

視覚的顕著性の一貫性を考慮した色覚異常のための画像再配色

岡山県立大学大学院 中釜 匠海

◆ 症状

- 通常と異なる色の見え方
- 色を識別する能力が低下

◆ 原因

- 一部の錐体を持たない or 分光感度のずれ

◆ タイプ

- P型 L錐体に異常
 - D型 M錐体に異常
 - T型 S錐体に異常(T型の発生は非常に稀)
- } 赤緑色弱

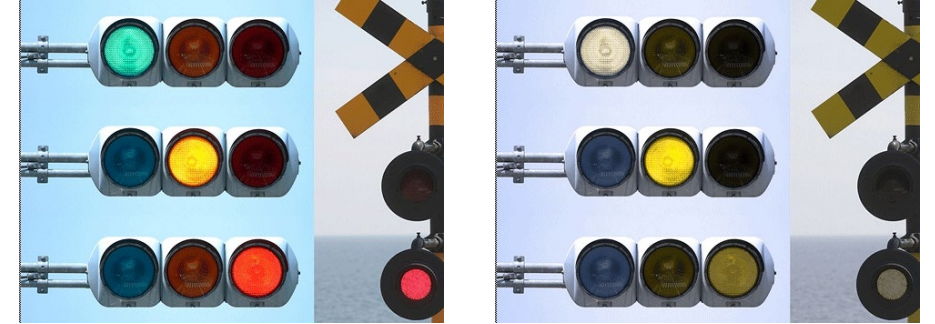


正常色覚者の見え



色覚異常者の見え

- 色覚異常者は色を識別する能力が低下
- 色の差により表現される情報を見落としてしまう
- 現在のところ治療法はない



色覚異常者の見え



色覚異常者の色の識別性を補うための画像再配色手法



CVD原画像



原画像



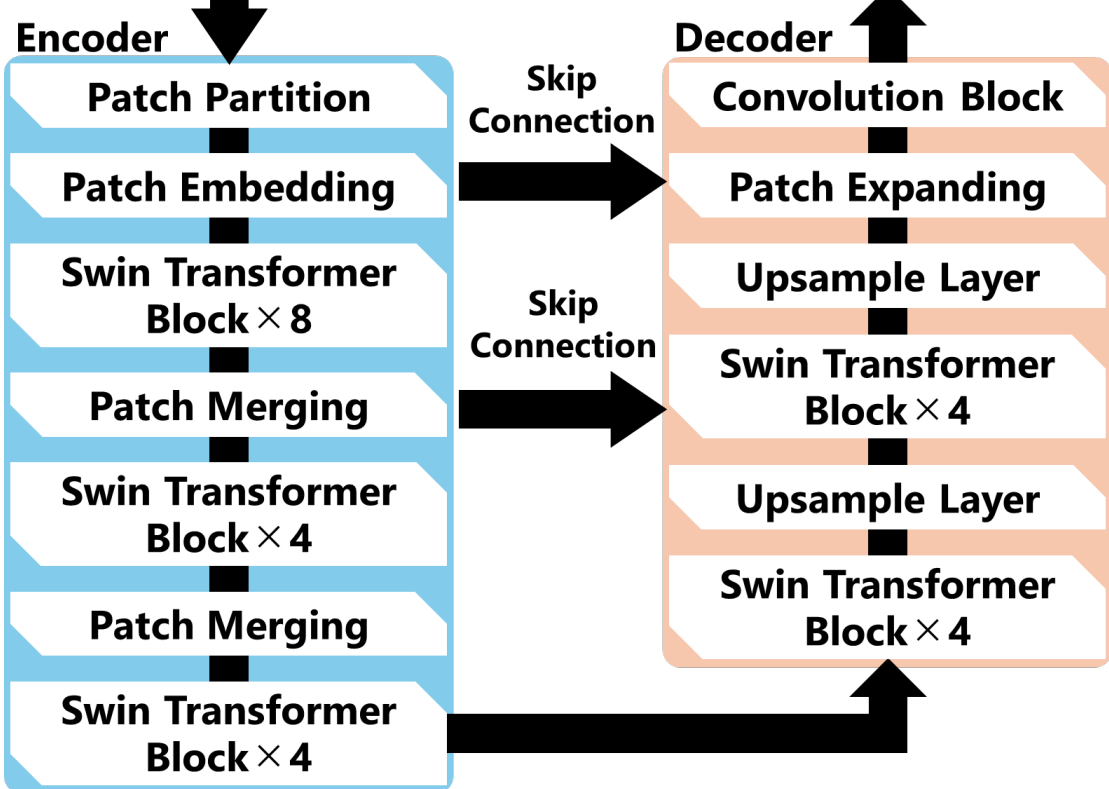
再配色画像



CVD再配色画像

◆ Image recoloring for color vision deficiency compensation using Swin transformer [1]

入力画像 CVDの見え 再配色画像



目的：コントラスト強調と自然さ保持の両立

- **コントラスト強調**
色覚異常者が知覚する色の識別性を向上
- **自然さ保持**
元の画像の色からの逸脱を抑制
- 特徴抽出器としてSwin transformerを使用した再配色モデル**CVD-swin**を提案
- コントラスト損失 L_C と自然さ損失 L_N を設計し再配色モデルを教師無し学習

