

Détection Statistique de Similarité d'images basée sur la Carte des Dissimilarités Locales

Moustapha DIAW¹, Florent RETRAINT², Agnès DELAHAIES¹, Jérôme LANDRÉ¹, Frédéric MORAIN-NICOLIER¹

¹Laboratoire CReSTIC, IUT TROYES
9 Rue Quebec, 10430 Rosières-prés-Troyes, France

²Laboratoire LIST3N, UTT, TROYES
12 Rue Marie Curie, 10300 Troyes, France

moustapha.diaw@univ-reims.fr, florent.restraint@utt.fr

Résumé – En raison de l'augmentation considérable de l'utilisation des images numériques dans la vie quotidienne, de nombreuses applications nécessitent une étude sur leur similarité. Le principal défi est de trouver une méthode simple et efficace pour comparer et classer des paires d'images en classes similaires et dissimilaires. Dans cet article, nous introduisons une nouvelle technique de comparaison de paires d'images basée sur la Carte des Dissimilarités Locales (CDL). La distribution de Weibull est utilisée pour modéliser la distribution empirique des niveaux de gris de la CDL. Nous caractérisons la CDL par le paramètre de forme et le paramètre d'échelle de la distribution de Weibull. Ces deux paramètres sont pertinents pour discriminer les CDL des paires d'images. Un détecteur statistique basé sur un test d'hypothèse est également introduit pour classer les paires d'images similaires et dissimilaires. Les résultats expérimentaux confirment l'utilité et l'efficacité du détecteur proposé. Notre approche donne une Aire Sous la Courbe de 95.01% pour des images réelles.

Abstract – Due to the considerable increase of the use of digital images in every day life, many applications require a study on their similarity. The main challenge is to find a simple and efficient method to compare and classify image pairs into similar and dissimilar classes. In this paper, we introduce a new technique for image pairs comparison based on the Local Dissimilarity Map (LDM). The Weibull distribution is used to model the empiric gray level distribution of the LDM. We then characterize the LDM by the shape parameter and the scale parameter of the Weibull distribution. In fact, these two parameters are relevant to discriminate the LDM of image pairs. A statistical detector based on the hypothesis testing framework is also introduced to classify image pairs into similar and dissimilar. Experimental results confirm the convenience and the effectiveness of the proposed detector. Our approach gives an Area Under Curve of 95.01% for real images.

1 Introduction

En raison de l'augmentation considérable de l'utilisation des images numériques dans la vie quotidienne, de nombreuses applications nécessitent une étude sur leur similarité. L'objectif de ce travail est donc de proposer une méthode de comparaison et de classification de paires d'images. Dans cet article, nous nous intéressons aux images contenant des structures. Dans les années antérieures, de nombreuses mesures de similarités ont été utilisées dans le domaine du traitement d'images (comparaison, compression, qualité, etc). Parmi elles, on trouve l'indice de similarité structurelle (SSIM) [1] et l'indice de similarité des caractéristiques (FSIM) [2].

Dans ce travail, nous nous sommes intéressés à la Carte des Dissimilarités Locales [3]. Elle permet une évaluation bien définie des dissimilarités locales entre les images, ce qui n'est pas le cas des autres méthodes. Cette mesure est construite sur des images binaires et une extension a été introduite pour l'adapter aux images en niveaux de gris [4]. Ensuite, une distribution statistique de Weibull a été proposée pour modéliser la distribution empirique des valeurs en niveaux de gris de la CDL [4].

Par ailleurs, des travaux antérieurs ont proposé des détecteurs statistiques basés sur des tests d'hypothèses pour classer des images. Dans [5], Nguyen *et al.* ont utilisé la loi gamma pour caractériser les textures des images. Des détecteurs statistiques basés sur des tests d'hypothèses sont également présentés pour classer les codes à barres imprimés authentiques et contrefaits.

Dans [6], Mahfoudi *et al.* ont proposé une méthode pour détecter la double compression vidéo afin de vérifier l'intégrité de la vidéo. La méthode se concentre sur les coefficients de la transformée en cosinus discret pour détecter une double compression. Ces coefficients sont ensuite caractérisés par une loi de Laplace et des tests d'hypothèses statistiques sont proposés pour déterminer si une vidéo a été compressée deux fois ou non.

Dans cet article, nous caractérisons la CDL par une loi de Weibull à deux paramètres et un test du rapport de vraisemblance est proposé pour déterminer si deux images sont similaires ou non.

