



HAL
open science

Apprendre à coloriser des images

Aurélie Bugeau, Nicolas Papadakis

► **To cite this version:**

Aurélie Bugeau, Nicolas Papadakis. Apprendre à coloriser des images. Interstices, 2019. hal-02187269

HAL Id: hal-02187269

<https://hal.science/hal-02187269>

Submitted on 23 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Apprendre à coloriser des images

Aurélie Bugeau, Nicolas Papadakis

La colorisation d'archives est couramment utilisée dans le milieu audiovisuel pour restaurer et rendre plus réaliste des vidéos anciennes. Elle nécessite un travail manuel fastidieux. Comme pour le coloriage (voir Figure 1), un algorithme de colorisation doit respecter deux règles simples :

- 1) Choisir un crayon : une couleur doit être déterminée pour chaque zone de l'image
- 2) Ne pas déborder : les couleurs doivent s'arrêter aux contours des objets.

Quelles sont les raisons techniques et pratiques qui empêchent d'automatiser ce procédé ?



Figure 1. Coloriages d'enfants : on souhaiterait créer un algorithme automatique de colorisation s'inspirant plutôt de l'exemple de droite.

La colorisation consiste à convertir une image en noir et blanc — en réalité, il s'agit d'une image aux 256 nuances de gris — en une image réaliste en couleurs. Ce processus est couramment utilisé dans l'industrie cinématographique, afin de rendre les images d'archives plus attrayantes aux yeux du grand public. Les couleurs doivent être déterminées correctement en fonction de l'objet à coloriser, mais aussi de son environnement. Si une voiture peut raisonnablement être de couleur rouge, violette ou jaune, il est plutôt rare de croiser une personne à la peau verte dans la rue en dehors des fêtes d'Halloween ou de Mardi gras. Les algorithmes de colorisation considèrent l'image en niveaux de gris comme un canal de luminance. La luminance est le flux lumineux émis ou réfléchi par chaque objet de la scène et perçue par la caméra. Dans le cas d'une image en couleur, la partie du signal vidéo correspondant à la couleur est appelée la chrominance. Cette caractéristique est représentée informatiquement par deux canaux supplémentaires (Figure 2).

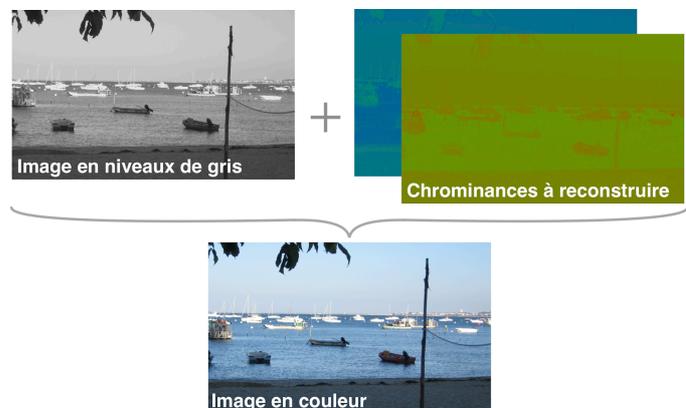


Figure 2. Processus de colorisation d'une image en niveaux de gris : deux canaux de chrominance doivent être estimés pour produire une image en

Une conversion dans l'encodage standard Rouge, Vert, Bleu (RVB ou RGB en anglais) peut alors être réalisée pour afficher l'image sur un écran. La colorisation revient ainsi à choisir les valeurs des deux canaux de chrominance pour chaque pixel de l'image. Si à une couleur correspond un unique niveau de gris, la réciproque est fautive : à chaque niveau de gris correspond une palette spécifique de chrominances possibles (voir la figure 3 ci-dessous). Il existe différentes approches pour lever cette ambiguïté : demander un coloriage partiel à l'utilisateur, fournir une image d'exemple à l'algorithme ou utiliser des millions d'images pré-colorisées pour apprendre la couleur la plus adaptée au contenu d'une image en noir et blanc.

