

Combinaison linéaire optimale de métriques pour la colorisation d’images.

Fabien PIERRE^{1,2}, Jean-François AUJOL¹, Aurélie BUGEAU², Vinh-Thong TA^{2,3}

¹Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France, CNRS, IMB, UMR 5251, F-33400 Talence, France.

²Univ. Bordeaux, LaBRI, UMR 5800, F-33400 Talence, France, CNRS, LaBRI, UMR 5800, F-33400 Talence, France.

³Bordeaux INP, LaBRI, UMR 5800, F-33402 Talence, France.

fabien.pierre@math.u-bordeaux.fr, jean-francois.aujol@math.u-bordeaux.fr
aurelie.bugeau@labri.fr, vinh-thong.ta@labri.fr

Résumé – Cet article propose une approche originale pour construire une métrique optimale. La méthode décrite calcule une combinaison linéaire de métriques, afin de créer une nouvelle distance capable de mieux comparer des patches. L’utilisateur fournit un ensemble d’images d’apprentissage, ainsi que les métriques de base. Le cadre théorique et un algorithme efficace sont décrits pour résoudre ce problème. Des expériences numériques en colorisation d’images, sont réalisées pour démontrer l’efficacité de cette nouvelle approche.

Abstract – This paper proposes a novel approach to compute an optimal metric. This approach performs a linear combination of distances, in order to compute a new metric which can better compare patches. The user provides a set of training images, as well as the basic distances. A theoretical framework and an efficient algorithm are described to solve this problem. Numerical experiments in image colorization are performed in order to demonstrate the interest of this new approach.

1 Introduction.

La comparaison de textures est un problème toujours ouvert en vision par ordinateur. Bien que de nombreux descripteurs aient été proposés dans la littérature, ([1], [10], etc), le descripteur universel parfait n’existe pas. En trouver un bon est une tâche difficile, même avec une vérité terrain [7]. En général, les approches classiques combinent différents types de descripteurs de textures. On cherche alors à comparer ces descripteurs avec des métriques.

Plusieurs types d’approches non-expérimentales permettent d’établir une méthode optimale pour la comparaison de descripteurs. Un premier type d’approche modifie la distance utilisée pour trouver les k plus proches voisins lors de la classification par vote k -NN. La métrique originale proposée pour mesurer la distance entre ces voisins est la norme Euclidienne [4]. [12] propose de remplacer cette distance par celle de Mahalanobis dont les paramètres sont appris via la minimisation d’un critère. Ce dernier permet de limiter les erreurs de classification possibles. Un deuxième type d’approche se base sur la théorie des probabilités et de l’information [9]. Il repose sur un modèle probabiliste pour fournir une métrique optimale définie comme combinaison linéaire de diverses distances de base. Un troisième type d’approche, proposée par [5], consiste à calculer une combinaison linéaire optimale de

métriques. Cette dernière est utilisée afin de résoudre le problème suivant : étant donné une image test et deux de référence, on détermine l’image de référence la plus proche de celle de test. Notre approche est proche de cette dernière méthode et est appliquée aux algorithmes basés sur les patches.

Une application simple est la colorisation d’images, qui consiste à transformer une image en niveau de gris en une image en couleur. Elle nécessite un *a priori* afin de produire un résultat réaliste. Celui-ci peut être fourni sous forme manuelle, par exemple [8, 14], ou sous la forme d’une image de référence, de sémantique proche [13, 6]. Cette dernière approche est connue sous le nom de *colorisation basée-exemple*. Sa principale difficulté est de trouver un bon descripteur de texture. En général, plusieurs sont combinés. Par exemple, les méthodes de [13] et [6] utilisent une pondération de métriques, pour comparer des morceaux d’image. Afin d’éviter cette pondération (choisie expérimentalement, en pratique), les auteurs de [11] utilisent un modèle variationnel pour combiner les résultats obtenus par plusieurs métriques entre patches. Les colorisations obtenues séparément avec plusieurs métriques différentes sont localement sélectionnées en fonction de la régularité souhaitée.