◆ 再配色前後の視覚的顕著性の一貫性の保持を考慮していない

● 視覚的顕著性 (visual saliency)

- どこが視覚的に注目を集めやすいかを表す指標
- 日常生活において重要な領域であることが多い
- 顕著性推定モデルで推定されるSaliency mapで可視化



Saliency map

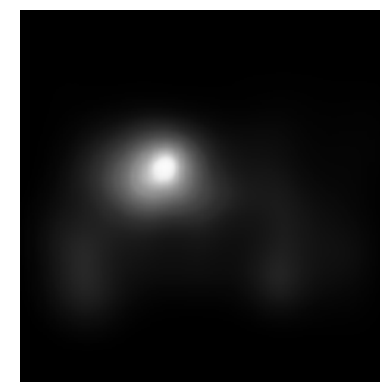
● 入力画像に対する色覚正常者と再配色画像に対する色覚異常者の視覚的顕著性が変化



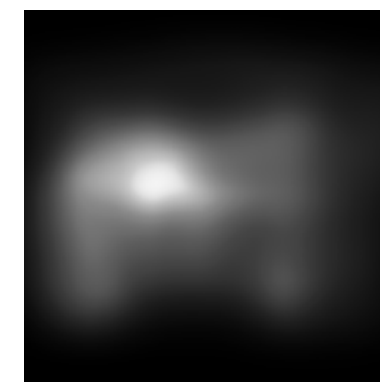
入力画像



CVD再配色画像



入力画像



CVD再配色画像

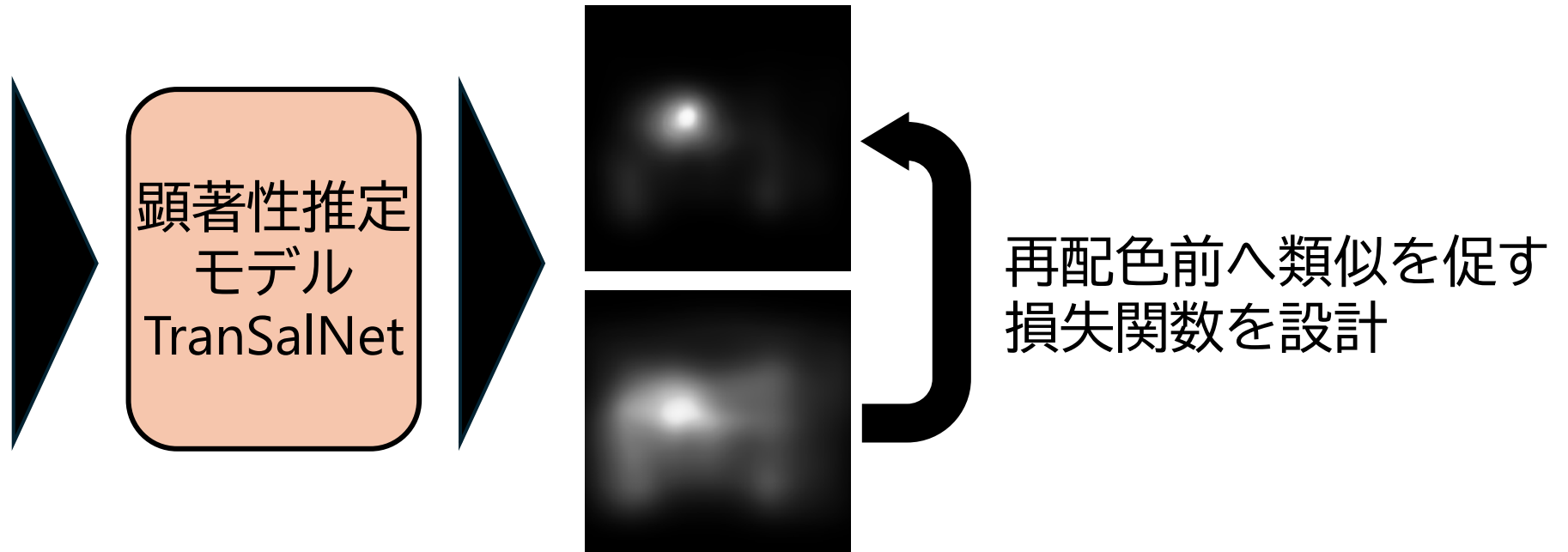
- ◆ 視覚的顕著性を考慮した色覚異常のための画像再配色手法の提案
 - 再配色モデルCVD-swinに顕著性推定モデルを統合
 - 既存の自然さ損失とコントラスト損失に加え, 新たに顕著性損失を設計
 - 再配色前後の視覚的顕著性の一貫性を保持

