

approches afin de rendre la méthode de détection plus robuste. De plus, il est intéressant de combiner les approches par dictionnaire définissant des modèles sur le patch global et les approches ponctuelles définissant des modèles localement sur les patches.

Les différentes approches combinées sont :

- la NMF (\hat{Y}_{f,N_2}^{NMF})
- les curvelets (\hat{Y}_{f,N_2}^C)
- l'estimation de densité indépendante par noyau gaussien ($\hat{Y}_{f,N_2}^{K_N}$)
- l'estimation de densité conditionnelle par rapport à tous les voisins d'ordre 1 ($\hat{Y}_{f,N_2}^{K_V}$)
- l'estimation de densité conditionnelle par rapport au voisinage directionnel ($\hat{Y}_{f,N_2}^{K_{\vec{V}}}$).

Nous définissons *methods* comme étant l'ensemble des différentes méthodes énoncées ci-dessus.

La décision à partir de cette combinaison d'approche correspond à un vote, un point est considéré comme inusuel si la majorité des approches le considère comme tel.

$$\hat{Y}_{f,N_2}^i = \begin{cases} 1 & \text{si } \frac{1}{\text{card}(\text{methods})} \sum_{k \in \text{methods}} \hat{Y}_{f,N_2}^{i,k} \geq 0.5 \\ 0 & \text{sinon} \end{cases}$$

Les figures 6.14, 6.15, 6.16, 6.17 et 6.18 donne les résultats de détections des différentes approches et de l'approche combinée pour les patches sur lesquels les résultats ont été illustrés tout au long du mémoire. Les différentes sous-images de chacune de ces figures correspondent au patch, à la décision effectuée à partir de chacune des approches ainsi qu'à la décision effectuée à partir d'un vote sur ces différentes approches. La figure 6.14 concerne un patch normal tandis que les autres figures contiennent différents types de signatures atypiques.

La combinaison des différentes approches de détection permet d'affiner la décision finale. Les différentes méthodes ne détectent pas à tort les mêmes points. Les signatures atypiques restent détectées par les différentes approches, mais des éléments supplémentaires sont également détectés comme des portions de raies pour la NMF ou des points mal représentés par les autres modèles tels que les points voisins aux signatures atypiques. Ce mélange des différentes méthodes permet de détecter plus précisément les points inusuels ainsi que la structure des signatures atypiques. Il semble être prometteur pour la détection des signatures atypiques sur les spectrogrammes.

