbf{C}, \mathbf{T} \rangle \stackrel{\text{def}}{=} \sum_{i,j} \mathbf{C}_{i,j} \mathbf{T}_{i,j}, \tag{5}$$

where $\mathbf{C}_{i,j} \stackrel{\text{def}}{=} \mathbf{C}(\boldsymbol{d}_i^x, \boldsymbol{d}_j^y)$ is the ground cost to move atom $\boldsymbol{d}_i^x$ to $\boldsymbol{d}_j^y$, and the map $\mathbf{T}$ satisfies

$$U(\boldsymbol{\mu}, \boldsymbol{\nu}) \stackrel{\text{def}}{=} \{\mathbf{T} \in R_{+}^{n \times m} | \mathbf{T}\mathbf{1}_m = \boldsymbol{\mu}, \mathbf{T}^\top \mathbf{1}_n = \boldsymbol{\nu}\}, \tag{6}$$

Problem Formulation

Definition (OT over Sparse Dictionaries)

Let $\boldsymbol{\mu} := \frac{1}{|\alpha|}\sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{d}_i^x}$, $\boldsymbol{\nu} := \frac{1}{|\beta|}\sum_{j=1}^{m} b_j \delta_{\boldsymbol{d}_j^y}$. The Kantorovich OT problem over dictionaries, parameterized by $\mathbf{T} \in \mathbb{R}_+^{n\times m}$, is formulated as

$$
\begin{aligned}
\min_{\mathbf{T}} \quad & \langle \mathbf{C}, \mathbf{T} \rangle := \sum_{i,j} \mathbf{C}_{ij}\mathbf{T}_{ij} \\
\text{s.t.} \quad & \mathbf{D}^x\mathbf{A} = \mathbf{X}, \quad \|\boldsymbol{a}_i\|_0 \le K_1, \quad \Longleftarrow \text{Sparse coding for } \mathbf{X} \\
& \mathbf{D}^y\mathbf{B} = \mathbf{Y}, \quad \|\boldsymbol{b}_j\|_0 \le K_2, \quad \Longleftarrow \text{Sparse coding for } \mathbf{Y} \\
& \mathbf{A}\mathbf{1}_N = \boldsymbol{\alpha}, \quad \mathbf{B}\mathbf{1}_M = \boldsymbol{\beta}, \quad \Longleftarrow \text{Probabilities of } \mathbf{D}^x \text{ and } \mathbf{D}^y \\
& \mathbf{T}\mathbf{1}_m = \boldsymbol{\mu}, \quad \mathbf{T}^\top\mathbf{1}_n = \boldsymbol{\nu}, \quad \Longleftarrow \text{OT constraints}
\end{aligned}
\tag{OT1}
$$

where $\mathbf{C}_{ij} := C(\boldsymbol{d}_i^x, \boldsymbol{d}_j^y)$ is the cost of moving atom $\boldsymbol{d}_i^x$ to atom $\boldsymbol{d}_j^y$.

Outline

Model Relaxation and Solution

Definition (Relaxed Form)

By multiplying both sides of the third-line constraints in problem (OT1) by $\mathbf{D}^x$ and $\mathbf{D}^y$, respectively, one has the following relaxed form,

$$\begin{aligned}
&\min_{\mathbf{T}} \langle \mathbf{C}, \mathbf{T} \rangle \stackrel{\text{def}}{=} \min_{\mathbf{T}} \sum_{i,j} \mathbf{C}_{i,j} \mathbf{T}_{i,j} \\
s.t. \quad & \mathbf{D}^x \mathbf{A} = \mathbf{X}, \quad ||\boldsymbol{a}_i||_0 \leq K_1, \quad \Longleftarrow \text{Sparse coding for } X \\
& \mathbf{D}^y \mathbf{B} = \mathbf{Y}, \quad ||\boldsymbol{b}_j||_0 \leq K_2, \quad \Longleftarrow \text{Sparse coding for } Y \qquad \text{(OT2)} \\
& \color{red}{\mathbf{D}^x \boldsymbol{\alpha} = \mathbf{X}\mathbf{1}_N}, \quad \color{red}{\mathbf{D}^y \boldsymbol{\beta} = \mathbf{Y}\mathbf{1}_M}, \quad \Longleftarrow \text{Probability of } \mathbf{D}^x, \mathbf{D}^y \\
& \mathbf{T}\mathbf{1}_m = \boldsymbol{\mu}, \quad \mathbf{T}^\top \mathbf{1}_n = \boldsymbol{\nu}, \quad \Longleftarrow \text{OT constraints}
\end{aligned}$$

where $\mathbf{C}_{\boldsymbol{i},\boldsymbol{j}} \stackrel{\text{def}}{=} \mathbf{C}(\boldsymbol{d}_i^x, \boldsymbol{d}_j^y)$ is the cost function to move atom $\boldsymbol{d}_i^x$ to $\boldsymbol{d}_j^y$.