Le reste de cet article est structuré comme suit. La section 2 introduit la CDL des images binaires et des images en niveaux de gris. La section 3 présente la distribution statistique utilisée

pour modéliser la CDL. La section 4 s'intéresse au détecteur statistique basé sur le test du rapport de vraisemblance. En section 5 sont présentés les résultats et les discussions. Enfin, la section 6 conclut notre article.

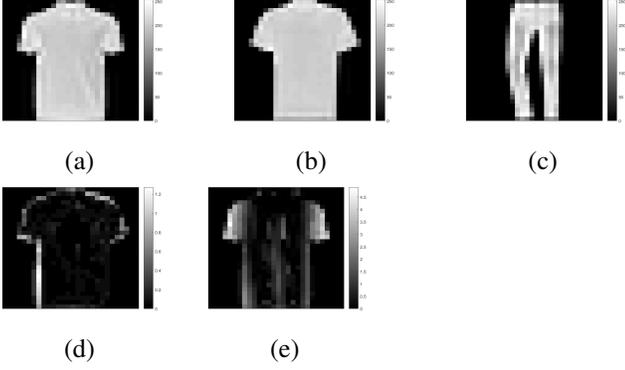


FIGURE 1 – (a), (b) et (c) sont des images en niveaux de gris. (d) CDL des deux images similaires (a) et (b) . (e) CDL des deux images dissimilaires (a) et (c).

2 Carte des Dissimilarités Locales

La Carte des Dissimilarités Locales (CDL) est une mesure basée sur la distance de Hausdorff qui permet de caractériser les différences locales entre deux images binaires [3]. Elle est définie par :

$$\text{CDL}(A_1, B_1)(p) = B_1(p) dt_{A_1}(p) + A_1(p) dt_{B_1}(p), \quad (1)$$

où A_1 et B_1 sont deux images binaires, $p \in A_1 \triangle B_1$ et dt_{A_1} est la transformée en distance euclidienne de A_1 .

Une extension de la transformée en distance euclidienne aux images en niveaux de gris a été proposée par Molchanov et Teran [7]. Le travail de Molchanov nous a permis d'adapter la CDL aux images en niveaux de gris. Celles-ci sont d'abord divisées en images binaires à l'aide de seuils. La CDL en niveaux de gris est ensuite calculée comme la somme des CDL des sous-images binaires. La CDL de deux images en niveaux de gris (GCDL) de \mathbb{R}^2 à \mathbb{R}^2 est définie par :

$$\text{GCDL}(A, B)(p) = \frac{1}{N} \sum_{i=1}^N \text{CDL}(A_i, B_i)(p), \quad (2)$$

où N est le nombre de seuils, t_i , $i = 1, \dots, N$ est le seuil et $A_i = \{p \in A : A(p) \geq t_i\}$. Chaque seuil t_i est choisi selon un espacement régulier entre $m = \max(\min(A), \min(B))$ et $M = \min(\max(A), \max(B))$. Cet espacement régulier a donc un pas $s = (M - m)/N$. Dans la suite, $N = 15$ est choisi qui est suffisant pour donner de bons résultats avec moins de temps de calcul, comparé à l'article original ($N = 255$). La Figure 1 montre deux exemples de CDL d'images en niveaux de gris.

3 Modèle de la Carte des Dissimilarités Locales

Dans cette section, le comportement de la distribution empirique des valeurs des niveaux de gris de la CDL est étudié. Dans [8], théorème 1, les auteurs ont montré que la norme L_ρ ($\rho \geq 1$) suit une distribution de Weibull sous certaines conditions. Le calcul de la CDL dépend du calcul de la transformée en distance. En outre, les valeurs de pixels d'une transformée en distance vérifient les conditions du théorème 1 dans [8] qui sont l'indépendance, la corrélation et une limite supérieure. D'après les équations (1) et (2) et pour $\rho = 2$, la transformée en distance, la CDL des images binaires ainsi que la CDL des images en niveaux de gris suivent la distribution de Weibull de densité probabilité et de fonction de répartition données par :

$$f(t) = \frac{\alpha}{\beta} \left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha} \quad (3)$$

et

$$F(t) = 1 - e^{-\left(\frac{t}{\beta}\right)^\alpha}, \quad (4)$$

respectivement, pour $t > 0$, $\alpha > 0$ and $\beta > 0$. Le paramètre de forme est α et le paramètre d'échelle est β .

