Fusion en traitement d'images : spécificités et approches numériques

par **Isabelle BLOCH**
École nationale supérieure des télécommunications
Département Traitement du signal et des images
CNRS URA 820

et **Henri MAÎTRE**
École nationale supérieure des télécommunications
Département Traitement du signal et des images
CNRS URA 820

La fusion d'informations regroupe les techniques utilisées pour associer des informations variées sur un même problème. En traitement des images, la fusion d'informations se préoccupe de combiner au mieux des images d'origines différentes pour mieux connaître l'objet d'observation. La fusion est devenue un aspect important de traitement de l'information dans plusieurs domaines très différents, dans lesquels les informations à fusionner, les objectifs, les méthodes, et donc la terminologie, peuvent varier beaucoup, même si les analogies sont également nombreuses. L'ampleur que prend la fusion d'informations suit celle que prennent les technologies et le traitement de l'information en général.

Cet article vise à préciser le contexte et les concepts de la fusion dans le domaine du traitement du signal et des images (TDSI), à dégager des définitions et à présenter les grandes lignes des principales approches numériques. Nous ne présentons pas ici les approches à bases de règles, syntaxiques, logiques, ni les approches neuronales.

*Le paragraphe 1 présente une **définition générale**, les caractéristiques des données à prendre en compte dans un système de fusion, ainsi que les principales étapes. Le paragraphe 2 est consacré de manière plus précise aux **spécificités de la fusion en traitement des images**, en soulignant ce qui la distingue de la fusion dans d'autres domaines. Les principales **approches numériques** sont ensuite exposées, dans les paragraphes 3 pour les approches probabilistes et bayésiennes, 4 pour la théorie des fonctions de croyance, et 5 pour les méthodes floues et possibilistes. Enfin dans le paragraphe 6, nous discutons du traitement de l'information spatiale en fusion d'images.*

1. Généralités et définitions

1.1 Définitions

Nous adoptons dans cet article un sens large du terme « **information** ». En particulier, il couvre à la fois des données (par exemple, des mesures, des images, des signaux, etc.) et des connaissances (sur les données, sur le domaine, sur des contraintes, etc.) qui peuvent être génériques ou spécifiques.

■ La définition de la fusion d'informations que nous utiliserons tout au long de cet article est la suivante :

> **Définition 1**. **La fusion d'informations consiste à combiner des informations hétérogènes issues de plusieurs sources afin d'améliorer la prise de décision**.

Cette définition est suffisamment générale pour englober la diversité des problèmes de fusion que l'on rencontre en traitement du signal et des images. Son intérêt est qu'elle est focalisée sur les étapes de combinaison et de décision, ces deux opérations pouvant prendre des formes différentes suivant les problèmes et les applications.

Pour chaque type de problème et d'application, cette définition pourra être plus spécifique, en répondant à un certain nombre de questions : quel est le but de la fusion ? comment s'exprime la décision ? quelles sont les informations à fusionner ? quelles sont leurs origines ? quelles sont leurs caractéristiques (incertitude, relations entre les informations, génériques ou factuelles, statiques ou dynamiques, etc.) ? quelle méthodologie choisir ? comment évaluer et valider la méthode et les résultats ? quelles sont les difficultés principales, les limites ? etc.

Situons cette définition par rapport à celles proposées par quelques autres groupes de travail qui ont structuré le domaine de la fusion d'informations.

La définition 1 est un peu plus spécifique que celle de [26], qui est la suivante : « *regrouper des informations issues de plusieurs sources et exploiter l'information regroupée afin de répondre à des questions, prendre des décisions, etc.* ». Dans cette définition, qui est également focalisée sur la combinaison et les buts, les buts s'arrêtent souvent avant l'étape de décision, et ne sont pas restreints à l'amélioration de l'information globale. Ils peuvent inclure l'obtention d'un point de vue global, par exemple dans les problèmes de fusion d'opinions ou de préférences de personnes, qui est un des thèmes abordés dans ce projet, mais qui sort du cadre de cet article. Ici, l'amélioration de la connaissance fait référence au monde tel qu'il est, et non au monde tel qu'on voudrait qu'il soit comme c'est le cas en fusion de préférences par exemple.

Un des premiers efforts notables d'éclaircissement du domaine traduit un souci de préciser et de codifier la terminologie de la fusion de données dans une sorte de dictionnaire (« Data Fusion Lexicon ») [1]. La méthode proposée est dédiée essentiellement à des applications de défense (telles que la poursuite de cibles, la reconnaissance et l'identification automatiques de cibles) et est focalisée sur les fonctionnalités, en identifiant des processus, des fonctions et des techniques [67]. Elle insiste sur la description d'une hiérarchie d'étapes de traitement dans un système. La définition que nous utilisons ici s'oppose à celle-ci et choisit un autre point de vue, en s'attachant plus à la **description des méthodes de combinaison et de décision** qu'à celle d'un système. Elle est mieux adaptée à la variété des situations qui peuvent être rencontrées en TDSI. Elle est en ce sens plus générale.

Plus spécifiquement en imagerie satellitaire, la fusion de données a été définie comme un *cadre formel dans lequel s'expriment les données provenant de sources diverses ; elle vise à l'obtention d'informations de plus grande qualité* ; la définition exacte de « plus grande qualité » dépendra de l'application. Cette définition englobe la plupart des définitions proposées par plusieurs auteurs en imagerie satellitaire, qui sont regroupées dans [122]. La définition 1 va plus loin que celle-ci, jusqu'à la décision.

■ La fusion peut être comprise dans un sens plus ou moins large. D'autres concepts tels que l'**estimation**, la **révision**, l'**association de données**, la **fouille de données** sont parfois considérés comme des problèmes de fusion dans un sens large du terme. Précisons ces concepts.

● **Fusion et estimation** : l'estimation vise à combiner plusieurs valeurs d'un paramètre ou d'une distribution pour en déduire une valeur plausible. On retrouve donc les étapes de combinaison et de décision qui sont les deux ingrédients principaux de la définition 1. D'un autre côté, les méthodes de fusion numériques nécessitent souvent une étape préliminaire d'estimation des distributions à combiner (voir paragraphe 1.4), et l'estimation apparaît alors comme une des étapes d'un processus de fusion.

● **Fusion et révision ou mise à jour** : la révision, ou mise à jour, consiste à compléter ou modifier une information en fonction de nouveaux apports d'informations. Elle peut être considérée comme un des domaines de la fusion. Parfois, la fusion est considérée dans un sens plus restreint, où la combinaison est symétrique. La révision, elle, ne l'est pas, et fait une distinction entre l'information a priori et les nouvelles informations. Ici, nous considérons entre autres des processus dynamiques (en particulier en robotique), et il nous semble donc important d'inclure la révision et la mise à jour dans la fusion (par exemple pour des applications à la compréhension de l'environnement par un robot). La révision concerne l'ajout d'une nouvelle information permettant de modifier ou préciser l'information disponible auparavant sur le phénomène observé, alors que la mise à jour concerne une modification du phénomène conduisant à modifier l'information sur celui-ci (typiquement dans un processus temporel).

● **Fusion et association** : l'association de données est l'action qui permet de trouver parmi différents signaux, issus de deux sources ou plus, ceux qui sont émis par le même objet (source ou cible). Selon Bar-Shalom et Fortman [8], l'association de données est l'étape la plus difficile en poursuite de cibles multiples. Elle consiste à détecter et associer des mesures bruitées, dont les origines sont incertaines à cause de plusieurs facteurs, tels que les fausses alarmes aléatoires dans la détection, le fouillis, les cibles interférentes, les pièges et autres contre-mesures. Les modèles principaux utilisés dans ce domaine sont soit déterministes (à partir de tests d'hypothèses classiques), soit probabilistes (essentiellement bayésiens) [8][75][104]. La méthode la plus courante [8] repose sur le filtre de Kalman sous hypothèse gaussienne. Plus récemment, d'autres méthodes d'estimation ont été proposées, comme le modèle d'estimation interactive multiple (*Interactive Multiple Model estimator*, IMM), qui peut s'adapter à différents types de mouvement, réduire le bruit, tout en préservant une bonne précision des estimations des états [127]. Ainsi les problèmes traités par l'association de données sont assez différents de ceux couverts par la définition 1.

● **Fusion et fouille de données** : la fouille de données (*Data Mining*) consiste à extraire des parties intéressantes de l'information et des données, qui peuvent être par exemple des données spéciales (au sens d'une propriété particulière) ou rares. Elle peut être opposée à la fusion visant à expliquer, où des tendances générales sont recherchées, ou à la fusion visant à généraliser et à induire des connaissances plus génériques à partir de données. Ici, **nous écarterons la fouille de données des problèmes de fusion**.

1.2 Caractéristiques générales des données

Dans cette partie, nous décrivons succinctement les caractéristiques générales des informations à fusionner, qui doivent souvent être prises en compte dans un processus de fusion.

■ Une première caractéristique concerne le **type d'information** à fusionner. Il peut s'agir d'observations directes, de résultats de traitements sur ces observations, de connaissances plus génériques, exprimées sous forme de règles par exemple, ou d'avis d'experts. Ces informations peuvent être exprimées sous forme numérique ou sous forme symbolique (voir paragraphe 1.3). Une attention particulière doit être portée à l'**échelle utilisée** pour représenter les informations. Cette échelle n'a pas forcément de signification absolue, mais le minimum à garantir est que les informations puissent être comparées selon cette échelle, c'est-à-dire que les échelles induisent un ordre sur les populations. Cela se traduit par des propriétés de commensurabilité, voire de normalisation.

Le **niveau de l'information** qui va être fusionnée est également un aspect très important. On distingue généralement le bas niveau (typiquement les mesures originales) d'un niveau plus élevé nécessitant des étapes préliminaires de traitement, d'extraction de primitives et de structuration de l'information. Suivant le niveau auquel on se place, les contraintes ne sont pas les mêmes, ni les difficultés. Cette question se pose tout particulièrement en fusion d'images et sera détaillée dans le paragraphe 2.

D'autres **distinctions sur les types de données** sont également intéressantes à souligner, car elles donnent lieu à des modélisations et à des types de traitements différents. La distinction entre données **fréquentes** et données **rares** en fait partie. Les informations peuvent également être **factuelles** ou **génériques**. Les connaissances génériques peuvent être par exemple un modèle du phénomène observé, des règles générales, des contraintes d'intégrité. Les informations factuelles sont plus directement reliées aux observations. Souvent, ces deux types d'information ont des spécificités différentes. Les informations génériques sont en général moins spécifiques (et font office de « défaut ») que les informations factuelles qui sont directement relatives au phénomène particulier observé. Le défaut est considéré si l'information spécifique n'est pas présente ou pas assez fiable. Dans le cas contraire, et si les informations sont contradictoires, on préférera les informations les plus spéciques. Enfin, les informations peuvent être **statiques** ou **dynamiques**, et là encore leur modélisation et leur prise en compte diffèrent dans les deux cas.

Les informations manipulées dans un processus de fusion sont d'une part les **informations à fusionner**, et d'autre part des **informations supplémentaires** qui servent à guider ou à aider la combinaison. Il peut s'agir d'informations sur les informations à combiner telles que des informations sur les sources, sur leur dépendance, sur leur fiabilité, des préférences sur les informations à combiner, etc. Il peut s'agir également d'informations contextuelles, sur le domaine. Ces informations supplémentaires ne sont pas forcément exprimées dans le même formalisme que les informations à combiner (en général elles ne le sont pas), mais elles peuvent intervenir dans le choix de la modélisation des informations à fusionner.

■ Une des caractéristiques importantes de l'information en fusion est son **imperfection**. Celle-ci est toujours présente (sinon la fusion ne serait pas nécessaire). Elle peut prendre diverses formes, qui sont brièvement décrites ci-dessous. Notons que ces notions ne trouvent pas toujours de définition consensuelle dans la littérature. Nous en proposons ici des définitions qui sont assez intuitives et conviennent bien au problème de la fusion, mais qui ne sont sûrement pas universelles.

● **Incertitude** : l'incertitude est relative à la vérité d'une information, et caractérise son degré de conformité à la réalité [53]. Elle fait référence à la nature de l'objet ou du fait concerné, à sa qualité, à son essence ou à son occurrence.

● **Imprécision** : l'imprécision concerne le contenu de l'information et mesure donc un défaut quantitatif de connaissance, sur une mesure [53]. Elle concerne le manque d'exactitude en quantité, en taille, en durée, le manque de définition d'une proposition qui est ouverte à diverses interprétations ou qui a des frontières vagues et mal définies. Cette notion est souvent abusivement confondue avec celle d'incertitude, car les deux types d'imperfection sont souvent présents simultanément, et l'un peut induire l'autre. Il est important de les distinguer car ils sont souvent antagonistes, même si ces deux termes peuvent être inclus dans une acception plus large de l'incertitude. Inversement, des typologies en un plus grand nombre de classes ont été proposées [70].

● **Incomplétude** : l'incomplétude caractérise l'absence d'information apportée par la source sur certains aspects du problème. L'incomplétude des informations issues de chaque source est la raison principale qui motive la fusion. L'information fournie par chaque source est en général partielle, elle ne fournit qu'une vision du monde ou du phénomène observé, en n'en mettant en évidence que certaines caractéristiques.

● **Ambiguïté** : l'ambiguïté exprime la capacité d'une information de conduire à deux interprétations. Elle peut provenir des imperfections précédentes, par exemple de l'imprécision d'une mesure qui ne permet pas de différencier deux situations, ou de l'incomplétude qui induit des confusions possibles entre des objets ou des situations qui ne peuvent être séparés selon les caractéristiques mises en évidence par la source. Un des objectifs de la fusion est de lever les ambiguïtés d'une source grâce aux informations apportées par les autres sources ou par les connaissances supplémentaires.

● **Conflit** : le conflit caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires et donc incompatibles. Les situations conflictuelles sont fréquentes dans les problèmes de fusion, et posent toujours des problèmes difficiles à résoudre. Tout d'abord, la détection des conflits n'est pas forcément facile. Ils peuvent facilement être confondus avec d'autres types d'imperfection, ou même avec la complémentarité des sources. Ensuite, leur identification et leur typologie est une question qui se pose souvent, mais de manière différente suivant le domaine. Enfin, leur résolution peut prendre différentes formes. Elle peut reposer sur l'élimination de sources non fiables, sur la prise en compte d'informations supplémentaires, etc. Dans certains cas, il peut être préférable de retarder la combinaison et d'attendre d'autres informations susceptibles de lever les conflits, ou même de ne pas fusionner du tout.

■ D'autres caractéristiques de l'information sont plus positives, et sont exploitées pour **limiter les imperfections** :

● **Redondance** : la redondance est la qualité de sources qui apportent plusieurs fois la même information. La redondance entre les sources est souvent observée, dans la mesure où les sources donnent des informations sur le même phénomène. Idéalement, la redondance est exploitée pour réduire les incertitudes et les imprécisions.

● **Complémentarité** : la complémentarité est la propriété des sources qui apportent des informations sur des grandeurs différentes. Elle vient du fait qu'elles ne donnent en général pas d'informations sur les mêmes caractéristiques du phénomène observé. Elle est exploitée directement dans le processus de fusion pour avoir une information globale plus complète et pour lever les ambiguïtés.

Les outils probabilistes permettant de modéliser et de mesurer l'information, la redondance et la complémentarité seront décrits dans le paragraphe 3.1.

1.3 Fusion numérique/fusion symbolique

De nombreuses discussions ont eu lieu dans la communauté de la fusion sur la dualité entre fusion **numérique** et fusion **symbolique**. L'objectif de ce paragraphe n'est pas de retracer en détail ces discussions mais plutôt de présenter les différents niveaux auxquels cette question peut être posée. Par une judicieuse description de ces niveaux, il est souvent possible de réduire ces débats. Les trois niveaux que nous distinguons ici sont ceux du type de données, du type de traitement appliqué aux données et du rôle des représentations. Ils sont détaillés dans les paragraphes suivants.

1.3.1 Données et informations

• Par information **numérique**, nous entendons une information qui est directement donnée sous forme de nombre. Ces nombres peuvent représenter des mesures physiques, des niveaux de gris dans une image, une intensité de signal, la distance donnée par un télémètre, ou encore la réponse à un opérateur numérique de traitement. Ils peuvent être soit directement lus dans les données à fusionner, soit attachés au domaine ou à la connaissance contextuelle.

• Par information **symbolique**, nous entendons toute information qui est donnée sous forme de symboles, de propositions, de règles, etc. De telles informations peuvent être attachées aux données à fusionner, ou à la connaissance du domaine (par exemple, des propositions sur les propriétés du domaine concerné, une information structurelle, des règles générales sur le phénomène observé, etc.).

• La classification des informations et des données en numériques et symboliques ne peut pas toujours être faite de manière binaire. En effet, les informations peuvent être également **hybrides**, et des nombres peuvent représenter des codages d'informations de nature non numérique. C'est typiquement le cas de l'évaluation d'une donnée ou d'un traitement, de la quantification de l'imprécision ou de l'incertitude. Dans de tels cas, les valeurs absolues des nombres sont souvent peu importantes, et c'est surtout leur place dans une échelle, ou leur ordre si plusieurs grandeurs sont évaluées, qui joue un rôle prépondérant. Le terme « hybride » désigne alors des nombres utilisés comme des symboles pour représenter une information, mais avec une quantification, qui permet ensuite de les manipuler de manière numérique. Ces nombres peuvent être utilisés pour des informations symboliques aussi bien que numériques.

1.3.2 Traitements de l'information

• En ce qui concerne le traitement de l'information, on appelle traitement **numérique** tout calcul sur des nombres. En fusion d'informations, cela couvre toutes les approches qui combinent des nombres par un calcul formel. Il est important de remarquer que ce type de traitement ne fait pas forcément d'hypothèse sur le type d'information qui est représentée par des nombres. À l'origine, les informations peuvent être de nature aussi bien numérique que symbolique.

• Les traitements **symboliques** incluent le calcul formel sur des propositions (par exemple des approches de type logique ou les grammaires, sur lesquelles on peut trouver plus de détails dans [26]), prenant éventuellement en compte des connaissances numériques. Les approches structurelles, telles que des approches à base de graphes, qui sont largement utilisées en reconnaissance des formes structurelles (en particulier pour la fusion), peuvent être mises dans cette classe.

• Nous appelons traitement **hybride** les méthodes où la connaissance a priori est utilisée d'une manière symbolique pour contrôler des traitements numériques, par exemple en déclarant des règles propositionnelles qui suggèrent, permettent ou au contraire interdisent certaines opérations numériques. Typiquement, une proposition qui définit dans quels cas deux sources sont indépendantes peut être utilisée pour choisir la manière dont des probabilités sont combinées.