Optimization Solution

Wasserstein distance $W_2(\alpha,\beta)$:

- Considering the $p=2$ Wasserstein distance (e.g., $\mathbf{C}_{i,j} \overset{\text{def}}{=} \|\mathbf{d}_i^x - \mathbf{d}_j^y\|_2^2$), and the relaxed OT problem can be rewritten in the following regularized form,

$$\begin{aligned}
\underset{\mathbf{T},\mathbf{A},\mathbf{B},\mathbf{D}^x,\mathbf{D}^y}{\operatorname{argmin}} \ \gamma \sum_{i,j} T_{i,j} \|\boldsymbol{d}_i^x - \boldsymbol{d}_j^y\|_2^2 \\
+\lambda_x \|\mathbf{X} - \mathbf{D}^x\mathbf{A}\|_F^2 + \lambda_y \left\|\mathbf{Y} - \mathbf{D}^y\mathbf{B}\right\|_F^2 \\
+\tau_x \left\|\mathbf{X}\mathbf{1}_M - \mathbf{D}^x\boldsymbol{\alpha}\right\|_F^2 + \tau_y \left\|\mathbf{Y}\mathbf{1}_M - \mathbf{D}^y\boldsymbol{\beta}\right\|_F^2, \qquad (7)
\end{aligned}$$

$$\text{s.t.} \quad \mathbf{T} \in \Pi(\boldsymbol{\mu},\boldsymbol{\nu}), \quad \mathbf{D}^x \in \mathcal{D}_1, \quad \mathbf{D}^y \in \mathcal{D}_2, \quad ||\boldsymbol{a}_i||_0 \le K_1, \quad ||\boldsymbol{b}_j||_0 \le K_2.$$

where γ, $\lambda_{x(y)}$ and $\tau_{x(y)}$ are the positive regularization parameters.

Alternating Minimization:

- **Step 1:** Fix $\mathbf{T}, \mathbf{D}^x, \mathbf{D}^y$, and $\boldsymbol{\alpha}, \boldsymbol{\beta}$, solving $\mathbf{A}, \mathbf{B}$ for sparse coding,

$$\min_{\boldsymbol{a}_i} ||\mathbf{X} - \mathbf{D}^x\mathbf{A}||_F^2, \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq K_1. \tag{8}$$

- **Step 2:** Fix $\mathbf{T}$ and $\mathbf{A}$ and $\mathbf{B}$ (thus $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$), solving $\mathbf{D}^x \in \mathcal{D}_1, \mathbf{D}^y \in \mathcal{D}_2$ for dictionaries learning,

$$\underset{\mathbf{D}^x \in \mathcal{D}_1}{\operatorname{argmin}} \quad \lambda_x ||\mathbf{D}^x\mathbf{A} - \mathbf{X}||_F^2 + \tau_x ||\mathbf{D}^x\boldsymbol{\alpha} - \mathbf{X}\mathbf{1}_M||_2^2 + \gamma \sum_{i,j} T_{i,j} ||\boldsymbol{d}_i^x - \boldsymbol{d}_j^y||_2^2. \tag{9}$$

- **Step 3:** Fix $\mathbf{D}^x, \mathbf{D}^y$, and $\mathbf{A}, \mathbf{B}$ (thus,$\boldsymbol{\alpha}, \boldsymbol{\beta}$), solving $\mathbf{T}$ for transport mapping,

$$\underset{\mathbf{T}}{\operatorname{argmin}} \sum_{i,j} T_{i,j} ||\boldsymbol{d}_i^x - \boldsymbol{d}_j^y||_2^2, \quad s.t. \quad \mathbf{T}\mathbf{1}_n = \boldsymbol{\mu}, \ \mathbf{T}^\top \mathbf{1}_m = \boldsymbol{\nu}. \tag{10}$$

Sparse Coding: (Orthogonal) Matching Pursuit

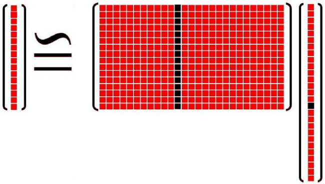
- **Step 1:** Fix $\mathbf{T}, \mathbf{D}^x, \mathbf{D}^y$, and $\boldsymbol{\alpha}, \boldsymbol{\beta}$, solving $\mathbf{A}, \mathbf{B}$ for sparse coding,

$$\min_{\boldsymbol{a}_i} ||\mathbf{X} - \mathbf{D}^x \mathbf{A}||_F^2, \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq K_1.$$

✓ For each signal $\boldsymbol{x}^i$, find the best matching $\boldsymbol{d}_k^x$,

✓ Given the found atoms $\{\boldsymbol{d}_1^x, \boldsymbol{d}_2^x, ..., \boldsymbol{d}_k^x\}$, find the next one to best fit the residual,

✓ Stop the algorithm when the criterion met.



