Chapitre 5

Classification automatique des objets vidéo

Chaque jour, me fixer 5 actions prioritaires.

5.1 Introduction

Dans le contexte des problématiques et des motivations des deux chapitres précédents, nous présentons dans ce chapitre une approche de classification automatique des objets suivis. La motivation plus particulière que nous avons ici est de répondre au besoin suivant : considérons un utilisateur qui doit identifier puis répertorier tous les objets d'une vidéo et les grouper dans des classes; si un outil permet d'opérer dans un premier temps une construction approximative de ces classes, l'utilisateur n'aura plus qu'à les éditer plus finement, phase pour laquelle une aide pourrait encore être proposée (mais que nous n'explorons pas ici). Les données sont les paramètres gaussiens de toutes les classes d'apparences intra-plan des objets suivis, estimés dans l'espace de descripteurs classiques – l'histogramme de couleurs par exemple. La mise en correspondance entre deux objets suivis est effectuée dans cet espace de paramètres, et le but à atteindre est d'estimer la partition optimale de ces données, c'est-à-dire les classes d'équivalences inter-plans d'objets suivis. Le critère de partitionnement classique que l'on utilise est que chaque classe doit rassembler des objets aussi similaires que possible et que les classes doivent aussi être distinctes que possible les unes des autres.

5.1.1 Quelques points techniques à résoudre

Pour pouvoir classifier automatiquement les objets, des réponses satisfaisantes doivent être apportées aux questions suivantes :

1. Comment apparier deux objets modélisés par des mélanges gaussiens (ayant des structures différentes) ? Faut-il calculer une distance globale entre les densités de

mélanges gaussiens ou bien des distances individuelles entres leurs composantes gaussiennes (les classes d'apparences intra-plan des objets suivis)?

2. Quelle technique de classification automatique est applicable sur ces données? Si chaque objet suivi est représenté par un ou plusieurs points multidimensionnels dans l'espace de paramètres gaussiens, l'algorithme EM est-il applicable sur cette nouvelle distribution? Une classification hiérarchique pourra-t-elle être appliquée avec succès et un coût minimal pour choisir le meilleur nombre de classes d'objets suivis?

5.1.2 Solution adoptée

Les problèmes abordés ci-dessus seront analysés en détail plus tard. L'approche proposée dans ce chapitre consiste dans un premier temps à retenir comme distance entre deux objets suivis une distance - de Kullback , ou de Bhattacharyya - minimale entre leurs composantes gaussiennes. Ceci est justifié par le fait que deux objets suivis provenant de la même classe doivent avoir aux moins deux classes d'apparences intra-plan semblables i.e. deux composantes gaussiennes qui peuvent être identifiées. Dans un deuxième temps, la classification ascendante hiérarchique est appliquée, pour fournir une suite de partitions emboîtées en se basant sur la matrice de proximités calculée entre les objets suivis. Lors de la construction d'une hiérarchie le choix d'une mesure (indice d'agrégation) appropriée entre les classes formées et la détermination du nombre de classes restent deux handicaps. Nous fournissons à l'auteur de la vidéo hyperliée une technique interactive pour sélectionner le nombre de classes, en mettant en service des outils graphiques et visuels pour juger la qualité de la classification et pour corriger manuellement les résultats.

5.1.3 Organisation du chapitre

Après une présentation de l'état de l'art dans la section 5.2, la section 5.3 décrit en détail notre approche de classification hiérarchique des objets suivis. L'expérimentation est réalisée sur la séquence vidéo "Avengers-2" (section 2.5) de 1938 images et de 2749 apparences d'objets correspondant à 29 objets suivis différents. Cette partie sera décrite dans la section 5.4. Le choix du nombre de classes est effectué d'une manière interactive par l'utilisateur expert. La section 5.4.3 décrit l'algorithme que nous avons proposé pour évaluer les résultats expérimentaux obtenus sur la base d'objets de "Avengers-2". Cette évaluation montre un taux de bonne classification automatique autour de 80%, un pourcentage qui est bien acceptable pour une initialisation qui serait reprise par l'auteur de la vidéo hyperliée. Ce taux est aussi satisfaisant vu la grande variation de l'apparition inter-plans des objets de la même classe sémantique. Une telle technique est une solution de l'approche de classification supervisée que nous avons discutée dans le chapitre précédent : l'interaction de l'utilisateur se limite au choix du nombre de classes à la place de la sélection manuelle des modèles d'objets suivis. L'interprétation des résultats ainsi qu'une étude comparative seront présentées dans la section 5.4.4. La section 5.5 conclut ce chapitre.

5.2 État de l'art

Du fait que la classification elle-même est un domaine de recherche très ancien, motivé par les divers domaines d'applications (biologie, finance, imagerie médicale, vidéo numérique, ...), plusieurs classificateurs ont été testés en particulier sur la catégorisation d'images [32] [96] [102] : algorithme de centres mobiles, K-plus proches voisins, arbre de décision, réseaux de neurones, mélanges gaussiens, classification hiérarchique, etc.

Pun et Squire [96] fabriquent des index qui pointent sur des classes d'images, et les représentent selon une arborescente avec les images aux feuilles. L'arbre peut aussi être vu comme un arbre de décision pour la recherche d'images à partir des descripteurs. La composition des classes se fait par une analyse hiérarchique ascendante classique.

Carson et al. [20] proposent une nouvelle représentation des images. Chaque image est décomposée en plusieurs régions nommées *blobs*; chaque blob est cohérent dans l'espace de couleurs et de textures. Tous les blobs des données d'apprentissage provenant de 14 catégories d'images sont classifiés dans 180 blobs "canoniques" en utilisant le modèle gaussien diagonal. Un vecteur de scores est associé à chaque image mesurant sa similarité avec chaque blob canonique. Ensuite, ces vecteurs de scores sont utilisés pour apprendre le classificateur de l'arbre de décision. Sur la même base d'images, l'arbre de décision est aussi testé sur l'histogramme de couleurs au lieu de blobs. Comme il est mentionné dans [62] (page 69), cette comparaison montre que l'histogramme de couleurs fournit des résultats meilleurs que les blobs. Plusieurs explications ont été données pour cette dégradation de performance de blobs : (1) les blobs canoniques ne sont pas suffisamment descriptifs pour distinguer entre les catégories d'images et ils peuvent être mal fabriqués par le moyen du modèle gaussien diagonal; (2) l'apprentissage de l'arbre de décision prend en compte les blobs non pertinents ce qui dégrade les résultats; et (3) les 14 catégories d'images se chevauchent entre elles, en terme des régions, ce qui cause les difficultés.