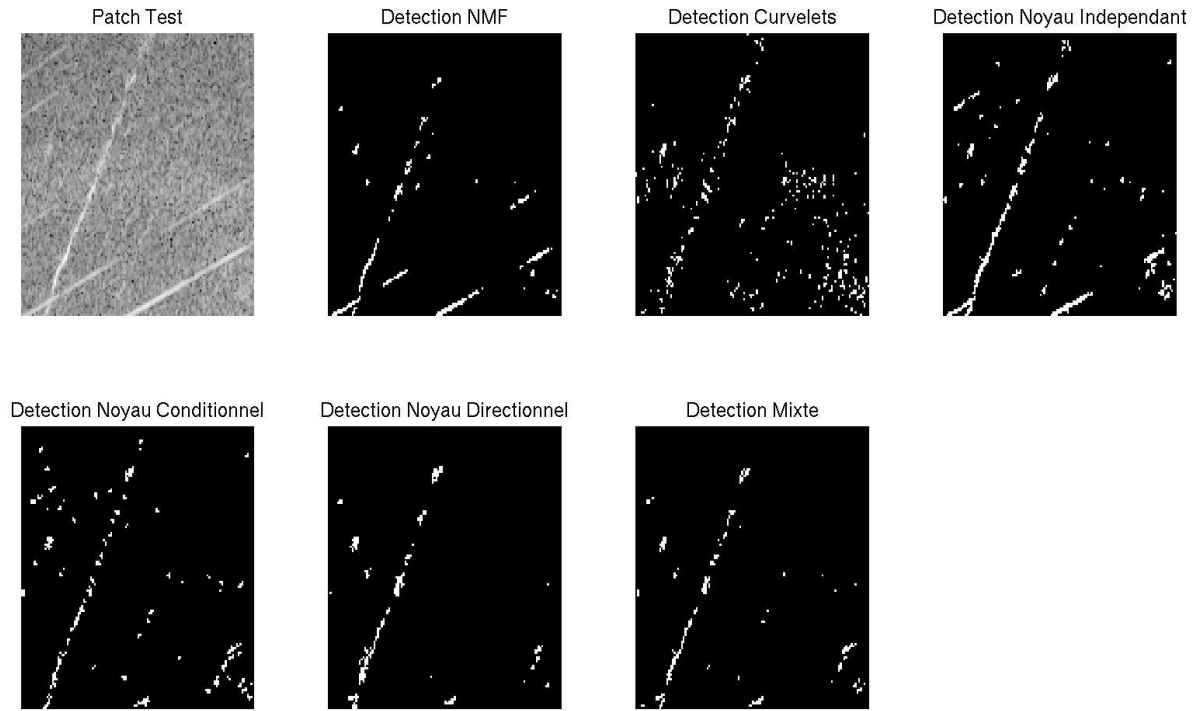


FIGURE 6.14 – Détection du patch sans signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{v}}}$, ainsi que de l'approche mixte \hat{Y} . La seule signature détectée sur ce patch correspond à la raie N_1 se décalant sur les différents spectrogrammes.

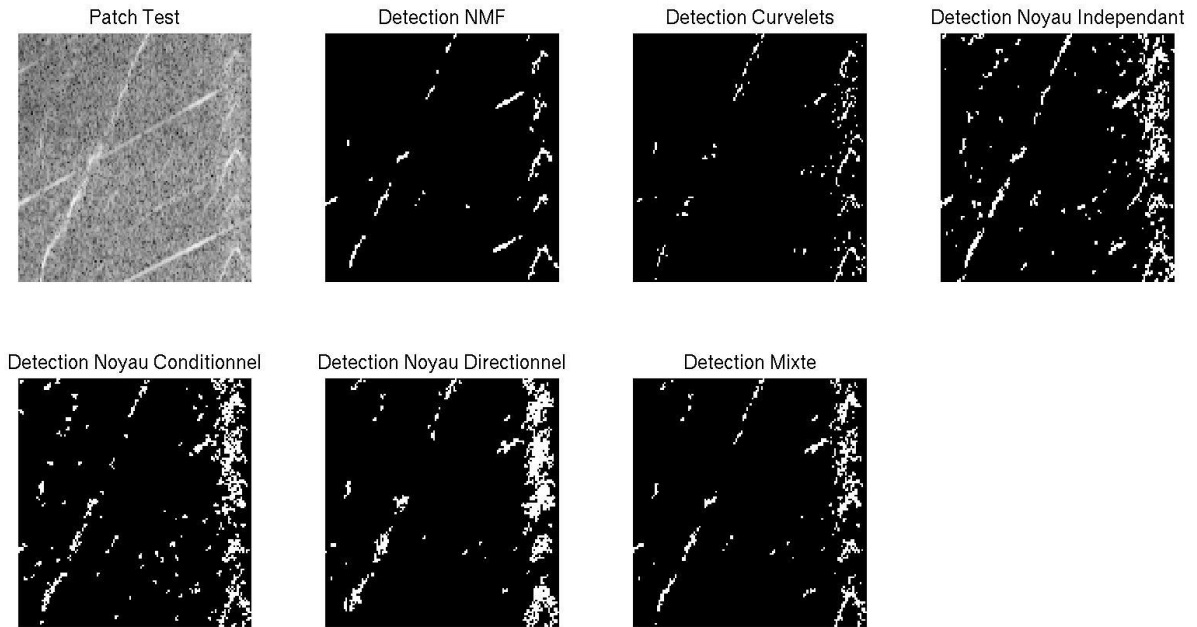


FIGURE 6.15 – Détection du patch contenant un premier type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{v}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