Sparse Coding: (Orthogonal) Matching Pursuit

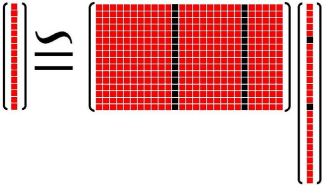
- **Step 1:** Fix $\mathbf{T}, \mathbf{D}^x, \mathbf{D}^y$, and $\boldsymbol{\alpha}, \boldsymbol{\beta}$, solving $\mathbf{A}, \mathbf{B}$ for sparse coding,

$$\min_{\boldsymbol{a}_i} ||\mathbf{X} - \mathbf{D}^x\mathbf{A}||_F^2, \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq K_1.$$

✓ For each signal $\boldsymbol{x}^i$, find the best matching $\boldsymbol{d}_k^x$,

✓ Given the found atoms $\{\boldsymbol{d}_1^x, \boldsymbol{d}_2^x, ..., \boldsymbol{d}_k^x\}$, find the next one to best fit the residual,

✓ Stop the algorithm when the criterion met.



Sparse Coding: (Orthogonal) Matching Pursuit

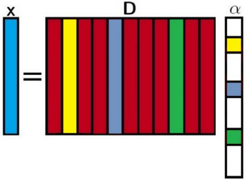
- **Step 1:** Fix $\mathbf{T}, \mathbf{D}^x, \mathbf{D}^y$, and $\boldsymbol{\alpha}, \boldsymbol{\beta}$, solving $\mathbf{A}, \mathbf{B}$ for sparse coding,

$$\min_{\boldsymbol{a}_i} ||\mathbf{X} - \mathbf{D}^x \mathbf{A}||_F^2, \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq K_1.$$

✓ For each signal $\boldsymbol{x}^i$, find the best matching $\boldsymbol{d}_k^x$,

✓ Given the found atoms $\{\boldsymbol{d}_1^x, \boldsymbol{d}_2^x, ..., \boldsymbol{d}_k^x\}$, find the next one to best fit the residual,

✓ Stop the algorithm when the criterion met.



Optimization Solution

Dictionary Update: Extended K-SVD Algorithm

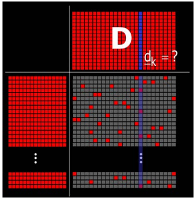
- **Step 2:** Fix $\mathbf{T}$ and $\mathbf{A}$ and $\mathbf{B}$ (thus $\boldsymbol{\alpha}$ and β), solving $\mathbf{D}^x \in \mathcal{D}_1, \mathbf{D}^y \in \mathcal{D}_2$ for dictionaries learning,

$$\underset{\mathbf{D}^x \in \mathcal{D}_1}{\operatorname{argmin}} \quad \lambda_x \|\mathbf{D}^x \mathbf{A} - \mathbf{X}\|_F^2 + \tau_x \|\mathbf{D}^x \boldsymbol{\alpha} - \mathbf{X}\mathbf{1}_M\|_2^2 + \gamma \sum_{i,j} T_{i,j} \|\boldsymbol{d}_i^x - \boldsymbol{d}_j^y\|_2^2.$$

✓ Select the examples using the column $\boldsymbol{d}_k$,

✓ Fix all $\mathbf{D}^{\mathrm{x}}$ and $\mathbf{A}$ apart from the k-th column, and seek both $\boldsymbol{d}_k^x$ and the k-th column of $\mathbf{A}$ to better fit the residual, by analogy for $\boldsymbol{d}_k^y$,

✓ Closed-form solution or solve the sub-problem with SVD algorithm.



Optimization Solution

Dictionary Update: Extended K-SVD Algorithm

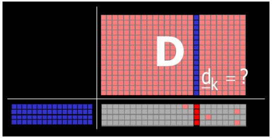
- **Step 2:** Fix $\mathbf{T}$ and $\mathbf{A}$ and $\mathbf{B}$ (thus $\boldsymbol{\alpha}$ and β), solving $\mathbf{D}^x \in \mathcal{D}_1, \mathbf{D}^y \in \mathcal{D}_2$ for dictionaries learning,

$$\underset{\boldsymbol{d}_k^x}{\operatorname{argmin}} \quad \lambda_x \left\| \mathbf{E}_R^k - \boldsymbol{d}_k^x \boldsymbol{A}_R^k \right\|_F^2 + \tau_x \left\| \mathbf{F}^k - \boldsymbol{d}_k^x \alpha_k \right\|_F^2 + \gamma \sum_j T_{k,j} \| \boldsymbol{d}_k^x - \boldsymbol{d}_j^y \|_2^2 .$$

✓ Select the examples using the column $\boldsymbol{d}_k$,

✓ Fix all $\mathbf{D^x}$ and $\mathbf{A}$ apart from the k-th column, and seek both $\boldsymbol{d}_k^x$ and the k-th column of $\mathbf{A}$ to better fit the residual, by analogy for $\boldsymbol{d}_k^y$,

✓ Closed-form solution or solve the sub-problem with SVD algorithm.



