

TP2 Rapport

Jose Miguel Alfaro Castillo

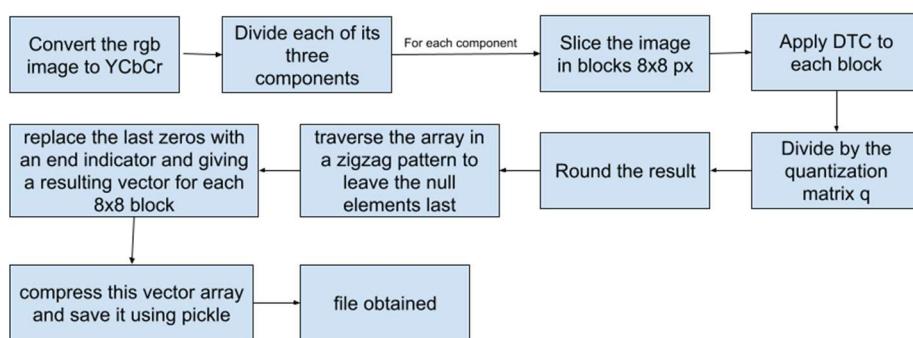
Introduction

Ce TP a été divisé en trois parties principales : d'abord, la recherche et la sélection d'un algorithme de compression d'images utilisant la DHT ; ensuite, l'entraînement de différents algorithmes de machine learning pour classer des chiffres manuscrits de 0 à 9, puis la répétition du processus après un prétraitement des images à l'aide du HOG (Histogram of Gradient Feature), afin de comparer les deux résultats et déterminer lequel offre les meilleures performances. Enfin, la dernière partie traite du processus de génération d'une mosaïque de vignettes (thumbnails), à partir d'une large image. Trois approches différentes sont considérées, en variant les caractéristiques utilisées pour chaque bloc de l'image ainsi que la manière de calculer la distance entre chaque bloc et les vignettes disponibles. Ce TP a été divisé en trois parties principales : d'abord, la recherche et la sélection d'un algorithme de compression d'images utilisant la DHT ; ensuite, l'entraînement de différents algorithmes de machine learning pour classer des chiffres manuscrits de 0 à 9, puis la répétition du processus après un prétraitement des images à l'aide du HOG (Histogram of Gradient Feature), afin de comparer les deux résultats et déterminer lequel offre les meilleures performances. Enfin, la dernière partie traite du processus de génération d'une mosaïque de vignettes (thumbnails), à partir d'une large image. Trois approches différentes sont considérées, en variant les caractéristiques utilisées pour chaque bloc de l'image ainsi que la manière de calculer la distance entre chaque bloc et les vignettes disponibles.

Méthodologie

1. Compression d'images avec DHT

La première étape consistait à définir le processus de compression d'une image à l'aide de l'algorithme DHT. Ce processus est illustré dans l'image suivante :



Le processus de décompression est l'inverse de celui-ci, en utilisant la même matrice q que précédemment. Pour disposer d'un point de comparaison, cet algorithme a d'abord été réalisé avec une matrice q par défaut, utilisée pour la compression JPEG.

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Afin de trouver une matrice q offrant de meilleurs résultats que celle du JPEG (appelée ici q_JPEG), un expériment aléatoire a été proposé : on définit une matrice $Q(i,j)=a+b(i+j)$. Sur 200 itérations, a prend une valeur aléatoire entre 3 et 100, et b entre 1 et 40. La première étape consistait à définir le processus de compression d'une image à l'aide de l'algorithme DHT. Ce processus est illustré dans l'image suivante : [...] Le processus de décompression est l'inverse de celui-ci, en utilisant la même matrice q que précédemment. Pour disposer d'un point de comparaison, cet algorithme a d'abord été réalisé avec une matrice q par défaut, utilisée pour la compression JPEG. [...] Afin de trouver une matrice q offrant de meilleurs résultats que celle du JPEG (appelée ici q_JPEG), un expériment aléatoire a été proposé : on définit une matrice $Q(i,j)=a+b(i+j)$. Sur 200 itérations, a prend une valeur aléatoire entre 3 et 100, et b entre 1 et 40.