入力画像



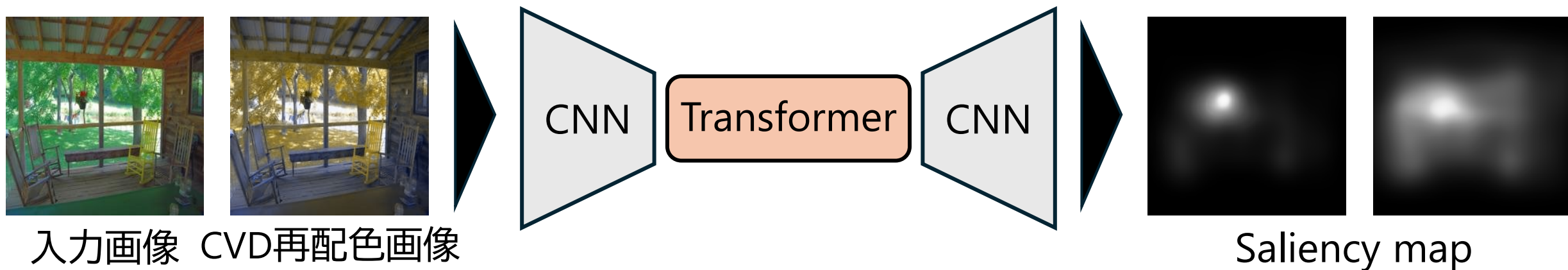
CVD

再配色画像



◆ TranSalNet(2022) [2]

- 顕著性推定の分野において高精度な推定モデル
- 特徴抽出器にCNNとTransformerを採用
- 色覚正常者の視線データを真値として学習した色覚正常者用の顕著性推定モデル
色覚異常者用の顕著性推定モデルは存在しない
CVDシミュレーション後の画像を入力することで視覚特性の差異を補完



目的：入力画像とCVD再配色画像間の視覚的顕著性の一貫性を保つ

- **Saliency L1 loss**

$$L_{SL} = \frac{1}{N} \sum_{i=1}^N |S(I_i) - S(O'_i)|$$

$S(I_i)$ ：入力画像のsaliency mapの画素*i*

$S(O'_i)$ ：CVD再配色画像のsaliency mapの画素*i*

N ：saliency mapの全画素

- **Saliency composite loss**

$$L_{SC} = (-2.0 \times CC) + (-1.0 \times SIM) + 10.0 \times KLD$$

CC, SIM, KLD ：2つのsaliency map間の分布類似度の評価指標

- **最終的な損失関数**

$$L = \alpha L_C + \beta L_N + \gamma L_S \quad \alpha, \beta, \gamma : \text{重み}$$

目的：再配色による視覚的顕著性の一貫性, 自然さ保持, コントラスト強調を評価

- 比較する再配色モデル

- 既存手法を用いて学習したCVD-Swin
コントラスト損失($\alpha = 0.5$), 自然さ損失($\beta = 0.5$)
- 提案手法を用いて学習したCVD-Swin
既存手法 + 顕著性損失 L_{SL} or L_{SC} (γ を変更)