Dans ce papier, nous montrons que l’utilisation de métriques isolées peut échouer, tandis que leur combinaison linéaire peut donner un bon résultat. Nous propo-

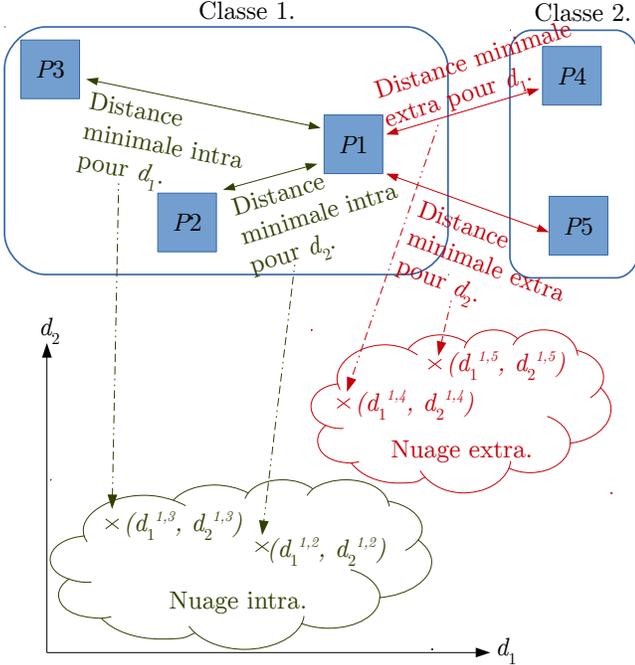


FIGURE 1 – Formation des nuages de points dans le cas de deux classes et de deux métriques. $(d_1^{i,j}, d_2^{i,j})$ représente le point de coordonnées $d_1(P_i, P_j)$ et $d_2(P_i, P_j)$.

sons une méthode permettant de créer une nouvelle métrique en calculant une combinaison linéaire d'une liste de distances existantes. Notre méthode s'inspire des *Support Vector Machine (SVM)* [2] qui sont utilisés pour séparer des nuages de points par un hyperplan, afin de discriminer des données. Cette séparation de l'espace permet de classer une nouvelle entrée en calculant de quel côté de l'hyperplan celle-ci se situe. La méthode *soft margin* de [3] est une extension permettant de trouver un hyperplan optimal lorsque les deux nuages de points ne se séparent pas parfaitement.

Dans la suite, nous commençons par modéliser le problème sous forme de séparation de nuages de points, puis nous proposons un modèle pour calculer effectivement la métrique optimale. Enfin, nous décrivons une application à la colorisation d'images, afin de démontrer l'intérêt de cette méthode innovante.

2 La distance optimale.

Dans cette section, nous proposons une modélisation du problème du choix de la métrique optimale.

2.1 Modélisation du problème.

Dans la plupart des problèmes de vision par ordinateur, les algorithmes doivent distinguer des objets avec la même sensibilité que le système visuel humain. On consi-

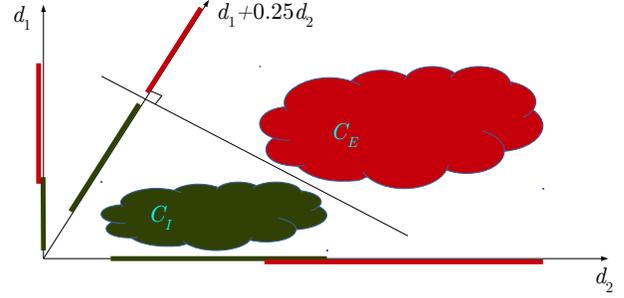


FIGURE 2 – Exemple de représentation de nuages de distances *intra* et *extra* patch. Ni la métrique d_1 ni la d_2 ne sont capables de séparer les nuages, mais la combinaison $d_1 + 0.25d_2$ en est capable. Les traits rouges et verts représentent les nuages *intra* et *extra* patch monodimensionnels pour chaque métrique séparée.

dère, dans la suite, des morceaux d'image carrés, d'une taille prédéfinie, que l'on appellera patch. L'utilisation de descripteurs est établie par une distance entre eux qui est faible quand le système visuel humain considère que le contenu lui semble proche. Au contraire, cette distance doit être élevée si le système visuel humain perçoit ces éléments comme étant différents. Cette hypothèse est le point de départ de notre méthode qui requiert l'utilisation d'une vérité terrain fournie par l'utilisateur.

Dans la suite, nous proposons un modèle permettant de calculer une métrique qui est adaptée à ces données. Pour apprendre cette métrique, on suppose que l'utilisateur fournit des ensembles de patch (en sélectionnant une partie d'une image, par exemple). Il faut que la distance minimale, pour la métrique retenue, entre un patch et tous les autres, soit atteinte idéalement avec un autre patch de sa classe.

