

Détection d'objet par mesure de dissimilarités locales

Frédéric MORAIN-NICOLIER, Jérôme LANDRÉ, Su RUAN

Centre de Recherche en STIC (CReSTIC)
9 rue de Québec, IUT, 10026 Troyes Cedex, France
frederic.nicolier@univ-reims.fr

Résumé – Cette communication porte sur la recherche de la position d'un gabarit de référence dans une image, par mise en correspondance. Le détecteur proposé s'appuie sur des cartes de dissimilarités locales qui permettent de saisir les écarts entre deux images. La mesure de dissimilarités locales revient à prendre en compte la non-symétrie de la notion de similarité, ce qui permet une détection fine. Cette non-symétrie est reliée à la notion de similarité proposée dans un cadre plus général par Tversky. Le détecteur que nous proposons est rapide à calculer. Nous montrons sur un exemple qu'il renvoie potentiellement moins de faux positifs qu'une autre méthode de référence (*chamfer matching*). Des simulations confirment la bonne robustesse du détecteur.

Abstract – In previous works, we have proposed a local dissimilarity map (LDM) in order to compare images. In this research, we show how the LDM can be applied in the field of shape recognition. A global dissimilarity measure (GDM) is obtained from the LDM. This versatile measure allows to measure symmetric as well as asymmetric similarities. Its properties are related to the human similarity judgement from Tversky results. A matcher is derived by summing the values of the LDM. The obtained matcher is compared to the chamfer matching. A qualitative illustration shows the good behaviour of our matcher. In particular, less false-positives are observed. A more intensive test allows to conclude the good robustness of our matcher with respect to models degradations.

1 Introduction

La reconnaissance d'un objet dans une image constitue une étape majeure dans un processus d'interprétation visuelle. La reconnaissance de logos ou la compression vidéo sont quelques-unes des nombreuses applications de ce sous-ensemble des méthodes de mise en correspondance [6]. Le schéma classique d'une reconnaissance débute par une segmentation ou une binarisation permettant d'extraire les pixels d'un objet. Puis une comparaison d'attributs avec ceux d'un objet de référence conduit à une prise de décision. Une méthode éprouvée a été proposée par Borgefors : le *chamfer matching* [2]. L'algorithme consiste à minimiser une distance généralisée entre les contours d'un gabarit et d'une image.

Dans des travaux antérieurs [1], nous avons introduit une mesure de dissimilarité locale permettant la comparaison d'images binaires. Une carte de dissimilarité locale (CDL) est ainsi construite en utilisant une fenêtre glissante de taille auto-adaptative.

Nous proposons de construire un détecteur d'objet basé sur la CDL. La détection est effectuée à partir d'un gabarit. Ce gabarit est une image contenant un exemplaire de l'objet à trouver. Ce gabarit est comparé localement aux informations de l'image grâce à la CDL. Ce détecteur de dissimilarités locales est comparé au détecteur *chamfer matching* à l'aide un exemple illustratif et de données simulées .

2 Reformulation de la carte de dissimilarités locales

Parmi les distances entre ensembles, la distance de Hausdorff (DH) a souvent été utilisée en reconnaissance de formes [3]. La DH est définie par

$$D_H(A, B) = \max(h(A, B), h(B, A)), \quad (1)$$

avec $h(A, B) = \max_{a \in A} (\min_{b \in B} d(a, b))$ où d est une distance sous-jacente. Cette distance est globale. Nous avons proposé une mesure locale puis une carte de dissimilarités locales (CDL) construite à l'aide d'un algorithme itératif [1]. L'idée générale est que la taille d'une fenêtre locale croît jusqu'à être suffisamment grande pour saisir convenablement les caractéristiques locales. Nous avons également proposé une expression de la CDL qui permet un calcul rapide :

$$CDL_{A,B}(p) = |A(x) - B(x)| \max(td_A(x), td_B(x)), \quad (2)$$

td_X étant la transformée en distance de l'image X . Cette équation peut être simplifiée, pour des images binaires :

$$CDL_{A,B}(p) = Btd_A + Atd_B. \quad (3)$$

Cette nouvelle équation présente un grand intérêt car elle élimine l'opérateur \max et la valeur absolue. Il ne reste alors que des opérations linéaires.