1.3.3 Représentations

• Ainsi que le montrent les deux paragraphes précédents, les représentations et leurs types peuvent jouer des rôles très différents. Des représentations **numériques** peuvent être utilisées pour des données intrinsèquement numériques mais aussi pour l'évaluation et la quantification de données symboliques. Les représentations numériques en fusion d'informations sont utilisées de manière importante pour quantifier l'imprécision, l'incertitude ou la fiabilité de l'information (cette information peut être de nature numérique aussi bien que symbolique), et donc pour représenter plutôt de l'information sur les données à combiner que ces données elles-mêmes. Ces représentations seront plus détaillées dans les parties sur les méthodes numériques de fusion. Les représentations numériques sont aussi souvent utilisées pour des degrés de croyance attachés à des connaissances numériques ou symboliques, et pour des degrés de cohérence ou d'incohérence (ou de conflit) entre les informations (le cas le plus typique est peut-être celui de la fusion de bases de données ou de réglementations). Notons que le même formalisme numérique peut être utilisé pour représenter différents types de données ou de connaissances [15] : l'exemple le plus évident est celui de l'utilisation des probabilités pour représenter des données aussi différentes que des fréquences ou des degrés de croyance subjectifs [40].

• Les représentations **symboliques** peuvent être utilisées dans des systèmes logiques, ou des systèmes à base de règles, mais aussi en tant que connaissance a priori ou connaissance contextuelle ou générique pour guider un traitement numérique, en tant que support structurel par exemple en fusion d'images, et bien sûr en tant que sémantique attachée aux objets manipulés.

• Dans de nombreux exemples, une forte **dualité** peut être observée entre les rôles des représentations numériques et symboliques, et peut être exploitée dans le cas de fusion de sources hétérogènes.

1.4 Systèmes de fusion et types d'architecture

En général, la fusion n'est pas une tâche simple. Elle peut se décomposer de manière schématique en plusieurs tâches. Nous les décrivons succinctement ici, car elles serviront de guide à la description des outils théoriques dans les paragraphes 3 à 6. Considérons un problème général de fusion pour lequel on dispose de ℓ sources $S_1, S_2, ..., S_\ell$ (par exemple des images I_j), et pour lequel le but est de prendre une décision dans un ensemble de n décisions possibles $d_1, d_2, ..., d_n$.

■ Les **principales étapes** à résoudre **pour construire le processus de fusion** sont les suivantes :

1. Modélisation : cette étape comporte le choix d'un formalisme, et des expressions des informations à fusionner dans ce formalisme. Cette modélisation peut être guidée par les informations supplémentaires (sur les informations et sur le contexte ou le domaine). Supposons pour fixer les idées que chaque source S_j fournisse une information représentée par M_i^j sur la décision d_i. La forme de M_i^j dépend bien sûr du formalisme choisi. Elle peut être par exemple une distribution dans un formalisme numérique, ou une formule dans un formalisme logique.

2. Estimation : la plupart des modélisations nécessitent une phase d'estimation (par exemple toutes les méthodes utilisant des distributions). Là encore les informations supplémentaires peuvent intervenir.

3. Combinaison : cette étape concerne le choix d'un opérateur, compatible avec le formalisme de modélisation retenu, et guidé par les informations supplémentaires.

4. Décision : c'est l'étape ultime de la fusion, qui permet de passer des informations fournies par les sources au choix d'une décision d_i.

Nous ne donnons pas plus de détails sur ces étapes ici, car cela nécessiterait d'entrer dans les formalismes et les aspects techniques. Cela fait l'objet des paragraphes 3, 4 et 5.

■ La **manière** dont ces étapes sont agencées définit le **système de fusion et son architecture**.

● Dans le cas idéal la décision est prise à partir de **tous les** M_i^j, pour **toutes les sources** et **toutes les décisions**. C'est ce qu'on appelle la **fusion globale**. Dans ce modèle global, aucune information n'est négligée. La complexité de ce modèle et de sa mise en œuvre a conduit à développer des systèmes simplifiés, mais aux performances plus limitées [27].

● Un deuxième modèle consiste ainsi à prendre d'abord des décisions locales au niveau de chaque source séparément. Dans ce cas, **une décision** $d(j)$ est prise à partir de **toutes les informations** issues de **la source** S_j **seulement**. Cette décision est dite **décentralisée**. Puis, dans une deuxième étape, ces décisions locales sont fusionnées en une décision globale. Ce modèle s'impose lorsque les sources ne sont pas disponibles simultanément. Il permet des réponses rapides, grâce à des procédures spécifiques à chaque source, et peut être très facilement adapté à l'introduction de sources supplémentaires. Ce type de modèle s'appuie avec profit sur des techniques de contrôle adaptatif, et utilise souvent des architectures distribuées. Il est aussi appelé **fusion de décisions** [42][115]. Son principal inconvénient provient de ce qu'il tient mal compte des relations entre capteurs et des corrélations ou dépendances possibles entre sources. De plus, ce modèle conduit très facilement à des décisions locales contradictoires ($d(j) \neq d(k)$ pour $j \neq k$) et la résolution de ces conflits implique des arbitrages de niveau supérieur, difficiles à optimiser, puisque l'information de départ n'est plus disponible. Des modèles de ce type sont souvent mis en œuvre dans des applications en temps réel, par exemple dans le domaine militaire.

● Un troisième modèle, « **orthogonal** » au précédent, consiste à combiner par une opération F tous les M_i^j relatifs à la **même décision** d_i, pour obtenir une **forme fusionnée** $M_i = F(M_i^1, M_i^2, ..., M_i^\ell)$. Puis une décision est prise sur le résultat de cette combinaison. Ici, aucune décision intermédiaire n'est prise, et l'information est manipulée dans le formalisme choisi jusqu'à la dernière étape, diminuant ainsi les contradictions et les conflits. Ce modèle, comme le modèle global, est un modèle centralisé qui nécessite de disposer simultanément de toutes les sources. Plus simple que le modèle global, il est moins souple que le modèle distribué, rendant plus difficile l'ajout éventuel de sources d'informations.

● Enfin, un modèle intermédiaire, **hybride**, consiste à choisir de manière adaptative les informations nécessaires pour un problème donné en fonction des spécificités des sources. Ce type de modèle copie souvent l'expert humain et fait intervenir des connaissances symboliques sur les sources et sur les objets. Il est donc très utilisé dans les systèmes à base de règles. Des architectures de type multi-agents sont bien adaptées à ce modèle.

1.5 Fusion en TDSI et fusion dans d'autres domaines

La fusion en TDSI a des spécificités qui doivent être prises en compte à toutes les étapes de la construction d'un processus de fusion. Ces spécificités nécessitent également de modifier et de complexifier certains outils théoriques, souvent issus d'autres domaines. C'est typiquement le cas de l'information spatiale en fusion d'images ou en robotique. Ces spécificités seront détaillées pour le cas de la fusion en image dans le paragraphe 2.

La quantité des données à traiter et leur hétérogénéité sont souvent plus importantes que dans d'autres domaines (problèmes de combinaison d'opinions d'experts par exemple). Cela induit un niveau de complexité supplémentaire, à prendre en compte non seulement dans la modélisation, mais aussi dans les algorithmes.

Les données sont surtout objectives (données par des capteurs), ce qui les oppose aux données subjectives telles qu'elles peuvent être données par des individus. Elles contiennent cependant une certaine part de subjectivité (par exemple dans le choix des capteurs ou des sources d'informations, ou encore des paramètres d'acquisition). La subjectivité peut également intervenir dans l'expression des buts. Les données objectives sont en général dégradées, soit à cause de l'imperfection des systèmes d'acquisition, soit à cause des traitements qui leur sont appliqués.

En fait, une des difficultés principales vient du fait que les types de connaissances manipulées sont très hétérogènes. Il ne s'agit donc pas que de mesures et d'observations (qui elles-mêmes peuvent être hétérogènes), mais aussi de cas généraux, d'exemple typiques, de modèles génériques, etc.

Les différences principales par rapport à d'autres domaines d'application de la fusion d'informations viennent tout d'abord du fait que la question essentielle qui est posée (donc le but de la fusion) n'est pas la même. En TDSI, il s'agit essentiellement, selon la définition 1, d'améliorer la connaissance du monde (tel qu'il est). Cela suppose la notion d'existence d'une vérité, même si l'on n'en a qu'une vision partielle et déformée, ou difficile d'accès. Cela diffère de la fusion de préférences (comment on veut que le monde soit), de la fusion de réglementations (comment le monde devrait être), de problèmes de vote, où typiquement la vérité n'existe pas, etc. [26].

2. Spécificités de la fusion d'informations en traitement des images

Nous décrivons ici les spécificités de la fusion en traitement d'images. Les définitions générales proposées dans le paragraphe 1 sont précisées et illustrées dans ce cadre particulier. Nous insisterons sur la spécificité des images et de leur représentation par rapport aux problèmes de fusion, ce qui les distingue de la plupart des autres domaines d'application de la fusion.

Nota : pour un panorama des différents domaines, on pourra se référer à [119], étude bibliographique présentée sous forme d'analyse statistique, qui présente les développements effectués de 1997 à 1999 selon le domaine d'application, les objectifs de la fusion, l'architecture des systèmes et le cadre mathématique.

2.1 Objectifs de la fusion en traitement des images

■ Les premières applications dans lesquelles sont apparues des images dans les systèmes de fusion concernent des modèles fonctionnels, hérités de ceux développés pour la fusion de mesures en contrôle de processus, pour expliquer ou prédire des systèmes plus complexes, tels que des configurations météréologiques ou des systèmes écologiques. Ces modèles, souvent d'une très grande complexité, relient des contributions très différentes telles que des observations à échelle microscopique (évapo-transpiration par exemple) ou macroscopique (variations saisonnières par exemple). La fusion des informations est alors effectuée essentiellement au niveau des modèles physiques. Certaines des mesures impliquées peuvent être des images (par exemple une image radar pour estimer les zones d'eau).

Puis les images ont pris une place de plus en plus importante dans des problèmes de fusion en vision robotique, pour le guidage des véhicules, ou pour la surveillance. Le but est alors de fournir une description précise de l'espace en termes de géométrie et de mouvement, pour garantir un suivi 3D parfait de tout objet entrant dans le champ de vision d'un capteur, et prédire sa trajectoire. La théorie du contrôle reste encore à la base de ces systèmes, et le traitement des images n'est pas la composante essentielle.

La fusion réelle d'images est apparue plus tard, depuis une dizaine d'années. Ses objectifs principaux sont d'utiliser plusieurs images de la même scène, obtenues par des capteurs différents, pour fournir une description et une compréhension complètes de la scène, pas seulement en termes de position et de géométrie, mais surtout en termes de sémantique et d'interprétation.

La fusion d'images n'est pas un avatar des modes qui secouent le traitement de l'information depuis vingt ans, c'est la réponse naturelle des utilisateurs à la profusion des images produites par les systèmes contemporains d'investigation, que ce soit en imagerie aérienne et satellitaire ou en génie biologique et médical [27]. Prenons deux exemples (il en existe bien sûr de similaires dans d'autres applications).

Exemples

• L'**observation de la Terre par satellites** croît selon deux tendances : la première nous donne des images dont la résolution est de plus en plus fine et la cadence de plus en plus élevée ; la seconde propose des signaux de plus en plus variés : à la profusion des canaux visibles spectrométriques répond la variété des bandes radar (L, C, X, P avec leurs diverses polarisations) et des bandes infrarouges. S'ajoutent, lors de l'exploitation de ces données, les informations collectées au sol et disponibles à travers les systèmes d'information géographique (cartes, cadastres, plans d'occupation des sols, relevés géologiques ou agronomiques, modèles altimétriques, etc.). Les décisions du photo-interprète relèvent de l'ensemble de ces sources. Son expertise, acquise par de longues études relayées par de multiples expériences, permet d'associer de façon très efficace l'indice détecté dans une image et sa confirmation dans une seconde s'il sait « lire » cette seconde et « rapprocher » les deux événements. La prolifération des images rend ces deux tâches de plus en plus difficiles par la création d'une combinatoire galopante, et par l'impossibilité d'acquérir une expertise réelle dans toutes les modalités disponibles.

• Une situation assez semblable apparaît en **imagerie médicale**. Les techniques d'imagerie se diversifient : rayons X, résonance magnétique, imagerie nucléaire, imagerie ultrasonore, chacune pouvant se décliner selon des modalités diverses en fonction des protocoles d'acquisition. Selon une pratique ancienne, le milieu médical confie chaque type d'image à un expert qui porte un diagnostic partiel sur la modalité de sa spécialité, puis les spécialistes échangent leurs expériences et de cette confrontation naît le diagnostic final. Le souci de regrouper toutes les sources d'images sur une même console conduit à l'introduction progressive des systèmes d'archivage et de consultation intégrés à l'hôpital (les PACS des anglo-saxons [*Picture Archiving and Communication Systems* : systèmes d'archivage et d'échange d'images]). Mais sans les outils d'aide à la décision associés à ces systèmes, ces autoroutes informatiques et ces bases de données spécialisées demeurent encore d'une pratique modeste, tant l'exploitation de nombreuses images variées est une tâche difficile pour un seul expert.

Nous voyons sur ces deux exemples les **raisons profondes de la nécessité de la fusion de données** : apporter des outils d'aide à la décision capables d'intégrer, au fur et à mesure de leur mise au point, des modes d'imagerie nouveaux, en préparant et en raccourcissant l'étape d'élaboration de l'expertise humaine, gérer une profusion grandissante d'images sans sacrifier les apports potentiels d'une combinatoire complexe, c'est-à-dire extraire de chaque image toute l'information qu'elle peut apporter au-delà même de ce que les experts humains savent aujourd'hui en tirer.

■ Dans ce contexte, qu'appelle-t-on plus précisément la **décision** en fusion d'images, et quelles formes peut-elle prendre ? Tout d'abord, précisons que nous excluons ici tout ce qui concerne le recalage, la composition colorée d'images multisources, la visualisation, et les cas simples où il existe une métrique sur l'espace des données multisources, même si ces problèmes sont importants, et souvent indispensables en tant qu'étape préliminaire à la fusion (comme le recalage dans certains cas). L'objectif du recalage est de superposer exactement les pixels correspondant à un même objet observé dans les diverses modalités. Cette étape peut être facilitée s'il existe un référentiel absolu reconnu pour décrire la scène. Les techniques de recalage sont nombreuses et utilisent des principes variés (on pourra consulter [79] et [83] pour une synthèse des méthodes en imagerie aérienne et satellitaire et en imagerie médicale respectivement). Elles sont bien maîtrisées dans de nombreux cas, mais il demeure des domaines reconnus où le recalage automatique est peu fiable. C'est le cas lorsque les capteurs sont très hétérogènes et que les déformations géométriques qui affectent les images sont complexes. Si le recalage géométrique des images est universellement pratiqué aujourd'hui avant toute étape de fusion, il n'est pas clair que ce soit une étape nécessaire ni même souhaitable. En fait, l'utilisateur n'a jamais besoin que de connaître l'application d'une image dans l'autre ; mettre en œuvre cette application au moyen d'un rééchantillonnage et de filtrage dégrade généralement les signaux sans autre bénéfice que d'autoriser un contrôle visuel de la superposition des images. Si l'étape de recalage semble donc généralement échapper à la fusion d'images (puisqu'il peut être fait auparavant de façon totalement indépendante), il demeure des circonstances où ce recalage doit être conçu dans la phase de fusion, c'est en particulier le cas du recalage symbolique où l'on associe à une image des descriptions de haut niveau des objets qui la composent : une ville dans l'association carte – image satellitaire, ou une aire fonctionnelle du cerveau dans l'association atlas – image de résonance magnétique. Dans cette situation, on ne conçoit pas aujourd'hui de mise en correspondance pixel par pixel, mais une association entre représentations de niveau élevé qui passe alors par les étapes de modélisation, puis de décision qui sont les composants de la fusion. Dans la suite, **nous supposerons que le recalage a été mené à son terme**.

Citons les **exemples de décision les plus typiques en fusion d'images** :

● **Test ou évaluation d'une hypothèse** : il peut s'agir par exemple de la présence ou non d'un objet donné dans une image (par exemple un véhicule sur une route, une sténose dans un vaisseau). Ces objectifs de détection sont parfois combinés avec le suivi des objets détectés, et on se rapproche alors des problèmes de suivi traités en signal et en robotique.

● **Classification** : la classification multisources est certainement l'application la plus courante en fusion d'images, que ce soit en imagerie aérienne et satellitaire ou en imagerie médicale. La décision consiste à attribuer à une classe chaque élément sur lequel on raisonne, qu'il s'agisse du pixel au niveau le plus bas, de régions ou d'objets à des niveaux plus élevés, en fonction des informations fournies par toutes les sources sur cet élément.

● **Segmentation** : il s'agit là d'un objectif plus focalisé que la classification, en cherchant à extraire le plus précisément possible des objets déterminés. Il peut s'agir par exemple de la détermination du réseau routier, en fusionnant le tracé précis fourni par une image aérienne, et la structure topologique du réseau et sa géométrie approximative fournies par une carte.

● **Reconnaissance d'objets** : par exemple la reconnaissance d'une zone d'activation fonctionnelle en imagerie cérébrale peut être obtenue par fusion d'images fonctionnelles (TEP ou IRM fonctionnelle) et d'images anatomiques (IRM) permettant de restreindre la zone de recherche au cortex par exemple. Cette focalisation permet d'une part de guider la reconnaissance, et d'autre part de préciser la localisation anatomique de la zone activée. Le résultat de la reconnaissance contient alors à la fois une composante fonctionnelle et une composante anatomique.

TEP : tomographie par émission de positons ; IRM : imagerie par résonance magnétique

● **Interprétation de scène** : ce type d'objectif regroupe les précédents en cherchant non seulement à classifier, segmenter et reconnaître les différentes structures présentes dans la scène, mais aussi leurs relations, interactions, permettant d'avoir une vision de la structure globale de la scène.

● **Reconstruction 3D** : il peut s'agir par exemple de la fusion de modalités complémentaires telles que l'échographie endovasculaire et les angiographies X pour reconstruire la géométrie et la morphologie 3D de segments vasculaires, ou encore de la fusion de données laser et d'images stéréoscopiques pour la reconstruction 3D d'une ville ou de bâtiments.

● **Détection de changements** : ce type de décision concerne typiquement les images acquises à des dates différentes, qu'il s'agisse d'une carte et d'une image, ou d'images multidates pour le suivi des cultures ou le suivi d'une pathologie. Il peut s'agir également de séquences d'images multisources (de cadence plus élevée que les images multidates).

● **Mise à jour de la connaissance sur un phénomène ou une scène** : contrairement au cas précédent, ici la décision consiste à utiliser les informations provenant de différentes sources (éventuellement multidates) pour modifier ou compléter une connaissance précédente, par exemple compléter un réseau routier avec les nouveaux ronds-points pour mettre à jour une carte.

Parmi ces types de décision, certains se retrouvent dans d'autres domaines. Par exemple les tests d'hypothèses se retrouvent en fusion en traitement du signal, la mise à jour des connaissances et la détection en robotique, etc.