On considère N patches $\{p_i\}$ répartis en C classes. Les patches de la classe c sont indexés par l'ensemble d'indices I_c . Supposons que l'on dispose d'une métrique entre patches notée d . Pour chaque patch p_i tel que $i \in I_c$ on considère l'indice j_i^I tel que :

$$j_i^I = \operatorname{argmin}_{j \in I_c, j \neq i} d(p_i, p_j). \quad (1)$$

On appelle *distance intra* du patch p_i et on note dI_i la valeur $d(p_i, p_{j_i^I})$. Considérons également l'indice j^* tel que :

$$j_i^E = \operatorname{argmin}_{j \notin I_c} d(p_i, p_j). \quad (2)$$

On appelle *distance extra* du patch p_i et on note dE_i la valeur $d(p_i, p_{j_i^E})$. En calculant les distances *intra* et *extra* pour chaque patch, on obtient deux ensembles de valeurs que l'on appelle nuages de points *intra* et *extra* dans la suite. Si l'on considère K distances, on peut alors former des nuages de points de dimension K . Pour chaque métrique d_k et chaque patch p_i on calcule les indices j_i^I et j_i^E que l'on regroupe dans des ensembles d'indices \mathcal{I}_k et \mathcal{E}_k .

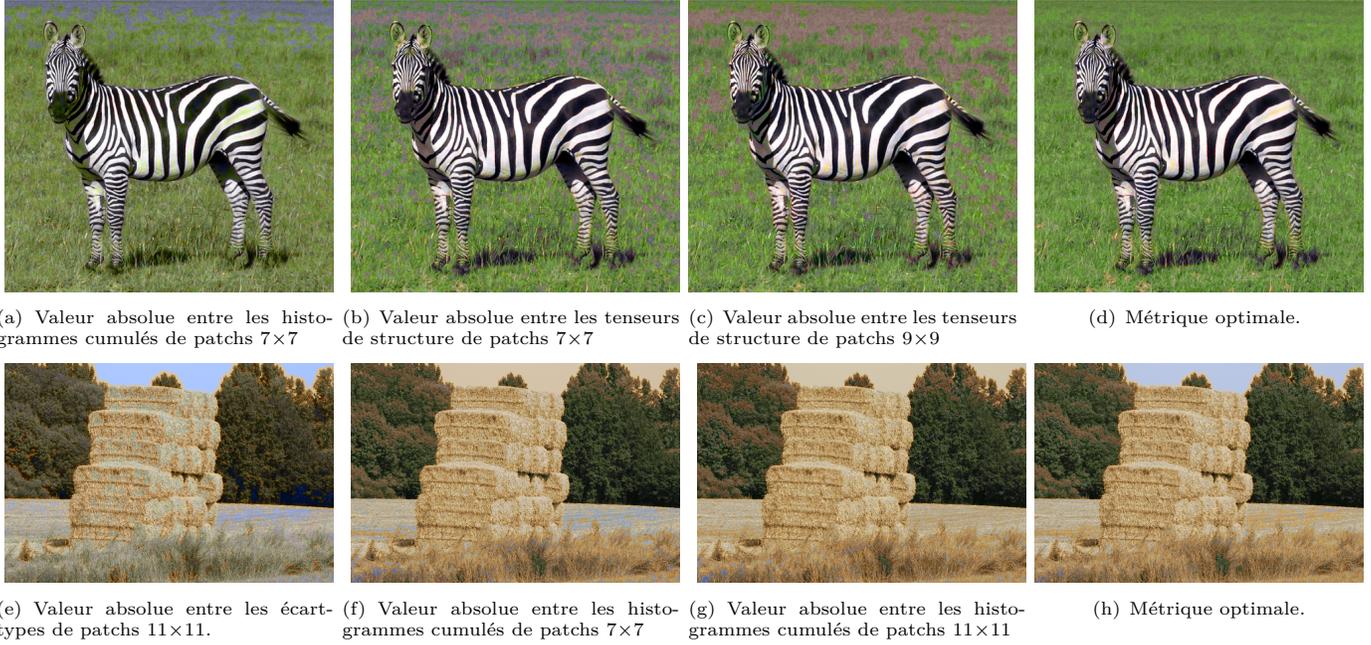


FIGURE 3 – Colorisation basée-exemple avec plusieurs types de métriques, comparés à la métrique optimale. La métrique optimale donne un résultat plus réaliste.