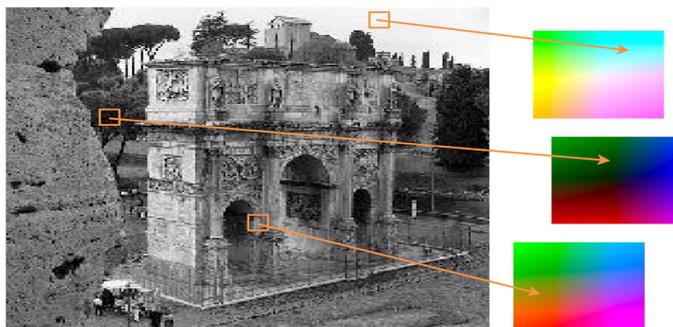


Figure 3. Ensemble des couleurs possibles pour des pixels de différents niveaux de gris.

Colorisation partielle de l'utilisateur et diffusion

Cette approche, directement inspirée du coloriage, consiste à faire dessiner par un utilisateur un trait de couleur pour chaque zone de l'image. Les couleurs sont ensuite propagées jusqu'aux frontières des objets (cf. figure 4 ci-dessous). Pour cela, les contours de l'image de luminance sont utilisés pour définir une force et une orientation privilégiée de diffusion des couleurs.

Des résultats satisfaisants sont obtenus sur des scènes relativement simples, mais le procédé demande une intervention manuelle longue dès que l'image contient de nombreux objets et détails. Cette technique, basée sur la règle de « ne pas déborder », est trop dépendante de l'utilisateur chargé de « choisir la couleur » de chaque région de l'image.

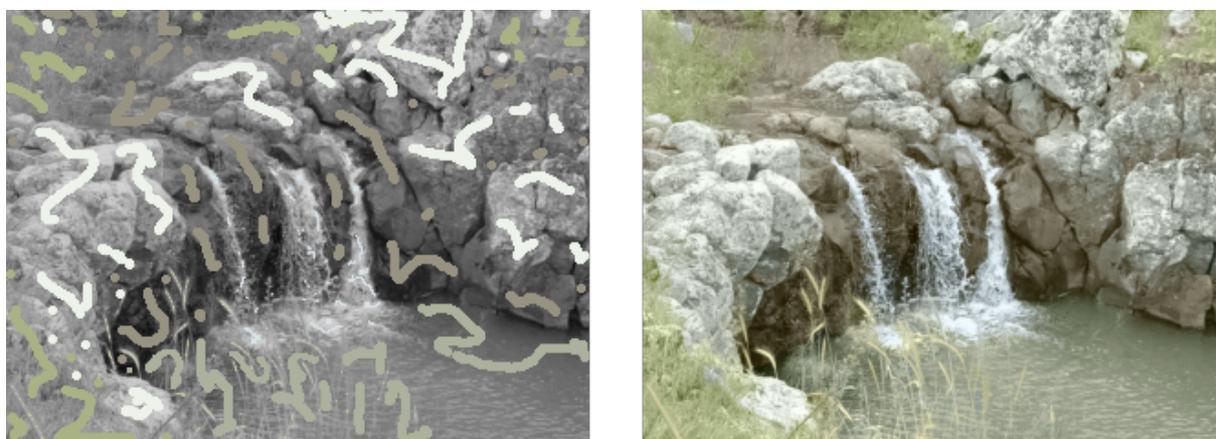


Figure 4. Processus de colorisation par diffusion.
À gauche : traits de couleurs d'un utilisateur. À droite : résultat de colorisation.

Colorisation par l'exemple

Comme détaillé dans un [précédent article](#), l'intervention de l'utilisateur peut se limiter à fournir à l'algorithme une image en couleur comme modèle. L'idée est de copier/coller les couleurs de cette image exemple dans l'image à coloriser. L'image exemple est tout d'abord convertie en niveau de gris. Pour chaque petite région carrée — ou patch — de l'image à coloriser, on cherche la région la plus proche, en termes de luminance et/ou de [texture](#), dans l'image exemple « grisée ». Les couleurs de cette région sont alors copiées dans l'image initiale (voir figure 5 ci-dessous). La comparaison entre deux patches se fait soit en mesurant directement la différence de luminance pixel à pixel, soit après calcul d'un descripteur de chacun des patches. Différents descripteurs peuvent être considérés : moyenne/variance de luminance, positions et orientations des contours ou des gradients, etc.