Parmi ces différents problèmes de décision, certains sont proches de la combinaison d'experts. En effet, chaque image peut être considérée comme un expert donnant son opinion selon ses compétences. Mais en général l'information dans les problèmes de fusion d'experts est plus éparse qu'en image. L'apprentissage est donc plus difficile car les données sont moins nombreuses, mais en revanche l'utilisateur a souvent moins de contraintes sur les coûts algorithmiques des méthodes. En image, la quantité des données à fusionner est à la fois un avantage pour l'apprentissage et un inconvénient pour les charges de calcul.

Par rapport aux problèmes de fusion portant sur l'agrégation et l'optimisation multicritère, une différence essentielle est que ces problèmes cherchent une solution qui satisfasse au mieux un ensemble de contraintes plus ou moins souples. Au contraire, les images procurent (de manière plus ou moins explicite) un degré de satisfaction (de l'appartenance à une classe par exemple, qui peut alors être considérée comme un critère) et la décision consiste plutôt à choisir la meilleure (meilleure classe par exemple).

2.2 Les situations de fusion

Selon les applications, les problèmes de fusion peuvent se produire dans des situations différentes, dans lesquelles les types d'informations ne sont pas les mêmes. Les principales situations de fusion en traitement d'images sont les suivantes.

● **Plusieurs images du même capteur** : il s'agit par exemple de plusieurs canaux du même satellite, d'images multiéchos en IRM, ou encore de séquences d'images pour des scènes en mouvement. Les données sont alors relativement homogènes car elles correspondent à des mesures physiques similaires.

● **Plusieurs images de capteurs différents** : c'est le cas le plus fréquent, où les principes physiques différents des capteurs permettent d'avoir des points de vue complémentaires sur la scène. Il peut s'agir d'images ERS et SPOT, d'images IRM et ultrasonores, etc. L'hétérogénéité est alors beaucoup plus importante, les différents capteurs ne parlant pas du même aspect du phénomène. Les images donnent chacune une vision partielle et ne sont pas informatives sur les caractéristiques auxquelles elles ne sont pas dédiées (par exemple une IRM anatomique ne donne pas d'information fonctionnelle, et une image TEP a une résolution trop limitée pour que l'anatomie y soit vue de manière suffisamment précise).

ERS : European Remote Sensing Satellite ; SPOT : satellite pour l'observation de la Terre

● **Plusieurs informations extraites d'une même image** : il s'agit de situations dans lesquelles on extrait divers types d'informations d'une image à l'aide de plusieurs détecteurs, opérateurs, classifieurs, etc., qui s'appuient sur des caractéristiques différentes des données, cherchent à extraire des objets différents, rendant ainsi les informations à fusionner souvent très hétérogènes. Les informations extraites peuvent concerner le même objet (fusion de détecteurs de contours par exemple) ou des objets différents et on cherche alors une interprétation globale de la scène et une cohérence entre les objets. Elles peuvent être de niveaux différents (très locales ou plus structurelles quand on s'intéresse aux relations spatiales entre objets).

● **Images et autre source d'information** : par autre source d'information, on entend par exemple un modèle, qui peut être soit particulier comme une carte, soit générique comme un atlas anatomique, des bases de connaissances, des règles, des informations issues d'experts, etc. Les informations sont à nouveau de types très différents, à la fois dans leur nature, et dans leur représentation initiale (images quand il s'agit d'une carte ou d'un atlas numérisé, mais aussi descriptions linguistiques, bases de données, etc.).

2.3 Caractéristiques des données en fusion d'images

Les spécificités de la fusion en traitement d'images ne permettent que difficilement d'exploiter les progrès réalisés dans d'autres domaines de la fusion d'informations.

■ Une des raisons est **la complexité des données et des connaissances**, qui rendent impossible la recherche d'un système global pour regrouper en une relation unique toutes les composantes de l'image.

La complexité vient d'une part du **volume d'informations** à traiter (par exemple, une seule image IRM du cerveau occupe un volume de 8 à 16 megaoctets). Ces gros volumes de donnés, sur lesquels des apprentissages statistiques sont souvent possibles, sont une des raisons qui expliquent la large utilisation de méthodes probabilistes et statistiques en fusion d'images.

La complexité vient d'autre part de la **forte hétérogénéité des informations** à combiner, qu'il s'agisse d'images issues de capteurs différents, d'images et de modèles, ou de caractéristiques différentes extraites d'une ou plusieurs images. Cette hétérogénéité est présente à la fois dans la nature des informations et dans leur représentation. Les données peuvent être **fréquentes**, par exemple lorsqu'il s'agit des cas typiques rencontrés dans une application donnée, pour lesquels il est possible d'avoir des données statistiques par exemple. Elles peuvent aussi être **rares** (par exemple des images pathologiques) et dans ce cas il est beaucoup plus difficile de les modéliser de manière statistique. La combinaison des deux types de données est fréquemment rencontrée en traitement d'images. De plus, elles peuvent être soit **factuelles** (typiquement, une photographie d'une scène à un instant donné) ou **génériques** (modèle, règles, connaissances générales sur l'application). Combiner des informations de spécificités différentes pose souvent des problèmes de conflit à résoudre. En traitement d'images, ce n'est pas une tâche facile, car l'information factuelle n'est pas toujours suffisamment sûre et précise pour qu'on puisse systématiquement lui donner la priorité sur des informations moins spécifiques et plus génériques qui peuvent admettre des exceptions.

La combinaison des informations est souvent guidée ou contrainte par des **informations supplémentaires** (sur l'information à combiner, sur le contexte et le domaine d'application). C'est aussi une source de forte hétérogénéité. Parmi les exemples d'informations supplémentaires sur l'information, on peut citer la **fiabilité** d'une source, soit globale, soit conditionnelle aux objets observés. Ce cas est très souvent rencontré en classification d'images multisources, où une image peut être fiable pour une classe mais pas pour une autre. Donnons maintenant quelques **exemples d'informations supplémentaires** sur le domaine et le contexte :

— les fleuves sont sombres dans le canal XS3 de SPOT (information liant le type d'acquisition à une observation) ;

— le LCR (liquide céphalorachidien) est sombre dans les images IRM en T1 ;

— les routes se croisent pour former des carrefours (contrainte d'intégrité).

Ces informations génériques sont utilisées pour guider le processus de fusion. Le dernier exemple est un cas typique de règle avec exception. La règle donne le cas le plus général, mais elle n'est plus vraie dans le cas d'impasses par exemple.

La **fusion active** est un des moyens pour réduire la complexité en choisissant à chaque instant les meilleures informations à fusionner. Ce choix peut être effectué en fonction d'un résultat partiel de fusion obtenu dans une étape antérieure, de mesures d'informations, d'informations externes susceptibles de guider la fusion, de l'identification d'ambiguïtés à lever, etc.

■ Les informations fusionnées en traitement d'images sont fortement entachées d'**imperfections** (incertitude, imprécision, incomplétude, ambiguïté, conflit, etc., selon la distinction proposée dans le paragraphe 1.2). Ces imperfections trouvent leur origine à différents niveaux, depuis les phénomènes observés jusqu'aux traitements.

Exemples

• La transition douce entre tissus sains et pathologiques est une **imprécision** due au phénomène physiologique. De la même façon, des caractéristiques similaires entre deux tissus différents se retrouvent sur les images mesurant ce type de caractéristique et se traduisent par une **incertitude** sur l'appartenance d'un point particulier à un tissu ou à l'autre, cette incertitude étant due à la fois au phénomène et au capteur.

• La délocalisation de l'information spatiale, due au regroupement en un même pixel de l'information contenue dans tout un volume, est due au capteur et à sa résolution, et constitue une **imprécision** sur la localisation de l'information sur l'image (effet de volume partiel).

• Les phénomèmes de Gibbs au niveau des transitions nettes, qui apparaissent en IRM ou en imagerie radar par exemple, sont des sources d'**imprécisions** dues aux algorithmes de reconstruction numérique des images.

• La représentation d'informations (symboliques) sous forme schématique (par des cartes ou des atlas) est source à la fois d'**imprécision** et d'**incertitude**.

Imprécisions et incertitudes sont ensuite renforcées dans les primitives extraites des images et sur lesquelles s'appuie la fusion. L'exemple le plus familier est celui de la détection de contours par des filtrages gaussiens à différentes échelles : en augmentant l'écart-type de la gaussienne, on gagne en certitude sur la présence de contours mais on perd en précision sur leur localisation. Cet **antagonisme entre précision et certitude** a bien été identifié comme un trait caractéristique de la démarche en reconnaissance des formes [110]. De cet antagonisme naissent souvent des contradictions en fusion d'images puisque l'on dispose de plusieurs mesures sur un même événement : si ces données sont précises, alors elles sont probablement incertaines, et elles risquent donc d'être en contradiction les unes avec les autres ; si l'on renforce leur certitude, cela se fait souvent au prix de plus d'imprécision, rendant les données peu informatives si cette imprécision devient trop importante. Il conviendra donc qu'un système de décision en fusion gère explicitement incertitude et imprécision pour éviter les incohérences.

L'imprécision n'est pas une caractéristique propre aux données, mais elle peut également être attachée aux **objectifs** et aux **buts**, en particulier lorsqu'ils sont exprimés sous forme linguistique vague.

■ Enfin, le **caractère spatial** de l'information, spécifique au traitement d'images, mérite une attention particulière. Son introduction dans les méthodes de fusion, souvent inspirées d'autres domaines dans lesquels ne figure pas ce caractère spatial, n'est pas immédiate mais pourtant nécessaire afin de garantir la cohérence spatiale des résulats. L'imprécision est également présente à ce niveau. À bas niveau, il s'agit de problèmes de recalage, ou de volume partiel par exemple. À plus haut niveau, il s'agit par exemple de relations entre objets qui peuvent être intrinséquement vagues et mal définies (relations du type « à gauche de » par exemple).

■ Comme dans les autres applications de la fusion, la **redondance** et la **complémentarité** entre les images à fusionner sont des atouts pour réduire les imperfections telles que l'incertitude et l'imprécision, pour lever des ambiguïtés, pour compléter l'information, pour résoudre des conflits.

Voici quelques **exemples de complémentarité** en fusion d'images.

• **Sur l'information elle-même** : parties cachées qui peuvent être différentes dans des images de profondeur ou des images aériennes.

• **Sur le type d'information** : information anatomique ou information fonctionnelle pour le même sujet dans des modalités d'imagerie différentes.

• **Sur la qualité de l'information** : deux images de même type mais avec des paramètres d'acquisition différents qui peuvent donner des informations de qualités différentes pour différentes structures.

La redondance provient de manière plus évidente du fait que ce sont des images de la même scène que l'on cherche à combiner. Pour certains problèmes de fusion, comme les études de groupe en imagerie fonctionnelle, la redondance (quelles sont les zones activées chez tous les sujets ?) et la complémentarité (où sont les différences ?) sont elles-mêmes sujet d'étude.

■ Le **conflit** est un problème très délicat, comme dans d'autres applications de la fusion d'informations, comme cela a été discuté dans le paragraphe 1.2. En image, les exemples de conflit qui ne sont qu'apparents et facilement confondus avec la complémentarité se trouvent dans les cas où une image n'est pas capable de différencier deux classes alors qu'une autre les distingue. Les imprécisions et incertitudes sont également sources de conflit. Par exemple la mauvaise localisation d'un contour peut provoquer un conflit entre plusieurs détecteurs de contours. Les conflits dus à des spécificités différentes des informations à combiner sont fréquents dans les applications de fusion d'images et de modèles.

Exemples

• La reconnaissance de structures cérébrales par fusion d'images IRM et de données issues d'un atlas anatomique doit faire face à la variabilité entre les individus, souvent non représentée dans l'atlas, ou encore à la présence possible de pathologies dans des images de patients, non présentes dans le modèle générique.

• Des problèmes similaires se rencontrent dans les problèmes de fusion d'images aériennes et satellitaires et de cartes numériques. Ici les conflits peuvent provenir du tracé imprécis de la carte, de modifications dans la scène non portées sur une carte plus ancienne, etc.

2.4 Contraintes

Les contraintes spécifiques au traitement d'images à prendre en compte dans un processus de fusion sont de plusieurs ordres.

Du point de vue de l'architecture du système de fusion, il est rare que les systèmes décentralisés soient imposés. Le cas le plus fréquent est celui de la fusion *off-line* où l'on dispose simultanément de toutes les informations. Des systèmes centralisés peuvent alors être utilisés.

Les **contraintes de temps réel** sont assez peu fréquentes, sauf dans les applications de surveillance ou multimédias, dans lesquelles elles sont amenées à prendre une place grandissante. Notons que ces contraintes sont beaucoup plus fortes dans d'autres domaines, en robotique par exemple.

En revanche, les **contraintes de cohérence** spatiale sont très fortes et constituent un point important des recherches en fusion d'images. De plus en plus de travaux sont donc consacrés à la prise en compte de l'information spatiale, soit au niveau local via le contexte spatial, soit à un niveau plus structurel via des relations spatiales entre les structures ou les objets de la scène.

Le fait que les données soient volumineuses peut imposer des **contraintes de temps de calcul**. Ainsi les manipulations au niveau des pixels sont limitées à des opérations simples. Des opérations plus sophistiquées nécessitent souvent l'introduction d'informations de plus haut niveau, et une représentation plus structurelle de l'information.

La complexité et la quantité des données à manipuler impose souvent que soit effectué **un choix** dans les informations et les connaissances à fusionner. Ce choix est bien sûr guidé principalement par des critères de pertinence par rapport aux objectifs de décision, mais aussi par des critères portant sur la facilité d'accès et de représentation des informations et des connaissances, et sur leur qualité.

Enfin, un problème crucial est celui de l'**évaluation**. Dans les problèmes de fusion d'images, la « vérité » ou la « bonne » solution existe en général plus ou moins, mais elle est le plus souvent difficile d'accès. Ainsi l'évaluation et la validation d'une méthode de fusion ne peut se faire que sur des simulations, des acquisitions sur fantômes (*), ou par comparaison avec une décision manuelle. Cette situation s'oppose à celle rencontrée dans les problèmes de vote ou de choix social, dans lesquels la vérité n'existe pas, mais où on cherche à trouver une « meilleure » solution au sens d'un compromis, ou de critères d'équité et d'éthique.

(*) Un fantôme est une maquette utilisée pour remplacer l'être humain lors de la mise au point des systèmes d'imagerie médicale.

2.5 Aspects numériques et symboliques en fusion d'images

Si le côté **numérique** de l'information manipulée en traitement d'images est évident, le côté symbolique mérite qu'on s'y arrête un peu plus.

Les **informations symboliques** peuvent être attachées aux données à fusionner (par exemple une information graphique dans une carte ou un atlas anatomique, des attributs calculés sur les données ou sur des objets extraits préalablement des données), ou attachées à la connaissance du domaine. Typiquement, les informations sur le domaine sont souvent représentées par des règles, des représentations structurelles telles que des graphes, souvent utilisés en reconnaissance des formes dans les images, des contraintes à prendre en compte dans les algorithmes. Une **information structurelle** peut par exemple spécifier qu'un réseau routier peut être représenté par un graphe en utilisant des routes et des carrefours, ou exprimer sous forme propositionnelle des règles générales sur la scène telles que « les ventricules cérébraux sont toujours à l'intérieur de la matière blanche », etc. Les informations structurelles peuvent également être représentées de manière iconique, donc proche des données images, par exemple dans le cas de cartes numérisées ou d'atlas anatomiques. On combine alors la représentation géométrique des structures et la nature ou la sémantique de ces structures. Mais les informations sur le domaine peuvent être également **purement numériques** quand il s'agit par exemple de caractéristiques de l'acquisition comme les longueurs d'onde utilisées en imagerie satellitaire, ou les temps d'acquisition en imagerie médicale. Des **représentations hybrides**, dans lesquelles des nombres sont traités comme des symboles avec une quantification, sont utilisées en fusion d'images, pour quantifier la qualité d'un détecteur, pour évaluer une donnée symbolique ou la confiance dans une mesure, la fiabilité d'une image pour certaines classes ou structures, etc. Du point de vue des traitements, cet aspect hybride se manifeste par exemple lorsqu'une proposition établissant que la reconnaissance d'une structure ne dépend que du contexte local conduit à une modélisation appropriée dans un cadre markovien. Les représentations symboliques sont ainsi utilisées en tant que connaissance a priori ou connaissance contextuelle ou générique pour guider un traitement numérique. Elles servent aussi de support structurel par exemple en fusion d'images et de cartes [92].

3. Fusion probabiliste et bayésienne

3.1 Mesures d'information

Lorsqu'on dispose d'un ensemble de ℓ images I_j , une première tâche consiste souvent à le transformer en un sous-ensemble plus réduit, donc de traitement plus simple, sans perdre d'information.

L'approche de l'analyse en composantes principales est souvent employée, qui projette chaque image sur les vecteurs propres de la matrice de corrélation, et permet ainsi d'obtenir ℓ nouvelles images décorrélées, classées par ordre décroissant d'énergie. Une troncature aux ℓ' ($\ell > \ell'$) premières images permet souvent de conserver l'essentiel de l'énergie de l'ensemble d'origine.

Mais, en pratique, cette méthode montre très vite ses limites en traitement d'images car elle ne permet de prendre en compte ni les dépendances complexes entre images, ni les variations spatiales de dépendance.

Pour exprimer l'apport d'information dû à l'ajout d'une nouvelle image $I_{\ell+1}$ à un ensemble déjà connu $\{I_1, ..., I_\ell\}$, on préfère l'approche proposée par Shannon et reposant sur les notions d'**information** et d'**entropie** [73][80][81].

À partir de la probabilité jointe des ℓ premières images $p(I_1, ..., I_\ell)$ (estimée le plus souvent par dénombrement, par exemple à partir de l'histogramme multidimensionnel des niveaux de gris des images), on définit l'**entropie** (ou information moyenne par pixel) des ℓ premières images par :

$$H(I_1, ..., I_\ell) = -\sum p(I_1, ..., I_\ell) \ln p(I_1, ..., I_\ell) \quad (1)$$

et l'entropie apportée par la $(\ell + 1)^e$ image s'exprime, soit en fonction des entropies, soit en fonction des probabilités par :

$$\begin{aligned} H(I_{\ell+1}|I_1, ..., I_\ell) &= H(I_1, ..., I_{\ell+1}) - H(I_1, ..., I_\ell) \\ &= -\sum p(I_1, ..., I_{\ell+1}) \ln p(I_{\ell+1}|I_1, ..., I_\ell) \end{aligned} \quad (2)$$

Pour deux images, on définit ainsi la **redondance** entre elles par :

$$R(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2) \quad (3)$$

Nota : cette redondance ne peut malheureusement pas être étendue à plus de deux images sans perdre potentiellement sa propriété de positivité.