Optimization Solution

Dictionary Update: Extended K-SVD Algorithm

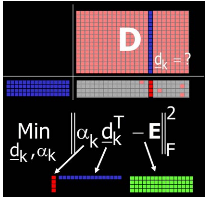
- **Step 2:** Fix $\mathbf{T}$ and $\mathbf{A}$ and $\mathbf{B}$ (thus $\boldsymbol{\alpha}$ and β), solving $\mathbf{D}^x \in \mathcal{D}_1, \mathbf{D}^y \in \mathcal{D}_2$ for dictionaries learning,

$$\boldsymbol{d}_k^x = \frac{\lambda_x \mathbf{E}_R^k (\boldsymbol{A}_R^k)^T + \tau_x \alpha^k \mathbf{F}^k + \gamma \sum_j (T_{k,j} \boldsymbol{d}_j^y)}{\lambda_x \boldsymbol{A}^k (\boldsymbol{A}_R^k)^T + \tau_x (\alpha^k)^2 + \gamma \sum_j T_{k,j}}.$$

✓ Select the examples using the column $\boldsymbol{d}_k$,

✓ Fix all $\mathbf{D^x}$ and $\mathbf{A}$ apart from the k-th column, and seek both $\boldsymbol{d}_k^x$ and the k-th column of $\mathbf{A}$ to better fit the residual, by analogy for $\boldsymbol{d}_k^y$,

✓ Closed-form solution or solve the sub-problem with SVD algorithm.



Optimal Transport over dictionaries

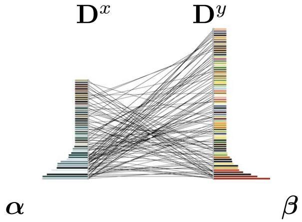
- **Step 3:** Fix $\mathbf{D}^x, \mathbf{D}^y$, and $\mathbf{A}, \mathbf{B}$ (thus,$\boldsymbol{\alpha}, \boldsymbol{\beta}$), solving $\mathbf{T}$ for transport mapping,

$$\underset{\mathbf{T}}{\operatorname{argmin}} \sum_{i,j} T_{i,j} \|\boldsymbol{d}_i^x - \boldsymbol{d}_j^y\|_2^2,$$
$$s.t. \quad \mathbf{T}\mathbf{1}_n = \boldsymbol{\mu}, \ \mathbf{T}^\top \mathbf{1}_m = \boldsymbol{\nu}.$$

✓ Standard optimal transport problem,

✓ Linear programming (small-scale problem),

✓ Efficient solver.



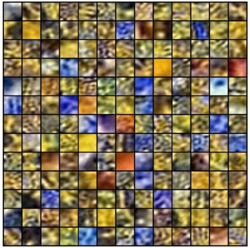
Outline

Experimental Results: Color Transform

Learned Dictionaries:

Trey Ratcliff

Content image



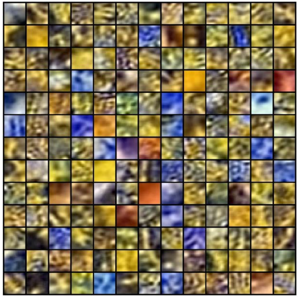
H., Hristina et al.

Style image

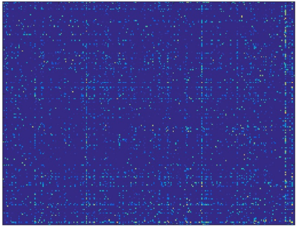
Learned dictionaries of different images

Experimental Results: Color Transform

Transport Mapping:



Content Dictionary



Optimal $\Longleftrightarrow$ Transfer

Style Dictionary

Transport mapping over learned dictionaries

Reconstructed Result:

Trey Ratcliff

Content image

Transferred image

H., Hristina et al.

Style image

The proposed image color transform effect.

Experimental Results: Color Transform

Visual Comparison:

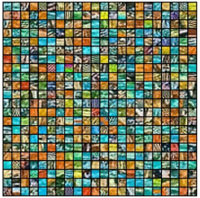
Content Image | MKL [12] | ROT [5] | Our Results | Reference Image

Optimal image color transform over sparse dictionaries.

Learned Dictionaries:

Arturo Pacheco Altamirano

Content image



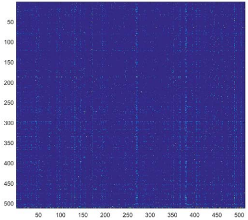
Franz Richard Unterberger

Style image

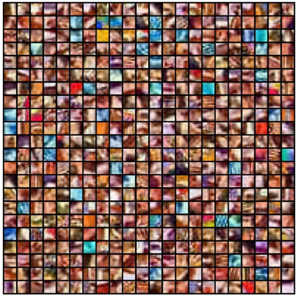
Learned dictionaries of two different images

Transport Mapping:

Content Dictionary



Optimal $\Longleftrightarrow$ Transfer



Style Dictionary

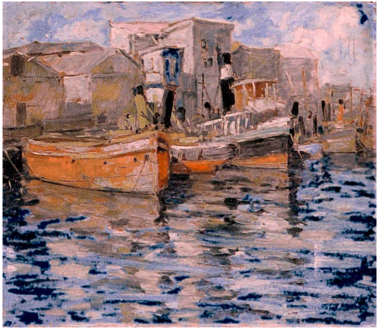
Optimal transport mapping over learned dictionaries

Experimental Results: Photorealistic Style Transfer

Reconstruction Result:

Arturo Pacheco Altamirano

Content image



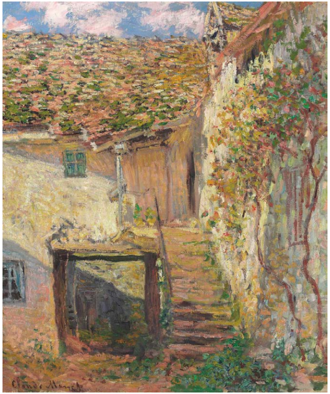
Transferred image

Franz Richard Unterberger

Style image

Optimal image style transfer over sparse dictionaries.

Experimental Results: Photorealistic Style Transfer



Content image

Our Results

Reference Image

Optimal image style transfer over sparse dictionaries.

Experimental Results: Photorealistic Style Transfer

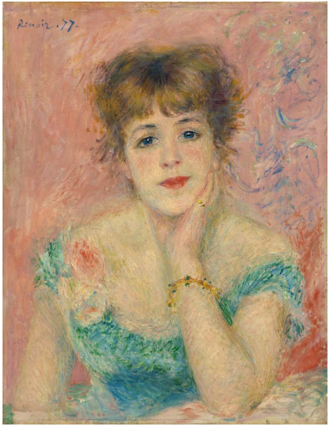
Content image

Our Results

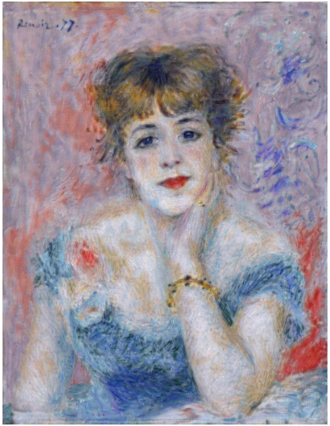
Reference Image

Optimal image style transfer over sparse dictionaries.

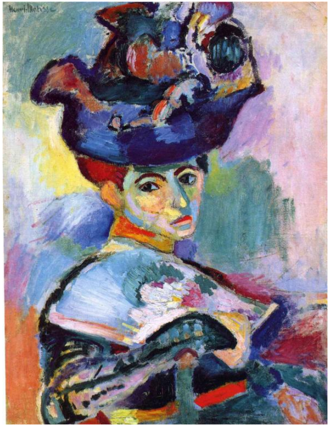
Experimental Results: Photorealistic Style Transfer



Content image



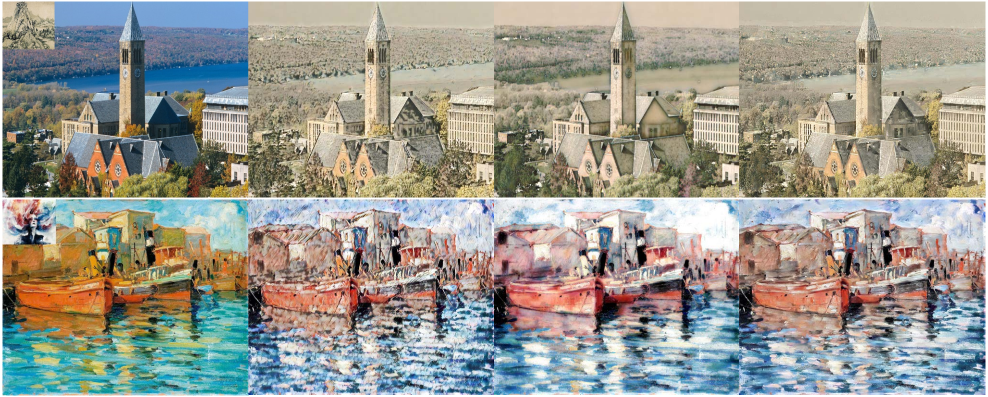
Our Results



Reference Image

Optimal image style transfer over sparse dictionaries.

Experimental Results: Photorealistic Style Transfer

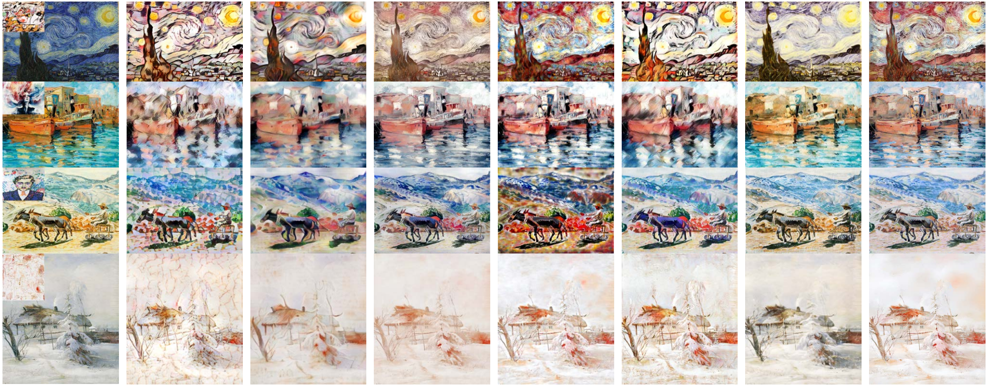


Input AdaIN [8] WCT2 [17] Our Results

Comparison of photorealistic style transfer with deep learning methods.