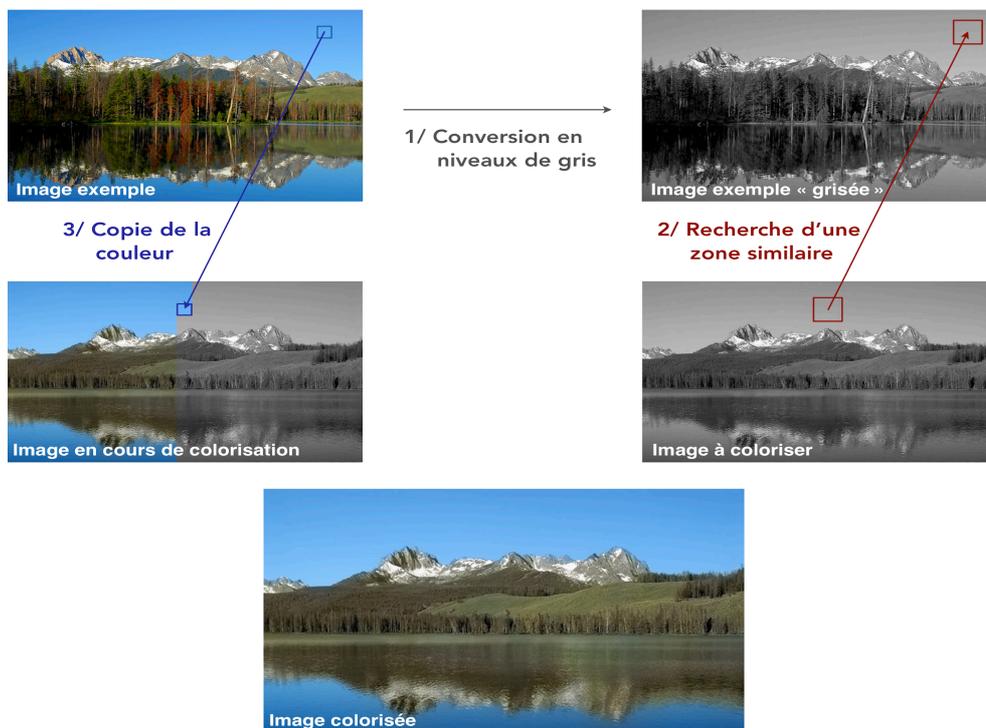


Figure 5. Processus de colorisation par l'exemple.

Le procédé est adapté aux scènes texturées (régions de couleur non uniforme), à condition d'avoir sélectionné une image exemple appropriée. Il permet de « choisir une couleur » automatiquement, mais ne respecte pas nécessairement la règle de « ne pas déborder », puisqu'il n'encourage pas la superposition des contours contenus dans les images de chrominance et de luminance. Pour éviter ce problème, les objets de l'image peuvent être préalablement segmentés puis colorisés un à un.

Une autre solution est de définir puis optimiser un modèle mathématique global cherchant un compromis entre (i) la copie, pour chaque pixel, de la couleur du centre du patch le plus proche lui ayant été associé et (ii) le fait que deux pixels voisins aient des couleurs similaires, s'ils ne sont pas séparés par un fort contour de luminance. Ce modèle peut inclure plusieurs choix de couleurs pour chaque pixel, en utilisant différents descripteurs pour comparer les patches.

