Le tatouage de documents numériques
Cours 2

Marc Chaumont

12 novembre 2008

Plan

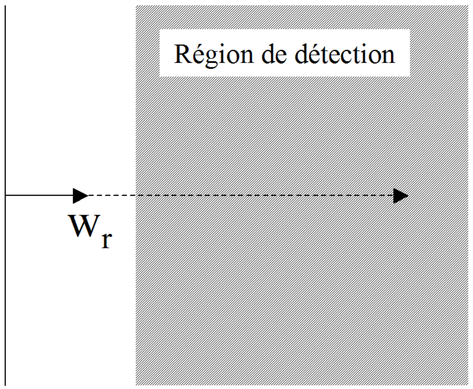
Définition

La corrélation linéaire entre deux vecteurs c_{w_n} et w_r de taille N est :

$$z_{lc}(c_{w_n}, w_r) = \frac{1}{N} c_{w_n}.w_r = \frac{1}{N} \sum_{i=1}^{N} c_{w_n}[i].w_r[i]$$

Interprétation géométrique

Lorsque la corrélation entre un vecteur c_{w_n} et une marque w_n dépasse un seuil τ, on dira que les deux vecteurs sont correlés. On peut alors voir la *région de détection* comme un **hyper-plan** perpendiculaire au vecteur marque. Tout signal c étant dans la région de détection est considéré comme tatoué.



Les défauts de la corrélation linéaire

Un des problèmes de la correlation linéaire est que le seuil de détection est fortement dépendant de l'amplitude des vecteurs. La mesure de corrélation linéaire est donc **non robuste aux attaques valumetriques** (exemple : reduction de la luminosité).

Un autre problème réside dans la difficulté à modéliser la probabilité de *random-work* faux positif (même quand la marque suit une distribution blanche Gaussienne).

Plan

Définitions

La corrélation normalisée permet de résoudre les problèmes liés à la corrélation linéaire. Le vecteur c_{w_n} et le vecteur w_r sont normalisés à 1 avant d'effectuer le produit scalaire :

$$z_{nc}(c_{w_n}, w_r) = \frac{c_{w_n}}{|c_{w_n}|}.\frac{w_r}{|w_r|} = \frac{1}{|c_{w_n}||w_r|}\sum_{i=1}^{N} c_{w_n}[i].w_r[i]$$

Le produit scalaire entre deux vecteurs c_{w_n} et w_r est égal au produit des longueurs des vecteurs par le cosinus de l'angle θ entre les deux vecteurs : $c_{w_n}.w_r = |c_{w_n}||w_r|cos(\theta)$. Ainsi la correlation normalisée est égale au cosinus de l'angle entre c_{w_n} et w_r :

$$z_{nc}(c_{w_n}, w_r) = cos(\theta)$$

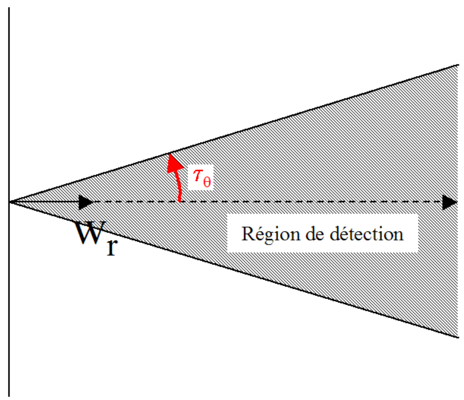
Interprétation géométrique

Lorsque la corrélation normalisée entre un vecteur c_{w_n} et une marque w_r dépasse un seuil τ, on dira que les deux vecteurs sont corrélés. On peut alors voir la *région de détection* comme un hyper-cone. Appliquer un seuil sur la corrélation est équivalent à appliquer un seuil sur l'angle entre les deux vecteurs c_{wn} et w_n :

$$\frac{c_{w_n}.w_r}{|c_{w_n}|.|w_r|} > \tau_{nc} \Leftrightarrow \theta < \tau_\theta$$

avec $\tau_\theta = cos^{-1}(\tau_{nc})$.

Interprétation géométrique



Mesure équivalente

De manière proche, certains auteurs utilisent la mesure suivante :

$$z_1(c_{wn}, w_r) = \frac{c_{wn}.w_r}{|c_{wn}|}$$

qui est équivalente à la mesure de corélation normalisée à un facteur près (la norme du vecteur w_r).

Plan

1. Les différentes mesures de corrélation
   - La corrélation linéaire
   - La corrélation normalisée
   - Coefficient de corrélation

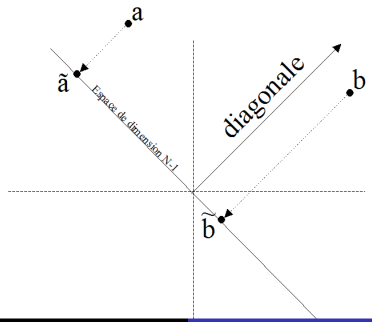
Définition

Le coefficient de corrélation est obtenu en soustrayant la moyenne de c_{w_n} (resp. w_r) à c_{w_n} (resp. w_r) et ensuite de calculer la corrélation normalisée :

$$\begin{aligned} \tilde{c}_{w_n} &= c_{w_n} - \overline{c}_{w_n} \\ \tilde{w}_r &= w_r - \overline{w}_r \\ z_{cc}(c_{w_n}, w_r) &= z_{nc}(\tilde{c}_{w_n}, \tilde{w}_r) \end{aligned}$$

Interprétation géométrique

L'interprétation géométrique du *coefficient de corrélation* dans espace de dimension N est la même que l'interprétation géométrique de la *corrélation normalisée* dans un espace $N-1$. La soustraction de la moyenne est en effet équivalente à projeter les vecteurs sur un vecteur diagonal.
Ci-dessous, l'illustration du passage 2D vers 1D obtenue en soustrayant à a (resp. b) sa moyenne.



Mesure équivalente

De manière proche, certains auteurs utilisent la mesure suivante :

$$z_2(c_{w_n}, w_r) = \frac{c_{w_n}.w_r}{s_{c_{w_n}}}$$

avec $s_{c_{wn}}$ l'ecart-type du vecteur c_{wn} :

$$s_{c_{wn}} = \sqrt{\frac{1}{N}\sum_i^N (c_{w_n}[i] - \overline{c}_{w_n}[i])^2} = \frac{1}{\sqrt{N}}|\tilde{c}_{wn}|$$

On a alors :

$$z_2(c_{wn}, w_r) = \sqrt{N}\frac{c_{wn}.w_r}{|\tilde{c}_{wn}|}$$

Si w_r a une moyenne nulle on a $w_r = \tilde{w}_r$ et $c_{wn}.\tilde{w}_r = \tilde{c}_{wn}.\tilde{w}_r$[1]. La mesure z_2 est alors équivalente au coefficient de corrélation au facteur $\sqrt{N}|\tilde{w}_r|$ près.

[1] $(x-\overline{x}).y = x.y - \overline{x}.\sum_i y[i] = x.y - \overline{x}.N.\overline{y} = x.y$ si $\overline{y} = 0$