Outline

Experimental Results: Visual Comparison



Input AdaIN [8] WCT [10] PWCT [11] WCT^2 [17] $StyTr^2$ [3] QuantArt [7] Ours

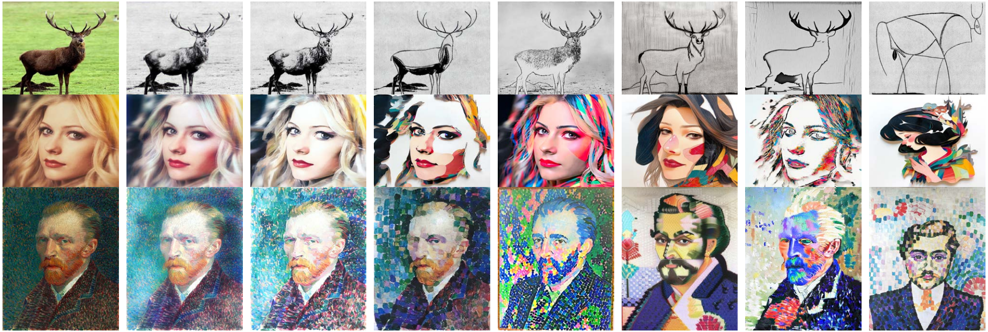
Comparison of photorealistic style transfer with deep learning methods.

Experimental Results: Quantitative Evaluation

Table: Quantitative evaluation of image style transfer. The top two results are highlighted in **bold** and <u>underlined</u>, respectively.

Methods	Metrics		AdaIN[8]	WCT[10]	PWCT[11]	WCT^2[17]	StyTr^2[3]	QuantArt[7]	Ours
Content Fidelity	SSIM (edge)	↑	0.5785	0.6174	0.6773	0.6160	0.4944	<u>0.7590</u>	**0.7934**
	IIT loss	↓	61.35	44.82	35.34	<u>34.51</u>	50.02	41.38	**32.87**
	LPIPS loss	↓	0.5260	0.5645	<u>0.2284</u>	0.2885	0.4532	0.4257	**0.2146**
Style Fidelity	Gram loss	↓	1.64	1.87	1.45	**1.28**	1.51	2.17	<u>1.37</u>
Realisticness	FID metric	↓	278.45	272.57	153.81	152.49	246.75	**146.50**	<u>152.20</u>
User Study	Color Fidelity	↓	2.59	4.82	3.97	3.04	<u>1.93</u>	3.43	**1.81**
	Structural Fidelity	↓	4.94	4.16	3.21	2.83	3.48	<u>2.01</u>	**1.40**
	Aesthetics	↓	3.55	4.87	3.75	3.12	**1.31**	2.32	<u>2.02</u>

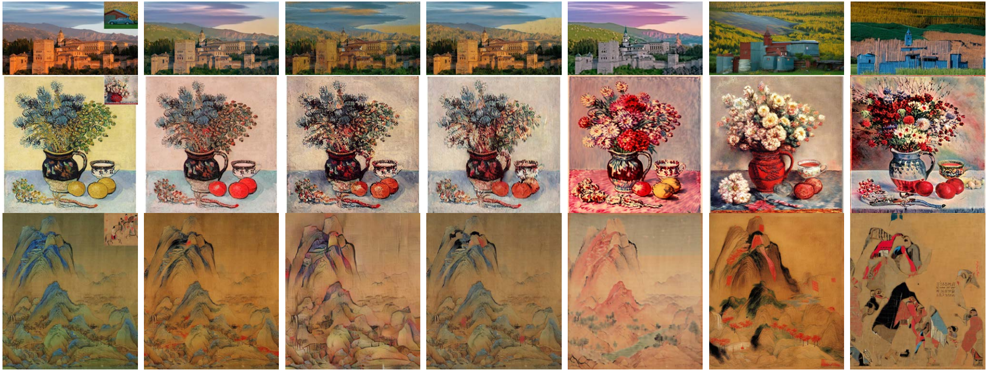
Visual Comparison:



Contents | Ours | StyleID [2] | AD [19] | CSGO [15] | SADis [13] | Styleshot [6] | References

Visual comparison with recent diffusion–based methods.