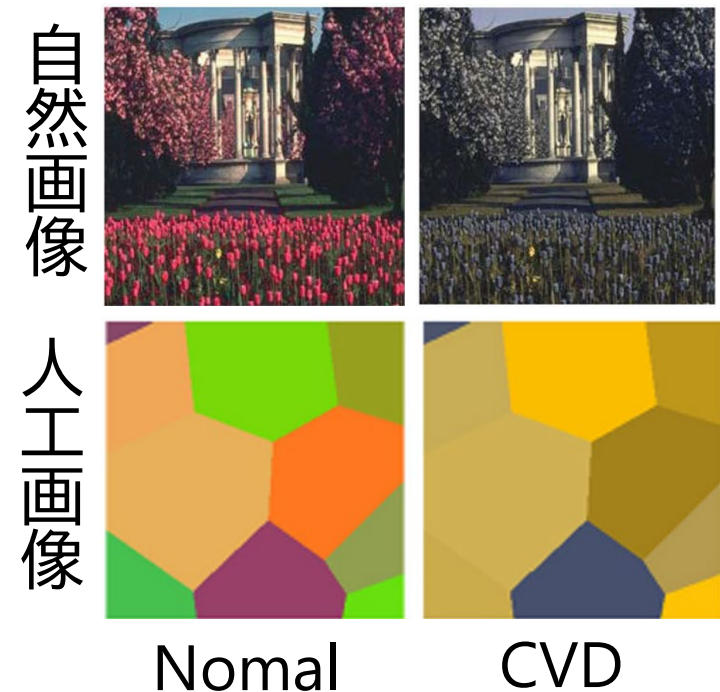
- 色覚異常：D型

- データセット

色覚異常の識別不能色を含むデータセット

D型色覚者用の画像を使用

学習用：80,000枚 評価用：10,000枚

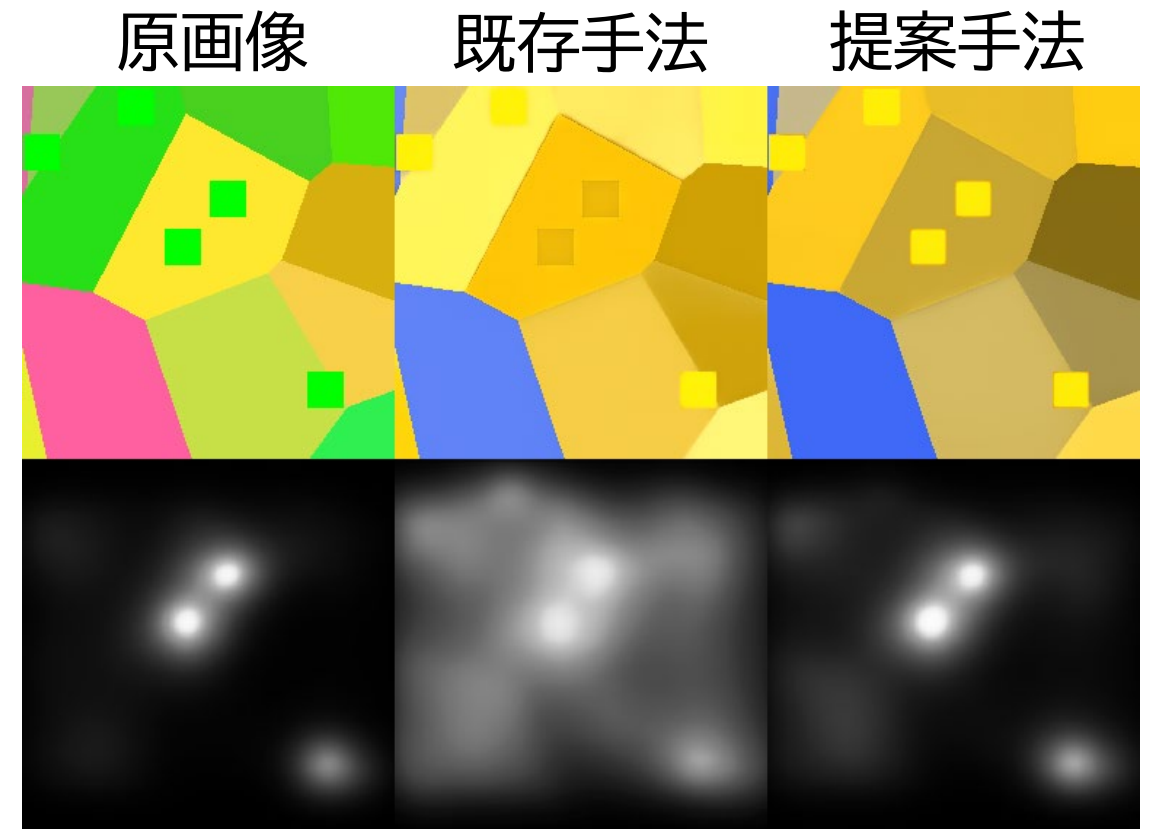
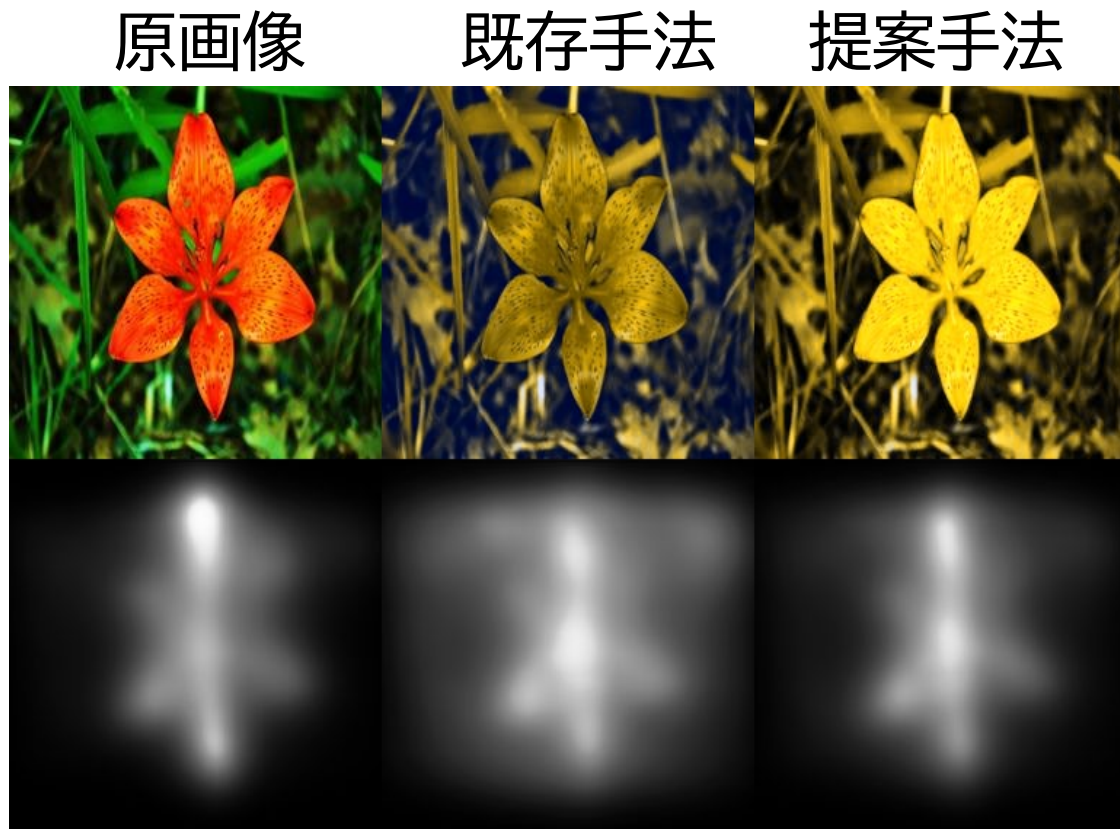