Le travail récent effectué par Cadez et al. [17] sur la détection de l'anémia est le plus proche de notre approche en termes de données intermédiaires à classifier : les paramètres des mélanges gaussiens. Cadez et al. classifient les patients en deux catégories : patient normal et patient avec un manque de fer. Dans un premier temps, les données de bas niveau (40000 cellules sanguines) extraites de chaque patient sont modélisées par un mélange gaussien. Dans un second temps, la distribution des paramètres gaussiens estimées auparavant est de nouveau modélisée par un mélange gaussien de deux composantes (deux classes de patients). Cette approche est nommée hiérarchique car le mélange gaussien est appliqué deux fois. La deuxième fois, le mélange gaussien est utilisé pour la discrimination entre les deux types de patients. Des bons résultats sont obtenus par le mélange gaussien comparé avec d'autres classificateurs.

Cet aspect des données à plusieurs niveaux est aussi présent dans notre démarche. Cependant nos données posent quelques problèmes techniques : la taille de l'échantillon c'est-à-dire le nombre d'apparences d'un objet suivi est limité, les mélanges gaussiens des différents objets suivis n'ont pas des structures similaires (nombre de composantes par exemple) et le nombre de classes d'objets n'est pas connu a priori.

5.3 Notre approche

Cette section reprend les deux questions posées dans l'introduction de ce chapitre, et répond pour chacune d'entre elles dans les différentes étapes de l'approche de classification détaillées dans la suite. Dans le paragraphe suivant on présente un rappel du problème de la variabilité des objets suivis ainsi que de la solution proposée pour ce problème, tout en mettant en clair les données exploitées par la technique de classification. Ensuite, nous détaillerons la deuxième étape de l'approche qui consiste à mettre en correspondance les objets suivis modélisés par des densités de mélanges gaussiens en calculant les distances de Kullback et de Bhattacharyya . En se basant sur la matrice de proximités ainsi obtenue, l'algorithme de classification ascendante hiérarchique sera appliqué. C'est dans cette dernière étape que la sélection interactive du nombre de classes sera effectuée.

5.3.1 Hiérarchie des données

Les expérimentations menées dans les chapitres précédents ont montré que souvent la distribution d'un objet suivi dans l'espace de descripteurs de bas niveau est multimodale, à cause de la grande variation de l'apparence intra-plan de l'objet suivi et du manque de robustesse et d'invariance des descripteurs existant à ces changements de l'image. La modélisation de cette distribution par un mélange gaussien, dont la structure est déterminée automatiquement en fonction du degré de variabilité de l'objet suivi, consiste à identifier les classes d'apparences intra-plan de cet objet.

Ceci crée une nouvelle représentation de l'objet suivi dans l'espace de paramètres (μ, Σ). Ici, (μ, Σ) désigne le centre et la matrice de variances d'une composante gaussienne du mélange. Dans cet espace la représentation de l'objet suivi est plus compacte que celle dans l'espace initial de descripteurs; seulement quelques points multidimensionnels décrivent l'objet suivi (le nombre de ces points est équivalent au nombre des classes d'apparences de l'objet suivi). En revanche, la dimension de l'espace (μ, Σ) est égale à $d \times (d+1)$, donc beaucoup plus grande que celui de l'espace de descripteurs $\mathbb{R}^d$. La figure 5.1 illustre la distribution de l'enfant suivi de la figure 3.2 dans l'espace de l'histogramme de couleur et dans l'espace de paramètres gaussiens (μ, σ).

Pour résumer cette étape, les objets suivis représentés initialement dans l'espace de descripteurs $\mathbb{R}^d$ sont modélisés et ensuite représentés dans le nouvel espace de paramètres gaussiens $\mathbb{R}^{d\times(d+1)}$. Du fait que les objets suivis ne sont pas représentés en général par des points uniques dans ce nouvel espace, plusieurs algorithmes de classification, le mélange gaussien par exemple, ne peuvent pas être appliqués directement sur ces données.

5.3.2 Distance entre les objets suivis

Comme nous venons de voir, les données représentant les objets suivis dans l'espace de paramètres ne vérifient pas l'hypothèse d'unicité de description des individus pour une stratégie de classification classique; notamment plusieurs vecteurs de descripteurs (les paramètres (μ, Σ)) sont associés à chaque individu (un objet suivi).

Cette section discute le calcul d'une distance entre ces données comme une solution du problème de construction de classe sous ces conditions. Une formalisation de la distance