Démonstration. La preuve est assez rapide. CDL_1 (resp. CDL_2) est la CDL issue de l'éq. 2 (resp. eq. 3). Pour un pixel p , les alternatives sont :

1.

$$\begin{aligned} A(p) = B(p) &\Rightarrow A(p) - B(p) = 0 \\ &\Rightarrow \text{CDL}_1(p) = \text{CDL}_2(p) = 0 \end{aligned}$$

2.

$$\begin{aligned} A(p) = 1 \text{ et } B(p) = 0 &\Rightarrow A(p) - B(p) = 1 \\ A(p) = 1 &\Rightarrow \text{td}_A(p) = 0 \text{ (par définition)} \\ \Rightarrow \text{CDL}_1(p) &= \max(\text{td}_A(p), \text{td}_B(p)) \\ \Rightarrow \text{CDL}_1(p) &= \text{td}_B(p) \\ \text{CDL}_2(p) &= B(p)\text{td}_A(p) + A(p)\text{td}_B(p) \\ &= \text{td}_B(p) \\ \Rightarrow \text{CDL}_1(p) &= \text{CDL}_2(p) \end{aligned}$$

3. $A(p) = 1$ et $B(p) = 0$: voir 2. □

La transformée en distance associée à chaque pixel nul d'une image, la distance au pixel non-nul le plus proche. Elle peut être très rapidement calculée par l'utilisation de la distance du chanfrein, qui représente une bonne approximation de la distance euclidienne. Elle permet en outre de calculer la transformée en distance, en ne faisant parcourir à un masque de taille réduite, que deux fois l'image [4].

3 Utilisation de la CDL comme détecteur de motifs

L'objet est représenté par son gabarit, une image binaire G , et l'image I contenant des objets est une image binaire. Le problème est de savoir si G est présent autour d'un pixel (x, y) de I . Nous désignons par $I_{(x,y)}$ la portion de I , centrée en (x, y) et de taille identique à G . Il s'agit alors de construire un indicateur de similarité entre $I_{(x,y)}$ et G . La CDL étant représentative de la dissimilarité entre deux images, nous choisissons comme détecteur la somme des distances de la CDL entre $I_{(x,y)}$ et g . Si cette somme vaut zéro, cela signifie que les deux images sont identiques et que le gabarit est présent à la position (x, y) . Plus la valeur du détecteur est grande, plus les pixels du gabarit et de I autour de (x, y) sont distincts. Le détecteur ainsi décrit est (à partir de l'éq. 3) :

$$\begin{aligned} D_{I,G}(x, y) &= \sum_k \sum_l \text{CDL}_{I_{(x,y)}, G}(k, l) \quad (4) \\ &= \sum_{k,l} G(p)\text{td}_{I_{(x,y)}}(p) \\ &\quad + \sum_{k,l} I_{(x,y)}(p)\text{td}_G(p) \quad (5) \end{aligned}$$

Cette équation peut s'interpréter comme un *chamfer-matching* symétrique. Le *chamfer-matching* consiste à chercher la minimisation d'une distance généralisée entre deux ensembles

de points de contours [2]. La gabarit M est traduit pour parcourir td_I , la transformée en distance de I . A chaque translation, le gabarit G_t est superposé à td_I . La moyenne des valeurs de distances touchées par le gabarit G_t fournit une valeur représentative (*chamfer-score* : CS) de la ressemblance (locale). Ainsi pour une translation t donnée :

$$\text{CS}(I, G) = \frac{1}{N} \sum_p G(p)\text{td}_I(p). \quad (6)$$

CS mesure comment l'image est dissimilaire au gabarit G , ou encore comment l'image ressemble au gabarit. Par une comparaison des équations (4) et (6), il est assez clair que $D_{I,G}$ mesure comment G est similaire à I , mais aussi comment I est similaire à G . Tversky a montré que le jugement humain de la similarité est asymétrique [5]. Par exemple, la similarité de la Corée du Nord à la Chine n'est pas la même que la similarité de la Chine à la Corée du Nord. La mesure globale que nous proposons en (4) est symétrique. L'introduction de pondérations $\alpha, \beta \in [0, 1]$ conduit à une formulation flexible, la mesure de dissimilarité globale (MDG) :

$$\begin{aligned} \text{MDG}(A, B) &= \alpha \sum_p B(p)\text{td}_A(p) \\ &\quad + \beta \sum_p A(p)\text{td}_B(p). \quad (7) \end{aligned}$$

Une mesure symétrique est obtenue avec $\alpha = \beta$. Des mesures asymétriques sont obtenues dans les autres cas. En particulier, $\alpha = 1$ et $\beta = 0$ permet de retrouver l'équation du *chamfer-matching* (eq. 6).