- 視覚的顕著性の一貫性の評価指標（入力画像とCVD再配色画像のsaliency map）
 - CC**(↑) : **C**orrelation **C**oefficient
2つのマップの線形的な相関を測定
 - SIM**(↑) : **SIM**ilarity
2つの確率分布化されたマップ間の重なりを測定
 - KLD**(↓) : **K**ullback-**L**eibler **D**ivergence
正解分布から予測分布への情報量の逸脱度を測定
- コントラスト強調の評価指標（入力画像とCVD再配色画像）
 - TCC**(↑) : **T**otal **C**olor **C**ontrast
2画像間のコントラスト値の比率
- 自然さ保持の評価指標（入力画像と再配色画像）
 - CD**(↓) : **C**hromatic **D**istance
2画像間のLab色空間における距離

	CC (↑)	SIM (↑)	KLD (↓)	TCC (↑)	CD (↓)
既存手法 $\gamma=0$	0.9630	0.9086	0.0693	0.8719	14.14
$L_{SL}, \gamma=2.0$	0.9682	0.9149	0.0609	0.8857	16.74
$L_{SL}, \gamma=2.5$	0.9685	0.9150	0.0612	0.8795	16.62
$L_{SC}, \gamma=0.1$	0.9665	0.9123	0.0680	0.8545	11.82
$L_{SC}, \gamma=0.5$	0.9670	0.9127	0.0674	0.8568	11.93

- CC, SIM, KLDの結果より, 提案手法は L_{SL}, L_{SC} ともに視覚的顕著性の一貫性が向上
- TCCの結果より, L_{SL} では既存手法と比較してコントラストを強調
- CDの結果より, L_{SC} では既存手法と比較して自然さを保持

◆ Saliency L1 Loss ($\gamma = 2.5$) を採用したモデルの実験結果の一例



- 提案手法は既存手法と比較して入力画像の視覚的顕著性を忠実に保つ傾向

- ◆ 視覚的顕著性を考慮した色覚異常のための画像再配色手法の提案
 - 提案手法
 - ・ 顕著性推定モデルを再配色モデルCVD-swinに統合
 - ・ 再配色前後の視覚的顕著性の一貫性を保持する顕著性損失を設計
 - 実験結果
 - ・ 視覚的顕著性の一貫性の向上を確認
- ◆ 今後の展望
 - 色覚異常者用の顕著性推定モデルの作成と高精度化
 - 色覚異常者用の顕著性モデルをCVD-swinに統合
 - 損失関数の検討

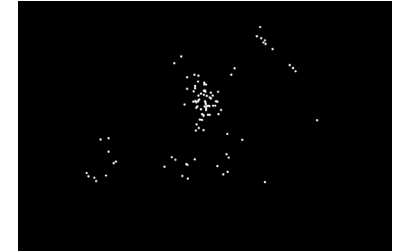
- 色覚異常者用の顕著性推定モデルの問題点

色覚異常用のデータ数が少ない(100セットのみ)