On définit, de même la **complémentarité** de l'image I_2 par rapport à l'image I_1, c'est-à-dire la quantité moyenne d'information qu'il faut ajouter à I_2 pour retrouver I_1 :

$$C(I_1|I_2) = H(I_1|I_2) \quad (4)$$

ce qui conduit à la relation :

$$H(I_1) = R(I_1, I_2) + C(I_1|I_2) \quad (5)$$

Des approches analogues peuvent être envisagées dans un cadre non probabiliste, en s'appuyant par exemple sur l'entropie floue [77]. Le formalisme est pour l'instant moins développé dans cette direction.

En fusion d'images, on utilisera des images fortement redondantes pour confirmer une décision incertaine et des images complémentaires pour élargir le champ des décisions. Des images complémentaires peuvent conduire soit à des décisions conflictuelles soit à des décisions consensuelles.

En traitement d'images, la notion d'entropie a été élargie pour caractériser non seulement la dispersion des mesures sur l'espace de mesure, mais également la cohérence spatiale des mesures en prenant en compte des probabilités d'occurrence de configurations particulières de pixels, soit dans le cadre de classifications [81][82], soit dans celui de champs markoviens [118][121].

La notion d'**entropie globale** n'est pas toujours bien adaptée aux problèmes de fusion, et la notion d'**entropie conditionnelle aux classes à reconnaître**, par exemple, est souvent préférable : elle permet une analyse plus fine de l'information qu'apporte chaque image pour chaque classe et est donc mieux adaptée aux problèmes pour lesquels une source est meilleure pour certaines classes et moins bonne pour d'autres. Bien que la définition formelle de tels concepts ne pose pas de difficulté particulière, ils ne sont pas encore beaucoup utilisés en fusion et c'est vraisemblablement un point qui mériterait d'être approfondi.

3.2 Modélisation et estimation

La théorie la plus exploitée dans la littérature est de loin la **théorie des probabilités**, associée à la **théorie bayésienne de la décision** [57]. L'information y est modélisée par une **probabilité conditionnelle**, par exemple, la probabilité pour qu'un pixel appartienne à une classe particulière, étant donné les images disponibles. Ainsi, la mesure introduite dans le paragraphe 1.4 s'écrit-elle sous la forme :

$$M_i^j(x) = p(x \in C_i|I_j) \quad (6)$$

Cette probabilité est calculée à partir de **caractéristiques** $f_j(x)$ **de l'information extraites des images**. Il peut s'agir, dans les cas les plus simples, du niveau de gris du pixel considéré, ou d'informations plus complexes nécessitant des traitements préliminaires. L'équation (6) ne dépend alors plus de toute l'image I_j mais s'écrit sous la forme simplifiée :

$$M_i^j(x) = p(x \in C_i|f_j(x)) \quad (7)$$

En traitement des images, en l'absence de modélisations fonctionnelles fortes des phénomènes observés, les probabilités $p(f_j(x)|x \in C_i)$, ou plus généralement $p(I_j|x \in C_i)$ (qui représente la probabilité, conditionnelle à la classe C_i, de l'information fournie par l'image I_j), sont apprises par dénombrement sur des zones de test (ou par apprentissage sur ces zones des paramètres d'une loi donnée) et on en déduit les probabilités des équations (6) et (7) par application de la règle de Bayes.

L'avantage essentiel des méthodes probabilistes vient de ce qu'elles reposent sur une base mathématique solide et ont été l'objet de nombreux travaux. Elles proposent donc un éventail d'outils très riche permettant aussi bien la modélisation (par exemple par des familles de lois paramétriques aux propriétés bien étudiées) que l'apprentissage des modèles (pour des lois paramétriques ou non paramétriques) (voir par exemple [34][74][78]). Elles proposent également des règles d'usage soit théoriques (bornes, valeurs asymptotiques) soit heuristiques (tests d'hypothèses, critères de validité, tables de confiance). Enfin la modélisation probabiliste, soutenue par l'interprétation fréquenciste largement répandue dans le monde de la physique et du traitement du signal, est actuellement un concept universellement partagé qui sert naturellement de base de comparaison aux autres modélisations.

Mais les méthodes probabilistes sont également l'objet de critiques. Tout d'abord, si elles représentent bien l'incertain qui entache l'information, elles ne permettent pas aisément de représenter son imprécision, et elles conduisent souvent à confondre ces deux notions. Ensuite, elles nécessitent que, lors de l'apprentissage, des contraintes très strictes soient vérifiées par les mesures (imposées par les axiomes de base des probabilités) et par l'ensemble de classes considéré (exhaustivité). Ces contraintes peuvent rendre l'apprentissage très délicat [comment caractériser des zones qui ne soient pas du blé en imagerie aérienne ? (ce problème est un exemple du problème plus général que l'on rencontre en classification et en reconnaissance de formes : en général le complémentaire d'une classe n'est pas une classe)], ou, si le problème à traiter est complexe, conduire pratiquement à des incohérences car l'utilisateur ne peut alors prendre en compte tout le réseau des dépendances probabilistes (cas des boucles logiques [96]). L'apprentissage de lois de probabilités nécessite, outre la vérification d'hypothèses contraignantes, un nombre de données important. Typiquement l'apprentissage non paramétrique d'une loi multidimensionnelle dans des images ou des zones de taille limitée n'est souvent pas pertinent car les grandeurs à déterminer peuvent être plus nombreuses que les mesures, et on se tourne alors vers des modèles paramétriques, qui à leur tour nécessitent des hypothèses sur la forme des lois.

3.3 Combinaison dans un cadre bayésien

Dans le modèle bayésien, la fusion peut être effectuée de manière équivalente à deux niveaux :

— soit au niveau de la **modélisation**, et on calcule alors des probabilités de la forme :

$$p(x \in C_i|I_1, ..., I_\ell) \quad (8)$$

à l'aide de la règle de Bayes :

$$p(x \in C_i|I_1, ..., I_\ell) = \frac{p(I_1, ..., I_\ell|x \in C_i)p(x \in C_i)}{p(I_1, ..., I_\ell)} \quad (9)$$

où les différents termes sont estimés par apprentissage ;

— soit par la **règle de Bayes** elle-même, où l'information issue d'un capteur vient mettre à jour l'information sur x estimée d'après les capteurs précédents (c'est la seule forme utilisable si les informations sont disponibles successivement et non simultanément) :

$$p(x \in C_i | I_1, ..., I_\ell) = \frac{p(I_1 | x \in C_i) p(I_2 | x \in C_i, I_1) ... p(I_\ell | x \in C_i, I_1, ..., I_{\ell-1}) p(x \in C_i)}{p(I_1) p(I_2 | I_1) ... p(I_\ell | I_1, ..., I_{\ell-1})} \quad \textbf{(10)}$$

Très souvent, étant donné la complexité de l'apprentissage à partir de plusieurs capteurs et la difficulté d'obtenir des statistiques suffisantes, ces équations sont simplifiées sous l'hypothèse d'**indépendance**. Là encore, des critères ont été proposés pour vérifier la validité de cette hypothèse. Les formules précédentes deviennent alors :

$$p(x \in C_i | I_1, ..., I_\ell) = \frac{\prod_{j=1}^{\ell} p(I_j | x \in C_i) p(x \in C_i)}{p(I_1, ..., I_\ell)} \quad \textbf{(11)}$$

Cette équation fait apparaître clairement le type de combinaison des informations, sous la forme d'un produit, donc une fusion conjonctive. Il est notable que la probabilité a priori joue exactement le même rôle dans la combinaison que chacune des sources, auxquelles elle est combinée également par un produit.

L'avantage de cette théorie, du point de vue de la combinaison, est qu'elle repose sur de solides bases mathématiques, et peut être utilisée pour la mise à jour de réseaux complexes de connaissances [96][97]. Elle permet d'introduire des informations qui s'expriment facilement sous forme de probabilités, telles que le contexte spatial dans le cadre des champs de Markov (voir § 6) ou la qualité des informations exprimée comme la probabilité pour qu'une mesure soit fiable [65].

Cependant, elle est contrainte, comme pour la modélisation, par les axiomes des probabilités, et son utilisation en pratique nécessite souvent des hypothèses simplificatrices (comme l'indépendance) rarement vérifiées. Elle nécessite de plus l'estimation des probabilités a priori $p(x \in C_i)$, qui est souvent délicate et est primordiale dans les cas où l'on a peu d'informations (distributions très plates des probabilités conditionnelles). Si, dans le cas du traitement d'images, les probabilités conditionnelles peuvent être souvent bien estimées par apprentissage à partir de fréquences d'occurrence, ce n'est en général pas le cas des probabilités a priori. Leur évaluation sort du cadre des probabilités fréquencistes et fait souvent appel à des concepts plus subjectifs. La forme conjonctive de la fusion bayésienne conduit souvent en pratique à un effondrement des probabilités des événements qui sont déduits d'une longue chaîne de déduction. Enfin, elle ne permet pas de modéliser l'ignorance pour la prendre en compte dans la combinaison.

3.4 Combinaison vue comme un problème d'estimation

Une autre manière de voir la fusion probabiliste consiste à considérer que chaque source donne une probabilité (d'appartenance à une classe) par exemple, et que la fusion consiste à combiner ces probabilités pour trouver la probabilité globale d'appartenance à la classe. Cette vision revient à considérer la fusion comme un problème d'estimation, et permet d'utiliser des opérateurs de combinaison différents du produit. En particulier les méthodes de moyenne ou moyenne pondérée, de médiane, de consensus sont souvent employées [38][39][58]. Des estimateurs robustes peuvent également être employés, afin de limiter ou supprimer l'influence des valeurs aberrantes (*outliers*). Enfin, des méthodes issues de la théorie des variables régionalisées [87], telles que le krigeage ou le krigeage universel, pourraient également être utilisées dans ce cadre.

3.5 Décision

La dernière étape concerne la décision, par exemple le choix de la classe à laquelle appartient un point. Cette décision binaire peut être assortie d'une mesure de la qualité de cette décision, pouvant éventuellement conduire à la rejeter. La règle la plus utilisée pour la décision probabiliste et bayésienne est le maximum a posteriori :

$$x \in C_i \text{ si } p(x \in C_i | I_1, ..., I_\ell) = \max\{p(x \in C_k | I_1, ..., I_\ell), 1 \leq k \leq n\} \quad \textbf{(12)}$$

mais de très nombreux autres critères ont été développés par les probabilistes et les statisticiens, pour qu'ils s'adaptent au mieux aux besoins de l'utilisateur et au contexte de sa décision : maximum de vraisemblance, maximum d'entropie, marginale maximale, espérance maximale, risque minimal, etc. Cependant, la grande variété de ces critères laisse l'utilisateur à nouveau démuni devant la justification d'un choix et l'éloigne de l'objectivité recherchéee initialement par ces méthodes.

4. Fusion dans la théorie des fonctions de croyance

La théorie des fonctions de croyance (ou théorie de Dempster-Shafer) date des années 1970 mais son utilisation en fusion d'images est relativement récente. Pourtant les premières applications sont prometteuses, et nous montrons dans ce paragraphe quelles sont les caractéristiques de cette théorie qui justifient que l'on s'y intéresse, aussi bien du point de vue de la représentation des connaissances et de leurs imperfections (imprécision, incertitude, ambiguïté, ignorance, conflit) que de leur combinaison.

4.1 Modélisation

La théorie des fonctions de croyance permet, de manière analogue à la théorie des possibilités comme nous le verrons dans le paragraphe 5, **de représenter à la fois l'imprécision et l'incertitude**, à l'aide de **fonctions de masse** m, de **plausibilité** *Pls* et de **croyance** *Bel* [66][108][112]. Les fonctions de masse sont définies sur **tous les sous-ensembles de l'espace de discernement** D (contenant par exemple les classes auxquelles on s'intéresse) et pas simplement sur les singletons comme les probabilités qui ne mesurent que la probabilité d'appartenance à une classe donnée.

Posons $D = \{C_1, C_2, ..., C_n\}$ où chaque C_i désigne une hypothèse en faveur de laquelle une décision peut être prise (typiquement une classe dans un problème de classification multisources).

■ Une **fonction de masse** est définie comme une fonction de 2^D (ensemble des parties de D) dans [0, 1]. En général on impose $m(\emptyset) = 0$, et une normalisation de la forme :

$$\sum_{A \subset D} m(A) = 1 \quad \textbf{(13)}$$

qui garantit une sorte de commensurabilité entre plusieurs jeux de masse.

La contrainte $m(\emptyset) = 0$ correspond à une hypothèse de monde clos, dans lequel toutes les solutions possibles sont effectivement représentées dans D (ce qui suppose que l'on est capable de les énumérer). Si l'on relâche cette contrainte et que l'on accepte d'avoir une masse strictement positive sur $\emptyset$, cela correspond alors à une hypothèse de monde ouvert, dans lequel des solutions hors de D sont envisageables.

Un élément focal est un sous-ensemble A de D tel que $m(A) > 0$. La réunion des éléments focaux est appelée noyau.

■ Une **fonction de croyance** *Bel* est une fonction totalement croissante définie de 2^D dans [0, 1] :

$$\forall A_1 \in 2^D, \ldots, A_k \in 2^D, \quad Bel(\cup_{i=1\ldots k} A_i) \geqslant \sum_{I \subset \{1\ldots k\}, I \neq \emptyset} (-1)^{|I|+1} Bel(\cap_{i \in I} A_i) \tag{14}$$

où $|I|$ désigne le cardinal de I, et telle que $Bel(\emptyset) = 0$, $Bel(D) = 1$.

Étant donné une fonction de masse m, la fonction *Bel* définie par :

$$\forall A \in 2^D, Bel(A) = \sum_{B \subset A, B \neq \emptyset} m(B) \tag{15}$$

est une fonction de croyance. Inversement, à partir d'une fonction de croyance définie comme une fonction totalement croissante [inégalité (14)] telle que $Bel(\emptyset) = 0$, $Bel(D) = 1$, on peut définir une fonction de masse par :

$$\forall A \in 2^D, m(A) = \sum_{B \subset A} (-1)^{|A-B|} Bel(B) \tag{16}$$

Cette fonction de masse vérifie alors l'équation (15).

La fonction de croyance mesure la confiance totale que l'on a dans un sous-ensemble A.

■ Une **fonction de plausibilité** *Pls* est également une fonction de 2^D dans [0, 1] définie par :

$$\forall A \in 2^D, Pls(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - Bel(A^C) \tag{17}$$

La plausibilité mesure la **confiance maximale** que l'on peut avoir dans A. Cette fonction a une interprétation naturelle dans le modèle des croyances transférables [112] où l'on considère que l'apport d'information peut permettre de transférer des croyances sur des sous-ensembles plus précis. La plausibilité représente alors la croyance maximale que l'on pourrait potentiellement affecter à un sous-ensemble A si l'on apprend par exemple que la solution se trouve dans A (toute la confiance mise dans un sous-ensemble B intersectant A est alors transférée sur A afin de mettre à 0 la confiance sur A^C).

■ On a les **propriétés suivantes** :

$$\forall A \in 2^D, Pls(A) \geqslant Bel(A) \tag{18}$$

$$\forall A \in 2^D, Bel(A) + Bel(A^C) \leqslant 1 \tag{19}$$

$$\forall A \in 2^D, Pls(A) + Pls(A^C) \geqslant 1 \tag{20}$$

$$\forall A \in 2^D, Bel(A) + Bel(A^C) = 1 - Bel(A) = Pls(A) \tag{21}$$

L'intervalle $[Bel(A), Pls(A)]$ est appelé **intervalle de confiance** et sa longueur est une mesure de l'ignorance que l'on a sur un événement A et son complémentaire.

Si l'on affecte des masses uniquement aux **hypothèses simples** ($m(A) = 0$ pour $|A| > 1$), alors les trois fonctions m, *Bel* et *Pls* sont égales et sont une probabilité. Dans les cas plus complexes, ce n'est pas le cas et il n'y a pas d'équivalence directe avec des probabilités. L'analogue des fonctions de crédibilité et de plausibilité pourrait être obtenue par exemple à partir de probabilités conditionnelles à des comportements pessimistes et optimistes respectivement, mais leur formalisation serait beaucoup plus délicate que ce que propose la théorie des fonctions de croyance.

La possibilité d'affecter des masses aux **hypothèses composées**, et donc de travailler sur 2^D plutôt que sur D constitue un des avantages de cette théorie. Elle permet en effet une modélisation très souple et très riche, en particulier de l'ambiguïté ou de l'hésitation entre classes.

■ Citons **quelques exemples de situations** dans lesquelles la fusion par la théorie des fonctions de croyance peut être employée :

— dans les cas limites (que l'on peut considérer comme idéaux) où l'on connaît toute l'information sur le problème à traiter ;
— lorsqu'une source donne des informations seulement sur certaines classes : par exemple certaines images TEP donnent des informations sur les limites du cerveau mais pas de la tête ;
— lorsqu'une source n'est pas capable de différencier deux classes : la théorie des fonctions de croyance permet alors de considérer la disjonction de ces deux classes, sans introduire d'information arbitraire forçant leur séparation ;
— lorsque l'on veut modéliser les effets de volume partiel, donc typiquement représenter l'appartenance d'un pixel ou d'un voxel à plusieurs classes ;
— lorsque l'on veut représenter la fiabilité globale d'une source : cela peut être réalisé en affectant une masse non nulle à D ;
— dans les cas où la fiabilité d'une source dépend des classes (par exemple des images cérébrales fonctionnelles ne donnent pas d'informations très fiables sur l'anatomie, alors qu'au contraire les images IRM sont très fiables pour ce type de classe) ;
— dans les cas où l'on veut introduire des informations a priori : même si ces informations ne sont pas représentables facilement par des probabilités, elles peuvent être introduites si elles induisent un moyen de choisir les éléments focaux (en particulier les disjonctions d'hypothèses), de définir ou de modifier les fonctions de masse.

4.2 Estimation de fonctions de masse

L'estimation des fonctions de masse est un problème difficile, qui n'a pas de solution universelle. La difficulté est augmentée ici si l'on veut affecter des masses aux hypothèses composées [60][76]. En TDSI, on peut les construire à trois niveaux : au niveau le plus haut (souvent abstrait et symbolique), la représentation de l'information est utilisée d'une manière similaire à ce qui est fait en intelligence artificielle et les masses sont affectées à des propositions, et souvent données par des experts [6][63][94]. Ce type d'information n'est le plus souvent pas directement dérivé de mesures dans les images, et les méthodes correspondantes ne sont donc pas spécifiques du TDSI. À un niveau intermédiaire, les masses sont calculées à partir d'attributs et peuvent s'appuyer par exemple sur des modèles géométriques [2][35][41][120]. Ce niveau est bien adapté à des problèmes de reconnaissance des formes à partir de modèles mais il est difficile de l'utiliser pour des problèmes de fusion sur des structures complexes sans modèle. Au niveau du pixel, beaucoup de méthodes sont envisageables, et la plupart s'appuient sur des méthodes de reconnaissance des formes statistique.