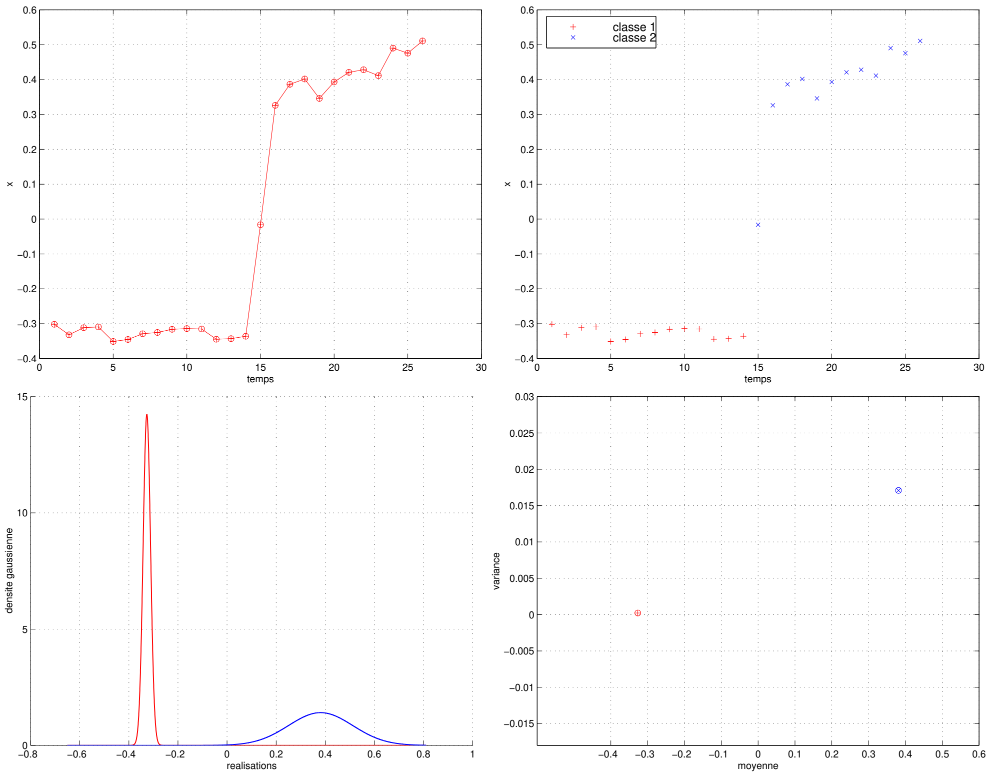


FIG. 5.1: *(a) Distribution de l'enfant suivi de la séquence Ajax (figure 3.2) : le premier axe factoriel de l'histogramme de couleur par rapport a l'axe du temps; (b) la partition correspondante décomposée en deux classes gaussiennes (obtenue par l'algorithme EM); (c) illustration de leurs densités gaussiennes; (d) nouvelle représentation de l'enfant suivi dans l'espace* (μ, σ).

entre deux objets suivis est décrite d'abord. Ensuite nous détaillons les deux distances de Kullback et Bhattacharyya .

5.3.2.1 Forme générale

Définition 5.3.1 *Nous considérons, une distance entre deux objets suivis,* $d_{\ell m}$*, comme étant le minimum des distances entre leurs classes d'apparences intra-plan,*

$$d_{\ell m} = \arg\min_{k,j}(\delta_{kj}) \tag{5.1}$$

avec $1 \leq k \leq K$ *et* $1 \leq j \leq J$*,* K *et* J *représentent le nombre de classes d'apparences intra-plan de deux objets suivis* ℓ *et* m *respectivement, et* δ_{kj} *représente la distance de Kullback ou de Bhattacharyya entre le couple* (k, j) *de classes d'apparences intra-plan. La formulation de ces distances est donnée dans le paragraphe 5.3.2.2.*

Justification de cette formulation Il est clair que deux objets suivis similaires ont très probablement des apparences intra-plan assez proches. Du fait que les apparences intra-plan de chaque objet suivi sont regroupées dans des classes homogènes et qu'une loi de mélange gaussiens les réunit, deux façons pour mettre en correspondance deux objets suivis peuvent être considérées :

1. *Appariement global*: de deux objets suivis en calculant une distance globale entre leurs densités de mélanges gaussiens.

2. *Appariement local*: de deux objets suivis en calculant des distances locales entre leurs composantes gaussiennes i.e. leurs classes d'apparences intra-plan.

Belongie et al. [7] choisissent la première voie pour apparier une image requête avec une image de la base; pour ce faire, ils utilisent dans leur système d'indexation d'images fixes la distance de Kullback calculée entre les densités de mélanges gaussiens. C'est dans l'espace $I\!R^5$, où les pixels d'une image sont représentés par leurs coordonnées spatiales et leurs trois canaux de couleurs RGB, qu'ils cherchent à partitionner l'image en régions homogènes par l'application de l'algorithme EM. L'explication que nous pourrions donner à cette décision est que deux images différentes peuvent avoir plusieurs (petites) régions semblables, et donc un calcul global de la distance peut empêcher une telle confusion et bien expliquer la divergence entre les deux images appariées. En revanche, une composante gaussienne de nos données représente une apparence intra-plan d'un objet suivi. Donc, le fait d'avoir deux classes d'apparences proches est une condition nécessaire et suffisante pour considérer que les deux objets suivis correspondant sont aussi similaires. Il suffit de l'identification de deux apparences pour conclure, et c'est le premier argument pour une similarité basée sur cette correspondance partielle des lois de mélanges.

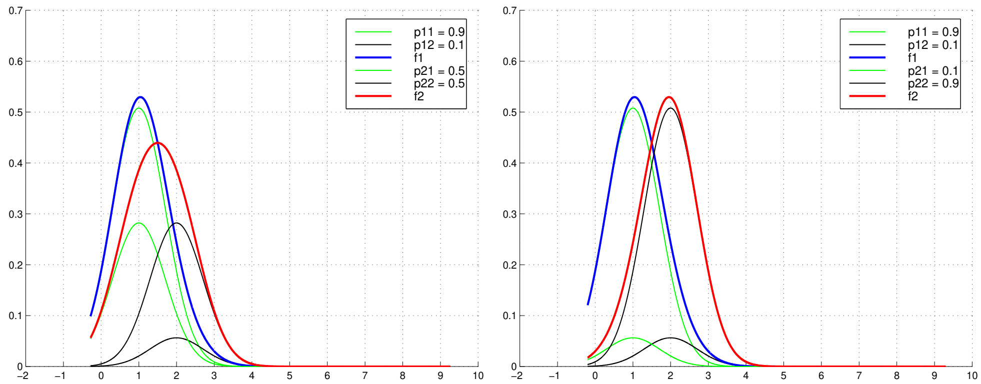


FIG. 5.2: *Illustration de l'effet de la variation du paramètre de proportion des composantes gaussiennes sur la forme générale de la densité de mélange. La densité de référence a pour paramètres* $(p_{11}, p_{12}, \mu_{11}, \mu_{12}, \sigma_{11}, \sigma_{12}) = (0.9, 0.1, 1, 2, 0.5, 0.5)$*. La densité de mélange modifiée a les mêmes paramètres que la densité de référence sauf les proportions qui vaulent* $(0.5, 0.5)$ *dans le premier exemple (figure du gauche) et* $(0.1, 0.9)$ *dans le second exemple (figure du droite).*