2. Classification d'images

Pour cette partie, 70 000 images étiquetées du jeu de données mnist_784 ont été utilisées. L'objectif était d'entraîner plusieurs modèles de machine learning afin d'identifier celui qui est le plus performant pour cette base. Les données ont d'abord été normalisées entre 0 et 1. Trois méthodes vues en cours ont été testées : KNN, Arbre de Décision, et SVM, dont les performances ont été comparées. [...] Dans une deuxième étape, un prétraitement HOG (Histogram of Oriented Gradients) a été appliqué. Pour comprendre son fonctionnement, il a d'abord été appliqué à une image PPM, puis à tout le jeu de données mnist_784. Les performances avant et après HOG ont été comparées. Pour cette partie, 70 000 images étiquetées du jeu de données mnist_784 ont été utilisées. L'objectif était d'entraîner plusieurs modèles de machine learning afin d'identifier celui qui est le plus performant pour cette base. Les données ont d'abord été normalisées entre 0 et 1. Trois méthodes vues en cours ont été testées : KNN, Arbre de Décision, et SVM, dont les performances ont été comparées. [...] Dans une deuxième étape, un prétraitement HOG (Histogram of Oriented Gradients) a été appliqué. Pour

comprendre son fonctionnement, il a d'abord été appliqué à une image PPM, puis à tout le jeu de données mnist_784. Les performances avant et après HOG ont été comparées.

3. Large Mosaic

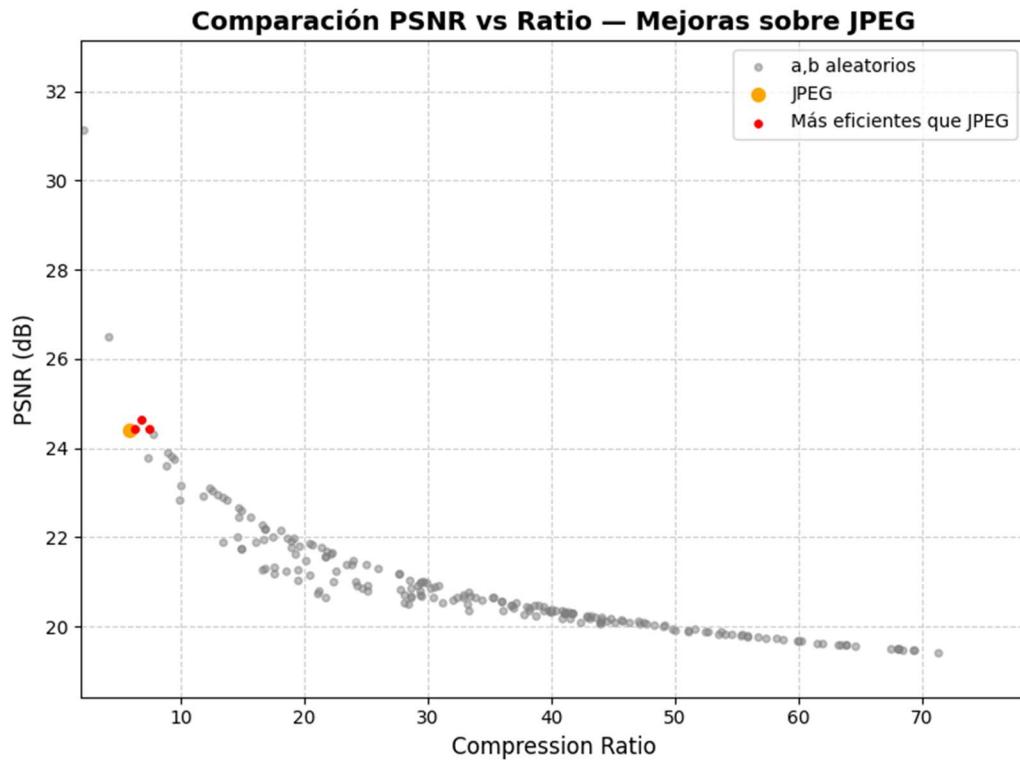
L'objectif est de générer une mosaïque composée de petites images (thumbnails), tout en conservant la structure générale d'une grande image de référence. Une image de grande taille a été sélectionnée et divisée en blocs de 32×32 . Les caractéristiques des vignettes ont ensuite été définies selon trois approches : la moyenne RGB, l'histogramme HSV, et une méthode combinée. Deux distances ont été utilisées : Euclidienne et Cityblock. Les trois résultats ont été comparés pour analyser l'influence des caractéristiques et des distances sur la qualité finale.