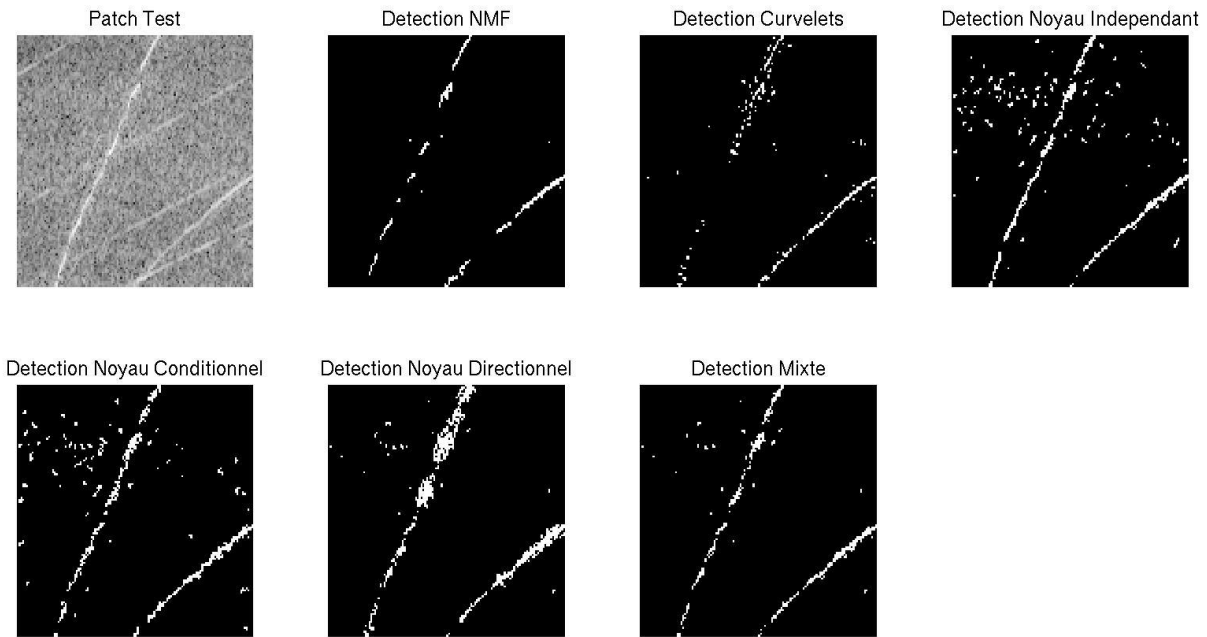


FIGURE 6.16 – Détection du patch contenant un second type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{v}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

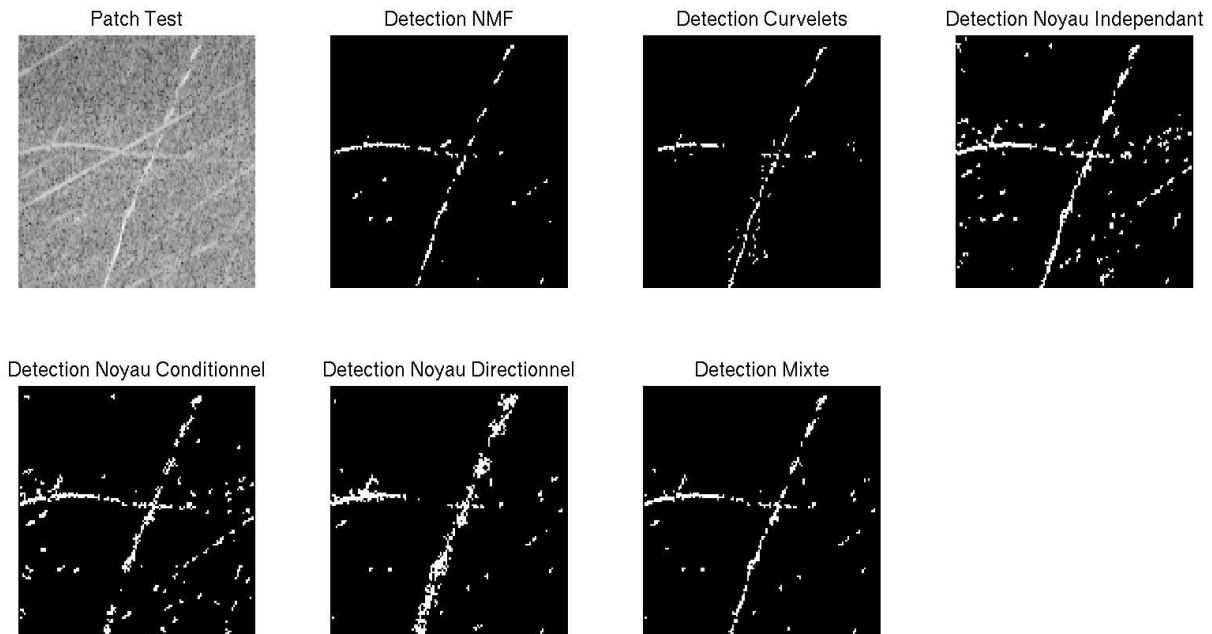


FIGURE 6.17 – Détection du patch contenant un troisième type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{v}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