La manière la plus simple que l'on puisse imaginer consiste à calculer les masses sur les singletons dans une image I_j par :

$$m_j(\{C_i\})(x) = M_i^j(x) \tag{22}$$

où $M_i^j(x)$ est estimée le plus souvent comme une probabilité. Les masses sur tous les autres sous-ensembles de D sont alors nulles. Il est clair que ce modèle est très réducteur et n'exploite pas les caractéristiques intéressantes de la théorie des fonctions de croyance. Beaucoup d'approches s'appuient toutefois sur un tel modèle initial, puis répartissent les masses sur l'ensemble des hypothèses composées, ou n'utilisent que certaines hypothèses composées, dans une démarche simplificatrice et souvent très heuristique [74][102][120][132]. Mais d'autres approches peuvent également être envisagées. Dans la suite, nous présentons **quelques modèles de la littérature**.

■ **Modification de modèles probabilistes**

● Le modèle le plus simple et le plus souvent utilisé consiste à utiliser la **technique d'affaiblissement** [108]. Les nouvelles masses m' sont calculées à partir des masses initiales m de la manière suivante (l'indice j représentant la source d'information ainsi que l'élément x sur lequel on raisonne sont omis ici) :

$$m'(\{C_i\}) = \alpha m(\{C_i\}) \tag{23}$$

$$m'(D) = 1 - \alpha + \alpha m(D) \tag{24}$$

où $\alpha \in [0, 1]$ est le coefficient d'affaiblissement. Dans le cas où les masses initiales sont apprises sur les singletons seulement, par exemple à partir de probabilités, alors $m(D) = 0$ et $m'(D) = 1 - \alpha$. Cette technique est souvent utilisée pour affaiblir une source en fonction de sa fiabilité, et permet d'affecter une masse à D qui sera faible si la source est fiable et importante si la source ne l'est pas. Dans les cas extrêmes, la valeur $\alpha = 0$ est utilisée pour une source qui n'est pas fiable du tout, et toute la masse est alors affectée à D, ce qui représente l'ignorance totale. La valeur $\alpha = 1$ est utilisée pour une source fiable dans laquelle toute la masse est affectée aux singletons et où il n'y a aucune ambiguïté entre classes.

Ce type de modèle est très simple. L'apprentissage des masses sur les singletons peut bénéficier des techniques classiques d'apprentissage statistique. Cependant les disjonctions d'hypothèses ne sont pas modélisées, ce qui limite beaucoup la portée de ce modèle.

● Deux modèles d'inspiration probabiliste ont été proposés par A. Appriou [4], et prennent en compte d'autres disjonctions que D. Ces modèles supposent une **estimation initiale de probabilités conditionnelles** $p(f(x)|C_i)$ (où $f(x)$ désigne les caractéristiques de x extraites de la source et sur lesquelles s'appuie la fusion), notées plus simplement $p(x|C_i)$. La fonction de masse associée à une source est calculée par combinaison de fonctions de masse associées à chaque singleton, définies dans le premier modèle par :

$$m^i(\{C_i\})(x) = \frac{\alpha_i R p(x|C_i)}{1 + R p(x|C_i)} \tag{25}$$

$$m^i(D \backslash \{C_i\})(x) = \frac{\alpha_i}{1 + R p(x|C_i)} \tag{26}$$

$$m^i(D)(x) = 1 - \alpha_i \tag{27}$$

où α_i est un coefficient d'affaiblissement lié à la classe C_i, qui permet de prendre en compte la fiabilité de la source pour cette classe en particulier (et non plus globale comme dans le modèle précédent), et R est un coefficient de pondération des probabilités. Si $R = 0$, seule la fiabilité de la source est prise en compte, sinon les données sont également prises en compte.

Dans le deuxième modèle, les masses associées à chaque singleton sont définies par :

$$m^i(\{C_i\})(x) = 0 \tag{28}$$

$$m^i(D \backslash \{C_i\})(x) = \alpha_i(1 - R p(x|C_i)) \tag{29}$$

$$m^i(D)(x) = 1 - \alpha_i + \alpha_i R p(x|C_i) \tag{30}$$

Ce modèle correspond au cas où $p(x|C_i)$ nous donne une information essentiellement sur ce qui n'est pas C_i.

La masse associée à la source est ensuite calculée comme $\oplus_i m^i$, où $\oplus$ est la somme orthogonale de Dempster (voir § 4.3). Ce modèle est bien adapté dans les cas où l'on apprend facilement une classe contre toutes les autres, ce qui est fréquent en reconnaissance des formes dans les images, ou dans les cas où chaque classe est déterminée à partir d'un détecteur adapté (par exemple un détecteur de routes dans une image aérienne permet de définir la probabilité d'appartenance à la route par rapport à celle des toutes les autres classes, mais n'est pas capable de distinguer ces autres classes).

● Dans [49], les disjonctions sont définies en fonction d'un **critère de significativité des probabilités conditionnelles**. Si une seule probabilité $p(x|C_i)$ est significative (ce qui nécessite de définir des seuils), alors un modèle simple de masse portant sur les singletons est utilisé. Si plusieurs probabilités sont significatives, les disjonctions des hypothèses correspondantes sont également prises en compte. Par exemple si trois valeurs sont significatives, et telles que $p(x|C_i) > p(x|C_j) > p(x|C_k)$, la fonction de masse est construite par :

$$m(\{C_i\})(x) = p(x|C_i) - p(x|C_j) \tag{31}$$

$$m(\{C_i \cup C_j\})(x) = p(x|C_j) - p(x|C_k) \tag{32}$$

$$m(\{C_i \cup C_j \cup C_k\})(x) = p(x|C_k) \tag{33}$$

puis les masses sont normalisées. Si aucune probabilité n'est significative, la masse porte entièrement sur D.

Modification de modèles de distances

Une approche de type reconnaissance des formes est proposée dans [45]. Si chaque classe C_i est représentée par un prototype (ou un centre) x_i, une fonction de masse associée à chaque classe peut-être définie, dans laquelle C_i et D sont les seuls éléments focaux :

$$m^i(\{C_i\})(x) = \alpha \exp[-\gamma d^2(x, x_i)] \tag{34}$$

$$m^i(D)(x) = 1 - \alpha \exp[-\gamma d^2(x, x_i)] \tag{35}$$

Les paramètres α et γ permettent de jouer sur la quantité d'ignorance et la forme des fonctions de masse. La distance $d^2(x, x_i)$ permet d'affecter une masse d'autant plus importante que x « ressemble » à la classe C_i. Les m^i sont ensuite combinées selon la règle de Demspter pour avoir une masse prenant en compte l'information sur toutes les classes.

Cette approche peut également être appliquée aux k plus proches voisins. La distance est alors la distance de x à l'un de ses voisins, et la masse est affectée selon le modèle précédent à la classe à laquelle appartient ce voisin et à D. Les fonctions calculées pour chacun des voisins de x sont ensuite combinées par la règle de Dempster.

A priori sur les éléments focaux composés (disjonctions)

Dans de nombreuses applications, il est possible de disposer d'informations a priori qui permettent de déterminer de manière supervisée quels sont les éléments focaux à prendre en compte. Ces méthodes ont été utilisées par exemple dans [17][90][91][116]. Dans [17], des images du cerveau sont combinées pour détecter des pathologies. Les fonctions de masse sont estimées à partir des niveaux de gris [25] et les classes non distinguées dans certaines images par leurs niveaux de gris sont regroupées en disjonctions. Dans [116], les résultats de détecteurs de plusieurs structures sont fusionnés pour interpréter une image radar. Ce sont alors les capacités des détecteurs à différencier ou non différentes classes de structures qui permettent de définir les éléments focaux et les disjonctions de classes à prendre en compte. Dans [90][91], des attributs extraits d'images de différents capteurs sont combinés pour différencier des mines d'objets inoffensifs, dans un programme de déminage humanitaire. Les mesures à combiner peuvent être caractéristiques d'une classe ou de l'espace de discernement complet. Par exemple, la profondeur des objets permet d'affecter une masse aux objets inoffensifs si elle est importante, mais ne permet pas de distinguer les types d'objets si elle est faible et la masse est alors affectée à D.

Ce type d'approche est très efficace si l'on dispose de telles informations, mais elle reste supervisée, et donc applicable à des problèmes où le cardinal de D reste raisonnable.

■ **Apprentissage des éléments focaux composés**

Les méthodes d'apprentissage des éléments focaux s'appuient souvent sur des classifications préalables effectuées dans chaque source séparément. Typiquement, de matrices de confusion peuvent être extraites les classes confondues selon une source, dont la réunion constituera un élément focal de la fonction de masse attachée à cette source.

De manière complètement non supervisée, les intersections entre les classes détectées dans une source et celles détectées dans une autre source peuvent définir les singletons de l'espace de discernement, les classes détectées dans chaque source devenant alors des disjonctions [86].

Des mesures de dissonance et consonance sont proposées dans [88]. L'idée consiste à modifier une fonction de masse initiale portant uniquement sur les singletons en affaiblissant les masses sur les singletons en fonction du degré de consonance de ceux-ci, et en créant des masses sur des disjonctions de deux classes en fonction du degré de dissonance entre ces classes. Cette méthode a été appliquée à la fusion de plusieurs classifieurs. La consonance d'une classe est calculée d'après le nombre d'éléments affectés à cette classe par tous les classifieurs, et la dissonance d'après le nombre d'éléments classés différemment.

Dans le cas où les éléments sont caractérisés par une mesure dans un espace à une dimension (typiquement représentés par un histogramme), les masses sur les hypothèses composées peuvent être définies dans les zones de recouvrement ou d'ambiguïté entre deux classes voisines. Une autre méthode, s'inspirant des méthodes de seuillage hiérarchique, est proposée dans [103], où chaque pic de l'histogramme correspond à un singleton. Puis l'histogramme est progressivement seuillé à des hauteurs décroissantes, et des disjonctions sont créées lorsque des maxima se regroupent. Cette méthode est à rapprocher des arbres de composantes utilisés par exemple en morphologie mathématique avec le concept de topologie des coupes, ainsi que des intervalles de confiance et de leurs liens avec les distributions de possibilité.

■ **Introduction de disjonctions par morphologie mathématique**

Sans se restreindre à des espaces de représentation de dimension 1, la méthode proposée dans [18] permet de calculer des masses sur des disjonctions par érosions et dilatations de masses définies dans un premier temps sur des singletons. Les propriétés de ces opérations morphologiques permettent de les interpréter comme des croyances et plausibilités, dont les masses sont ensuite déduites.

4.3 Combinaison conjonctive

■ Soit m_j ($j = 1 \ldots \ell$) la fonction de masse définie pour l'image (ou de manière plus générale la source) j. La **combinaison conjonctive des fonctions de masse** est effectuée selon la règle orthogonale de Dempster [108] :

$$\forall A \subset D, (m_1 \oplus m_2 \oplus \ldots \oplus m_\ell)(A) = \sum_{B_1 \cap \ldots \cap B_\ell = A} m_1(B_1) m_2(B_2) \ldots m_\ell(B_\ell) \tag{36}$$

Des justifications axiomatiques de cette règle peuvent être trouvées dans [112]. Les différences entre ces axiomes et ceux de Cox [40] (qui permettent de justifier les règles des probabilités) expliquent les origines des différences entre les deux théories [13].

Dans cette équation non normalisée, la masse affectée par la combinaison à l'ensemble vide est en général non nulle. Elle s'interprète souvent comme le conflit entre les sources. Notons que cette mesure de conflit n'est pas une mesure absolue mais dépend de la modélisation effectuée (en particulier de la répartition des masses sur les différents sous-ensembles de D). Le conflit peut avoir deux sources essentielles : soit les sources ne sont pas fiables, soit elles donnent des informations sur des phénomènes différents. Dans le premier cas, il est acceptable de combiner les sources, et une solution pour prendre en compte le conflit est d'affaiblir les sources en fonction de leur fiabilité. Dans le deuxième cas, la combinaison n'a pas de sens. Des méthodes de regroupement des sources selon les phénomènes qu'elles observent ont été proposées, visant à combiner les sources à l'intérieur de chaque groupe uniquement. Ces groupes sont calculés de sorte à minimiser le conflit dans chaque groupe [90][106].

■ Dans une hypothèse de monde ouvert, une masse non nulle sur l'ensemble vide peut également représenter une solution non prévue dans D. Sous l'hypothèse du monde fermé, où tout ce qui est possible est représenté dans D, cette interprétation n'est pas acceptable, ce qui conduit à normaliser le résultat de la combinaison sous la forme :

$$\left.\begin{aligned} &(m_1 \oplus m_2 \oplus \ldots \oplus m_\ell)(A) = \frac{\displaystyle\sum_{B_1 \cap \ldots \cap B_\ell = A} m_1(B_1) m_2(B_2) \ldots m_\ell(B_\ell)}{1 - \displaystyle\sum_{B_1 \cap \ldots \cap B_\ell \neq 0} m_1(B_1) m_2(B_2) \ldots m_\ell(B_\ell)} \\ &\text{et} \\ &(m_1 \oplus m_2 \oplus \ldots \oplus m_\ell)(\emptyset) = 0 \end{aligned}\right\} \tag{37}$$

si le dénominateur de l'équation (38) est non nul, c'est-à-dire si :

$$k = \sum_{B_1 \cap \ldots \cap B_\ell = \emptyset} m_1(B_1) m_2(B_2) \ldots m_\ell(B_\ell) < 1 \tag{38}$$

Cette quantité (qui mesure le conflit entre les sources) est donc directement prise en compte dans la combinaison sous forme de **facteur de normalisation**. Elle représente la masse qui serait affectée à l'ensemble vide si l'on n'avait pas cette normalisation (équation (37)). Il est important de prendre en compte cette valeur pour juger de la qualité de la combinaison : celle-ci peut ne pas avoir grand sens en cas de fort conflit et conduire à des décisions criticables.

Prenons un **exemple** simple où $D = \{C_1, C_2, C_3\}$ et deux fonctions de masse n'ayant que les singletons comme éléments focaux et les valeurs suivantes :

$$m_1(C_1) = 0,9 \quad m_1(C_2) = 0,0 \quad m_1(C_3) = 0,1$$

$$m_2(C_1) = 0,0 \quad m_2(C_2) = 0,9 \quad m_2(C_3) = 0,1$$

Leur fusion non normalisée donne :

$$m_1 \oplus m_2(C_1) = 0,0$$
$$m_1 \oplus m_2(C_2) = 0,0$$
$$m_1 \oplus m_2(C_3) = 0,01$$
$$m_1 \oplus m_2(\emptyset) = 0,99$$

En normalisant, on obtient :

$$m_1 \oplus m_2(C_1) = 0,0$$
$$m_1 \oplus m_2(C_2) = 0,0$$
$$m_1 \oplus m_2(C_3) = 1,0$$

Toute la masse est alors focalisée sur C_3 qui est la seule classe où les deux sources sont d'accord, mais pour dire que ce n'est pas une solution plausible. La normalisation a masqué le conflit. La forme non normalisée est le plus souvent préférable en cas de conflit. Ici, elle permet d'affecter la masse essentiellement à l'ensemble vide. Dans cet exemple, l'origine du conflit peut venir de l'hypothèse de monde ouvert, d'une faible fiabilité d'au moins une des deux sources, ou du fait qu'une source voit un objet dans la classe C_1 et la seconde source voit un autre objet dans la classe C_2.

■ D'autres méthodes que la normalisation ont été proposées pour éliminer la masse sur l'ensemble vide. Par exemple, cette masse est affectée à D dans [124]. Dans [54], une méthode plus fine est proposée : par exemple si les éléments focaux A_1 et A_2 de deux sources sont en conflit ($A_1 \cap A_2 = \emptyset$), alors le produit $m_1(A_1)m_2(A_2)$ est affecté à $m(A_1 \cup A_2)$. Cela suppose qu'au moins une des deux sources est fiable mais qu'on ne sait pas laquelle, et la forme disjonctive du résultat est l'attitude la plus prudente.

■ Examinons maintenant les **propriétés de la règle de combinaison**. Elle est **commutative** et **associative**. La fonction de masse définie par :

$$m_0(D) = 1 \text{ et } \forall A \subset D, A \neq D, m_0(A) = 0 \qquad \textbf{(39)}$$

est élément neutre pour la combinaison. Cette masse représente une source complètement non informative, qui ne distingue aucun élément de D. Qu'elle ne joue aucun rôle dans la combinaison correspond donc bien à l'intuition. La définition de cette fonction de masse remplace le principe d'indifférence utilisé en probabilités (équirépartition des probabilités sur tous les éléments), et représente mieux l'absence d'information.

La loi $\oplus$ n'est **pas idempotente**.

Considérons à nouveau l'**exemple précédent**, mais avec cette fois les fonctions de masse suivantes :

$$m_1(C_1) = 0,7 \quad m_1(C_2) = 0,2 \quad m_1(C_3) = 0,1$$

$$m_2(C_1) = 0,7 \quad m_2(C_2) = 0,2 \quad m_2(C_3) = 0,1$$

Leur fusion non normalisée donne :

$$m_1 \oplus m_2(C_1) = 0,49$$
$$m_1 \oplus m_2(C_2) = 0,04$$
$$m_1 \oplus m_2(C_3) = 0,01$$
$$m_1 \oplus m_2(\emptyset) = 0,46$$

En normalisant, on obtient :

$$m_1 \oplus m_2(C_1) = 0,91$$
$$m_1 \oplus m_2(C_2) = 0,07$$
$$m_1 \oplus m_2(C_3) = 0,02$$

Cet exemple illustre la non-idempotence de la règle de combinaison. Les valeurs les plus fortes sont renforcées et les plus faibles diminuées. Il est important également de noter que le conflit entre deux fonctions de masse identiques est non nul, et qu'il est d'autant plus fort que la masse est répartie sur les singletons.

À l'origine, cette règle de combinaison était réputée applicable seulement sous l'hypothèse d'indépendance des sources. Il a été montré [100][101] que la règle est encore applicable sans cette hypothèse, en s'appuyant sur l'analogie avec les ensembles fermés aléatoires. De manière moins technique et plus philosophique, l'indépendance dans le cadre des fonctions de croyance ne doit pas être comprise au sens statistique, mais dans un sens plus « cognitif » [114]. Imaginons par exemple des experts dont on veut combiner les opinions. Ils ne sont vraisemblablement pas indépendants statistiquement (s'ils sont experts du même domaine), mais on peut attendre d'eux qu'ils le soit cognitivement, c'est-à-dire que chacun se forge une opinion sans consulter les autres. C'est à ce type d'indépendance que s'applique la règle de Dempster, ce qui se traduit par la non-idempotence de la règle, conduisant à un renforcement des fonctions de masse identiques. Sous hypothèse de dépendance, on souhaiterait au contraire avoir une règle idempotente. Nous reviendrons sur ces considérations dans la théorie des ensembles flous.

