

Thèse de doctorat : colorisation automatique d'images et de vidéos à l'aide de réseaux de neurones et méthodes variationnelles

Fabien Pierre (fabien.pierre@loria.fr) et Frédéric Sur (frederic.sur@loria.fr).

L'objectif de cette thèse est d'investiguer des nouveaux paradigmes basés sur les réseaux de neurones et des approches variationnelles pour coloriser automatiquement des vidéos d'archives tournées en noir et blanc. Les méthodes automatiques par réseaux de neurones produisent des images dont on peut améliorer le résultat à l'aide de régularisation par méthode variationnelle. L'objectif principal de ce projet est d'intégrer la régularisation par méthode variationnelle dans la phase d'apprentissage du réseau de neurones. La thèse s'inscrit dans le cadre du projet ANR JCJC Arcé porté par Fabien Pierre.



Image en niveau de gris



Colorisation par CNN [9]



Colorisation par approche CNN et méthode variationnelle combinées [4].

Mots-clés : Édition d'image - Vision par ordinateur - Apprentissage automatique- Deep learning - Python - TensorFlow

Lieu de la thèse : LORIA, Inria Nancy Grand Est, projet Tangram.

Objectif et méthodologie du stage

L'objectif de cette thèse est de dépasser les performances des méthodes existantes de colorisation automatique de vidéos. Ces méthodes présentent des résultats spectaculaires à première vue, mais sont affectées par des artefacts numériques réhivitoires pour l'utilisation dans une chaîne de production. Citons par exemple l'utilisation de moyennes de couleur qui donnent un rendu terne, la difficulté de bien localiser les couleurs qui fait "baver" la colorisation (comme pointé par la flèche rouge dans l'exemple ci-dessus), ou l'intégration de l'estimation du mouvement pour propager la colorisation temporellement dans une vidéo. Une première contribution est la méthode décrite dans [4] qui combine un modèle variationnel [6] avec un algorithme basé sur un réseau de neurones [9]. Cette méthode rend les résultats de colorisation moins ternes en évitant les moyennes de couleur tout en les régularisant par minimisation de la variation totale.

Dans un premier temps, une phase exploratoire se focalisera sur l'apport de la régularité du résultat des réseaux de neurones. On explorera à cet effet les *Deep Image Prior* (DIP) [8] afin de régulariser l'image au lieu d'utiliser la variation totale. Une piste possible sera un travail sur la modification de la

fonction de coût des réseaux qui pourra intégrer la variation totale de l'image finale. Dans le même esprit, on pourra également chercher à réduire le biais de la méthode variationnelle [1, 7].

Dans le cadre de la colorisation de vidéos anciennes, on comparera des architectures de réseaux de neurones de colorisation trame par trame [3] et en blocs de trames [2], notamment pour voir s'il y a des avantages notamment dans le cadre des occultations.

Dans un second temps, on pourra investiguer la structure non-euclidienne de la couleur (par exemple l'espace cylindrique HSI) et utiliser une moyenne non-euclidienne de la couleur [5] dans le cadre des réseaux de neurones.

Compétences requises

Le candidat aura une formation mathématiques-informatique et devra présenter des compétences pour l'implémentation d'algorithmes en Python. Il devra être familier de l'apprentissage automatique (une expérience sur un projet en apprentissage automatique ou la participation à un enseignement de bon niveau dans ce domaine est exigé) et devra présenter un intérêt pour les éléments théoriques attenants à ce sujet. Il sera très utile au candidat d'avoir acquis des notions classiques de traitement d'image telles que la transformation dans des espaces couleur, du filtrage, des méthodes basées sur des équations aux dérivées partielles, l'apprentissage automatique sur des bibliothèques telles que Tensorflow ou Keras.

Références

- [1] Charles-Alban Deledalle, Nicolas Papadakis, Joseph Salmon, and Samuel Vaiter. Clear : Covariant least-square refitting with applications to image restoration. *SIAM Journal on Imaging Sciences*, 10(1) :243–284, 2017.
- [2] Chenyang Lei, Yazhou Xing, and Qifeng Chen. Blind video temporal consistency via deep video prior. *Advances in Neural Information Processing Systems*, 33 :1083–1093, 2020.
- [3] Simone Meyer, Victor Cornillère, Abdelaziz Djelouah, Christopher Schroers, and Markus Gross. Deep video color propagation. *Proceedings BMVC 2018*, pages 1–12, 2018.
- [4] Thomas Mouzon, Fabien Pierre, and Marie-Odile Berger. Joint CNN and Variational Model for Fully-automatic Image Colorization. In *SSVM 2019 - Seventh International Conference on Scale Space and Variational Methods in Computer Vision*, Hofgeismar, Germany, June 2019.
- [5] Frank Nielsen and Rajendra Bhatia. *Matrix information geometry*. Springer, 2013.
- [6] Fabien Pierre, Jean-François Aujol, Aurélie Bugeau, Nicolas Papadakis, and Vinh-Thong Ta. Luminance-chrominance model for image colorization. *SIAM Journal on Imaging Sciences*, 8(1) :536–563, 2015.
- [7] Fabien Pierre, Jean-François Aujol, Charles-Alban Deledalle, and Nicolas Papadakis. Luminance-guided chrominance denoising with debiased coupled total variation. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 235–248. Springer, 2017.
- [8] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446–9454, 2018.
- [9] Richard Zhang, Phillip Isola, and Alexei A Efros. Colorful image colorization. In *European conference on computer vision*, pages 649–666. Springer, 2016.