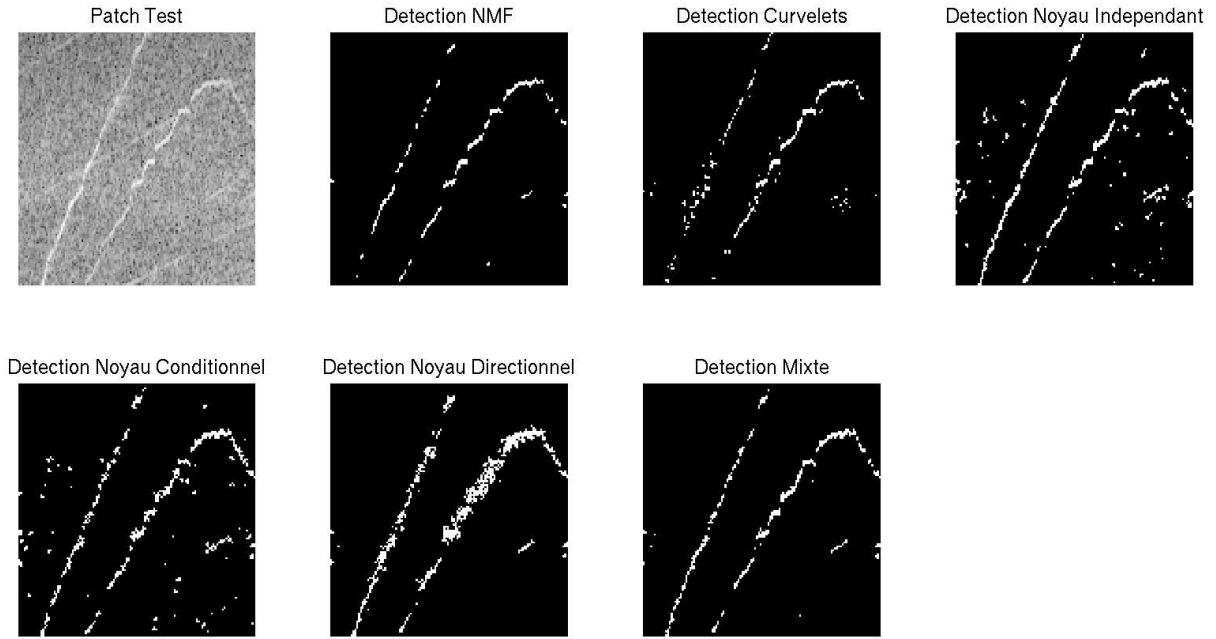


FIGURE 6.18 – Détection du patch contenant un quatrième type de signature inusuelle à partir de la \hat{Y}^{NMF} , \hat{Y}^C , \hat{Y}^{K_N} , \hat{Y}^{K_V} et $\hat{Y}^{K_{\vec{V}}}$, ainsi que de l'approche mixte \hat{Y} . Les éléments détectés correspondent à la raie N_1 se décalant sur les spectrogrammes et les signatures atypiques.

6.6 Conclusions

Les approches conditionnelles tiennent compte de la structure des signatures inusuelles. En effet, les signatures atypiques recherchées correspondent à des raies, les points voisins ont donc leur importance et il est donc pertinent de tenir compte du voisinage dans la mise en place du modèle de normalité. Cela permet de considérer la structure locale des données. Le rayon de voisinage ne doit pas être trop important car l'influence du voisinage sur un point à fréquence et régime donnés diminue en fonction de la distance au point étudié. La structure du voisinage est également à prendre en compte. Nous avons étudié le voisinage direct (d'ordre 1) où chaque intensité vibratoire voisine était considérée de manière uniforme. Cependant cette structure n'est pas en adéquation avec la structure du patch étant donné que les signatures correspondent à des raies et possèdent donc une orientation. La prise en compte de tous les voisins entraîne l'utilisation de voisins non pertinents. Nous avons également étudié un voisinage directionnel autour du point étudié où le voisinage correspond à des formes linéaires ayant une certaine orientation.