Lorsque les fonctions m, Bel, et Pls sont des probabilités (c'est-à-dire lorsque les seuls éléments focaux sont des singletons), la loi de combinaison de Dempster est cohérente avec les lois classiques des probabilités. Cela fait donc apparaître les probabilités comme la limite de la théorie des croyances lorsqu'il n'y a pas d'ambiguïté ni d'imprécision et que seule l'incertitude des données doit être prise en compte.

La règle de Demspter a un comportement conjonctif, puisqu'elle donne des éléments focaux qui sont les intersections des éléments focaux des fonctions de masse initiales. Elle renforce donc la focalisation, et diminue la longueur des intervalles de confiance $[Bel, Pls]$.

■ En pratique, le **calcul de la combinaison** s'effectue en établissant la table d'intersection des éléments focaux. Par exemple, si m_1 porte sur $C_1 \cup C_2$ (typiquement dans le cas d'une image qui n'est pas capable de différencier ces deux classes) et C_3, et m_2 sur C_1 et $C_2 \cup C_3$, les éléments focaux de $m_1 \oplus m_2$ sont donnés par la table d'intersection suivante :

	$C_1 \cup C_2$	C_3
C_1	C_1	$\emptyset$
$C_2 \cup C_3$	C_2	C_3

Les éléments focaux ne sont alors plus que les singletons et l'ensemble vide. Cet exemple illustre comment la combinaison conjonctive réduit l'imprécision et résout (ou diminue en général) l'ambiguïté de chaque source.

Supposons maintenant que la modélisation inclut une part d'ignorance et qu'une masse non nulle soit affectée à D dans les deux sources. La table d'intersection donne alors :

	$C_1 \cup C_2$	C_3	D
C_1	C_1	$\emptyset$	C_1
$C_2 \cup C_3$	C_2	C_3	$C_2 \cup C_3$
D	$C_1 \cup C_2$	C_3	D

Cette fois, l'ambiguïté n'est que partiellement réduite, et il reste une masse non nulle sur les éléments imprécis (disjonctions de classes). Le conflit en revanche diminue, ce qui est une propriété générale de la combinaison de masses affaiblies en renforçant D.

Prenons maintenant le cas particulier d'une source qui donne une information certaine sur un sous-ensemble B de D. Cette information se modélise de la manière suivante :

$$m_B(B) = 1 \text{ et } \forall A \subset D, A \neq B, m_B(A) = 0 \qquad \textbf{(40)}$$

Toutes les sources doivent alors être « conditionnées » par m_B, afin de prendre en compte que la vérité ne peut être que dans B. Le conditionnement se fait simplement en combinant une fonction de masse m avec m_B :

$$\forall A \subset D, m \oplus m_B(A) = \sum_{A = B \cap C} m(C) \qquad \textbf{(41)}$$

qui s'écrit également :

$$\forall A \subset D, A \not\subset B, m \oplus m_B(A) = 0 \qquad \textbf{(42)}$$

$$\forall A \subset D, A \subset B, m \oplus m_B(A) = \sum_{X \subset B^C} m(A \cup X) \qquad \textbf{(43)}$$

Le conditionnement correspond au modèle de croyances transférables [112] : la connaissance de B conduit à transférer toute la masse sur les sous-ensembles inclus dans B. Ainsi la croyance initialement affectée à un sous-ensemble $A = A_1 \cup A_2$ (avec $A_1 \subset B$ et $A_2 \subset B^C$) représentait le fait que la vérité pouvait être n'importe où dans A. La connaissance de B permet maintenant de préciser l'information et de réduire A à A_1. En quelque sorte, la croyance diffuse dans A est concentrée dans la seule partie qui est incluse dans B.

Dans le cas général, ainsi que le montre la formule (36), la combinaison est de complexité exponentielle. En pratique, il est rare que tous les sous-ensembles de D soient à prendre en compte, et la complexité reste souvent plus raisonnable. Une complexité linéaire est obtenue si les masses sont modélisées selon la structure de Barnett [9], c'est-à-dire où les éléments focaux de chaque source sont uniquement les singletons et les compléments des singletons. Cette structure est adaptée à des problèmes de reconnaissance des formes dans lesquels chaque source est un détecteur qui permet de distinguer une classe contre toutes les autres. Mais elle n'est pas générale et ne s'applique pas aux sources qui nécessitent des éléments focaux qui soient des disjonctions quelconques.

4.4 Autres modes de combinaison

D'autres modes de combinaison, tels que des **modes disjonctifs ou de compromis**, sont possibles, en remplaçant l'intersection dans la formule (37) par une autre opération ensembliste. Par exemple, une fusion disjonctive est obtenue en prenant la réunion [114] :

$$\forall A \subset D, (m_1 \oplus_{\cup} m_2 \oplus_{\cup} \ldots \oplus_{\cup} m_\ell)(A) = \sum_{B_1 \cup \ldots \cup B_\ell = A} m_1(B_1) m_2(B_2) \ldots m_\ell(B_\ell) \quad (44)$$

Notons que cette combinaison ne peut pas faire apparaître de conflit. Elle élargit les éléments focaux et fournit donc une information moins précise que chacune des sources. Ce mode de fusion peut être intéressant quand on ne sait pas modéliser a priori les fiabilités des sources, leurs ambiguïtés et imprécisions. Par exemple, si une source est focalisée sur A et une autre sur B avec $A \cap B = \emptyset$, une manière de ne pas lever le conflit est de conclure que la vérité est dans $A \cup B$, ce que permet la fusion disjonctive.

Toutefois, dans la plupart des applications en fusion d'images, on cherche à obtenir une fonction de masse combinée plus focalisée que les fonctions de masse initiales. Ainsi, **on préfère la fusion conjonctive**, ce qui implique de prendre en compte les imprécisions, fiabilités, ambiguïtés de chaque source à l'étape de modélisation. Elle constitue alors l'étape la plus cruciale et qui requiert le plus d'attention.

4.5 Décision

Une fois calculées les fonctions de masse combinées, les fonctions de croyance et de plausibilité sont déduites par les équations (15) et (17). La dernière étape est celle de la décision, donc du choix d'un sous-ensemble de D maximisant un certain critère. Dans la suite, m, Bel et Pls désignent les fonctions de masse, de croyance et de plausibilité obtenues après combinaison.

Dans la théorie des fonctions de croyances, plusieurs règles de décision sont possibles, et la plupart sont appliquées au choix d'un singleton C_i.

- Le **maximum de plausibilité** :

$$x \in C_i \text{ si } Pls(C_i)(x) = \max\{Pls(C_k)(x), 1 \leqslant k \leqslant n\} \quad (45)$$

cette règle étant optimale au sens de critères d'inspiration probabiliste pour des fonctions de masse dérivées de probabilités [3].

- Le **maximum de crédibilité** :

$$x \in C_i \text{ si } Bel(C_i)(x) = \max\{Bel(C_k)(x), 1 \leqslant k \leqslant n\} \quad (46)$$

qui est équivalent au critère du maximum de plausibilité dans le cas où le résultat de la combinaison ne porte que sur les singletons.

- Le **maximum de crédibilité sans recouvrement des intervalles de confiance** (sans risque d'erreur) :

$$x \in C_i \text{ si } Bel(C_i)(x) \geqslant \max\{Pls(C_k)(x), 1 \leqslant k \leqslant n, k \neq i\} \quad (47)$$

cette dernière condition étant particulièrement stricte et pouvant ne conduire à aucune décision.

- Le **maximum de crédibilité avec rejet** [86] :

$$\begin{aligned}(x \in C_i) \text{ si } & Bel(C_i)(x) = \max\{Bel(C_k)(x), 1 \leqslant k \leqslant n\} \\ \text{et } & Bel(C_i)(x) \geqslant Bel(C_i^C)\end{aligned} \quad (48)$$

qui exprime que la décision doit être suffisamment non ambiguë puisque la condition sera vérifiée si la masse est très focalisée sur C_i.

- Le **maximum de probabilité pignistique**, celle-ci étant définie par [113] :

$$\forall C_j \in D, \; BetP(C_j) = \sum_{C_i \in A} \frac{m(A)}{|A|(1 - m(\emptyset))} \quad (49)$$

où $|A|$ désigne le cardinal de A, qui permet de repasser à un contexte probabiliste souvent souhaité pour la prise de décision (ou le pari) ou pour associer cette décision à d'autres critères probabilistes, par exemple dans le cadre des champs de Markov pour des critères de régularisation spatiale [116].

- Des **règles mixtes** ont également été proposées, dans lesquelles la plausibilité est utilisée pour certaines classes et la croyance pour d'autres. Cela permet de favoriser la détection des classes pour lesquelles on considère la plausibilité [90].

La décision peut également être prise **en faveur d'une disjonction**. Elle est alors imprécise, mais permet de prendre en compte des mélanges de classes ou des ambiguïtés subsistant après fusion. Ce type de décision est intéressant par exemple pour tenir compte de l'effet de volume partiel, et des voxels qui en sont affectés seront ainsi classés comme des voxels de mélange, plutôt que comme des voxels de classes pures, ce qui correspond bien à l'intuition [17]. La décision permet également d'indiquer les éléments pour lesquels la fusion ne suffit pas à lever les ambiguïtés et donc de suggérer l'acquisition de nouvelles informations, ainsi que l'exploite la fusion active [59][99].

- Enfin, des **règles de décision avec coût** ont été proposées [45]. Pour toute fonction f de D dans $\mathbb{R}$, les espérances inférieure et supérieure de f relativement à une fonction de croyance Bel, au sens de Dempster, sont définies par :

$$E_*(f) = \sum_{A \subset D} m(A) \min_{C_i \in A} f(C_i) \quad (50)$$

$$E^*(f) = \sum_{A \subset D} m(A) \max_{C_i \in A} f(C_i) \quad (51)$$

Les règles de décision avec coût s'obtiennent alors en prenant pour f une fonction qui exprime le coût de l'action α lorsque l'élément sur lequel porte la décision appartient à la classe C_i. Cette fonction de coût peut également être introduite dans une règle de décision avec coût probabiliste classique, en utilisant la probabilité pignistique. Ainsi la décision peut être optimiste si l'espérance inférieure est minimisée, pessimiste si l'espérance supérieure est minimisée, ou intermédiaire si la probabilité pignistique est utilisée.

5. Fusion floue et possibiliste

5.1 Modélisation

■ Parmi les techniques non probabilistes qui ont fait leur apparition depuis une dizaine d'années en traitement d'images, la **théorie des ensembles flous** fournit un très bon outil pour représenter explicitement des informations imprécises, sous la forme de **fonctions d'appartenance** [7][69][129]. La mesure $M_i^j(x)$ prend alors la forme :

$$M_i^j(x) = \mu_i^j(x), \tag{52}$$

où $\mu_i^j(x)$ désigne par exemple le degré d'appartenance de x à la classe C_i selon l'image I_j, ou la traduction d'une information symbolique exprimée par une variable linguistique (par exemple [44]).

On trouve dans la littérature essentiellement deux approches pour l'utilisation des ensembles flous en traitement d'image [14] : la première est de type plutôt symbolique et exprime sous forme de règles floues l'appartenance de certaines structures à une classe en fonction des mesures obtenues par traitement d'image ; la deuxième utilise les ensembles flous pour représenter directement les classes ou structures dans l'image, recouvrant spatialement les objets d'une mesure floue.

Considérons l'**exemple** de la classe « route » dans une image satellitaire. Dans la première approche, on décrira la route sous forme linguistique du type « une route est une structure plutôt allongée ». L'appartenance d'un objet à la classe route sera alors représentée par une fonction associant à sa longueur un degré dans [0, 1]. Un algorithme quelconque de détection de contours parallèles permettra alors d'affecter aux objets qu'il détecte un degré d'appartenance à la classe route suivant leur longueur. Dans la deuxième approche, la route sera directement représentée sur l'image par un ensemble flou, avec des degrés d'appartenance forts au centre de la chaussée et voisins de 0 dans les champs ou les forêts.

Ces fonctions ne souffrent pas des contraintes axiomatiques imposées aux probabilités et offrent donc une plus grande souplesse lors de la modélisation. Cette souplesse peut être considérée comme un inconvénient puisqu'elle laisse facilement l'utilisateur démuni pour définir ces fonctions. L'inconvénient des ensembles flous est qu'ils représentent essentiellement le caractère imprécis des informations, l'incertitude étant représentée de manière implicite et n'étant accessible que par déduction à partir des différentes fonctions d'appartenance.

■ La **théorie des possibilités** [53][130], dérivée des ensembles flous, permet de représenter à la fois **l'imprécision et l'incertitude**, par l'intermédiaire de **distributions de possibilités** π sur un ensemble S et de deux fonctions caractérisant les événements : la **possibilité** Π et la **nécessité** N, définies à partir de la distribution de possibilité pour un événement $A \subset S$ par :

$$\Pi(A) = \sup\{\pi(s), s \in A\} \tag{53}$$

$$N(A) = \inf\{(1-\pi(s)), s \notin A\} = 1 - \Pi(A^C) \tag{54}$$

où A^C désigne le complémentaire de A (l'événement contraire).

On introduit parfois une **contrainte de normalisation**, plus faible que dans les probabilités :

$$\sup\{\pi(x), x \in S\} = 1 \tag{55}$$

Cette contrainte correspond à une hypothèse de monde clos dans lequel au moins un élément de S est complètement possible.

Une distribution de possibilités s'interprète comme une fonction donnant le degré de possibilité pour qu'une variable prenne la valeur s, S étant le domaine des valeurs de la variable. La distribution π s'interprète alors comme la fonction d'appartenance au sous-ensemble flou de S des valeurs possibles pour cette variable. Dans le cadre de la fusion numérique d'images, une application possible de cette théorie consiste à prendre $S = D$ (l'ensemble des classes) et à définir la mesure M_i^j par :

$$M_i^j(x) = \pi_j^x(C_i) \tag{56}$$

c'est-à-dire comme le degré de possibilité pour que la classe à laquelle appartient x prenne la valeur C_i, selon l'image I_j. On définit ainsi une distribution de possibilités par image et par élément x. La possibilité et la nécessité pour une classe C_i s'écrivent alors :

$$\Pi_j(\{C_i\}) = \pi_j(C_i),\ N_j(\{C_i\}) = \inf\{(1-\pi_j(C_k)), C_k \neq C_i\} \tag{57}$$

Pour un sous-ensemble quelconque A de D, la possibilité et la nécessité sont calculées d'après les formules (53) et (54).

Cette modélisation suppose que les classes sont nettes, alors que le modèle flou défini par l'équation (52) suppose que les classes sont floues.

■ De manière plus générale, on trouve dans la littérature trois interprétations pour le flou, en termes de **plausibilités**, de **similarités**, de **préférences** [56]. On retrouve ces interprétations en fusion d'images. L'interprétation en termes de **plausibilités** se retrouve dans l'appartenance à une classe, dans la définition d'un objet spatial flou (objet de limites imprécises). L'interprétation en termes de **similarités** est celle de la définition d'une classe floue sur un espace de caractéristiques comme fonction de la distance à un prototype par exemple, de variables linguistiques représentant des informations ou des connaissances sur des objets spatiaux, ou encore de degrés de satisfaction d'une relation, d'une contrainte. Enfin, les **préférences** se retrouvent dans l'expression de critères de choix (par exemple pour des applications de planification en robotique), souvent liés à des contraintes ou des connaissances externes à l'image.

5.2 Définition des fonctions d'appartenance ou des distributions de possibilités

La construction des fonctions d'appartenance ou distributions de possibilités peut être effectuée de plusieurs manières.

Dans la plupart des applications, cette construction est faite soit en s'inspirant directement des méthodes d'apprentissage probabiliste, soit par des heuristiques, soit par des méthodes neuromimétiques permettant d'apprendre les paramètres de formes particulières de fonctions d'appartenance, soit enfin par la minimisation de critères de classification [11]. Décrivons maintenant les principales méthodes.

■ Une première méthode consiste à définir la fonction d'appartenance d'une classe floue à partir de la **fonction d'intensité** I (les niveaux de gris) de l'image :

$$\mu_i(x) = F_i[I(x)] \tag{58}$$

où F_i est une fonction à déterminer en fonction du problème.

Les fonctions les plus utilisées sont des fonctions de normalisation ou des fonctions en S [95] (ce qui revient à considérer que les parties claires de l'image ont une appartenance élevée à la classe), des fonctions Π (monomodales, elles associent la classe à une plage de niveaux de gris aux frontières imprécises), ou encore des fonctions multimodales.

Ces fonctions sont souvent déterminées de manière supervisée, mais elles peuvent également être apprises, par exemple à partir d'algorithmes de classification automatique tels que les C-moyennes floues [11] ou les C-moyennes possibilistes [72] (voir par exemple [12] pour une revue des algorithmes de classification floue utilisés en traitement d'images). L'inconvénient principal des

C-moyennes floues est que les fonctions d'appartenance ont une forme contre-intuitive : les valeurs d'appartenance à une classe ne sont pas décroissantes en fonction de la distance au centre de la classe. Ce problème est évité avec les C-moyennes possibilistes.

■ D'autres caractéristiques peuvent être utilisées. Par exemple, l'ensemble des contours d'une image peut être défini par un ensemble flou spatial dont la fonction d'appartenance est une fonction du gradient de l'image :

$$\mu_i(x) = F_i[\nabla I(x)], \tag{59}$$

où F est une fonction décroissante du gradient.

Si l'on dispose de détecteurs d'objets particuliers, la fonction d'appartenance à ces objets est alors définie comme une fonction de la réponse à ces détecteurs (le cas des contours fait partie de cette catégorie). Par exemple, un détecteur de routes peut fournir dans une image satellitaire une réponse d'amplitude croissante avec l'appartenance à la route.

Dans le cas de variables linguistiques, les formes des fonctions d'appartenance et leurs paramètres sont souvent définies par l'utilisateur.

L'imprécision spatiale sur la délimitation des classes (si les fonctions d'appartenance sont définies sur l'espace de l'image) peut être introduite à partir d'une détection préliminaire binaire des classes. On construit alors une fonction d'appartenance qui vaut 1 à l'intérieur de la région binaire à une certaine distance des bords, 0 à l'extérieur de cette région également à une certaine distance des bords, et décroissante entre ces deux limites. Par exemple, on peut modéliser une zone d'imprécision au bord de la classe comme la zone comprise entre l'érodé et le dilaté de cet objet, la taille de ces opérations dépendant de l'extension spatiale de l'imprécision que l'on veut représenter. Si R est la région binaire de départ, $E^n(R)$ son érodé de taille n et $D^m(R)$ son dilaté de taille m, la fonction d'appartenance à la classe floue peut être définie par :

$$\mu(x) = 1 \text{ si } x \in E^n(R), \tag{60}$$

$$\mu(x) = 0 \text{ si } x \in D^m(R)^C \tag{61}$$

$$\mu(x) = F[d(x, E^n(R))] \text{ sinon} \tag{62}$$

où F est une fonction décroissante de la distance de x à $E^n(R)$.