D'autre part, une distance globale entre deux densités de mélanges exploite tous les paramètres estimés de ces densités, c'est-à-dire les proportions, les moyennes et les matrices de variances des classes gaussiennes. Le paramètre de proportion représente le nombre des individus appartenant à une classe, et il est considéré comme étant la probabilité a priori d'une classe parmi la composition de la densité du mélange. Donc, ce paramètre a un effet important sur la forme – ou l'allure – de la densité de mélange. Par exemple, la figure 5.2 illustre l'effet de la variation de ce paramètre sur la forme de la densité du mélange gaussien, dans l'espace $I\!R^2$. Sur chaque figure, deux densités de mélanges, f_1 et f_2, chacune de deux composantes gaussiennes sont affichées. Les centres et les variances des composantes correspondantes sont égaux et seul les paramètres de proportions p_{ij} sont variés de telle manière que : $p_{11} + p_{12} = p_{21} + p_{22} = 1$. Cette petite expérience, montre clairement l'effet de la variation de la proportion, qui a causé une sorte de changement d'échelle entre les densités de chaque composante gaussienne et ensuite une translation de la densité du mélange. Ceci implique, une augmentation de l'écart (ou bien diminue l'intersection) entre les deux densités de mélanges. Sur cet exemple simple, le calcul d'une distance globale entre les densités f_1 et f_2 nous emmène à une décision que les deux populations (deux densités de mélanges) ne sont pas semblables; cette fausse conclusion n'est pas bonne car les deux populations ont des centres et des matrices de variances identiques. La figure 5.3 illustre la sensibilité de la distance de Kullback globale calculée entre les densités de mélanges gaussiens de l'exemple précédent (figure 5.2) à la variation du paramètre de proportion.

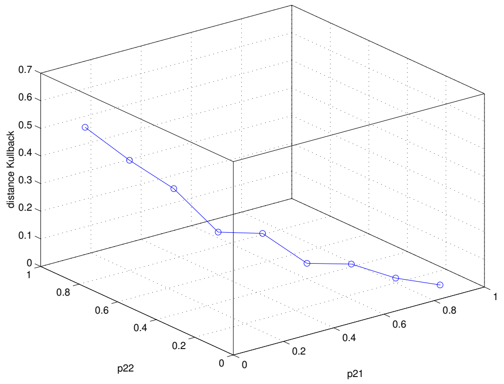


FIG. 5.3: *Illustration sur les données de l'exemple précédent de la sensibilité de la distance de Kullback globale à la variation du paramètre de proportion des composantes gaussiennes. La distance de Kullback est calculée entre la densité référence et celle dérivée de cette référence avec un changement des proportions seulement. Par exemple, entre les deux densités de mélanges gaussiens de la figure 5.2 (partie droite) la distance de Kullback mesure une divergence égale à* 0.4997 *qui est significative.*

En revanche, le calcul local d'une distance entre les composantes gaussiennes ne prend pas en compte le paramètre de la proportion i.e. le changement d'échelle entre les gaussiennes. Ceci est le deuxième argument pour lequel nous calculons une distance locale entre les composantes gaussiennes de deux objets suivis. En adoptant cette technique, les deux populations de l'exemple précédent, où la distance de Kullback vaut zéro tout au long de la variation des proportions, sont jugées entièrement similaires. Ceci est très différent du contexte du travail de Belongie et al. [7] le paramètre de proportion est important à prendre en compte, par exemple, lorsqu'il s'agit de distinguer entre une petite et une grande tache rouge dans deux images.

5.3.2.2 Appariement des classes d'apparences intra-plan

Nous avons défini auparavant la forme générale de la distance entre deux objets suivis (équation 5.1), qui est exprimée en fonction de distance locale entre les classes d'apparences intra-plan. La mise en correspondance entre deux classes d'apparences est un cas particulier d'un problème général de comparaison de deux échantillons, où le but est de mesurer la similarité et pas pour décider si les échantillons sont identiques ou différents. Dans notre cas nous disposons de deux distributions gaussiennes multivariées, f_r et f_q, dont leurs paramètres $\theta_r = (\mu_r, \Sigma_r)$ et $\theta_q = (\mu_q, \Sigma_q)$ sont connus. Dans la suite, nous rappelons du principe général des tests d'hypothèses utilisés souvent pour comparer deux distributions et ensuite nous présentons quelques distances adaptées à nos données.

Tests d'hypothèses. Les tests d'hypothèses sont souvent utilisés dans la comparaisons de deux ou plusieurs distributions. Un test d'hypothèse comporte trois étapes :

- définition du test, par exemple,

$$\begin{cases} H_0\text{: les deux distributions sont identiques } \textit{hypothèse nulle} \\ H_1\text{: les deux distributions sont différentes;} \end{cases}$$

- résolution du test c'est-à-dire la construction de la fonction de décision permettant d'associer à toute réalisation de l'échantillon $(X_1, \ldots, X_n)$ soit l'hypothèse H_0, soit l'hypothèse H_1;

- et enfin la décision à prendre à partir d'une réalisation de l'échantillon.