通常の推定モデルには学習に10,000セット使用



原画像



注視点マップ

- 色覚異常者用の顕著性推定モデルの現状

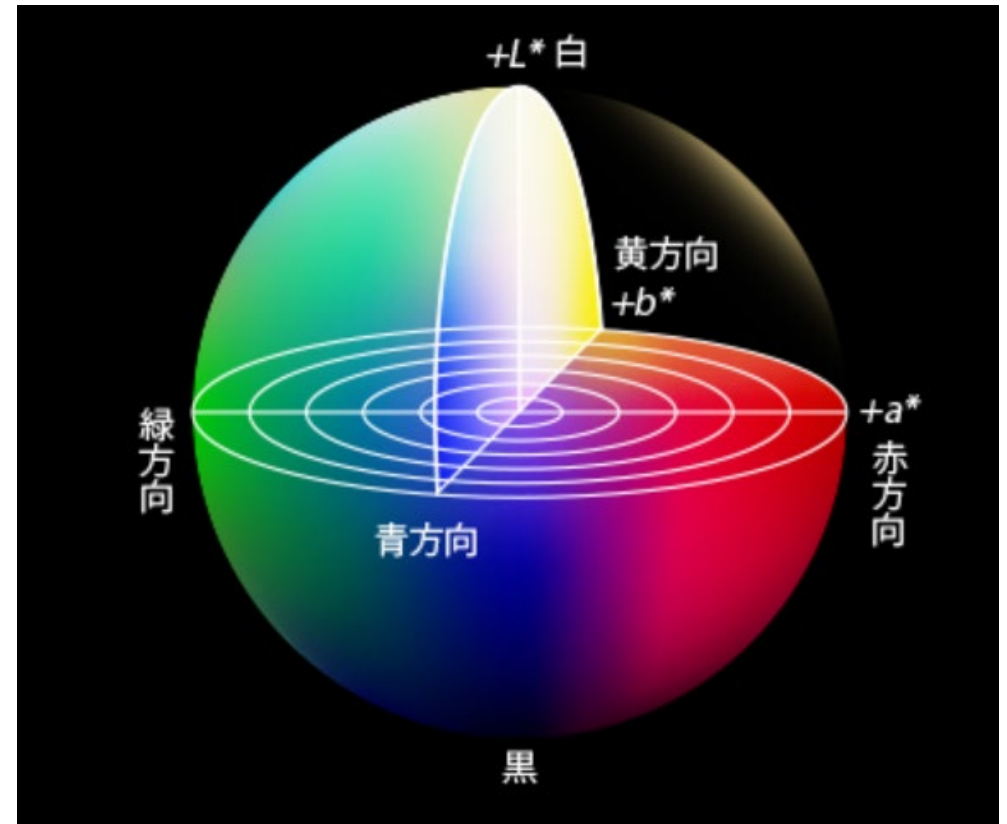
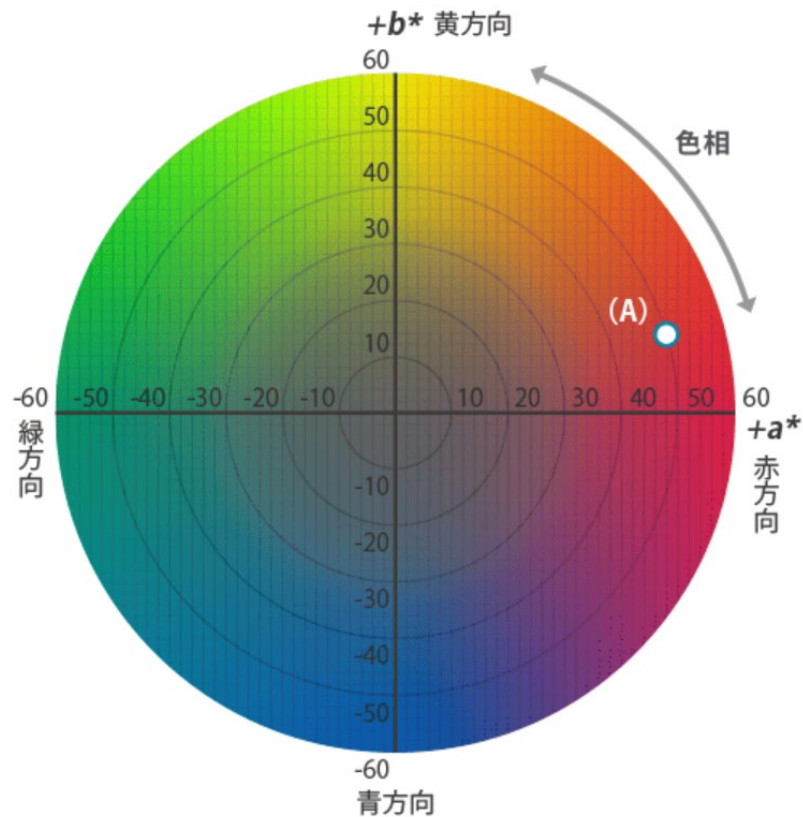
TranSalNetをFew-shot-learningを用いて少数データで学習

推定モデル	CC	SIM	KLD
TranSalNet	0.662	0.579	0.691

- 上記の推定モデルを統合した再配色モデルCVD-swinの実験結果

	CC	SIM	KLD
既存手法	0.810	0.738	0.549
$L_{SL}, \gamma = 0.1$	0.815	0.743	0.499
$L_{SC}, \gamma = 0.1$	0.813	0.742	0.536

- 物体の色を表すために最もよく利用されている表色系
- 明度を L^* , 色相と彩度を表す色度を a^* , b^* で表す



◆ 目的

● コントラスト損失

CVD再配色画像 o' のコントラストを入力画像 I と同レベルまで強調

● 自然さ損失

再配色による見え方の変化を最小限に抑える

◆ 画像間のコントラスト差の計算式

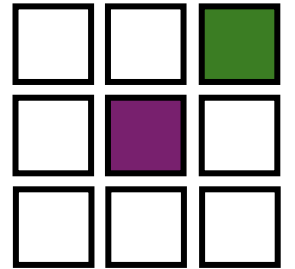
入力画像 I とCVD再配色画像 o' の相同画素間の色差の変化

$$CL(x, y) = \left| \left| \hat{c}'_x - \hat{c}'_y \right| - \left| c_x - c_y \right| \right|$$

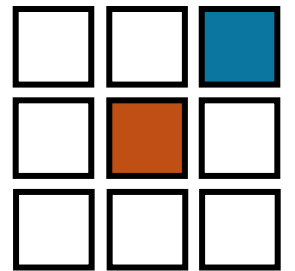
$c_x, c_y, \hat{c}'_x, \hat{c}'_y$: 入力画像と再配色画像(異)の相同画素

$|\cdot|$: ベクトルの L^1

入力画像



CVD再配色画像



● ローカルコントラスト損失

$$L_l = \sum_{x=1}^N \sum_{y \in \omega_x} \frac{CL(x, y)}{\|\omega_x\|}$$

ω_x : 画素 x を中心とするウィンドウ内の画素集合
 $\|\omega_x\|$: ω_x の画素数
 N : 画像全体の画素数

● グローバルコントラスト損失

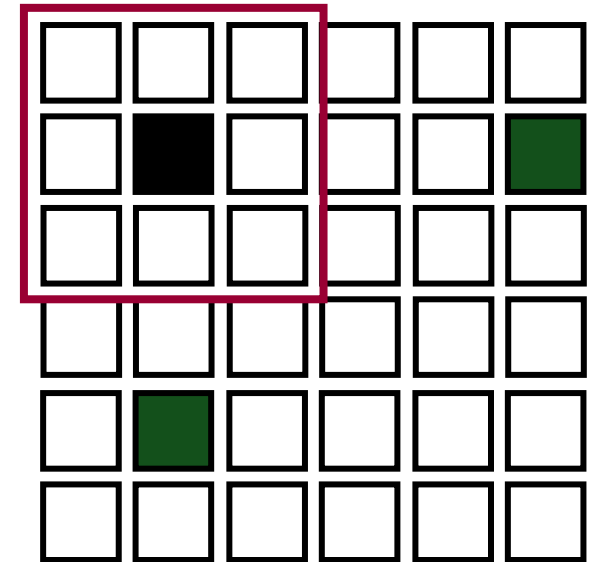
$$L_g = \frac{1}{\|\omega\|} \sum_{\langle x, y \rangle \in \omega} CL(x, y)$$