L'objectif est de générer une mosaïque composée de petites images (thumbnails), tout en conservant la structure générale d'une grande image de référence. Une image de grande taille a été sélectionnée et divisée en blocs de 32×32 . Les caractéristiques des vignettes ont ensuite été définies selon trois approches : la moyenne RGB, l'histogramme HSV, et une méthode combinée. Deux distances ont été utilisées : Euclidienne et Cityblock. Les trois résultats ont été comparés pour analyser l'influence des caractéristiques et des distances sur la qualité finale.

Analyse des résultats

1. Compression d'images avec DHT

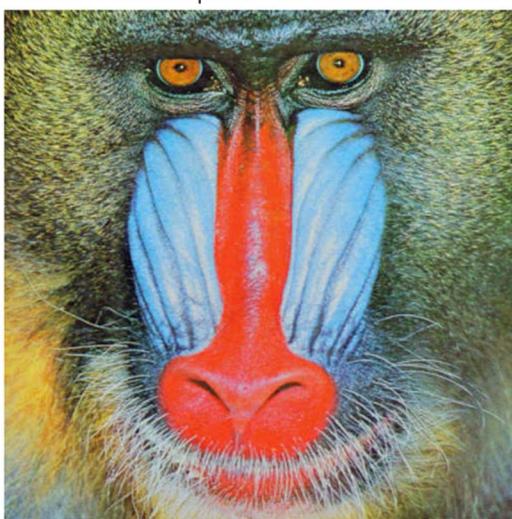
Pour qu'un résultat soit considéré comme plus efficace que JPEG, il doit atteindre un taux de compression supérieur tout en conservant une qualité (PSNR) égale ou meilleure. Trois configurations distinctes de a et b ont permis d'obtenir de meilleures performances que JPEG en termes de PSNR.



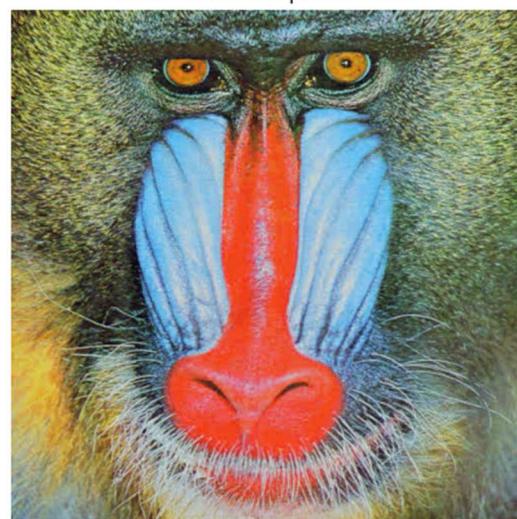
Visuellement, les différences restent peu perceptibles, mais l'image plus légère présente un PSNR supérieur. Pour qu'un résultat soit considéré comme plus efficace que JPEG, il doit atteindre un taux de compression supérieur tout en conservant une qualité (PSNR) égale ou meilleure. Trois configurations distinctes de a et b ont permis d'obtenir de meilleures performances que JPEG en termes de PSNR.