Apprentissage pour la colorisation

Les méthodes récentes de colorisation poussent plus loin le concept précédent en faisant appel à l'[apprentissage automatique](#). On montre à un programme un grand nombre d'exemples de colorisation, sous la forme de paires d'images « avant colorisation (niveaux de gris) / après colorisation (chrominance) ». En pratique, cette base de données de paires d'images est obtenue en convertissant des images couleurs au format luminance/chrominance. L'objectif est d'entraîner le programme, contenant un [réseau de neurones profonds](#), afin qu'il ajuste ses nombreux paramètres internes et puisse reproduire au mieux la colorisation sur l'ensemble des exemples qu'on lui a fourni (voir figure 6 ci-dessous). À partir d'une image en niveau de gris en entrée, le réseau de neurones profonds génère deux canaux de chrominances. Lors de la phase d'entraînement, ces derniers sont comparés avec les vraies chrominances. La comparaison peut être réalisée en mesurant la différence des chrominances pixel à pixel, en utilisant une distance sur les statistiques des chrominances, ou en considérant un autre réseau de neurones entraîné pour distinguer les vrais canaux de chrominance de faux. Cette mesure d'erreur de reconstruction, appelée fonction de perte, est itérativement rétro-propagée dans le réseau pour mettre à jour l'ensemble de ses paramètres.

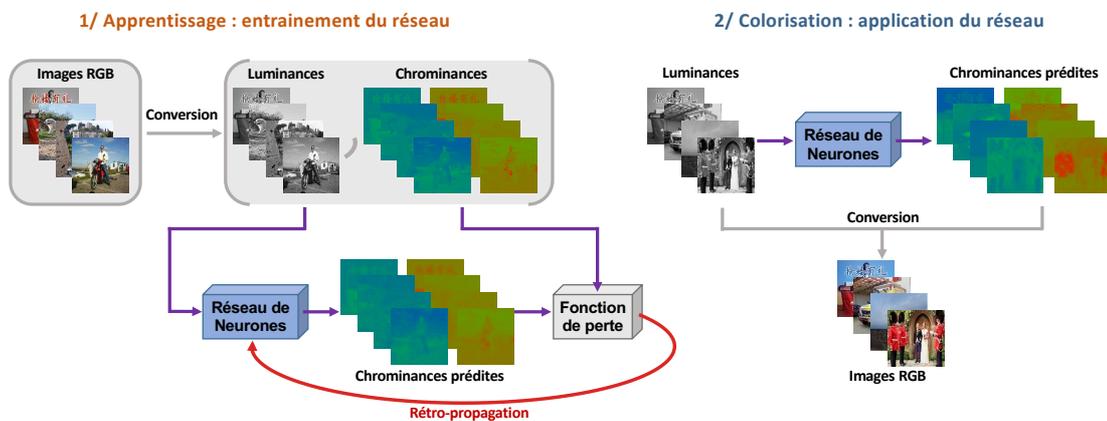


Figure 6. Processus de colorisation par apprentissage profond.

L'apprentissage automatique excelle à « choisir la couleur », notamment si des images similaires à celle à coloriser étaient présentes dans les données d'entraînement. Il n'y a cependant aucun contrôle sur l'image colorisée finale et les résultats obtenus sont imprévisibles. Il arrive ainsi parfois que les résultats ne soient pas réalistes. De plus, les contours des objets ne sont pas explicitement pris en compte et les couleurs peuvent « déborder » dans l'image colorisée (Figure 7).

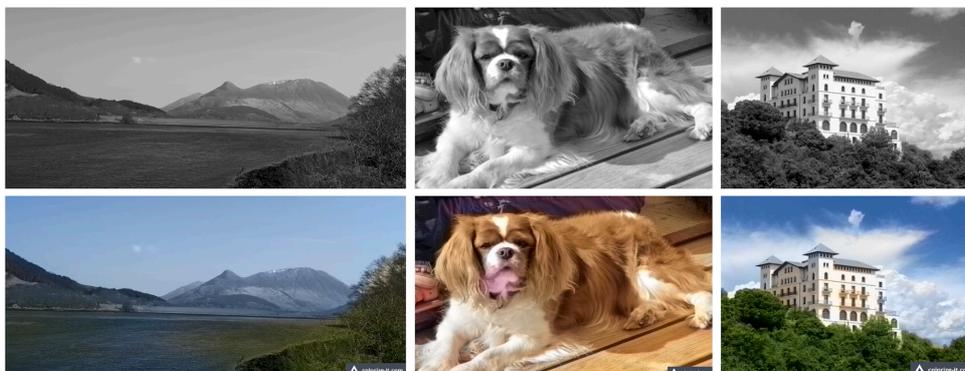


Figure 7. Exemples de colorisation par apprentissage profond. À gauche : les couleurs « débordent » ce qui mène à une eau verte et une montagne bleue. Au centre : résultat aberrant, le menton du chien est rose car une partie des images d'entraînement contenait des chiens avec la langue qui pend. À droite : la colorisation est excellente.