Visual Comparison:



Cont./Ref. Ours StyleID [2] AD [19] CSGO [15] SADis [13] Styleshot [6]

Visual comparison with recent diffusion–based methods.

Outline

It holds $\mathbf{A} = \mathbf{B}$ (also $\boldsymbol{\alpha} = \boldsymbol{\beta}$), thus $\mathbf{T} = \mathbf{I}$ and the problem reduces into

$$\begin{aligned} \underset{\mathbf{A},\mathbf{D}^x,\mathbf{D}^y}{\operatorname{argmin}} \; & \lambda_x \left\|\mathbf{X}-\mathbf{D}^x\mathbf{A}\right\|_F^2+\lambda_y \left\|\mathbf{Y}-\mathbf{D}^y\mathbf{A}\right\|_F^2+\gamma \sum_i \left\|\boldsymbol{d}_i^x-\boldsymbol{d}_i^y\right\|_2^2 \\ & +\tau_x \left\|\mathbf{X}\mathbf{1}_M-\mathbf{D}^x\boldsymbol{\alpha}\right\|_F^2+\tau_y \left\|\mathbf{Y}\mathbf{1}_M-\mathbf{D}^y\boldsymbol{\alpha}\right\|_F^2, \\ \text{s.t.} \quad & \mathbf{D}^x \in \mathcal{D}_1, \quad \mathbf{D}^y \in \mathcal{D}_2, \quad \left\|\boldsymbol{a}_i\right\|_0 \le K. \end{aligned} \tag{11}$$

HR Image | Bicubic | ScSR [16] | SCDL [14] | SRCNN [4] | SRGAN [9] | Ours

Visual comparison of image super-resolution between our method and representative methods.

More Extensions: Image Super-Resolution

Table: Comparison results of image super-resolution for scaling factor $s = 3$ on Set5 [1] and Set14 [18]. The top two results are highlighted in **bold** and <u>underlined</u>, respectively.

Dataset	Index	Bicubic	ScSR[16]	SCDL[14]	SRCNN[4]	SRGAN[9]	Ours
Set5	PSNR	30.39	31.42	31.59	**32.75**	31.68	<u>31.89</u>
	SSIM	0.8682	0.8821	0.8904	**0.9090**	0.8581	<u>0.9018</u>
Set14	PSNR	27.55	28.31	28.64	**29.28**	28.52	<u>28.93</u>
	SSIM	0.7742	0.7954	0.8063	**0.8209**	0.8057	<u>0.8104</u>

Note: Training is performed on a paired datasets, rather than on a signal paired images.

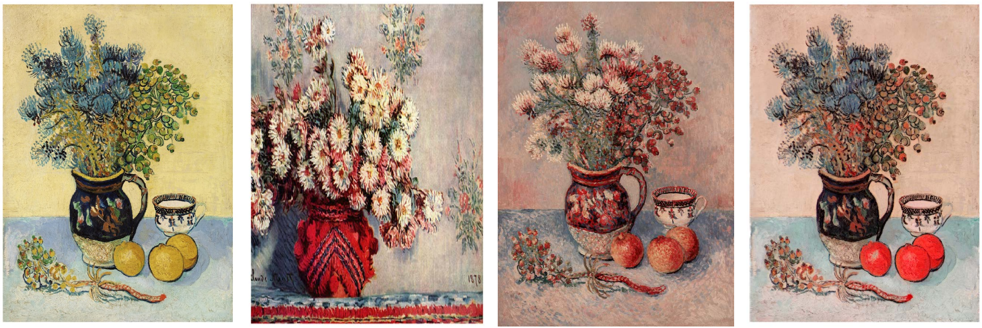
More Extensions: Video Style Transfer

Input Stylization 1 Stylization 2

Video style transfer effects under the specified reference images.

Note: Training is performed on multiple video frames and a reference image, rather than on a signal paired images.

The ChatGPT Effects



Input | Reference | ChatGPT | Ours

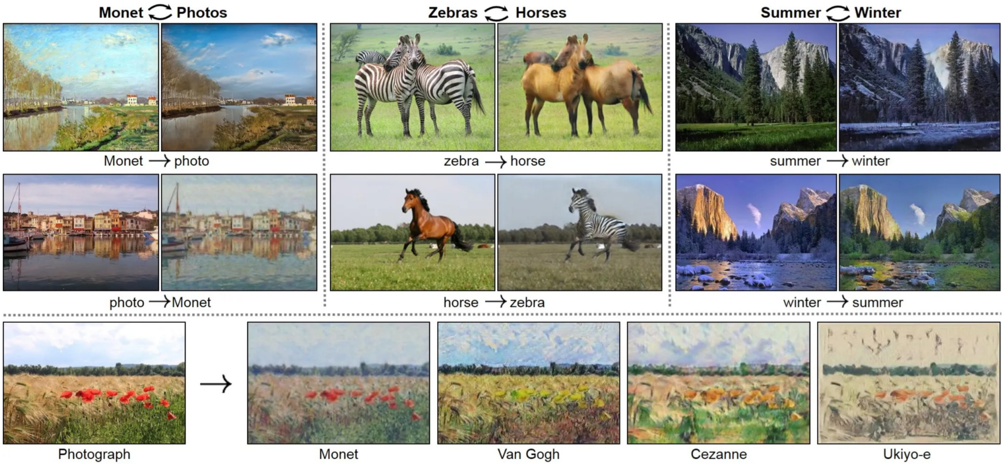
Visual comparison of the ChatGPT stylization and our method.