Comparison: JPEG vs. Best Custom Option

JPEG
PSNR=24.41 dB | Ratio=5.785378520932491



Best custom Q (a=17, b=3)
PSNR=24.65 dB | Ratio=6.83



Visuellement, les différences restent peu perceptibles, mais l'image plus légère présente un PSNR supérieur.

2. Classification d'images

Les trois modèles ont permis de classer directement les images MNIST. Le SVM à noyau RBF a obtenu la meilleure précision moyenne (97 %), suivi du KNN (96 %), tandis que l'arbre de décision (87 %) s'est montré plus rapide mais moins précis. L'utilisation du HOG a permis une légère amélioration, notamment pour SVM et KNN, car il fournit des caractéristiques plus compactes et moins bruitées.

Model	Average Accuracy	Standard Deviation
KNN (k=3)	0.971217	0.001026
Decision Tree (Gini)	0.870100	0.001932
SVM (RBF)	0.983133	0.000602

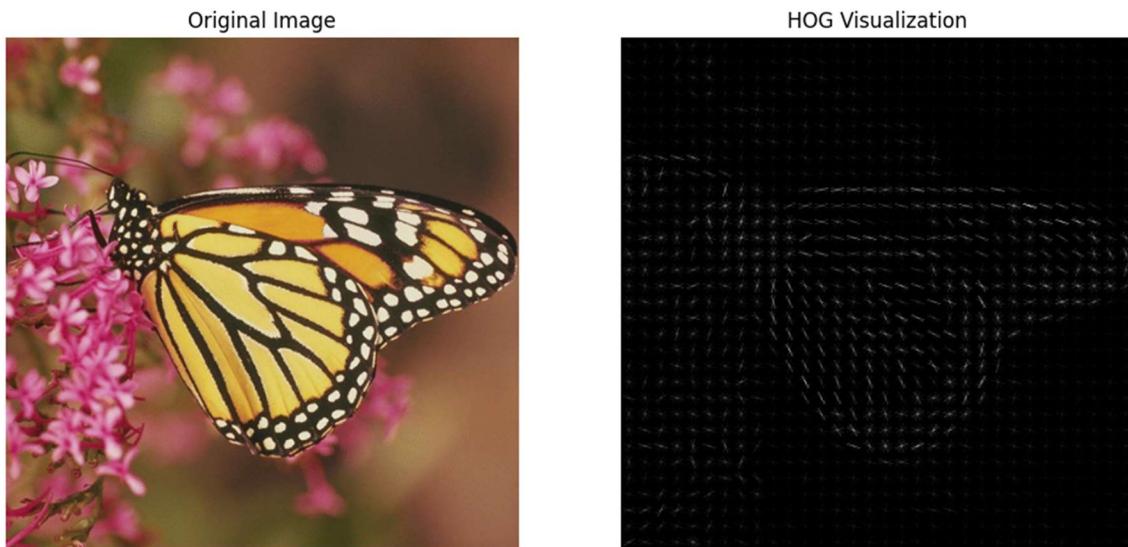


Table Results of training model with HOG

Model	Average Accuracy	Standard Deviation
KNN (k=3)	0.973717	0.001026
Decision Tree (Gini)	0.868717	0.001932
SVM (RBF)	0.989383	0.000602

Base results table without preprocessing vs. preprocessing with HOG

Model	Average Accuracy	Average Accuracy	Standard Deviation
KNN (k=3)	0.971217	0.973717	0.001026
Decision Tree (Gini)	0.870100	0.868717	0.001932
SVM (RBF)	0.983133	0.989383	0.000602

Les trois modèles ont permis de classer directement les images MNIST. Le SVM à noyau RBF a obtenu la meilleure précision moyenne (97 %), suivi du KNN (96 %), tandis que l'arbre de décision (87 %) s'est montré plus rapide mais moins précis. L'utilisation du HOG a permis une légère amélioration, notamment pour SVM et KNN, car il fournit des caractéristiques plus compactes et moins bruitées.

