

Une Statistique pour l'analyse de données multidimensionnelles

Application au filtrage d'images couleurs

C. de Runz¹, M. Herbin², F. Blanchard¹, L. Hussenet¹, V. Vrabie², P. Vautrot¹

1: CReSTIC-SIC, IUT de Reims Chalons Charleville, Rue des Crayères, BP 1035, 51687 Reims cedex 2, France
 tél : +33 3 26 91 84 58, email : cyril.de-runz@univ-reims.fr

2: Antenne CReSTIC-Châlons, chaussée du port, BP 541, 51012 Châlons-en-Champagne Cedex, France,
 tél : +33 3 26 91 81 98, email : michel.herbin@univ-reims.fr

Problème traité

L'obtention du meilleur représentant d'un échantillon est une question ouverte et dépend de la statistique utilisée dans le cadre de l'analyse multidimensionnelle.

Originalité

Nous proposons une statistique, dont le calcul repose sur les statistiques de rangs, et qui permet en outre d'introduire la notion de "données représentant le mieux son échantillon"

Résultat

En appliquant notre statistique au filtrage d'images couleurs, nous confirmons l'efficacité et l'intérêt de notre procédure visuellement et numériquement.

Principe

Soit un échantillon (x_1, x_2, \dots, x_n)

1. On détermine les **statistiques d'ordres** induites par les dissimilarités $d(x_i, x_j)$ et on note $Rg_{x_i}(x_j)$ le rang de la donnée x_j dans le classement par dissimilarité croissante à x_i .

2. On calcule, pour chaque donnée x_j , la moyenne des rangs $Rg_{x_i}(x_j)$ qu'elle a obtenu dans ces n classements. On note ce rang moyen $Rg(x_j)$ et on a :

$$Rg(x_j) = \frac{\sum_i Rg_{x_i}(x_j)}{n}$$

3. On extrait la donnée ayant le plus petit rang moyen, c'est à dire la donnée de l'échantillon la plus **représentative** dudit échantillon.

Nous avons donc défini une statistique dite de « Meilleur Rang Moyen » (MRM) qui associe à un échantillon l'élément qui le représente le mieux :

$$MRM : (x_1, x_2, \dots, x_n) \rightarrow \arg \min_{x_i, i=1..n} Rg(x_i)$$

Cette statistique est un **estimateur robuste de position** de l'échantillon qui fournit un **élément typique et représentatif de l'échantillon**.

Exemple



Image originale



Image Bruitée (5%)



Filtre Vecteur Médian

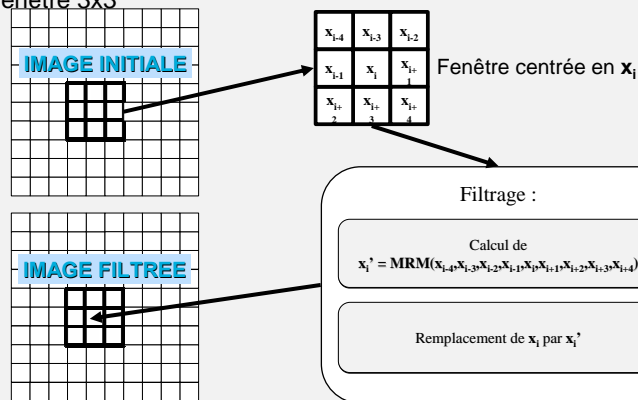


Filtre MRM

Application : Filtrage d'images couleurs

Principe : remplacer chaque pixel x_i par le pixel le plus représentatif de la fenêtre centrée en x_i .

Exemple : processus de filtrage d'une image couleur avec une fenêtre 3x3



Résultats

| Image | Bruit (%) | MAE Vecteur Median | MAE FMRM | MSE Vecteur Median | MSE FMRM |
|-------|-----------|--------------------|----------|--------------------|----------|
| House | 5 | 5.17 | 5.08 | 95.87 | 86.95 |
| House | 10 | 5.18 | 5.32 | 96.32 | 97.16 |
| Lena | 5 | 8.16 | 8.15 | 223.81 | 212.13 |
| Lena | 10 | 8.24 | 8.39 | 224.93 | 216.80 |

Conclusion

Nous avons construit une nouvelle statistique utilisant les statistiques de rangs. Ce nouvel outil permet de désigner, dans un échantillon, l'élément **représentant le mieux** cet échantillon.

La **robustesse** est une autre caractéristique de notre méthode. **Résistante aux valeurs extrêmes**, elle ne nécessite **pas non plus d'a priori sur la distribution** des données initiales.

Références

J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. In *Proceedings of the IEEE*, volume 78, pages 678–689, 1990.

H. A. David and H. N. Nagaraja. *Order Statistics*. Wiley, third edition, 2003.

P. Vautrot, L. Hussenet, and M. Herbin. A robust filtering method using owa filters : Application to color images. In *3rd European Conference on Colour in Graphics, Imaging and Vision*, University of Leeds, UK, jun 2006.