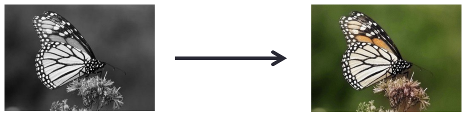
Image Coding With Neural Network-based Colorization

Diogo Lopes, João Ascenso, Catarina Brites, Fernando Pereira
Instituto Superior Técnico, Universidade de Lisboa - Instituto de Telecomunicações, Lisbon, Portugal

Introduction

- **Idea:** coding solution employing an automatic colorization to recover the chrominance components without transmitting any or just very few colour information
- **Motivation:**
  - High performance in the area of automatic colorization of grayscale images
  - Deep neural networks excel at colorization since it requires semantic understanding of the image

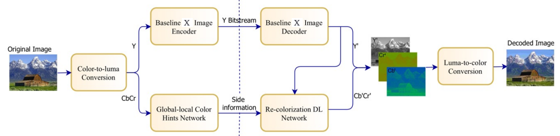


- Colorization process can be carried out at the decoder after luminance decoding:
  - Deep learning-based colorization network exploits the decoded luminance information to perform efficient luma-to-chroma estimation
  - Extraction, coding and transmission of some color information is needed to obtain accurate colorizations

Objective and High-level Framework

Proposes and evaluates the first image coding solution adopting a deep learning-based colorization process to recover the chrominance components at the decoder

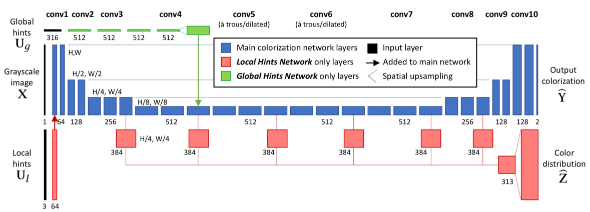
- To remove color ambiguity and boost color fidelity, chrominance values for specific pixel locations (color hints) are sent by the encoder to the decoder



- Methods are proposed for the selection of the color hints at the encoder, targeting the increase of the colorization performance

Real-Time User-Guided Colorization

- The state-of-the-art solution on user-guided colorization has been adopted since it is the best available:
  - R. Zhang, J. Zhu, P. Isola, X. Geng, A. Lin, T. Yu, A. Efros, "Real-Time User-Guided Image Colorization with Learned Deep Priors," in ACM Transactions on Graphics, vol. 36(4), 119:1-119:11, 2017.
  - Automatic colorization process with the capability of including color information from user inputs
- Can be adapted to include original image color information for some locations



Deep Learning Colorization-based Image Coding (DLCIC)

- Allows to reduce the bit rate associated to the coding of color information
- Includes an intelligent automatic selection of color hints, i.e. ground truth color information

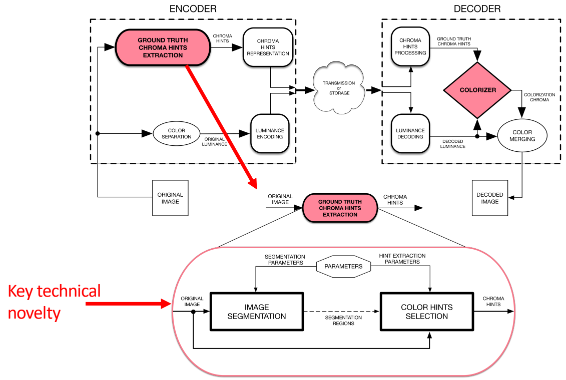
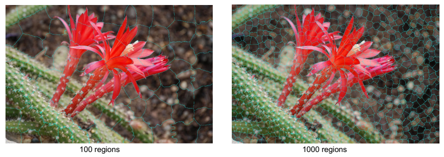


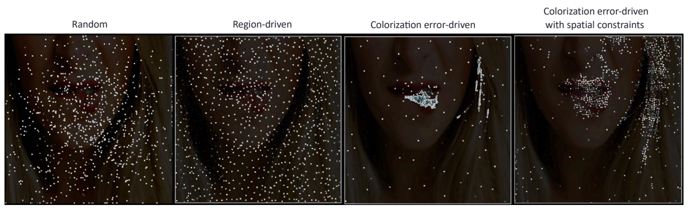
Image Segmentation

- Simple Linear Iterative Clustering (SLIC) algorithm to cluster the image pixels based on their spatial proximity and color similarity