Un raffinement peut-être apporté en suivant la suggestion de [2] sur le choix d'une somme quadratique des distances mesurées :

$$\begin{aligned} \text{MDGq}(A, B) &= \alpha \sum_p B(p)\text{td}_A^2(p) \\ &\quad + \beta \sum_p A(p)\text{td}_B^2(p). \quad (8) \end{aligned}$$

Si l'on souhaite localiser un gabarit G au sein d'une image I , il est donc nécessaire de calculer la dissimilarité entre chaque translation de G et la portion d'image correspondante de I , ce qui a été décrit en début de section. Ainsi une version améliorée de l'équation (4), basée sur la mesure de dissimilarité que nous venons de mettre au point, est :

$$D_{I,G}(x, y) = \text{MDGq}(I_{(x,y)}, M). \quad (9)$$

Il est clair que cette dernière équation n'est ni plus ni moins que la somme de deux inter-corrélations :

$$D_{I,G} = \alpha(\text{td}_I^2 \star G) + \beta(I \star \text{td}_G^2). \quad (10)$$

Les images étant binaires, on a $A^2 = A$. Ainsi une implémentation très-rapide de la mesure globale est réalisable en passant dans le domaine de Fourier. Le choix des valeurs de α et β est conditionné au besoin de l'application proprement dite.

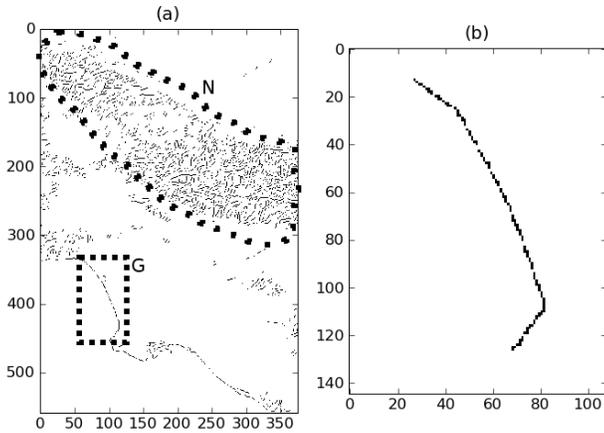


FIG. 1 – (a) Une image test. – (b) un gabarit test, la position idéale est étiquetée par la lettre G.

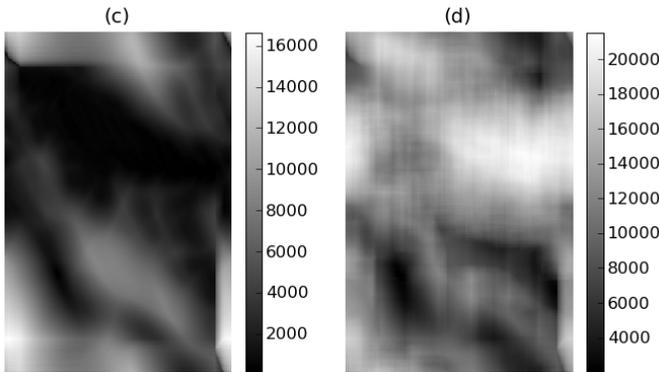


FIG. 2 – En (c), l’image obtenue par le détecteur *chamfer matching*. – En (d), celle obtenue avec le détecteur de dissimilarités locales. – Une bonne correspondance avec le gabarit est indiquée par des valeurs faibles, en foncé.

4 Un illustration et un test de robustesse du détecteur

Nous présentons ici une comparaison de notre détecteur, dit de dissimilarités locales (DL), et du détecteur de Borgefors. Dans l’équation (10), le détecteur de dissimilarités locales correspond au cas $\alpha = \beta = 1$ et le détecteur de Borgefors à $\alpha = 0, \beta = 1$. Les deux détecteurs sont ensuite testés sur leur robustesse à une dégradation de la qualité du gabarit.

Exemple illustratif Un exemple concret permet d’abord d’identifier le comportement des deux détecteurs. La figure 1(a) représente les contours d’une image d’une côte. La figure 1(b) est le gabarit du motif recherché. La position idéale est marquée par G dans la figure 1(a). Les figures 2(c) et 2(d) représentent respectivement les images obtenues grâce au détecteur de Borgefors et au détecteur DL. Une bonne correspondance se

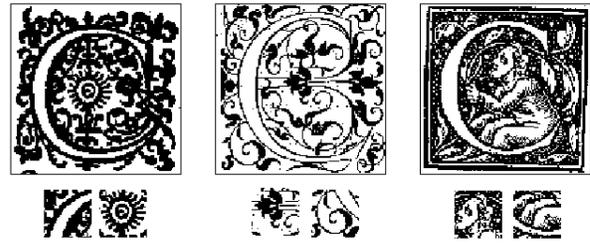


FIG. 3 – Trois images tests et quelques exemples de gabarits extraits.