3. Mosaïque d'images

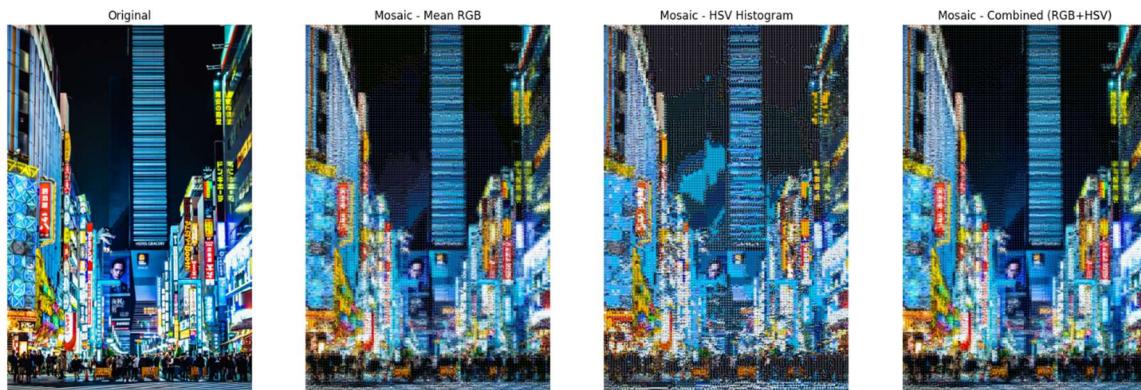
L'image initiale longue est la suivante :



À partir de l'image de référence et des vignettes issues de CIFAR-10, trois méthodes ont été testées. Le meilleur PSNR et la meilleure correspondance de couleur ont été obtenus avec la méthode combinée (RGB+HSV+Adjust).

Method	PSNR	SSIM	ΔE (avg)
Mosaic - Mean RGB	11.531	0.117	19.010
Mosaic - HSV Histogram	10.022	0.145	22.124
Mosaic - Combined (RGB+HSV+Adjust)	11.622	0.131	18.411

Le HSV a obtenu le meilleur SSIM mais un rendu visuel moins naturel. Le ΔE confirme que la méthode combinée est la plus fidèle à l'image originale. À partir de l'image de référence et des vignettes issues de CIFAR-10, trois méthodes ont été testées. Le meilleur PSNR et la meilleure correspondance de couleur ont été obtenus avec la méthode combinée (RGB+HSV+Adjust).



Le HSV a obtenu le meilleur SSIM mais un rendu visuel moins naturel. Le ΔE confirme que la méthode combinée est la plus fidèle à l'image originale.

Conclusions

La méthode de compression basée sur la DHT s'est révélée efficace. Les expérimentations aléatoires sur a et b ont permis d'identifier des configurations surpassant JPEG en PSNR. Il serait intéressant d'explorer des gammes de valeurs plus larges. Ces résultats montrent que la matrice de quantification standard du JPEG n'est pas universellement optimale. Pour la classification d'images, le SVM RBF a offert les meilleures performances, atteignant une précision moyenne de 98,9 %. L'ajout du descripteur HOG a légèrement amélioré les résultats. Pour la mosaïque d'images, la méthode Combinée (RGB + HSV + Adjust) a offert le meilleur compromis entre fidélité colorimétrique, structure et correspondance de couleur. La méthode de compression basée sur la DHT s'est révélée efficace. Les expérimentations aléatoires sur a et b ont permis d'identifier des configurations

surpassant JPEG en PSNR. Il serait intéressant d'explorer des gammes de valeurs plus larges. Ces résultats montrent que la matrice de quantification standard du JPEG n'est pas universellement optimale. Pour la classification d'images, le SVM RBF a offert les meilleures performances, atteignant une précision moyenne de 98,9 %. L'ajout du descripteur HOG a légèrement amélioré les résultats. Pour la mosaïque d'images, la méthode Combinée (RGB + HSV + Adjust) a offert le meilleur compromis entre fidélité colorimétrique, structure et correspondance de couleur.