Les deux paramètres de la loi de Weibull sont obtenus en utilisant l'estimation du maximum de vraisemblance.

4 Test du Rapport de Vraisemblance (TRV)

Nous avons montré dans la section précédente que les valeurs des niveaux de gris d'une CDL suivent une distribution de Weibull. Afin de vérifier si deux images sont similaires, nous allons proposer le test d'hypothèse suivant :

$$\begin{cases} H_0 : \mathbf{K} \sim \text{Weibull}(a, b_0) \\ H_1 : \mathbf{K} \sim \text{Weibull}(a, b_1), \end{cases} \quad (5)$$

où \mathbf{K} représente les valeurs des niveaux de gris de notre CDL et les paramètres a , b_0 , b_1 sont connus. Si deux images sont similaires, les valeurs des niveaux de gris de la CDL doivent suivre une distribution de Weibull de forme a et d'échelle b_0 , sinon elles vont suivre une distribution de Weibull de forme a et d'échelle b_1 . Nous définissons alors le rapport de log-vraisemblance comme suit :

$$\Lambda(\mathbf{K}) = \log \frac{\mathcal{L}_1(\mathbf{K})}{\mathcal{L}_0(\mathbf{K})} = \log \prod_{i=1}^k \frac{f_1(k_i)}{f_0(k_i)}, \quad (6)$$

où f_0 et f_1 sont respectivement les densités de \mathbf{K} sous H_0 et H_1 .

Le rapport de log-vraisemblance $\Lambda(\mathbf{K})$ est alors :

$$\Lambda(\mathbf{K}) = N \log\left(\frac{b_0^a}{b_1^a}\right) + \left(\frac{1}{b_0^a} - \frac{1}{b_1^a}\right) \sum_{i=1}^N k_i^a \quad (7)$$

$$= N \log\left(\frac{b_0^a}{b_1^a}\right) + \left(\frac{b_1^a - b_0^a}{b_1^a b_0^a}\right) \sum_{i=1}^N k_i^a. \quad (8)$$

Le Théorème Central Limite (TCL) donne :

$$\frac{1}{N} \sum_{i=1}^N k_i^a \sim \mathcal{N}(\mu, \sigma), \quad N \rightarrow \infty, \quad (9)$$

avec $\mu = \mathbf{E}(k_i^a)$ et $\sigma = \sqrt{\mathbf{Var}(k_i^a)}$

Sous les hypothèses H_r , $r \in \{0, 1\}$, k_i^a suit une loi exponentielle de paramètre b_r^{-a} . Nous avons alors $\mu = \mathbf{E}(k_i^a) = b_r^a$ et $\sigma = \sqrt{\mathbf{Var}(k_i^a)} = b_r^a$. Cela conduit à :

$$\sum_{i=1}^N k_i^a \sim \mathcal{N}(N b_r^a, \sqrt{N} b_r^a). \quad (10)$$

La combinaison des deux équations (8) et (10) implique sous H_r , $r \in \{0, 1\}$ que :

$$\Lambda_r(\mathbf{K}) \sim \mathcal{N}(\mu_r, \sigma_r), \quad (11)$$

avec

$$\mu_r = N \log\left(\frac{b_0^a}{b_1^a}\right) + N b_r^a \left(\frac{b_1^a - b_0^a}{b_1^a b_0^a}\right), \quad (12)$$

$$\sigma_r = \sqrt{N} b_r^a \left(\frac{b_1^a - b_0^a}{b_1^a b_0^a}\right). \quad (13)$$

On définit

$$\Lambda^*(\mathbf{K}) = \frac{\Lambda(\mathbf{K}) - \mu_0}{\sigma_0}. \quad (14)$$

Alors

$$\Lambda^*(\mathbf{K}) \sim \mathcal{N}\left(\frac{\mu_r - \mu_0}{\sigma_0}, \frac{\sigma_r}{\sigma_0}\right). \quad (15)$$

La statistique $\Lambda^*(\mathbf{K})$ suit donc une loi normale centrée réduite sous H_0 .