ω : ランダムに選択された画素ペアの集合
 $\|\omega\|$: ω の画素のペア数

◆ コントラスト損失

$$L_C = L_g + L_l$$

画像



● SSIM(構造的類似度)

入力画像 I と再配色画像 O に対して計算

$$SSIM(X, Y) = \frac{(2u_x u_y + c_1)(2\sigma_{xy} + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

X, Y : それぞれ画素 x, y を中心としたローカルウィンドウ

u_x, σ_x : X の画素の平均と標準偏差

u_y, σ_y : Y の画素の平均と標準偏差

σ_{xy} : X と Y の共分散

◆ 自然さ損失

$$L_N = 1 - SSIM(O, I)$$

■ 最終的な損失関数

$$L = \alpha L_C + \beta L_N$$

α, β : 重み

◆ CC (Correlation Coefficient)

- 2つのマップA, Bの線形的な相関を測定
- 1に近いほど, 2つのマップの形状や分布が類似していることを示す

$$\hat{A} = \frac{A - \mu_A}{\sigma_A} \quad CC(A, B) = \frac{\sum_i \hat{A}_i \hat{B}_i}{\sqrt{\sum_i \hat{A}_i^2 \sum_i \hat{B}_i^2}}$$

◆ SIM (SIMilarity)

- 2つの確率分布化されたマップS, G間の重なりを測定
- 1に近いほど, 2つのマップの分布がよく重なっていることを示す

$$G = \frac{GT}{\sum GT} \quad SIM(S, G) = \sum_i \min(G_i, S_i)$$

◆ KLD(Kullback-Leibler Divergence)

- 正解分布Gから予測分布Sへの情報量の差異を測定
- 0に近いほど, 2つのマップの分布が似ていることを示す

$$KLD(G||S) = \sum_i G_i \log \left(\frac{G_i}{S_i + \epsilon} \right)$$

◆ TCC(Total Color Contrast)

- 2つの画像(I, O)間のコントラスト値の比率を測定
- 1に近いほど, 2つの画像は同程度の色度コントラストを持つことを示す

$$TCC(I) = TCC_{Global} + TCC_{Local} \qquad TCC_{score} = \frac{TCC(I)}{TCC(O)}$$

◆ CD(Chromatic Distance)

- 2画像間のLab色空間上での距離を計算
- 値が小さいほど2画像間の人間が知覚する色のずれが小さいことを示す
- $\lambda = 0$ として明るさを除外した純粋な色のずれを評価

$$CD(i) = \sqrt{\lambda(l'_i - l_i)^2 + (a'_i - a_i)^2 + (b'_i - b_i)^2}$$

- **W-MSA (Window-based Multi-head Self-Attention)**

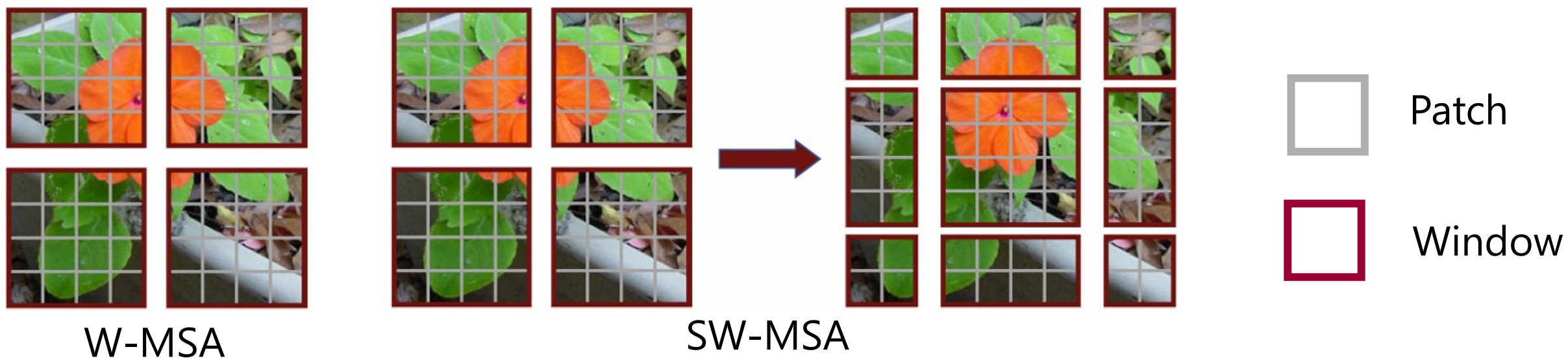
Window内のみでSelf-Attentionを計算

計算量の削減 & CNNのような局所的特徴も抽出可能

- **SW-MSA (Shifted W-MSA)**

WindowをずらしてW-MSAを実施

Window間の情報のやり取りを可能に



◆ Image recoloring for color vision deficiency compensation using Swin transformer [1]