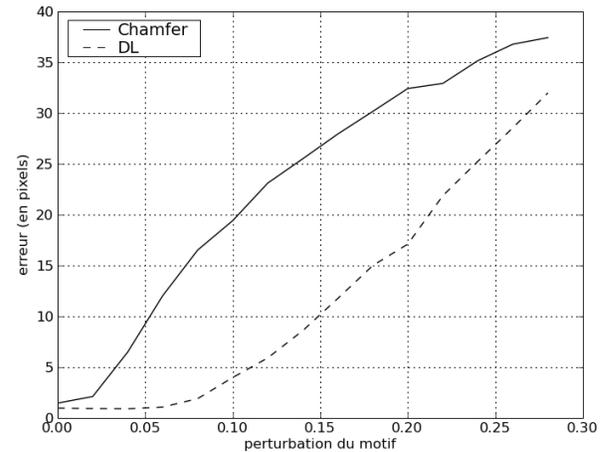


FIG. 4 – Courbes de robustesse des détecteurs *chamfer matching* et par dissimilarités locales (DL). Les courbes donnent l’erreur de détection (en localisation) en fonction du pourcentage de perturbation du gabarit extrait de l’image.

traduit par une valeur faible (en foncé).

Pour les deux détecteurs, le minimum est atteint pour le pixel de coordonnées $x = 88, y = 398$, ce qui correspond bien à la position idéale. Cependant, pour le détecteur de Borgefors, des petites valeurs sont trouvées pour la zone marquée N. Des petites valeurs indiquent la présence d’une variante du gabarit. L’interprétation est la suivante : parmi tous les points de cette zone, il existe un sous-ensemble de points qui correspond au gabarit cherché, la valeur du détecteur est donc faible. Les valeurs de la zone N peuvent être interprétées comme des faux-positifs potentiels, puisque correspondant à un minimum local du détecteur.

Le détecteur symétrique est plus discriminant, *c.f.* fig. 2(d). La zone N n’est pas détectée positivement du fait de la prise en compte de la double-similarité. De plus la position idéale est bien retrouvée.

Test de robustesse Nous proposons un test de la robustesse du détecteur DL en exploitant trois images de lettrines faisant

partie d'un des projets du laboratoire, *c.f.* fig. 3(a). Ces trois images sont de dimensions 256×256 . De chaque image, des gabarits sont extraits. L'opération consiste à extraire une image de taille 51×51 centrée en une position déterminée aléatoirement. Ce gabarit est alors perturbé en inversant un sous-ensemble de ses pixels. Les détecteurs *chamfer matching* et DL sont alors appliqués. La position estimée de l'objet est alors déterminée en recherchant le minimum absolu de l'image renvoyée par le détecteur. L'erreur est calculée en mesurant la distance entre la position réelle de l'objet (connue) et sa position estimée par le détecteur. Pour chaque image, un millier de gabarits est ainsi généré.

La figure 4(b) illustre la variation de la moyenne des erreurs obtenues pour les trois images, pour des perturbations de gabarit entre 0 et 30% des pixels. L'erreur *chamfer matching* devient assez importante alors que l'erreur du détecteur DL stagne à 1 pixel pour des perturbations inférieures à 7%. Au delà de 10% les deux détecteurs ne sont plus capables de retrouver la position du motif initial. Sur la base de ces simulations, le détecteur DL est significativement plus robuste que le *chamfer matching*.

5 Conclusion et perspectives

Nous avons proposé et testé un détecteur d'objet issu de la mesure de dissimilarités locales entre une image et un gabarit. Par rapport au *chamfer matching*, le détecteur renvoie potentiellement moins de faux positifs, ce qui est confirmé par des simulations. Nous prévoyons pour la suite des tests dans des conditions réelles avec vérités terrain, dans le cadre de la reconnaissance de logos.

Références

- [1] E. Baudrier, F. Nicolier, G. Millon, S. Ruan, "Binary-image comparison with local-dissimilarity quantification", *Pattern Recognition*, vol. 41, n. 5, pp. 1461–1478, jan. 2008
- [2] G. Borgefors, "Hierarchical chamfer matching : a parametric edge matching algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, n. 6, pp. 849–865, 1988
- [3] D.P. Huttenlocher, W.J. Rucklidge, "Comparing images using the hausdorff distance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, n. 9, pp. 850–863, 1993
- [4] U. Montanari, "A method for obtaining skeletons using a quasi-Euclidean distance", *Journal of the Association for Computing Machinery*, vol. 15, pp. 600-624, 1968.
- [5] A. Tversky, "Features of similarity", *Psychological Review*, vol. 84, n. 4, pp. 327-352, 1977
- [6] B. Zitova, J. Flusser, "Image registration methods : a survey", *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003