La détermination de la fonction de décision revient à définir la partition de l'espace des valeurs de l'échantillon en deux sous-ensembles :

- l'ensemble W des réalisations pour lequelles on rejette H_0 en faveur de H_1 : c'est la *région critique*;

- l'ensemble $\bar{W}$ des réalisations pour lesquelles on conserve H_0 : c'est la *région d'acceptation.*

Soient α la probabilité de rejeter H_0 en faveur de H_1 alors que H_0 est vraie, et β la probabilité de conserver H_0 alors que H_1 est vraie (erreur de première et de deuxième espèce). Une illustration graphique de ces deux erreurs est donnée dans la figure 5.4. Souvent, c'est le risque de première espèce considéré comme le plus important par Neyman et Pearson, qu'on cherche à minimiser soit donc à maximiser la probabilité $1 - \beta$, appelée

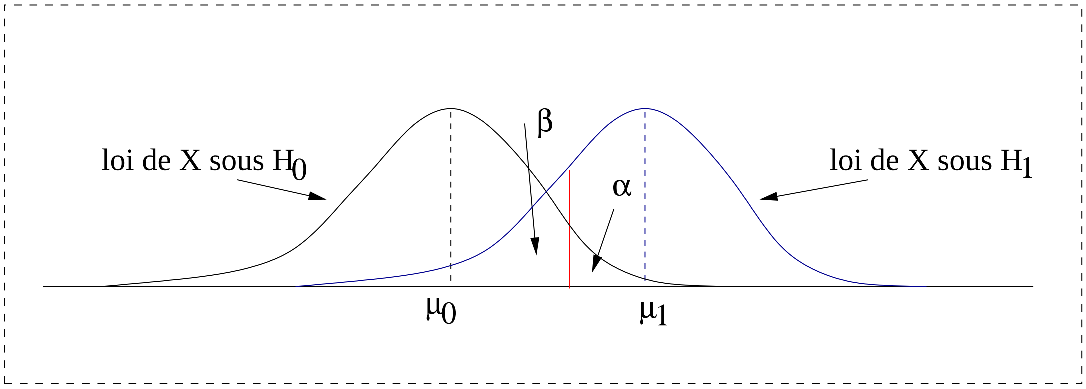


FIG. 5.4: *Les erreurs de première et deuxième espèce.*

puissance de test. La détermination de la fonction de décision revient alors à déterminer un sous espace W ayant, sous l'hypothèse H_0, une probabilité inférieure ou égale à $\alpha^\star$, qui est une valeur fixée a priori (souvent 5% ou 1%), appelée *niveau de signification du test* :

$$Prob[(X_1, \ldots, X_n) \in W \mid H_0] \leq \alpha^\star$$

Distance de Kullback Basée sur la théorie de l'information et des tests d'hypothèses, la distance de Kullback ([74] pages 3-6) est une distance bien adaptée pour la mesure de la divergence entre deux distributions quelconques, et en particulier entre deux distributions gaussiennes. Soient X une variable aléatoire dont la densité de probabilité dépend du paramètre θ, et les deux hypothèses H_0 et H_1 portant sur cette variable X, au vu d'une réalisation d'un échantillon $X_1, \ldots, X_n$, qu'elle provienne de la distribution gaussienne f_ℓ et f_m respectivement :

$$\begin{cases} H_0\colon \theta = \theta_r \\ H_1\colon \theta = \theta_q \end{cases}$$

En se basant sur le test paramètrique ci-dessus, la distance de Kullback est vue comme étant une fonction de décision permetttant d'associer à toute réalisation x de l'échantillon $X_1, \ldots, X_n$ soit l'hypothèse H_0 ou H_1. La divergence, $\delta(f_r, f_q)$, est ainsi définie par la somme de la moyenne des réalisations de θ_r qui conserve H_0 en faveur de H_1, notée par $I(0:1)$ et inversement.

$$\begin{aligned} \delta_k(f_r, f_q) =& I(0:1) + I(1:0) \\ =& \int f_r(x) log \tfrac{f_r(x)}{f_q(x)} d(\theta_r) + \int f_q(x) log \tfrac{f_q(x)}{f_r(x)} d(\theta_q) \\ =& \int log \tfrac{P(H_0|x)}{P(H_1|x)} d_{\theta_r}(x) - \int log \tfrac{P(H_0|x)}{P(H_1|x)} d_{\theta_q}(x) \end{aligned} \tag{5.2}$$

avec $P(H_i \mid x)$ est la probabilité conditionnelle d'avoir l'hypothèse H_i $(i = 1, 2)$ sachant la réalisation $X = x$.

L'évaluation de cette distance est réalisée par les procédures de Monte-Carlo qui consistent à générer des données simulées à partir des paramètres θ_0 et θ_1 et sur lesquelles on estime les probabilités a posteriori $(P(H_i \mid x))$. La distance de Kullback présentée sous la forme ci-dessus est égale à zéro lorsque $\theta_0 = \theta_1$, elle est aussi symétrique, mais par contre, elle ne vérifie pas la propriété triangulaire d'une métrique.

Distance de Bhattacharyya . Un calcul direct d'une distance entre deux distributions gaussiennes, portant seulement sur les paramètres gaussiens (μ, Σ), est réalisable par le moyen de la distance de Bhattacharyya ([38], page 99) :

$$\delta_b(f_r, f_q) = \underbrace{\frac{1}{8}(\mu_r - \mu_q)^T \left(\frac{\Sigma_r + \Sigma_q}{2}\right)^{-1} (\mu_r - \mu_q)}_{B1} + \underbrace{\frac{1}{2} log \left(\frac{\mid \frac{\Sigma_r + \Sigma_q}{2} \mid}{\sqrt{\mid \Sigma_r \mid\mid \Sigma_q \mid}}\right)}_{B2} \tag{5.3}$$

où | M | dénote le déterminant de la matrice M. Le premier terme dans cette expression, $B1$, est similaire à la distance de Mahalanobis et il mesure la distance entre deux populations causée par le décalage de la moyenne, ainsi que le deuxième terme $B2$ exprime la séparabilité entre les classes grâce à la différence entre les covariances.

5.3.3 Classification hiérarchique

Nous avons abordé dans la section précédente le calcul d'une distance entre les objets suivis dans l'espace de paramètres gaussiens. En se basant sur la matrice de distances $\mathcal{D} = (d_{\ell m})_{1 \leq \ell, m \leq n}$ ainsi obtenue entre les n objets suivis, la troisième étape de notre approche consiste à appliquer la classification ascendante hiérarchique pour regrouper les objets dans des groupes homogènes.