En vertu du lemme de Neyman-Pearson, le test δ le plus puissant pour le problème (5) est le test du rapport de vraisemblance :

$$\delta(\mathbf{K}) = \begin{cases} H_0 & \text{si } \frac{\mathcal{L}_1(\mathbf{K})}{\mathcal{L}_0(\mathbf{K})} < \tau \\ H_1 & \text{si } \frac{\mathcal{L}_1(\mathbf{K})}{\mathcal{L}_0(\mathbf{K})} \geq \tau. \end{cases} \quad (16)$$

On peut alors définir le test δ^* par :

$$\delta^*(\mathbf{K}) = \begin{cases} H_0 & \text{si } \Lambda^*(\mathbf{K}) < \tau^* \\ H_1 & \text{si } \Lambda^*(\mathbf{K}) \geq \tau^*, \end{cases} \quad (17)$$

où τ^* est une constante définie mathématiquement en fonction de la probabilité de fausse alarme prescrite.

Soit α_0 le taux de fausse alarme, le seuil de décision et la fonction de puissance du test δ^* sont définis par :

$$\tau^* = \Phi^{-1}(1 - \epsilon_0), \quad (18)$$

$$\beta_{\delta^*} = 1 - \Phi\left(\frac{\mu_0 - \mu_1 + \tau^* \sigma_0}{\sigma_1}\right), \quad (19)$$

où Φ et Φ^{-1} sont la fonction de répartition de la loi normale centrée et réduite et sa réciproque.

5 Résultats et Discussion

5.1 Ensemble de données

L'ensemble de données F-MNIST est utilisé pour nos applications. Nous avons pris aléatoirement 90 images dans F-MNIST. Ce qui a produit 4005 CDL dont 630 sont des CDL de paires d'images similaires et 3375 sont des CDL de paires d'images dissimilaires. Deux images sont similaires si elles appartiennent à une même classe, sinon elles sont dissimilaires. Puisque la CDL suit une distribution de Weibull, elle sera caractérisée par les deux paramètres (α, β) dans la suite de ce document. La figure 2 montre un exemple de données issues de la base F-MNIST (2a) et la représentation des paramètres d'échelle β en fonction des paramètres de forme α obtenus sur les 4005 CDL (2b). Il est à noter que dans l'espace des paramètres, les paires d'images similaires et celles dissimilaires appartiennent à des clusters distincts.

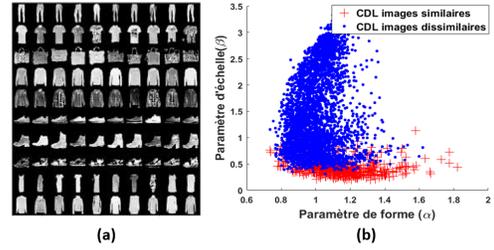


FIGURE 2 – Exemple de données issues de la base F-MNIST (a) et la représentation des paramètres β en fonction de α (b).

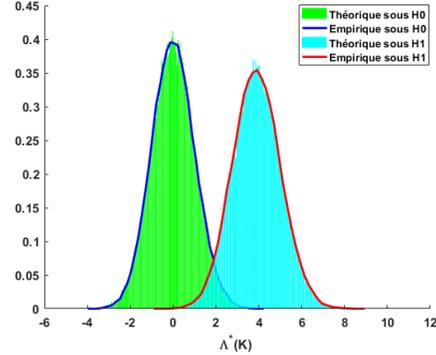


FIGURE 3 – Distributions théorique et empirique sous H_0 et H_1 avec $b_0 = 0.4$ et $b_1 = 0.6$.

5.2 Validation du détecteur

Pour s'assurer de la validité de notre détecteur TRV à classer des paires d'images en similaire et dissimilaire, nous l'avons d'abord testé sur des données simulées. Nous avons généré 1600 échantillons, une moitié pour l'hypothèse H_0 et l'autre moitié pour l'hypothèse H_1 . La figure 3 donne la comparaison entre la distribution théorique et la distribution empirique

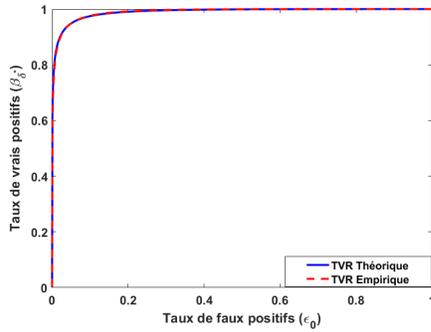


FIGURE 4 – Puissance théorique et empirique sous H_0 et H_1 avec $b_0 = 0.4$ et $b_1 = 0.6$.

sous H_0 et sous H_1 . On peut constater que la distribution empirique correspond bien à la distribution théorique donnée par l'équation (15) sous les 2 hypothèses. Cette même correspondance est également observée dans la figure 4 où la puissance théorique et la puissance empirique du test sont représentées en fonction du taux de fausse alarme α_0 . Pour $a = 1.5$, $b_0 = 0.4$ et $b_1 = 0.6$, la courbe ROC théorique et la courbe ROC empirique donnent des Aires Sous la Courbe (ASC) très proches, égales à 98%. La sous-section suivante présente l'application du test d'hypothèse à des données réelles.