Le futur de la colorisation

Les recherches actuelles se concentrent sur la combinaison des trois types d'approches présentées et notamment sur l'intégration d'une colorisation partielle de l'utilisateur dans les méthodes par l'exemple et par apprentissage profond. D'autres défis restent encore à relever.

Les deux premiers défis concernent tous deux la colorisation automatique de vidéos, qui n'en est qu'à ses débuts, avec des points de vue et donc des problèmes différents (algorithmique et efficacité). Comme pour le coloriage manuel, passer du dessin au dessin animé demande un traitement minutieux pour obtenir un résultat homogène entre les images successives. La stabilisation temporelle de la colorisation demande de maintenir une chrominance cohérente le long des trajectoires des objets. L'estimation du mouvement apparent entre images successives est nécessaire pour déterminer les trajectoires des différents objets au cours de la vidéo. Le calcul du mouvement n'est cependant pas complètement maîtrisé lorsqu'un objet bouge rapidement par rapport à la position de la caméra, et il est même parfois impossible en cas d'occultation, par exemple lorsque l'objet est subitement caché par d'autres éléments de la scène avant de réapparaître.

Une deuxième difficulté pratique concerne le développement d'algorithmes capables de traiter des données vidéos, dont le volume de stockage est important. Les mémoires actuelles des cartes graphiques, nécessaires pour l'entraînement et l'exécution rapides des réseaux de neurones, sont encore trop limitées pour pouvoir traiter des vidéos de plusieurs secondes et de haute définition.

Enfin, les algorithmes automatiques actuels ne sont pas adaptés aux documents d'archives, qui contiennent un grain particulier dû aux techniques d'acquisition de l'époque. Le nombre d'images et de vidéos d'archives publiques ayant été colorisées manuellement ou automatiquement, et validées par des experts, est limité. Il n'existe donc à ce jour aucune base de données d'exemples de taille suffisante pour entraîner correctement un algorithme d'apprentissage. Il est donc nécessaire d'aller chercher d'autres types d'informations a priori. Le débat sur la pertinence de la colorisation des archives, lié au respect de la véracité historique, a fait son chemin chez les historiens, et l'industrie cinématographique s'appuie désormais aussi sur leur expertise pour le choix des couleurs. Un enjeu passionnant pour le futur est de faire apprendre à la machine les connaissances des historiens et des coloristes, tout en donnant la possibilité aux artistes de contrôler manuellement les résultats de colorisation.

Bibliographie associée

- Pour la colorisation partielle de l'utilisateur respectant la 2ème règle : "Colorization using Optimization". par Anat Levin, Dani Lischinski, Yair Weiss, ACM Transactions on Graphics 2002. website : <https://www.cs.huji.ac.il/~yweiss/Colorization/>
- Pour la colorisation par l'exemple respectant la 1ère règle : "Transferring color to greyscale images." par Tomihisa Welsh, Michael Ashikhmin, Klaus Mueller, ACM Transactions on Graphics 2002. website : <https://www3.cs.stonybrook.edu/~mueller/research/colorize/>
- Pour les méthodes par l'exemple respectant les 2 règles : A. Bugeau, V.-T. Ta, N. Papadakis, "Variational Exemplar-Based Image Colorization", IEEE Transactions on Image Processing (TIP), Volume 33, Issue 1, 2014. <http://www.labri.fr/perso/bugeau/Research/imcolorization.html>
- Pour la colorisation par apprentissage : "Colorful Image Colorization", Richard Zhang, Phillip Isola, Alexei A. Efros, European Conference on Computer Vision (ECCV), 2016. website : <https://richzhang.github.io/colorization/>