Pour les deux approches, la détection des points inusuels s'effectue à partir de tests statistiques d'adéquation des points par rapport à la distribution. Cette distribution est définie pour chaque point à partir d'une estimation de densité conditionnelle par noyau gaussien calibrée sur des points normaux. La distribution estimée correspond alors au modèle de normalité sur lequel nous nous basons pour détecter les points atypiques. Le score de détection correspond à la p-valeur du test statistique dans le cas du voisinage direct et la p-valeur minimale suivant les différentes orientations pour le voisinage directionnel. L'anomalie est détectée lorsque ce score passe sous le seuil de détection. Un filtrage par voisinage complète l'approche pour filtrer les

points isolés faussement détectés.

Les résultats produits par cette approche sont très intéressants et permettent la détection des signatures inusuelles, mais également des raies liées à l'arbre BP (N_1). Les signatures liées à l'arbre HP et le bruit sont faiblement détectées. Ces deux approches restent tout de même sensibles aux intensités vibratoires fortes du patch. La prise en compte du voisinage directionnel permet de considérer des points pertinents vis à vis de la structure des signatures atypiques. Cependant elle entraîne également la détection des points voisins pouvant masquer la structure des signatures inusuelles, mais permettant la détection des zones du patch contenant des signatures atypiques.

La détection à partir du voisinage directionnel permet de trouver l'orientation principale de chaque point à partir de la densité conditionnelle minimale associée. Ainsi en analysant les composantes connexes des points détectés, il est possible d'obtenir l'histogramme des directions des points de ces composantes. Ces histogrammes obtenus caractérisent donc les structures connexes correspondant à l'ensemble ou une partie des signatures détectées. Pour de futurs travaux, il est possible d'étudier cet indicateur pour classifier les différentes signatures détectées.

Les résultats obtenus par cette approche sont donc très prometteurs et permettent une aide à la décision pour l'analyse des spectrogrammes vibratoires. Il est également possible d'utiliser la densité jointe du point et de son voisinage comme modèle de normalité plutôt que la densité conditionnelle.

Conclusions et perspectives

Conclusions

L'objectif de cette thèse a été de mettre en place un outil d'aide à la décision automatique pour la détection des signatures atypiques sur les spectrogrammes vibratoires des moteurs d'avions. Nous avons abordé cette étude en considérant la représentation de la normalité au sein de ces spectrogrammes. Les signatures vibratoires sont représentées par des raies sur les spectrogrammes. Les éléments à détecter sont le plus fréquemment caractérisés par la présence de raies non présentes sur les spectrogrammes normaux. A l'heure actuelle, les spectrogrammes sont analysés visuellement et annotés manuellement par les experts de Safran Aircraft Engines dans un format image des données.

Notre travail a débuté dans un premier temps par associer les spectrogrammes numériques aux annotations manuelles des experts (format image) afin de construire une base de données numérique de spectrogrammes annotés. Les annotations correspondant aux différentes zones atypiques présentes sur les spectrogrammes et contenant une signature inusuelle. Un aboutissement de ce travail a été la consolidation d'une base annotée unique, regroupant l'expertise issue des analyses précédentes dans un format numérique. Nous avons également catégorisé différents points d'une soixantaine de spectrogrammes en 4 différentes classes, les points situés sur les signatures vibratoires normales, les points situés sur les signatures atypiques, les points correspondant à du bruit et les points situés sur des signatures vibratoires normales mais décalés sur les différents spectrogrammes. Ceci a permis de créer une seconde base, plus restreinte que la première mais plus fine également (la première contenant une annotation par zone), contenant une annotation ponctuelle des spectrogrammes. Cette base a été utilisée pour présenter les résultats, mais également pour paramétrer les seuils de décision.

