

STRUCTURATION ET ANALYSE DE LA BASE DE DONNÉES

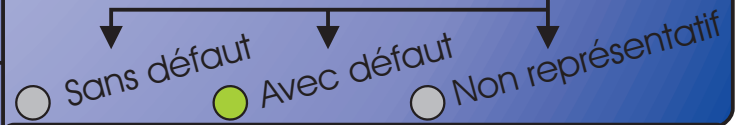
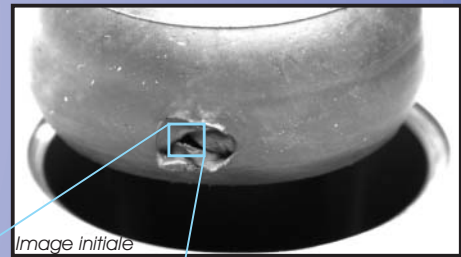
Présentation de modules opérationnels de prétraitement de l'information

La structuration et l'analyse de la base de données constituent une étape préalable à toute action neuronale. En optimisant la représentativité des données pour le problème spécifique posé, cette étape contribue à la qualité des résultats obtenus.

I - Construction active de la base de données

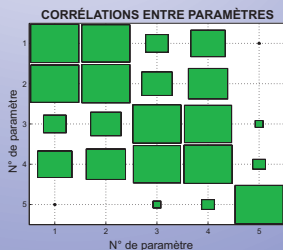
La qualité de la base de données constitue un élément clef dans la résolution du problème posé (en plus de présenter un intérêt fort pour la capitalisation des connaissances). Différents outils d'aide à la construction de la base de données ont été développés, notamment pour la vision artificielle (cas présenté ici) :

- définition d'une fenêtre mobile d'inspection de l'image
- positionnement de la fenêtre sur la zone d'intérêt
- détermination de la matrice correspondant au secteur inspecté
- extraction primaire de l'information ; détection des défauts
- visualisation de l'information extraite
- intégration de la décision du spécialiste en procédé
- enregistrement de la base de données

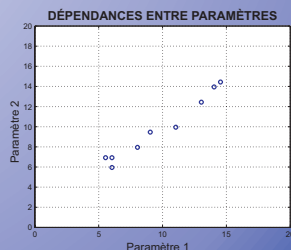


II - Analyse des corrélations et dépendances

L'analyse des corrélations et des dépendances multiparamétriques donne une indication de la qualité de la base de données. Les résultats de cette analyse peuvent amener à modifier la base de données pour en augmenter la représentativité (par exemple : suppression des données non pertinentes ou redondantes).

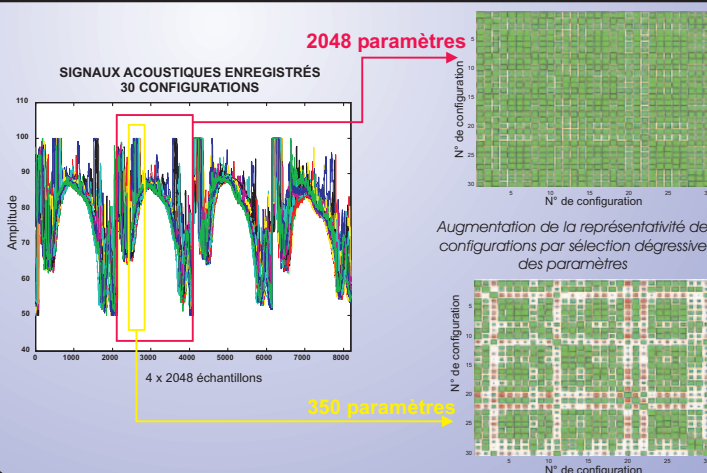


Corrélations entre paramètres (exemple de 5 paramètres)



Mise en évidence de la dépendance linéaire entre les paramètres 1 et 2

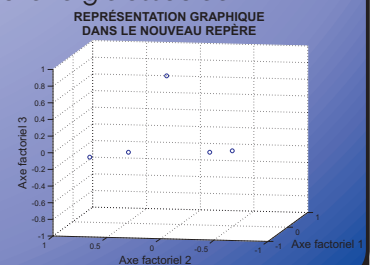
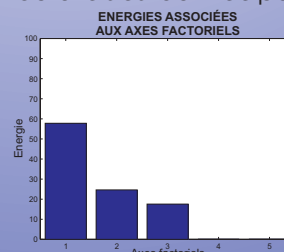
IV - Sélection dégressive des paramètres



III - Analyse en composantes principales (ACP)

L'Analyse en Composantes Principales est une méthode systématique permettant d'identifier des groupements dans un échantillon multidimensionnel. L'ACP crée de nouvelles variables (les composantes principales) synthétisant les principales variables d'origine, et fournit une représentation graphique permettant une interprétation facilitée des résultats. Sa mise en oeuvre est la suivante :

- recherche d'un nouveau repère aux dimensions réduites restituant le maximum d'information (changement d'axes),
- les axes du nouveau repère (axes factoriels) s'expriment comme une combinaison linéaire des axes d'origine et ne sont pas corrélés entre eux,
- la qualité de la représentation de chacun des axes factoriels est donnée par leur énergie associée.



L'analyse des données montre que la considération de l'ensemble de la base peut, dans certains cas, être nuisible à l'établissement de corrélations données / situation physique. Une solution consiste à segmenter la base de données afin d'identifier les secteurs représentatifs.