5.3 Performances sur des données réelles

Pour évaluer la performance de notre méthode, nous l'avons testé sur les valeurs de pixels des 4005 CDL des images de la base de données F-MNIST. La figure 5 donne la puissance du test en fonction du taux de fausse alarme pour notre approche de détection (TVR-CDL), notre approche avec apprentissage (SVM-CDL) et des méthodes de l'état de l'art (SVM-SSIM et SVM-FSIM). La validation croisée est appliquée pour apprendre et tester le modèle SVM. On peut remarquer que notre approche de détection donne la meilleure ASC (95.01%) en comparaison de la méthode d'apprentissage supervisé SVM-CDL (84.57%) et les autres méthodes de l'état de l'art (79.18% pour SVM-SSIM et 83.43% pour SVM-FSIM).

6 Conclusion et perspectives

Dans cet article, nous avons proposé une méthode basée sur la CDL pour détecter la similarité ou la dissimilarité entre deux images. Nous avons montré que la distribution empirique des valeurs en niveaux de gris des CDL est modélisée par une distribution de Weibull à deux paramètres. Le test du rapport de vraisemblance a été proposé pour déterminer si deux images sont similaires ou non. Nous avons montré que la puissance de détection de l'approche proposée était plus élevée que celles des autres méthodes de l'état de l'art.

Dans les travaux futurs, nous proposons d'utiliser le test du rapport de vraisemblance généralisé pour prendre en compte

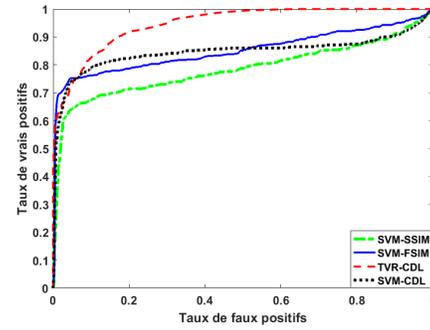


FIGURE 5 – Puissance du détecteur proposé (TRV) et des méthodes de l'état de l'art sur la base F-MNIST.

un paramètre b_1 inconnu.

Références

- [1] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli. *Image quality assessment : from error visibility to structural similarity*. Journal in IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [2] L. Zhang, L. Zhang, X. Mou and D. Zhang. *FSIM : A feature similarity index for image quality assessment*. IEEE transactions on image processing, vol.20, no.8, pp. 2378-2386, August 2011.
- [3] E. Baudrier, F. Nicolier, G. Millon, S. Ruan. *Binary-image comparison with local-dissimilarity quantification*. Pattern Recognition. Elsevier, vol.41 (5), pp. 1461-1478, 2008.
- [4] M. Diaw, A. Delahaies, J. Landré, F. Morain-Nicolier, and F. Retraint. *Fast process for classifying structural image pairs using Weibull parameters extracted from undersampled Local Dissimilarity Maps*. 2021 29th European Signal Processing Conference (EUSIPCO), pp. 631-635, doi : 10.23919/EUSIPCO54536.2021.9616133, 2021.
- [5] H. P. Nguyen, A. Delahaies, F. Retraint, F. Morain-Nicolier. *Face presentation attack detection based on a statistical model of image noise*. In IEEE Access, vol.7, pp.175429-175442, 2019.
- [6] G. Mahfoudi, F. Retraint, F. Morain-Nicolier and M. M. Pic. *Statistical H.264 Double Compression Detection Method Based on DCT Coefficients*. In IEEE Access, vol. 10, pp. 4271-4283, 2022, doi : 10.1109/ACCESS.2022.3140588.
- [7] I. S. Molchanov, P. Teran. *Distance transforms for real-valued functions*. Journal of Mathematical Analysis and Applications, vol. 278, pp. 472–484, 2003.
- [8] G. J. Burghouts, A. W. M. Smeulders, J. M. Geusebroek. *The distribution family of similarity distances*. 20th International Conference on Neural Information Processing, pp. 201-208, 2007.