■ La construction de distributions de possibilités peut également être effectuée à partir d'un **apprentissage probabiliste**, puis par une transformation de probabilité en possibilité. Plusieurs méthodes ont été proposées pour cela. L'avantage essentiel en traitement d'images est que l'on dispose souvent d'informations statistiques, en particulier l'histogramme de l'image, qui se prêtent bien à l'emploi de méthodes d'apprentissage statistiques. On obtient alors des distributions de probabilités p_k. Leur transformation en distributions de possibilités π_k (les deux distributions sont supposées discrètes, et $1 \leqslant k \leqslant K$) est effectuée en fonction de différents critères [48][51][71], tels que la préservation de l'ordre, des contraintes de normalisation, la conservation de l'incertain mesuré par l'entropie [71], la cohérence $p-\pi$, exprimée par [43] :

$$\forall k, \pi_k \leqslant p_k$$

qui n'est pas très satisfaisante (une classe peu probable peut être possible), ou [130] :

$$\sum_{k=1}^{K} p_k \pi_k = c$$

où c est une constante dans [0, 1],
ou encore une relation plus générale sur tous les sous-ensembles A [50] :

$$N(A) \leqslant P(A) \leqslant \Pi(A)$$

Une comparaison de ces approches peut être trouvée dans [71].

■ D'autres méthodes cherchent à estimer directement les fonctions d'appartenance à partir de l'**histogramme**, afin d'optimiser des critères d'entropie [36], ou de minimum de spécificité et de cohérence [37].

Dans tous les cas, ces méthodes cherchent une ressemblance entre l'histogramme et les fonctions d'appartenance ou les distributions de possibilités, et ne prennent pas en compte les interprétations spécifiques au flou qui invalident certaines de ces ressemblances. Par exemple, les queues des histogrammes correspondent aux classes peu représentées, donc avec des valeurs qui peuvent être très faibles, même si les points concernés appartiennent bien aux classes correspondantes. La méthode proposée dans [25] permet d'éviter ce problème grâce à un critère combinant la ressemblance des fonctions d'appartenance et de l'histogramme là où elle a un sens, et une forme a priori des fonctions correspondant à l'interprétation recherchée. Les paramètres des fonctions d'appartenance sont alors estimés pour optimiser ce critère, par une méthode de recuit simulé.

5.3 Combinaison

Un des intérêts de la théorie des ensembles flous et des possibilités, outre qu'elle impose peu de contraintes au niveau de la modélisation, est qu'elle offre une grande variété d'opérateurs de combinaison. Nous en présentons les principaux, puis nous donnons quelques indications permettant de choisir un opérateur de fusion en fonction de ses propriétés et de son comportement.

Une caractéristique importante, commune à toutes les théories, de ces opérateurs de combinaison est qu'ils fournissent un résultat de même nature que les fonctions de départ (propriété de fermeture) et qui a donc la même interprétation en termes d'imprécision et d'incertitude. Ainsi, ils permettent de ne prendre aucune décision binaire partielle avant la combinaison, ce qui pourrait conduire à des contradictions difficiles à lever. La décision n'est prise qu'en dernier lieu, sur le résultat de la combinaison.

5.3.1 Opérateurs

Dans la théorie des ensembles flous et des possibilités, de multiples modes de combinaison sont possibles [52][126]. Parmi les principaux opérateurs on trouve en particulier les t-normes, les t-conormes [89][107], les moyennes [64][125], les sommes symétriques, et des opérateurs prenant en compte des mesures de conflit ou encore de fiabilité des sources [47][55]. Rappelons les définitions de ces opérateurs. Dans ce qui suit, les lettres x, y, etc., désignent les valeurs à combiner, valeurs dans [0, 1] représentant μ_i^j ou $\pi_j(C_i)$.

■ **Normes et conormes triangulaires**

● Une **norme triangulaire** ou t-norme est une fonction $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ telle que :

1. t est **commutative** : $\forall (x, y) \in [0, 1]^2, t(x, y) = t(y, x)$;
2. t est **associative** : $\forall (x, y, z) \in [0, 1]^3, t[t(x, y), z] = t[x, t(y, z)]$;
3. 1 est **élément neutre** : $\forall x \in [0, 1], t(x, 1) = t(1, x) = x$;
4. t est **croissante** par rapport aux deux variables :

$$\forall (x, x', y, y') \in [0, 1]^4, (x \leqslant x' \text{ et } y \leqslant y') \Rightarrow t(x, y) \leqslant t(x', y') \;.$$

De plus, on a : $t(0, 1) = t(0, 0) = t(1, 0) = 0$; $t(1, 1) = 1$ et 0 est élément absorbant $(\forall x \in [0, 1], t(x, 0) = 0)$.

Exemple : les opérateurs min(x, y), xy, max (0, $x + y - 1$) sont des t-normes, de loin les plus usitées.

● Une **complémentation floue** est une fonction c de [0, 1] dans [0, 1] telle que :

1. $c(0) = 1$;

2. $c(1) = 0$;

3. c est **involutive** : $\forall x \in [0, 1], c(c(x)) = x$;

4. c est **strictement décroissante**.

Comme il est difficile de construire directement des fonctions involutives, il est intéressant de les caractériser par une forme générale de construction plus simple. Ainsi, les complémentations continues ont la forme générale suivante : $c(x) = \varphi^{-1}[1 - \varphi(x)]$ avec $\varphi : [0, 1] \to [0, 1]$, telle que $\varphi(0) = 0, \varphi(1) = 1$, φ est strictement croissante. Il existe de multiples fonctions φ vérifiant ces propriétés et il est facile d'en exhiber.

Exemple : l'exemple le plus simple est $\varphi(x) = x^n$ qui permet de construire la complémentation $c(x) = (1 - x^n)^{1/n}$. La forme la plus utilisée est $c(x) = 1 - x$, obtenue pour $n = 1$

● Une **t-conorme** est une fonction $T : [0, 1] \times [0, 1] \to [0, 1]$ telle que :

1. T est **commutative** : $\forall (x, y) \in [0, 1]^2, T(x, y) = T(y, x)$;

2. T est **associative** : $\forall (x, y, z) \in [0, 1]^3, T[T(x, y), z] = T[x, T(y, z)]$;

3. 0 est **élément neutre** : $\forall x \in [0, 1], T(x, 0) = T(0, x) = x$;

4. T est **croissante** par rapport aux deux variables.

De plus, on a : $T(0, 1) = T(1, 1) = T(1, 0) = 1$; $T(0, 0) = 0$ et 1 est élément absorbant $(\forall x \in [0, 1], T(x, 1) = 1)$.

Exemples : les plus classiques sont :

$\max(x, y)$, $x + y - xy$, $\min(1, x + y)$

● Une propriété importante est la **dualité entre les t-normes et les t-conormes par rapport à la complémentation** : pour toute t-norme t (respectivement t-conorme T), on peut construire une t-conorme T (respectivement t-norme t) par l'équation de la dualité

$$\forall (x, y) \in [0, 1]^2, T[c(x), c(y)] = c[t(x, y)]$$

Ainsi le min et le max sont des opérateurs duaux par exemple.

Il existe des formes génériques pour des t-normes et des t-conormes ayant des propriétés particulières, que nous ne détaillons pas ici.

Opérateurs de moyenne

Un opérateur de moyenne est une fonction $m : [0, 1] \times [0, 1] \to [0, 1]$ telle que :

1. le résultat de la combinaison est toujours compris entre le min et le max :

$$\forall (x, y) \in [0, 1]^2, \min(x, y) \leqslant m(x, y) \leqslant \max(x, y),$$

mais

$$m \neq \min \text{ et } m \neq \max ;$$

2. m est **commutative** ;

3. m est **croissante** par rapport aux deux variables.

La première propriété entraîne que m est toujours une opération **idempotente** :

$$\forall x \in [0, 1], m(x, x) = x$$

Ces opérateurs ne sont en général pas associatifs. Les seules moyennes qui le soient sont les médianes, où $m(x, y)$ est la valeur médiane de x, y et d'un paramètre α.

Une propriété plus faible que l'associativité est la **bissymétrie** :

$$\forall (x, y, z, t) \in [0, 1]^4, m[m(x, y), m(z, t)] = m[m(x, z), m(y, t)]$$

Les moyennes qui vérifient cette propriété, qui sont continues et strictement croissantes ont la forme générale suivante :

$$m(x, y) = k^{-1}\left[\frac{k(x) + k(y)}{2}\right]$$

où k est une fonction continue strictement croissante de [0, 1] dans [0, 1]. La fonction k peut s'interpréter comme un changement d'échelle ou de dynamique des valeurs à combiner. Celles-ci, une fois transformées par k, sont alors combinées par une simple moyenne arithmétique, puis on ramène le résultat à l'échelle initiale.

Les moyennes les plus classiques sont obtenues pour $k(x) = x^{\alpha}$. La moyenne **arithmétique** $\frac{x + y}{2}$ est obtenue pour $\alpha = 1$, la moyenne **quadratique** $\sqrt{\frac{x^2 + y^2}{2}}$ pour $\alpha = 2$, la moyenne **harmonique** $\frac{2xy}{x + y}$ pour $\alpha = -1$, la moyenne **géométrique** $\sqrt{xy}$ pour $\alpha = 0$. À la limite, quand α tend vers $-\infty$ ou $+\infty$, m tend vers le min ou le max.

On trouve également dans la classe des opérateurs de moyenne les moyennes **pondéréees**, les **OWA** (*ordered weighted average*) [125] et les **intégrales floues** [64].

Sommes symétriques

Les sommes symétriques sont définies par une propriété d'**autodualité**, qui correspond à l'invariance du résultat de l'opération par inversion de l'échelle des valeurs à combiner. Plus précisément, une somme symétrique est une fonction $\sigma : [0, 1] \times [0, 1] \to [0, 1]$ telle que :

1. $\sigma(0, 0) = 0$;

2. σ est **commutative** ;

3. σ est **croissante** par rapport à deux variables ;

4. σ est **autoduale** : $\forall (x, y) \in [0, 1]^2, \sigma(x, y) = 1 - \sigma(1 - x, 1 - y)$

On a de plus $\sigma(1, 1) = 1$.

Notons que l'autodualité s'oppose à la dualité mentionnée entre les t-normes et les t-conormes. Pour ces opérateurs, inverser l'échelle entraîne le changement de type d'opérateur.

La forme générale des sommes symétriques est donnée par :

$$\sigma(x, y) = \frac{g(x, y)}{g(x, y) + g(1 - x, 1 - y)}$$

où g est une fonction de $[0, 1] \times [0, 1]$ dans [0, 1], croissante, positive, continue telle que $g(0, 0) = 0$. Typiquement on peut prendre pour g une t-norme ou une t-conorme continue.

Voici quelques **exemples** de sommes symétriques :

$$\sigma_0(x, y) = \frac{xy}{1 - x - y + 2xy}$$

obtenue pour $g(x, y) = xy$ (associative),

$$\sigma_+(x, y) = \frac{x + y - xy}{1 + x + y - 2xy}$$

obtenue pour $g(x, y) = x + y - xy$ (non associative),

$$\sigma_{\min}(x, y) = \frac{\min(x, y)}{1 - |x - y|}$$

obtenue pour $g(x, y) = \min(x, y)$ (qui est une moyenne),

$$\sigma_{\max}(x, y) = \frac{\max(x, y)}{1 + |x - y|}$$

obtenue pour $g(x, y) = \max(x, y)$ (qui est une moyenne),

Opérateurs dépendant d'autres informations

Les opérateurs décrits précédemment se caractérisent par le fait que le résultat d'une combinaison ne dépend que des valeurs à combiner. D'autres opérateurs dépendent d'informations supplémentaires, telles que de facteurs de fiabilité ou du conflit. Tous les opérateurs décrits dans ce paragraphe ont été proposés dans [55].

Considérons par exemple le cas de la combinaison conjonctive de deux distributions de possibilités π_1 et π_2 définies sur D. Ce type de combinaison est bien adapté au cas où les distributions ont un recouvrement au moins partiel, c'est-à-dire que certaines classes sont données comme possibles par les deux sources. Si ce n'est pas le cas, les sources sont en conflit, et une mesure du conflit peut être :

$$conf(\pi_1, \pi_2) = 1 - \max_{c \in D} \min(\pi_1(c), \pi_2(c)) \quad \textbf{(63)}$$

qui représente 1 moins la hauteur de l'intersection entre les deux distributions (calculée par un min dans cette équation). La combinaison peut être normalisée par cette hauteur, mais cela masque le conflit : une possibilité de 1 est alors affectée aux classes données comme les plus possibles par les deux sources, même si cette possibilité est faible (on retrouve ici le même type de problème que celui mentionné dans le paragraphe 4.3 pour la combinaison conjonctive des fonctions de croyance). L'interprétation de cette quantité en termes de conflit correspond bien à l'intuition pour des distributions de possibilités en triangle ou en trapèze (monomodales de manière générale), mais n'est pas très bien adaptée à des formes quelconques où un seul point peut engendrer une forte valeur de conflit même si les deux distributions ne diffèrent qu'en ce point.

Dans le cas extrême où les distributions sont complètement conflictuelles, la combinaison conjonctive donne une distribution identiquement nulle. Une combinaison disjonctive est alors préférable, et permet de garder toutes les solutions données comme possibles par au moins une des deux sources. L'hypothèse sous-jacente est qu'au moins une des sources est fiable.

Afin de choisir automatiquement le comportement de la fusion en fonction du conflit, des opérateurs dépendant de celui-ci ont été proposés. En voici quelques exemples :

$$\max\left[\frac{\min(\pi_1, \pi_2)}{1 - conf(\pi_1, \pi_2)}, conf(\pi_1, \pi_2)\right] \quad \textbf{(64)}$$

$$\min\left[1, \frac{\min(\pi_1, \pi_2)}{1 - conf(\pi_1, \pi_2)} + conf(\pi_1, \pi_2)\right] \quad \textbf{(65)}$$

$$\max\left[\frac{\min(\pi_1, \pi_2)}{1 - conf(\pi_1, \pi_2)}, \min[\max(\pi_1, \pi_2)], conf(\pi_1, \pi_2)\right] \quad \textbf{(66)}$$

Les deux premières formes combinent la conjonction normalisée à une distribution constante ayant pour valeur le conflit, alors que la dernière forme permet de passer d'une combinaison strictement conjonctive lorsque le conflit est nul à une combinaison strictement disjonctive lorsque le conflit vaut 1. Cependant cet opérateur n'est pas associatif. Notons que le min pourrait être remplacé par une autre t-norme.

Lorsque les sources sont inégalement fiables et que l'on a une information sur cette fiabilité, le niveau du conflit entre deux sources indique dans quelle mesure on peut prendre en compte l'information de la source la moins fiable. Si par exemple π_1 est plus fiable que π_2, on peut considérer que si elles sont concordantes, π_2 peut apporter une information et rendre la fusion plus précise par conjonction. Si au contraire les deux sources sont en conflit, il est préférable de ne pas prendre en compte π_2. L'opérateur suivant modélise ce comportement :

$$\min[\pi_1, \max[\pi_2, conf(\pi_1, \pi_2)]] \quad \textbf{(67)}$$

Cela ne suppose que de connaître un ordre entre les fiabilités des sources.

Si on a de plus accès à des valeurs numériques de fiabilité (ce qui est beaucoup plus contraignant que l'hypothèse précédente), on peut alors transformer les distributions de possibilités en distributions ayant des fiabilités équivalentes. Soit w_j le coefficient de fiabilité de π_j. Si la source est complètement fiable, ce coefficient vaut 1, et il vaut 0 si la source n'est pas fiable du tout. La transformation de π_j se fait selon la formule :

$$\max(\pi_j, 1 - w_j) \quad \textbf{(68)}$$

ce qui revient à faire une disjonction entre π_j et une distribution constante de valeur $1 - w_j$. Ainsi, si la source est complètement fiable, la distribution correspondante n'est pas modifiée, alors que, si elle n'est pas du tout fiable, la distribution devient constante et égale à 1, ce qui représente l'ignorance (tout élément de D est complètement possible). Une fois les distributions transformées, elles peuvent alors être combinées de manière conjonctive.

D'autres opérateurs de ce type peuvent être trouvés dans [56], mais nous ne les détaillons pas plus ici.

Ces opérateurs peuvent être également utilisés de manière conditionnelle aux classes, pour prendre en compte les spécificités des sources pour chaque classe. Par exemple, deux sources peuvent être en conflit sur une classe mais pas sur les autres, une source peut être fiable pour certaines classes et pas pour d'autres, etc. Bien que ces idées ne soient pas encore beaucoup exploitées en fusion floue d'images, le cadre théorique le permet.

5.3.2 Choix des opérateurs

Le choix d'un opérateur peut se faire selon plusieurs critères pour la fusion d'images [16].

Un **premier critère** est le **comportement de l'opérateur**. Des comportements sévères, indulgents ou prudents se traduisent sous forme mathématique de conjonction, disjonction ou compromis. Soit x et y deux réels (dans [0, 1]) représentant les degrés de confiance à combiner. La combinaison de x et y par un opérateur F est dite :

- **conjonctive** si $F(x, y) \leqslant \min(x, y)$ (correspondant à un comportement **sévère**) ;
- **disjonctive** si $F(x, y) \geqslant \max(x, y)$ (comportement **indulgent**) ;
- **de compromis** si $x \leqslant F(x, y) \leqslant y$ si $x \leqslant y$, et $y \leqslant F(x, y) \leqslant x$ sinon (comportement **prudent**).

Cette distinction ne suffit pas à classer les opérateurs dont le comportement n'est pas toujours le même. Ainsi, la classification définie dans [16] ne décrit pas les opérateurs seulement comme conjonctifs ou disjonctifs, mais aussi en fonction de leur comportement selon les valeurs des informations à combiner. Ainsi, les trois classes proposées correspondent :

1) aux **opérateurs autonomes à comportement constant** (ACC) : le résultat ne dépend que des valeurs à combiner (le calcul ne fait intervenir aucune autre information) et le comportement est le même quelles que soient ces valeurs ;

2) aux **opérateurs autonomes à comportement variable** (ACV) : le comportement dépend des valeurs numériques des informations à fusionner ;

3) aux **opérateurs dépendant du contexte** (DC), par exemple d'une connaissance plus globale telle que la fiabilité des capteurs, ou encore le conflit entre les sources.

Les opérateurs de fusion floue se répartissent dans les trois classes. En effet, les t-normes, qui généralisent l'intersection ensembliste aux ensembles flous, sont des opérateurs ACC conjonctifs, puisque pour toute t-norme t, on a:

$$\forall (x, y) \in [0, 1]^2, t(x, y) \leqslant \min(x, y)$$

À l'opposé, les t-conormes, généralisant la réunion, sont des opérateurs ACC disjonctifs, puisque pour toute t-conorme T, on a :

$$\forall(x, y) \in [0, 1]^2, T(x, y) \geqslant \max(x, y)$$

Les opérateurs de moyenne sont également ACC et ont un comportement de compromis, puisqu'ils vérifient :

$$\forall(x, y) \in [0, 1]^2, \min(x, y) \leqslant m(x, y) \leqslant \max(x, y)$$

Notons que la fusion bayésienne, où l'opérateur impliqué est un produit, et la fusion des fonctions de croyances par la somme orthogonale de Dempster sont également conjonctives.