On compose ensuite les nuages de points dans l'ensemble \mathbb{R}^{+K} de la manière suivante :

$$C_I = \{(d_1(p_i, p_j), \dots, d_K(p_i, p_j))\}_{i \in 1..N, j \in \mathcal{I}_k, k \in 1..K}, \quad (3)$$

$$C_E = \{(d_1(p_i, p_j), \dots, d_K(p_i, p_j))\}_{i \in 1..N, j \in \mathcal{E}_k, k \in 1..K}. \quad (4)$$

La formation de ces nuages dans le cas de deux classes est résumé dans la Figure 1. Pour chaque patch et chaque métrique, on calcule la distance minimale avec les patches de la même classe et ceux des autres classe. Par exemple, pour $P1$, la distance minimale *intra* classe pour la distance d_1 est réalisée avec $P3$.

Pour avoir un bon appariement il faut que, quelque soit $i \in I_C$, l'indice j^* tel que :

$$j^* = \operatorname{argmin}_{j \in 1..N, j \neq i} d(p_i, p_j), \quad (5)$$

appartienne à I_C . Ainsi il faut :

$$\forall i \in 1..N, dE_i > dI_i. \quad (6)$$

Lorsque l'on dispose de C métriques, on peut les combiner de manière linéaire avec des coefficients positifs pour obtenir une nouvelle métrique. Pour que la combinaison soit pertinente, il faut que les projections des nuages K -dimensionnels sur une droite soient séparées. La recherche d'une métrique optimale revient à séparer les nuages de points C_E et C_i par un hyperplan.

La Figure 2 montre deux nuages *extra* C_E et *intra* C_I séparés en dimension 2 par une droite. Les nuages *intra* et *extra* des deux métriques de départ (segments de couleur sur les axes principaux) ne sont pas séparés. Puisqu'il existe une droite séparant les nuages, alors il existe

une combinaison linéaire pertinente (dans notre exemple $d_1 + 0.25d_2$) telle que les nuages *intra* et *extra* pour cette métrique soient séparés (segments de couleur sur l'axe $d_1 + 0.25d_2$).

Dans la section suivante, nous présentons sur une façon optimale de séparer les nuages *intra* et *extra*.

2.2 SVM contraint.

Dans cette section, nous proposons de séparer les nuages de points par un hyperplan.

Une technique, largement utilisée pour séparer des nuages de points, est la méthode des SVM introduite dans [2]. Cette technique repose sur la minimisation d'un critère quadratique. Dans le cas de nuages non-séparés, les auteurs de [3] introduisent la méthode de *soft margin hyperplane*. Puisque le problème de construction d'un hyperplan qui minimise le nombre d'erreurs d'appariement est NP-complet, les auteurs considèrent un problème relaxé.

Remarquons maintenant que la définition de la métrique optimale nécessite que les coefficients de l'hyperplan soient positifs afin qu'elle respecte l'inégalité triangulaire. Cette nouvelle contrainte est ajoutée au problème d'optimisation quadratique du SVM :

$$\begin{aligned} \min_{\alpha, b, \xi} \quad & \|\alpha\|_2^2 + \beta \sum_{i=1}^k \xi_i \\ \text{t.q.} \quad & y_i(\alpha^t x_i + b) \geq 1 + \xi_i \\ & \xi_i \geq 0, \alpha_j \geq 0. \end{aligned} \quad (7)$$

α est le vecteur des coefficients de l'hyperplan, b est le biais. y_i est un scalaire égal à -1 si x_i est dans le premier nuage et 1 s'il est dans le second. β est une constante. Il s'agit d'un problème d'optimisation quadratique contraint.

3 Colorisation basée-exemple.

Dans cette section, nous appliquons notre méthode pour construire une métrique optimale, dédiée à la colorisation d'images.

Une méthode, basée sur des patches, a été développée par les auteurs de [13]. La méthode fonctionne comme suit : pour chaque pixel de l'image à coloriser, le patch centré en ce pixel est comparé à un sous-ensemble de patches sous-échantillonné, de manière aléatoire, dans une image en couleur de référence, fournie par l'utilisateur, et préalablement convertie en niveau de gris. La couleur du pixel central correspondant au patch le plus proche est retenue. Le problème essentiel de cette méthode est de trouver une métrique pertinente pour comparer les patches.