Color Hint Selection

- Several intelligent and efficient color hint selection methods are proposed, from a rate-distortion performance point of view
- **No Hints,** where no hints are extracted and the colorization obtained at the decoder is only based on the decoded luminance
- **Random hints,** location of the color hints is chosen randomly, thus ignoring the image segmentation sub-module, following a Normal distribution located at the image center
- **Region-driven hints,** makes use of the regions created by the Image Segmentation module to extract more appropriate color hints depending on the image content.
  - For each segmented region, a hint location is selected corresponding to the pixel with the least color difference to all the remaining pixels in the region.
- **Colorization error-driven hints,** hints are selected according to the colorization error, i.e. hints are located in the areas where the colorization error is higher
  - Can be totally independent from the previous image segmentation or use it for bootstrapping



Test Conditions

- The proposed DLCIC solutions (using different hint extraction methods) are evaluated in terms of RD performance and subjective quality:
  - Four test images from the JPEG AI dataset with different characteristics
  - Peak Signal-to-Noise Ratio (PSNR) YCbCr, Structural Dissimilarity Index Measure (DSSIM)
- Coding solutions evaluated:
  - JPEG with a 4:2:0 color subsampling format (JPEG 420)
  - JPEG 400 NH (no hints): colorization process only based on the decoded luminance
  - JPEG 400 RH (random hints): colorization process based on the decoded luminance and some spatially random generated hints
  - JPEG 400 R-dH LD (region-driven hints), color hints locations correspond to pixels from each segmented color-uniform region of the image
  - JPEG 400 CE-dH SC-5px (colorizer error-driven hints), color hints locations are selected based on the colorization error with a minimum distance between hints of 5 pixels to avoid collapsing many hints in a small area
- For all the proposed DLCIC solutions, only the luminance is JPEG coded, varying the quality factors between 5% to 90% and 1000 color hints are considered.

Experimental Results

Table 1: BD-PSNR for the proposed DLCIC coding solutions (using as reference the JPEG 420 benchmark).

DLCIC SOLUTION	BD-PSNR YCbCr [dB]			
(1000 HINTS)	*Woman*	*Footballplayer*	*Buddhists*	*Harley*
JPEG 400 NH	-3.72	-2.35	-1.99	-1.94
JPEG 400 RH	-0.96	-1.78	-1.61	-0.55
JPEG 400 R-dH LD	-0.68	-1.17	-1.26	-0.46
JPEG 400 CE-dH SC-5px	-0.57	-0.89	-1.12	-0.34

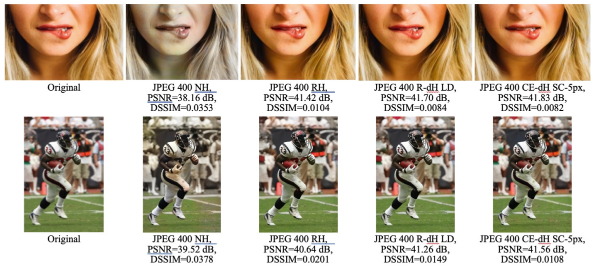


Figure 3: Comparative results for subjective inspection for Woman (top row) and Footballplayer (bottom row) test images.

Conclusions

- Objective quality assessment:
  - The JPEG 400 CE-dH SC-5px solution always outperforms the other DLCIC solutions
  - The colorization performance improvement with 'increased intelligence' color hints extraction methods is larger for *Footballplayer* image and lower for the *Harley* image
  - All DLCIC solutions perform worse than JPEG 420
- Subjective quality assessment:
  - Promising subjective quality is obtained with the DLCIC coding solutions
  - The lowest (decoded) quality is achieved with the JPEG 400 NH solution
  - The best (decoded) quality is achieved using 'more intelligent' and content aware DLCIC coding solutions in terms of colorization, as it can be observed for Footballplayer
- Objective quality metrics do not correlate well with perceived quality, especially since colorizations are not pixel-level faithful even if they are subjectively acceptable
- Future work must include formal subjective quality assessment