Une hiérarchie consiste en une suite de partitions emboîtées, depuis l'ensemble de tous les n objets $\{1, \ldots, n\}$ jusqu'aux singletons formés par les objets eux mêmes, $\{1\}, \ldots, \{n\}$, en passant par des divisions successives des sous-ensembles. La procédure la plus couramment utilisée pour produire automatiquement une hiérarchie est la classification ascendante hiérarchique (CAH). Celle-ci se base sur une matrice de dissimilarité (ou de similarité) entre les objets, symétrique et de taille $n \times n$.

5.3.3.1 Algorithme du CAH

Les étapes d'un processus de classification ascendante hiérarchique sont les suivantes [69] :

1. Initialisation : chaque objet constitue une classe, on commence donc avec n classes qui sont des singletons. Les dissimilarités entre les groupes sont au départ les dissimilarités entre les objets qui les constituent.

2. Trouver les deux classes les plus similaires et les fusionner en une classe – on se retrouve ainsi avec une classe de moins.

3. Calculer les dissimilarités entre la nouvelle classe et les autres classes.

4. Répéter les étapes 2 et 3 jusqu'à ce que tous les objets soient réunis dans une classe de taille n, ou bien jusqu'à ce que l'on ait obtenu le nombre de classes désiré.

Le calcul de la dissimilarité entre deux classes à l'étape 3 peut être réalisé de différentes façons. Par exemple, si l'on utilise le critère d'agrégation du lien ou saut minimum (single-link clustering en anglais), la dissimilarité entre deux classes est définie comme la plus petite dissimilarité entre les objets des deux classes. D'autres critères d'agrégation peuvent être utilisés, comme celui du lien ou saut maximum ou celui de la distance moyenne (complete-link et average-link clustering), ou bien le critère d'inertie de Ward [124] (*Ward's minimum variance clustering*). Cependant, le critère de Ward est souvent utilisé dans la cas où les observations sont des vecteurs de $I\!R^d$ et les similarités des distances euclidiennes [107], ce qui ne correspond pas à notre cas.

5.3.3.2 Motivation du choix de CAH

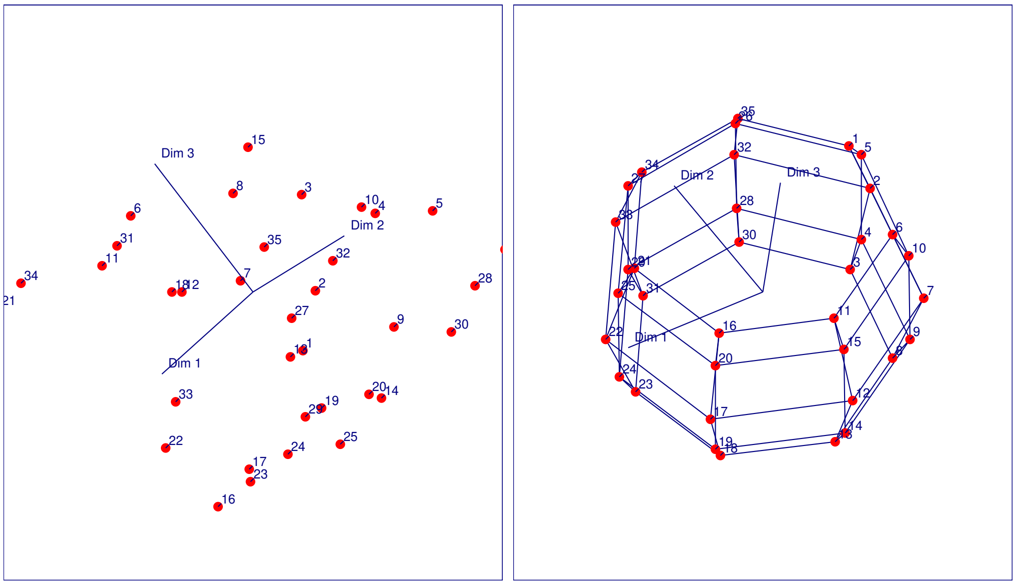
Dans le domaine de classification automatique de données, les méthodes de classification existantes (k-means, KNN, mélange gaussien, etc.), probabilistes et non probabilistes, peuvent être regroupées en deux familles – hiérarchique et partition – selon la structure de classes produite. Par exemple, le mélange gaussien est une méthode de classification probabiliste qui produit une partition unique des données, tandis que le CAH est hiérarchique et non probabiliste.

Nous avons employé la méthode de CAH pour les deux points suivants :

– *Nature des données exploitées :* Les données représentant les objets suivis dans l'espace de paramètres gaussiens ont une nature non habituelle pour les différentes méthodes de classifications, où chaque objets est représenté par plusieurs individus dans cet espace (voir section 5.3.1). Donc, on se retrouve avec une matrice de distances (de Kullback ou de Bhattacharyya) qui décrit la dissimilarités entre les objets suivis. Dans cette situation, la méthode de CAH est la seule applicable pour classifier automatiquement les objets. En revanche, d'autres méthodes de classification peuvent être appliquées sur la représentation spatiale de la matrice de distances fournie par exemple par la méthode de MDS [73]. Ainsi, l'objectif du MDS est de trouver pour les n objets suivis, notés par $x_1, \ldots, x_n$ et ayant comme matrice de distances $\mathcal{D}$, n points correspondants $y_1, \ldots, y_n$ de l'espace $I\!R^k$ (k petit, par exemple 2), avec une matrice de distances $\mathcal{D}_k$ qui soit la plus proche possible de $\mathcal{D}$ (voir figure 5.5). Il s'agit de minimiser la fonction de coût suivante :

$$\parallel \mathcal{D}(x_1, \ldots, x_n) - \mathcal{D}_k(y_1, \ldots, y_n) \parallel$$

Deux points nous ont amenés à renoncer à cette méthode non-linéaire : (1) la difficulté de trouver une représentation spatiale qui décrit parfaitement les données initiales, dans un espace de dimension réduite, et (2) la complexité en terme de calcul de la méthode de MDS pour trouver une telle représentation.