The ChatGPT Effects

Content Reference ChatGPT Ours

Visual comparison of the ChatGPT stylization and the proposed method.

Future Work: Image-to-Image Translation



Y. Zhu et al.

Future Work

Simultaneous Signal/Image Representation and Transformation

Sparse Representation $\Longleftrightarrow$ Optimal Transport

Signal/Image Representation

- Sparse representation (established)
- Non-negative sparse representation
- Convolutional sparse representation
- NN-based representation (encoder–decoder, diffusion, flow matching models ...)

Optimal Transport

- Vanilla OT (established)
- Sliced OT algorithm
- Entropy-regularized OT
- Other efficient OT variants

More expressive representation learning and efficient transport.

References I

[1] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi-Morel.
Low-complexity single-image super-resolution based on nonnegative neighbor embedding.
In *Proceedings of the British Machine Vision Conference*, pages 135.1–135.10. BMVA Press, 2012.

[2] J. Chung, S. Hyun, and J.-P. Heo.
Style injection in diffusion: A training-free approach for adapting large-scale diffusion models for style transfer.
arXiv preprint arXiv:2312.09008, 2024.

[3] Y. Deng, F. Tang, W. Dong, C. Ma, X. Pan, L. Wang, and C. Xu.
Stytr2: Image style transfer with transformers.
In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11326–11336, 2022.

[4] C. Dong, C. C. Loy, K. He, and X. Tang.
Image super-resolution using deep convolutional networks.
IEEE transactions on pattern analysis and machine intelligence, 38(2):295–307, 2015.

[5] S. Ferradans, N. Papadakis, G. Peyré, and J.-F. Aujol.
Regularized discrete optimal transport.
SIAM Journal on Imaging Sciences, 7(3):1853–1882, 2014.

References II

[6] J. Gao, Y. Liu, Y. Sun, Y. Tang, Y. Zeng, K. Zhang, and F. Zhao.
Styleshot: A snapshot on any style.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025.

[7] X. Huang, J. An, S. Wei, J. Deng, M. Yang, X. Xie, and J. Wang.
Quantart: Quantizing image style transfer towards high visual fidelity.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5947–5956, 2023.

[8] X. Huang and S. Belongie.
Arbitrary style transfer in real-time with adaptive instance normalization.
In *Proceedings of the IEEE international conference on computer vision*, pages 1501–1510, 2017.

[9] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al.
Photo-realistic single image super-resolution using a generative adversarial network.
In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[10] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang.
Universal style transfer via feature transforms.
Advances in neural information processing systems, 30, 2017.

References III

[11] Y. Li, M.-Y. Liu, X. Li, M.-H. Yang, and J. Kautz.
A closed-form solution to photorealistic image stylization.
In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 453–468, 2018.

[12] F. Pitié and A. Kokaram.
The linear monge-kantorovitch linear colour mapping for example-based colour transfer.
2007.

[13] J. Qin, J. Wu, Y. Chen, X. Li, C. Liu, X. Xiao, R. Wang, M. Zheng, and S. Pan.
Freecustom: Tuning-free customized image generation for multi-concept composition.
In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 9089–9099, 2025.

[14] S. Wang, L. Zhang, Y. Liang, and Q. Pan.
Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis.
In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2216–2223. IEEE, 2012.

References IV

[15] P. Xing, J. Zhu, H. Liu, Z. Cheng, C. Huang, Y. Xu, X. Zhu, and G. Shi.
Csgo: Content-style composition in text-to-image generation.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1655–1665, 2024.

[16] J. Yang, J. Wright, T. S. Huang, and Y. Ma.
Image super-resolution via sparse representation.
In *IEEE transactions on image processing*, volume 19, pages 2861–2873. IEEE, 2010.

[17] J. Yoo, Y. Uh, S. Chun, B. Kang, and J.-W. Ha.
Photorealistic style transfer via wavelet transforms.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9036–9045, 2019.

[18] R. Zeyde, M. Elad, and M. Protter.
On single image scale-up using sparse-representations.
In *Curves and Surfaces: 7th International Conference, Avignon, France, June 24-30, 2010, Revised Selected Papers 7*, pages 711–730. Springer, 2012.

[19] Y. Zhou, O. Makansi, M. Oquab, C. Buciluǎ, A. Kuznetsova, D. Chen, C. Couprie, Y. LeCun, and J. Verbeek.
Attention distillation: A unified approach to visual characteristics transfer.
arXiv preprint arXiv:2502.20235, 2025.

Thank you for your attention !