Les spectrogrammes correspondent à des données en très grande dimension où chaque point correspond à la mesure vibratoire à une certaine fréquence f et un certain régime N_2 . Les signatures atypiques au sein de ces spectrogrammes consistent en une partie du spectrogramme de dimension bien plus petite que ce dernier. Analyser les spectrogrammes dans leur globalité ne permet pas de mettre en évidence les potentielles signatures inusuelles car ces dernières se retrouvent noyées par les informations normales présentes dans les données. Ces dernières sont structurées avec des informations normales identiques physiquement sur les différents spectro-

grammes. Sur l'intégralité du spectrogramme, le rapport points atypiques sur points normaux est très faible, ne permettant pas une analyse globale du spectrogramme. Nous avons donc subdivisé le spectrogramme en patches définis par des plages de fréquences et de régimes, sur lesquels les différentes approches de détection d'anomalies ont été appliquées de manière indépendante. La proportion d'information inusuelle potentiellement présente sur un patch est bien plus importante qu'à l'échelle du spectrogramme. Cette approche est similaire à l'analyse visuelle des experts analysant les spectrogrammes par plage de fréquences. Cependant le nombre de données atypiques au niveau de ces différents patches n'est pas suffisant pour mettre en place des modèles d'anomalies sur chacun d'entre eux. De plus les signatures inusuelles présentes sur les différents patches sont trop variables pour pouvoir caractériser chaque type de signature par un modèle mathématique précis. Nous nous sommes donc portés sur des approches de détection de nouveautés, détection d'anomalies qui permettent de classer des données lorsqu'un des labels des données (généralement normal) est fortement prédominant dans la base de données. Nous avons donc caractérisé les comportements normaux de ces patches, c'est-à-dire les signatures vibratoires normales.

Nous avons étudié la normalité au sein des patches sous différents aspects. Dans un premier temps, globalement sur le patch en mettant en place des modèles de normalité caractérisant l'intégralité des patches à partir des approches par dictionnaire. Puis, dans un second temps localement en définissant un modèle de normalité pour chaque point des patches à travers la distribution des points normaux. La caractérisation de cette normalité permet de bonnes performances de détection des signatures atypiques sans aucun a priori sur ces dernières répondant ainsi à la problématique énoncée. Ces méthodes ont donné lieu à un programme informatique (en MATLAB) d'outil d'aide à la décision pour analyser les spectrogrammes sans a priori sur les signatures atypiques. Ce programme fonctionne en prenant en compte les résultats des différentes méthodes mises en place au cours de cette thèse.

Représentation globale par dictionnaire

La représentation de la normalité des patches des spectrogrammes a été étudiée pour deux dictionnaires différents. Le premier correspond à un dictionnaire fixe (non-adaptatif comme les atomes du dictionnaire sont définis par des fonctions et restent inchangés par rapport aux données), les curvelets. La normalité dans ce dictionnaire est caractérisée par une sélection des atomes des curvelets activés régulièrement sur les patches normaux. La reconstruction du patch est effectuée uniquement à partir des coefficients associés à ces atomes. Le dictionnaire des curvelets caractérise localement les signatures sur le patch. Le second dictionnaire est cette fois-ci data-driven (adaptatif avec des atomes dépendants des données), la NMF. Les atomes du dictionnaire sont directement appris sur les données normales. Ils caractérisent globalement les signatures normales des patches uniquement dues aux propriétés linéaires et positives de ce dictionnaire. La dimension du dictionnaire est calibrée sur les données normales comme le rang à partir duquel le gain de reconstruction n'est plus significatif. Ces deux dictionnaires donnent une représentation normale des patches qui permet une reconstruction de ces derniers sans signature atypique. Il

s'agit alors d'une estimation normale des patches.

Les reconstructions à partir de ces dictionnaires caractérisant la normalité donnent des estimations des caractères normaux des patches des spectrogrammes. Les signatures atypiques sont grossièrement reconstruites à partir du dictionnaire des curvelets dus aux atomes basses fréquences avec des artéfacts autour des signatures atypiques. Pour le dictionnaire appris de la NMF, les signatures inusuelles ne sont pas du tout reconstruites. Les approches par dictionnaire permettent de représenter les signatures atypiques au niveau des résidus de reconstruction. Il s'agit des points mal représentés par les dictionnaires et correspondent alors aux valeurs extrêmes positives au sein des résidus. Ces approches entraînent des fausses détections pour les deux dictionnaires qui sont limitées par l'application du filtrage de voisinage tenant compte de la décision au niveau des points voisins. Les signatures vibratoires liées à l'arbre basse pression sont en partie également détectées du fait de leur décalage sur les différents spectrogrammes. La combinaison des deux dictionnaires permet une meilleure considération visuelle des signatures atypiques.