(a) (b)

FIG. 5.5: *Représentation spatiale obtenue par la méthode du MDS (b) pour la matrice de distances euclidiennes calculées entre les individus du nuage de points initial (a).*

– *Interprétation interactive des résultats :* Un des intérêt des méthodes hiérarchiques est qu'elles permettent d'obtenir une classification à différents niveaux de détail (voir figure 5.6) : plus on descend dans la hiérarchie, plus la classification est fine. Au contraire des méthodes qui produisent des partitions, la structure obtenue par le CAH permet une interprétation relativement intuitive du résultat, et l'auteur de la vidéo interactive bascule entre les niveaux de la hiérarchie sans aucun coût de calcul ou d'estimation de la partition désirée.

5.3.3.3 Quelques problèmes ouverts

Le choix du critère d'agrégation (saut minimum, maximum, etc.) reste un problème délicat. Théoriquement, tous ces critères donnent la plupart du temps les mêmes résultats si les classes sont compactes et bien séparées. Par contre, si les classes sont trop proches, ou n'ont pas une forme hypersphérique, des résultats très différents peuvent être obtenus. De plus les méthodes de classification hiérarchique n'optimisent généralement pas de critère numérique explicite, ce qui rend difficile le choix entre différentes méthodes et leur évaluation par rapport à un modèle de données par exemple. Néanmoins, il existe quelques exceptions; par exemple, la classification hiérarchique ascendante par la stratégie du saut minimum trouve, parmi les ultramétriques inférieures à la dissimilarité de départ, celle qui lui ressemble le plus. Ceci revient à trouver l'arbre de longueur minimum reliant tous les objets.

Le deuxième problème du CAH est le choix automatique du nombre de classes (niveau

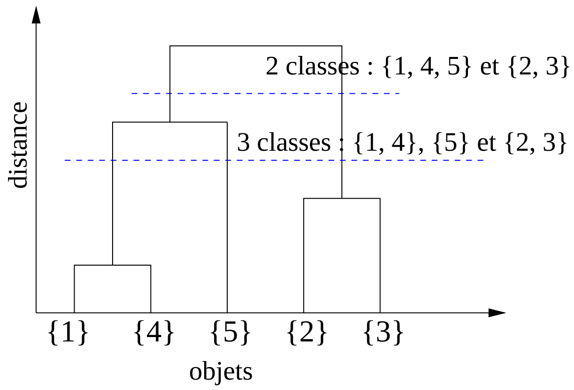


FIG. 5.6: *Exemple de hiérarchie : un dendrogramme correspondant à une classification hiérarchique de 5 classes initiales. Les traits pointillés montrent deux niveaux de détail que l'on peut choisir pour la classification finale.*

de coupure de la hiérarchie), où la théorie ne fournit pas des méthodes satisfaisantes. Nous abordons ce problème un peu plus en détail dans la section suivante.

L'algorithme général consiste à balayer à chaque étape un tableau de $\frac{n(n-1)}{2}$ distances afin d'en rechercher l'élément de valeur minimale, à réunir les deux objets correspondants, à mettre à jour les distances après cette réunion et à recommencer avec $n-1$ objets au lieu de n. La complexité d'un tel algorithme est de n^3 (ordre du nombre d'opérations à effectuer), qui est assez élevé. Cependant, diverses techniques ont été proposées pour accélérer les opérations [79]. Le lecteur peut s'adresser à la thèse de [99] pour plus de détails sur ce sujet qui sort du cadre de travail.

5.3.3.4 Sélection interactive du nombre de classes

Dans le domaine de la classification automatique ou non supervisée de données le nombre de classes est inconnu, et on ne sait pas de quelles classes proviennent les objets à classifier. Le fait d'avoir un critère automatique pour sélectionner le nombre de classes est très important, en particulier lorsqu'il s'agit d'un travail répétitif de classification de données. C'est par exemple le cas où nous avons cherché à classifier les apparences intra-plan de chaque objet suivis dans la vidéo par un mélange gaussien. Il est clair dans ce cas qu'un nombre fixe de classes d'apparences pour tous les objets suivis n'est pas le bon choix, vu que ces objets n'ont pas le même degré de variabilité.

Il existe quelques méthodes pour choisir le point de coupure de la hiérarchie formée par le CAH [66] [84]. Une étude comparative sur des données synthétiques a été menée par Milligan et Cooper [84]. Le critère le plus connu est basé sur la minimisation de la variance intra-classe et la maximisation de la variance inter-classe [97]. Il est connu que l'utilisation de la variance sur ce type de problème favorise la formation de classes hyper-sphériques. Une analyse de données ne vérifiant pas cette hypothèse implicite est donc biaisée. Sur l'exemple de la figure 5.7, donnée dans [100], la plupart de ces méthode détecterons vraisemblablement la présence de quatre classes dans les données. Les trois classes les plus denses sont fusionnées en une seule. Une méthode récente a été proposée par Ribert [100] qui consiste à couper la hiérarchie à plusieurs échelles, en prenant en compte le cas des classes de densités différentes. Cette méthode consiste à estimer un critère σ^2/μ^2 qui se fait sur les valeurs de l'histogramme calculé sur une hiérarchie. Par exemple, il calcule un histogramme à N intervalles sur une hiérarchie donnée. La variance et la moyenne seront estimées sur les N valeurs de l'histogramme (et non pas sur les valeurs des palliers de la hierarchie). Pour obtenir les valeurs critiques du critère, il faut établir un abaque f(*dimension de l'espace, nombre d'éléments*). Le principe est de générer des configurations avec la distribution attendue (ex. gaussienne) ne comprenant qu'une seule classe. Dans le cas où la dimension de l'espace est connue, il faut générer des bases synthétiques dans un espace de même dimension. Il est en effet important de prendre en compte ce paramètre car l'interprétation d'une hiérarchie peut changer en fonction de la dimension de l'espace de représentation. Cette méthode permet de résoudre quelques cas de densité variable (voir figure 5.7), où les méthodes de coupure unique (horizontale) échouent. L'inconvénient majeur (d'après l'auteur de cette méthode) réside dans la constitution d'une abaque de référence.