目的：コントラスト強調と自然さ保持の両立した再配色技術の実現

特徴抽出器としてSwin transformerを使用したU-netベースの再配色モデルを提案

コントラスト損失と自然さ損失を用いて再配色モデルを教師無し学習

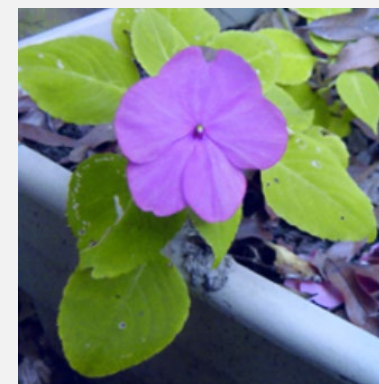


CVD
入力画像



入力画像

再配色モデル
CVD-swin



再配色画像



CVD
再配色画像

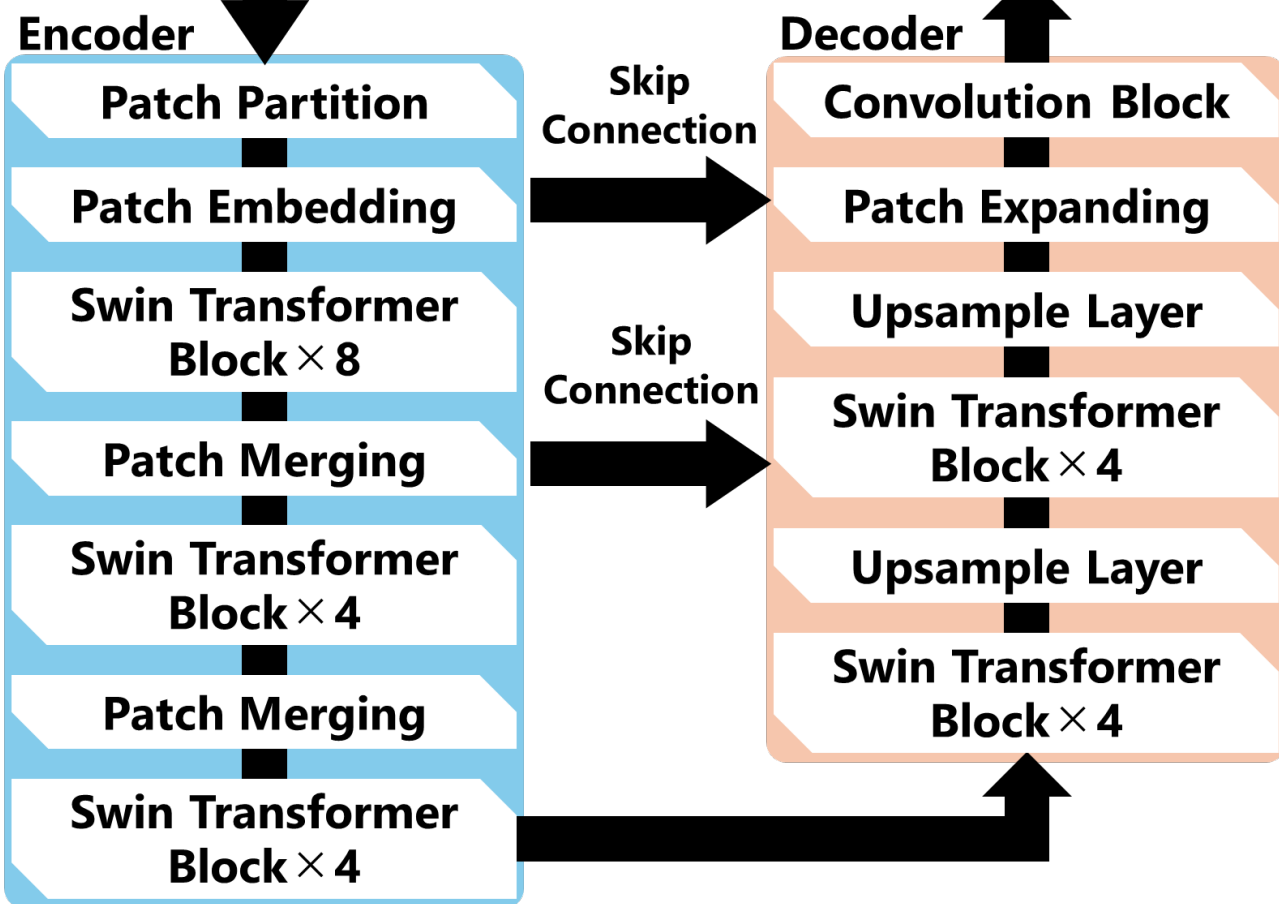
入力画像



CVDの見え



再配色画像



Encoder

- 入力画像から特徴量を抽出
- Patch Mergingにより解像度を下げる
- 階層的な特徴抽出

Decoder

- 補正された画像を生成
- Upsample Layerを解像度を上げる
- Skip connectionを使用