La Figure 3 montre le résultat d'une colorisation utilisant pour référence un ensemble de morceaux d'images. Les morceaux d'images sont également utilisés pour calculer la métrique optimale. Ils sont prélevés manuellement de plusieurs images en couleur, différentes de l'image elle-même, et répartis en classes. Par exemple, pour la colorisation du zèbre, on a extrait une classe contenant des morceaux de zèbre et une autre des morceaux d'herbe. Ainsi, la métrique optimale est capable de discerner le zèbre de l'arrière-plan, et produit un résultat pertinent. Pour l'image de paysage, trois classes sont utilisées, une pour le ciel, une pour les arbres et une pour la paille. Cela explique le fait que le ciel est bien colorisé avec la métrique optimale. On remarque qu'aucune des métriques de base ne fournit de résultat acceptable. La colorisation du zèbre par les métriques de base donne des résultats contenant des artefacts. Par exemple, l'herbe contient des taches bleues. L'image de paysage contient également des erreurs, par exemple, la présence de bleu sur les arbres, ou d'orange dans le ciel. La métrique optimale fournit un résultat pertinent et réaliste.

4 Conclusion.

Dans cet article, nous avons proposé un cadre pour combiner des métriques de manière optimale par rapport à une vérité terrain donnée. La métrique optimale ainsi calculée démontre une capacité accrue pour comparer des patches et fournir une colorisation d'images, pertinente et réaliste. Les applications possibles de cette méthode sont nombreuses, par exemple l'inpainting, la segmentation, la classification.

Remerciements

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'avenir Initiative d'excellence IdEx Bordeaux portant la référence ANR-10-IDEX-03-02.

Jean-François Aujol est membre de l'Institut Universitaire de France.

Références

- [1] H. Bay, T. Tuytelaars, et L. Van Gool. *Surf : Speeded up robust features*. European Conference on Computer Vision. 2006.
- [2] B. E. Boser, I. M. Guyon, et V. N. Vapnik. *A training algorithm for optimal margin classifiers*. Computational learning theory, pages 144–152. ACM, 1992.
- [3] C. Cortes et V. Vapnik. *Support-vector networks*. Machine learning, 20(3) :273–297, 1995.
- [4] T. Cover et P. Hart. *Nearest neighbor pattern classification*. IEEE Transactions on Information Theory, 13(1) :21–27, 1967.
- [5] A. Frome, Y. Singer, et J. Malik. *Image Retrieval and Classification Using Local Distance Functions*. Advances in Neural Information Processing Systems, 2007.
- [6] R. K. Gupta, A. Y.-S. Chia, D. Rajan, E. S. Ng, et H. Zhiyong. *Image colorization using similar images*. ACM International Conference on Multimedia, pages 369–378, 2012.
- [7] M. Herold, X. Liu, et K. C. Clarke. *Spatial metrics and image texture for mapping urban land use*. Photogrammetric Engineering & Remote Sensing, 69(9) :991–1001, 2003.
- [8] A. Levin, D. Lischinski, et Y. Weiss. *Colorization using optimization*. ACM Transactions on Graphics, 23(3) :689–694, 2004.
- [9] S. Mahamud et M. Hebert. *The optimal distance measure for object detection*. IEEE Conference on Computer Vision and Pattern Recognition, 2003.
- [10] T. Mäenpää. *The local binary pattern approach to texture analysis - extensions and applications*. PhD thesis, 2003. Dissertation. Acta Univ Oul C 187.
- [11] F. Pierre, J.-F. Aujol, A. Bugeau, N. Papadakis, et V.-T. Ta. *Luminance-chrominance model for image colorization*. SIAM Journal on Imaging Sciences, To appear, 2015.
- [12] K. Q. Weinberger, J. Blitzer, et L. K. Saul. *Distance metric learning for large margin nearest neighbor classification*. Advances in neural information processing systems, pages 1473–1480, 2005.
- [13] T. Welsh, M. Ashikhmin, et K. Mueller. *Transferring color to greyscale images*. ACM Transactions on Graphics, 21(3) :277–280, 2002.
- [14] L. Yatziv et G. Sapiro. *Fast image and video colorization using chrominance blending*. IEEE Transactions on Image Processing, 15(5) :1120–1129, 2006.