Dans la classe des opérateurs ACV, on trouve par exemple certaines sommes symétriques. De manière générale, toute somme symétrique associative σ (sauf les médianes) a le comportement suivant [53] :

— conjonctif si $\max(x, y) < 1/2$: $\sigma(x, y) \leqslant \min(x, y)$;

— disjonctif si $\min(x, y) > 1/2$: $\sigma(x, y) \geqslant \max(x, y)$;

— de compromis si $x \leqslant 1/2 \leqslant y$: $x \leqslant \sigma(x, y) \leqslant y$ (et l'inégalité contraire si $y \leqslant 1/2 \leqslant x$).

Les sommes symétriques non associatives ont également un comportement qui varie, mais selon des règles moins simples [16].

On trouve également dans la classe ACV les opérateurs proposés dans le système MYCIN pour combiner des facteurs de certitude [109].

Des exemples d'opérateurs DC se trouvent dans la théorie des possibilités. Nous avons présenté plus haut des opérateurs dépendant d'une mesure globale de conflit entre les deux sources d'information [55], applicables aux cas où l'une des deux informations est fiable sans qu'on sache laquelle, de telle sorte que :

— ils sont conjonctifs si les sources sont consonantes (de faible conflit) : dans ce cas, les deux sources sont nécessairement fiables, et donc l'opérateur peut être sévère ;

— ils sont disjonctifs si les sources sont dissonantes (de fort conflit) : une disjonction favorise alors l'ensemble des possibilités données par les deux sources ;

— ils se comportent comme un compromis dans les cas de conflit partiel : ces cas posant le plus de problèmes, les opérateurs sont alors « prudents ».

La difficulté est alors de trouver une bonne mesure de conflit. Celle proposée comme le maximum de l'intersection entre deux distributions de possibilités [55] n'est pas toujours bien adaptée aux problèmes de traitement d'images, en particulier pour la classification d'images multisources. Les distances floues (voir par exemple [21]) peuvent apporter des solutions à ce problème.

L'avantage des opérateurs DC pour le traitement d'images est indéniable. En effet, ils permettent de prendre en compte une grande variété de situations, dont plusieurs se produisent fréquemment simultanément en traitement d'images. En voici quelques exemples :

— les sources peuvent être conflictuelles lorsqu'elles donnent des informations sur un type d'événement (une classe par exemple) et consonantes pour une autre classe ;

— les sources peuvent avoir des fiabilités globales différentes ;

— une source peut être fiable pour une classe et peu fiable pour une autre, etc.

Malheureusement ces opérateurs sont encore, selon nous, trop peu développés en traitement d'images, et mériteraient des recherches spécifiques.

Cette classification, qui regroupe tous les opérateurs classiquement utilisés, constitue un premier critère de choix d'un opérateur pour une application spécifique.

■ Un **deuxième critère** est donné par les **propriétés des opérateurs et leur interprétation en termes de fusion de données** incertaines, imprécises, incomplètes ou encore ambiguës.

• Les propriétés de **commutativité** et d'**associativité** expriment que le résultat de la combinaison est indépendant de l'ordre dans lequel les informations sont combinées. Si la commutativité est satisfaite par tous les opérateurs couramment utilisés, l'associativité ne l'est pas systématiquement (les moyennes et les sommes symétriques ne sont en général pas associatives). Ces deux propriétés sont souvent posées comme les propriétés minimales que les opérateurs de fusion doivent satisfaire. Pourtant, le mode de raisonnement humain ne les respecte pas toujours. Par exemple, un photo-interprète commence souvent par construire une interprétation primaire d'une scène à partir d'une seule image, puis améliore cette inteprétation à l'aide des autres images, selon un processus qui n'est clairement pas commutatif.

• L'existence d'un **élément neutre** signifie qu'une source donnant cette valeur n'aura aucune influence sur le résultat de la combinaison, et représente une sorte d'indifférence de la source par rapport à l'information recherchée, ou encore une ignorance totale de celle-ci. Un tel élément existe pour les t-normes et les t-conormes.

• Autre élément particulier, un **élément absorbant** signifie qu'une source donnant cette valeur est complètement déterminante sur le résultat de la fusion. De tels éléments existent également pour les t-normes et les t-conormes.

• La propriété de **croissance** est généralement imposée aux opérateurs et correspond bien à l'intuition.

• Des **conditions aux limites**, définissant le comportement des opérateurs lorsque les informations à combiner prennent des valeurs extrêmes, permettent de garantir la compatibilité avec le cas binaire, où toutes les propositions sont soit justes soit fausses (cela correspond à la contrainte de compatibilité avec la logique déductive imposée par Cox pour définir une logique inductive [40]).

• La propriété de **continuité** satisfaite par la plupart des opérateurs garantit la robustesse de la fusion. Cependant, cette propriété n'est pas toujours nécessaire, puisque les phénomènes naturels (en particulier les phénomènes temporels) ne sont pas toujours continus.

• L'**idempotence** signifie que la donnée d'une information déjà disponible ne va pas changer le résultat de la fusion. Cette propriété n'est pas systématiquement imposée. Elle est vérifiée par les moyennes, par la t-norme min et la t-conorme max (et ce sont les seules). On peut vouloir au contraire que la combinaison de deux valeurs identiques renforce ou affaiblisse le résultat global. Prenons l'exemple de témoignages simultanés identiques. Si les témoins sont de connivence, il n'est pas surprenant qu'ils disent la même chose et on combinera donc les degrés de confiance associés de manière idempotente. Si au contraire ils sont complètement indépendants, on renforcera la crédibilité de ce qu'ils disent si on leur fait confiance, ou on l'affaiblira si on ne leur fait pas confiance. Notons que des règles de combinaison modélisant ces comportements sont connues depuis Bernoulli. De manière générale, on peut considérer que si les sources sont dépendantes (au sens cognitif), l'idempotence peut être imposée, alors que si elles sont indépendantes, des effets de renforcement peuvent être souhaitables.

• Dans le même ordre d'idées, la propriété de **nilpotence** sera imposée par exemple pour combiner des témoignages successifs, afin de modéliser la dégradation de l'information dans une chaîne de témoins qui ne sont pas complètement fiables. Par exemple pour certaines t-conormes, la satisfaction de cette propriété permet d'aboutir à un résultat égal à 1 en combinant un certain nombre de mesures non toutes nulles. Ce type de comportement peut être utile quand les informations résultent d'une longue chaîne de traitements.

• Les propriétés du **tiers exclu** et de **non-contradiction**, satisfaites pour certains opérateurs seulement, ont une interprétation reconnue en termes de raisonnement, dans le domaine de l'intelligence artificielle et du raisonnement approché. En traitement d'images, on trouve des exemples où le tiers exclu n'est pas souhaitable dès que l'on a besoin d'introduire de l'ignorance sur un événement et son complémentaire, et donc de relâcher la contrainte d'exhaustivité faite par exemple en probabilités.

La généralisation de tout ce qui précède à la combinaison de plus de deux informations ne pose pas de difficulté particulière (on retrouve en particulier les mêmes types de comportement, avec des règles un peu plus compliquées pour les opérateurs ACV), sauf pour les opérateurs non associatifs. La question principale pour de tels opérateurs est de savoir dans quel ordre combiner les informations. Plusieurs situations peuvent se produire :

— dans certaines applications, chaque information doit être combinée aux autres dès qu'elle est disponible (par exemple pour pouvoir prendre des décisions partielles à partir des données présentes à chaque instant) : l'ordre est alors fixé par l'ordre d'arrivée des informations ;

— l'ordre peut être imposé par des priorités sur les informations à prendre en compte, et des opérateurs ont été conçus pour répondre à de tels besoins (par exemple pour combiner des requêtes dans des bases de données) ;

— dans les autres situations, il faudra trouver des critères pour trouver un ordre approprié à l'application, en particulier lorsque les informations sont conflictuelles, car les résultats peuvent être très différents selon que les informations consonantes ou conflictuelles sont combinées en premier.

■ Enfin, l'**étude du comportement des opérateurs en termes de qualité de la décision à laquelle ils conduisent, et de réaction face aux situations conflictuelles** conduit à un **dernier critère de choix**. Un point important toutefois concerne le **pouvoir discriminant** des opérateurs. Les opérateurs fortement conjonctifs ou fortement disjonctifs (la t-norme et la t-conorme de Lukasiewicz par exemple) saturent tès vite à 0 ou à 1 et donc sont souvent peu discriminants.

Par **exemple**, avec la t-conorme $F(a, b) = \min(a+b, 1)$, on a : $F(0,5\,;0,5) = 1$, $F(0,1\,;0,9) = 1$ ou encore $F(0,8\,;0,8) = 1$ alors que ces trois situations ont des interprétations bien différentes.

La **capacité des opérateurs à combiner des informations quantitatives** (numériques) ou **qualitatives** (pour lesquels seul un ordre est connu) peut être également un **critère de choix**. Par exemple, le min, le max et tout filtre de rang sont intéressants à ce titre puisqu'ils peuvent combiner les deux types d'informations. En effet, le calcul de min (x, y) par exemple ne nécessite que de connaître un ordre entre x et y, mais ne nécessite pas de connaître leur valeur numérique. Les opérations ordinales sont de plus imposées si on veut qu'elles soient invariantes par une transformation croissante des degrés d'appartenance [56].

5.4 Décision

La règle principalement utilisée en fusion floue est le maximum des degrés d'appartenance :

$$x \in C_i \text{ si } \mu_i(x) = \max\{\mu_k(x), 1 \leqslant k \leqslant n\} \qquad \textbf{(69)}$$

où μ_k désigne la fonction d'appartenance à la classe k résultant de la combinaison.

La qualité de la décision est mesurée essentiellement selon deux critères :

— le premier porte sur la « **netteté** » de la décision : le degré d'appartenance maximal (ou plus généralement celui correspondant à la décision) est comparé à un seuil, choisi selon les applications (et éventuellement selon l'opérateur de combinaison choisi) ;

— le deuxième porte sur le caractère « **discriminant** » de la décision, évalué par comparaison des deux valeurs les plus fortes.

Dans le cas où ces critères ne sont pas vérifiés pour un élément x, celui-ci est placé dans une classe de rejet, ou reclassifié en fonction d'autres critères, spatiaux par exemple (voir § 6).

6. Introduction d'informations spatiales dans la fusion

L'information spatiale est fondamentale en traitement d'images. Son introduction dans les méthodes de fusion est cruciale, et nécessite souvent des développements spécifiques pour adapter les méthodes issues d'autres domaines. Un des objectifs les plus fréquents est de garantir que la décision soit spatialement cohérente. Par exemple, en classification multisource, on va chercher à éviter les points isolés dans une classe différente.

6.1 Au niveau de la modélisation

L'introduction de l'information spatiale au niveau de la modélisation est plus ou moins implicite suivant le niveau de représentation auquel on se place. Si l'on raisonne au niveau du pixel, l'information contenue dans un pixel ne contient pas d'information spatiale, et celle-ci doit alors être introduite explicitement. Le contexte spatial considéré est le plus souvent le voisinage local de chaque point. Une manière simple de la prendre en compte est de définir $M_i^j(x)$ en fonction des caractéristiques de x et également de ses voisins. Si l'on note $\mathcal{V}(x)$ le voisinage de x (contenant x), on définira $M_i^j(x)$ comme une fonction du type :

$$M_i^j(x) = F_i[f_j(y), y \in \mathcal{V}(x)] \qquad \textbf{(70)}$$

où $f_j(x)$ désigne les caractéristiques de y dans la source j. Ce type d'approche peut être vu comme un problème de filtrage spatial.

Si l'on raisonne au niveau de primitives (segments, contours, régions) ou au niveau des objets ou structures de la scène, l'information spatiale locale est implicitement prise en compte dans la représentation. Si la détection de ces éléments n'est pas précise ou si leur localisation ne l'est pas (par exemple à cause de l'imperfection du recalage), il est souvent souhaitable d'introduire cette imprécision spatiale explicitement dans la représentation, avant la fusion. La dilatation floue est une opération bien adaptée à cela [28][30][32]. Cela permet de réduire le conflit au moment de la fusion, et donc de choisir simplement et sans risque un mode de combinaison conjonctif.

De manière moins locale, les relations spatiales entre primitives constituent une information importante sur la structure de la scène [19][20][21][23][29] et elles peuvent avantageusement être prises en compte dans la fusion, comme source d'informations supplémentaire [22][24][61]. Dans ce cas, le contexte spatial $\mathcal{V}(x)$ d'un élément x est un ensemble de primitives ou d'objets dont on connaît les relations spatiales par rapport à x.

6.2 Au niveau de la décision

L'introduction de l'information spatiale au niveau de la décision est la plus facile. La méthode la plus courante consiste à établir dans un premier temps une règle de rejet (en fonction de la netteté et du caractère discriminant de la décision) puis à reclasser les éléments rejetés en fonction de leur contexte spatial. Par exemple, la reclassification peut être faite suivant la règle suivante :

$$x \in C_i \text{ si } |\{y \in \mathcal{V}(x), y \in C_i\}| \geqslant \frac{|\mathcal{V}|}{2} \qquad \textbf{(71)}$$

qui exprime qu'au moins la moitié des éléments du voisinage doit être dans C_i pour classer x dans C_i. Cette règle ne permet pas toujours d'affecter x à une classe. Une autre règle moins sévère considère juste la classe la plus représentée dans le voisinage :

$$x \in C_i \text{ si } \left|\{ y \in \mathcal{V}(x), y \in C_i\}\right| = \max_k \left|\{ y \in \mathcal{V}(x), y \in C_k\}\right| \qquad \textbf{(72)}$$

Ces règles s'appliquent quel que soit le niveau de représentation des éléments considérés.

Un exemple en classification floue peut être trouvé dans [31], mais la méthode est générale et peut être appliquée de manière similaire aux autres théories.

6.3 Au niveau de la combinaison

L'introduction de l'information spatiale au niveau de la combinaison est plus rare et plus délicate.

En fusion probabiliste, les champs de Markov offrent un cadre naturel pour cela. Dans l'expression de la règle de Bayes, c'est dans la probabilité a priori qu'intervient l'hypothèse markovienne. Cette probabilité est combinée aux probabilités conditionnelles aux classes par un produit. Cette remarque nous conduit à considérer que l'information spatiale constitue dans ce modèle une source de données au même titre que les autres.

C'est l'approche la plus courante, et elle a été appliquée à plusieurs niveaux de représentation. Au niveau local, du pixel, de nombreux exemples peuvent être trouvés dans la littérature (par exemple [5][46]). À un niveau plus structurel, les champs de Markov sont définis sur des graphes plus généraux que les graphes de pixels (les nœuds sont des primitives ou même des objets), et on trouve des exemples pour la détection de routes dans des images SAR [117], pour la segmentation d'images IRM du cerveau [62], pour la reconnaissance de structures du cortex cérébral [84][85], pour l'interprétation d'images aériennes [93], etc.

Dans les autres théories, il serait également possible de développer des approches similaires, toujours en considérant l'information spatiale comme une source de données supplémentaire.

C'est par exemple le cas des relations spatiales mentionnées plus haut considérées comme source supplémentaire d'informations : la reconnaissance d'un objet peut résulter de la fusion d'informations sur cet objet, et d'informations sur les relations qu'il doit avoir par rapport à d'autres objets. Le cadre des ensembles flous permet à la fois la représentation et la fusion de telles informations [24].

Autre exemple, dans [68], une fonction de masse est définie pour représenter le contexte spatial, et combinée à des fonctions de masse représentant les informations extraites des images par la règle de Dempster. Toutefois, encore peu de travaux existent dans ce domaine, qui mérite certainement d'être développé.

7. Conclusion

En fusion numérique pour le traitement d'images, les efforts des dernières années ont permis d'aboutir à une meilleure compréhension des différentes théories issues d'autres domaines telles que les fonctions de croyance et le flou. On sait ainsi maintenant quels sont les bons cadres d'application de ces théories, leurs atouts et limites en traitement d'images, représenter et modéliser l'information et les données numériques, symboliques ou structurelles dans chacun des formalismes, et effectuer leur combinaison. De nombreux développements ont vu le jour, en particulier pour les applications typiques de classification d'images multisources et de reconnaissance de structures ou d'objets dans les images.

Les **points restant encore à développer** concernent :

— la **gestion du conflit**, qu'il est souvent difficile de différencier de la complémentarité des sources, et pour lequel il n'est pas toujours facile de savoir s'il doit être résolu ou non ;

— la **prise en compte de l'origine des données et des connaissances**, ainsi que des relations entre les sources (souvent effectuée de manière supervisée, elle nécessite donc un peu d'expérience) ;

— le **choix des algorithmes** ;

— l'**évaluation des méthodes**, plus ou moins facile suivant que l'on a accès à la vérité ou non .

Nota : notons à ce propos que les essais de comparaison des approches numériques de fusion ont souvent donné des résultats contradictoires, et ont donc échoué. Nous pensons que la raison essentielle est que chaque problème s'exprime plus facilement dans une théorie que dans une autre, et que sa résolution par des outils non adaptés nécessite donc des distorsions de ces techniques et n'a pas beaucoup de sens.

L'introduction d'informations spatiales dans la fusion est un point important, pour lequel l'ensemble des méthodes existantes pourrait être encore étoffé.

Des travaux sur la combinaison de méthodes sont également prometteurs puisqu'ils visent à exploiter les avantages des différentes théories pour les faire coopérer. Cette combinaison peut s'appuyer sur les liens qui existent entre les différentes approches. Par exemple une probabilité peut être interprétée comme une fonction de masse particulière, une fonction de croyance dont les éléments focaux sont emboîtés peut s'interpréter comme une distribution de possibilités, une distribution de possibilités peut être interprétée comme des intervalles de confiance ou comme une famille de probabilités (voir par exemple [56] pour le détail des liens entre possibilités et probabilités), etc. Ainsi, des travaux ont déjà été effectués pour combiner l'imprécision représentée par des ensembles flous à l'incertitude probabiliste (par exemple [33][98][105] dans une approche de classification markovienne où les classes sont floues), pour combiner les champs de Markov et les fonctions de croyance [10] [68], ou pour raisonner avec des fonctions de croyance dont les éléments focaux sont flous [111][123][128][131].

Remerciements

Plusieurs éléments de discussion, en particulier dans les paragraphes 1 et 2, sont issus de réunions de deux groupes de travail : European Working Group Fusion, et le GT fusion du GDR ISIS. Les auteurs remercient les participants à ces groupes. Ils remercient également François Le Chevalier, qui est à l'origine de cet article et qui nous a prodigué ses encouragements.