La représentation de la normalité des patches dans des dictionnaires avec une décision basée sur l'erreur de reconstruction qui en est issue permet une bonne mise en évidence des signatures inusuelles. Les différentes signatures atypiques sont détectées indépendamment du type de signatures. Les approches par dictionnaires permettent également un débruitage des spectrogrammes au niveau de la reconstruction.

La détection est effectuée à partir des valeurs extrêmes des résidus intra-patch sans tenir compte du même patch des autres données. Cela entraîne une détection systématique de points sur tous les patches indépendamment de la présence de signatures atypiques sur ces derniers.

Représentation ponctuelle

La représentation des patches correspond à la distribution estimée à partir des points normaux pour chaque point du patch. Cette distribution correspond alors au modèle de normalité établie pour chaque point. L'estimation de la distribution est effectuée de manière non paramétrique à partir des méthodes à noyau. L'utilisation des méthodes non paramétriques permet de ne donner aucun a priori sur la distribution des points. Trois différentes approches ont été étudiées dans ce cadre. La première considérant tous les points du patch comme mutuellement indépendants. Cependant les points voisins des spectrogrammes sont corrélés et deux points appartenant à la même raie vibratoire sont dépendants. La seconde approche tient donc compte uniformément du voisinage du premier ordre autour du point étudié. La dernière approche considère le voisinage du point sous différentes directions afin de tenir compte des voisins liés à une certaine orientation uniquement. Lors de l'étude, nous avons étudié le noyau gaussien permettant une simplification des calculs mais attribuant une densité non nulle aux points négatifs (impossible pour les intensités vibratoires) et le noyau gamma, permettant de ne considérer que les points positifs mais complexifiant les calculs. Les résultats à partir des deux noyaux sont proches dans le cas indé-

pendant. Nous avons donc opté pour le noyau gaussien dans les approches conditionnelles pour la simplification des calculs.

Les points atypiques correspondent à des intensités vibratoires importantes par rapport aux mêmes vibrations exactement sur les autres données. Il s'agit donc des valeurs extrêmes de la distribution en chaque point. La détection est effectuée à partir de la p-valeur d'un test statistique d'adéquation à la loi estimée dans les cas indépendant et tenant compte uniformément du voisinage. La décision est donnée pour chaque point du spectrogramme comparativement à un seuil de décision. Dans le cas du voisinage directionnel, un test statistique d'adéquation est réalisé suivant chacune des directions, et donc a fortiori une p-valeur est calculée pour chacune d'entre elles. La décision est établie en considérant la p-valeur minimale suivant ces différentes directions. Il suffit donc de détecter une irrégularité sur une seule direction pour que le point soit considéré comme atypique. Ces différentes approches entraînent également de fausses détections qui sont réduites par l'application du filtrage de voisinage. Les signatures vibratoires liées à l'arbre basse pression sont également détectées.

Ces méthodes définies localement permettent d'obtenir de bons résultats de détections des points atypiques avec une mise en évidence de ces signatures. De plus, la considération du voisinage sous différentes directions permet d'associer une orientation aux points détectés correspondant à la vraisemblance minimale du point par rapport aux différents voisinages. L'histogramme des directions des composantes connexes de points détectés permet d'avoir une caractérisation des signatures atypiques pour possiblement les classifier par la suite.

Ces approches sont plus sensibles aux intensités vibratoires que les approches par dictionnaire. Cela se voit aussi bien visuellement que numériquement avec une détection plus importante des points issus des raies N_1 . Cependant, comme la détection s'effectue par comparaison avec les mêmes points des autres données, il n'y a pas de détection systématique pour chaque patch des spectrogrammes à la différence des approches par dictionnaire.

Fusion des approches

Les différentes approches proposées, aussi bien globalement sur les patches avec les dictionnaires que localement à partir de la modélisation ponctuelle des points, permettent toutes de détecter efficacement les signatures atypiques sur les patches des spectrogrammes. Les points faussement détectés engendrés sont dépendants des différentes approches.

Afin de les limiter, nous avons associé les détections des différentes approches par un système de vote. Un point est considéré comme atypique si plus de la moitié des approches le considère comme tel. Cette fusion permet de réduire les fausses détections tout en consolidant les bonnes détections. La combinaison des différentes approches sous la forme de vote permet donc d'obtenir un bon outil de détection automatique des signatures atypiques.