Vue la grande variation de l'apparence inter-plan des objets suivis à classifier, nous sommes quasiment convaincus que tous ces critères ne sélectionnerons pas le meilleur niveau de coupure de la hiérarchie.

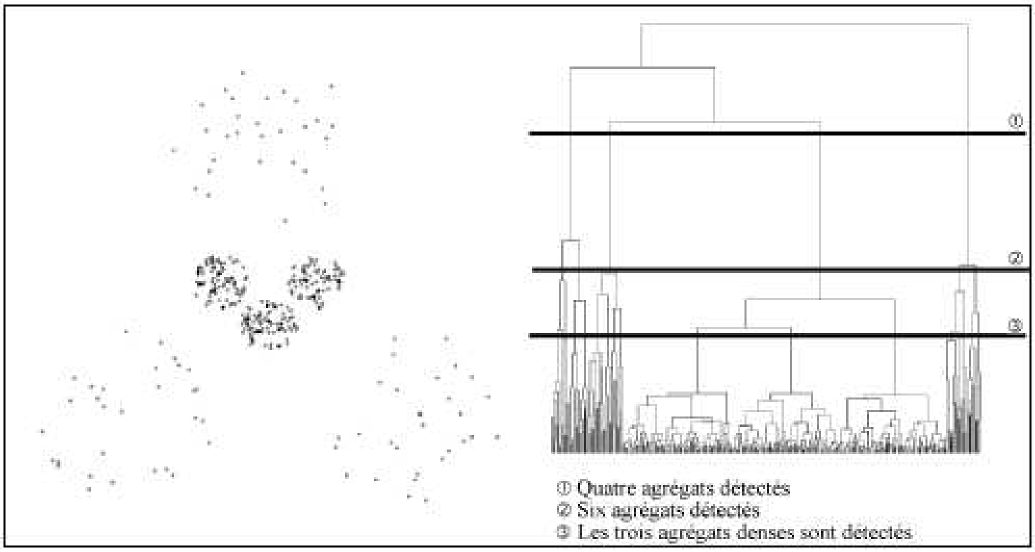


FIG. 5.7: *Exemple de densité variable de* 6 *agrégats (classes); échec de la coupure unique de la hiérarchie.*

Nous proposons à l'utilisateur expert une méthode interactive et visuelle pour sélectionner le nombre de classes des objets suivis. D'abord on filtre les niveaux de la hiérarchie fabriquée qui ont des distances de transition (distance relative) assez importante (par rapport à un seuil pré-défini), ensuite l'utilisateur expert choisit un niveau parmi eux, il visualise les classes obtenues à ce niveau, et enfin il fixe son choix pour un de ces niveaux.

Comme on ne peut pas s'attendre à un résultat parfait le système pour la construction de la hyper-vidéo fournit à l'utilisateur expert des outils interactifs qui lui permettent de corriger les résultats de la classification. Ces outils sont principalement du type "drag & drop" et permettent de déplacer par la souris un objet mal classé d'une classe à une autre déjà existant ou non. Plus de détails sur cet aspect interactif dans le chapitre 8.

5.4 Expérimentation

5.4.1 La séquence Avengers-2

Les expérimentations de l'approche proposée ont été menées sur la séquence Avengers-2 du film "Chapeau Melon et Bottes de Cuir". En appliquant les outils de segmentation de la vidéo présentés au chapitre 2, 2749 apparences d'objets correspondant à 29 objets suivis ont été localisées dans 1938 images. Ces apparences d'objets correspondent à 7 classes d'objets différentes (classes de *J. Steed, Licorne, voiture Ford-I, voiture Ford-II, Purdey, Prêtre* et *voiture Mercedes*). La figure 5.8 illustre quelques objets de cette séquence. Comme cette figure le montre, les objets suivis de Avengers-2 sont mobiles et d'apparences intra- et inter-plans très variables. A la différence avec la séquence Avengers-1 chaque objet suivi de Avengers-2 possède une trentaine d'apparences intra-plan au moins. Ce nombre d'apparences intra-plan par objet est relativement suffisant vis-à-vis de la dimension réduite de l'espace de descripteurs. Cette situation des données permet de tester l'approche proposée dans les meilleurs conditions d'estimation et de comparaison des paramètres gaussiens. Les classes d'apparences intra-plan des différents objets suivis seront modélisées par le même type de modèle de mélange gaussien afin qu'on soit capable de les comparer.

5.4.2 Paramètrage de l'approche

La modélisation de la variabilité intra-plan des objets suivis est une étape fondamentale de nos approches de classification. Les deux chapitres précédents ont évoqué cette étape du point de vue théorique et expérimentale. En effet, des restrictions sont posées sur les expérimentations présentées ici en vu des résultats obtenus auparavant : *meilleur espace de descripteurs* et *structure optimale du mélange gaussien.*

Descripteurs de bas niveaux L'adoption des descripteurs de couleurs dans nos expérimentations est un choix déjà discuté dans la section 3.3. Ici, nous calculons sur les apparences des objets suivis les histogrammes RGB, HSV, I, et H. Ces histogrammes sont quantifiés d'abord en 64, 64, 32 et 32 cellules respectivement. Ensuite, une analyse en composante principale est appliquée sur chaque nuage de descripteurs.

Données de classification Comme nous l'avons mentionné auparavant, les données sur lesquelles la classification hiérarchique sera appliquée sont les paramètres gaussiens des classes d'apparences intra-plan des 29 objets suivis de Avengers-2. Le nombre maximal de composantes gaussiennes d'une loi de mélange est fixé à 3 ($MaxNbC = 3$), et seul

Fig. 5.8: *Quelques apparences d'objets suivis de la séquence vidéo Avengers-2.*

le modèle gaussien général ($[\lambda_j Cj]$) est employé. D'une part ce modèle gaussien estime la frontière optimale entre les classes d'apparences intra-plan d'un objet suivi, et d'autre part le risque d'estimation instable est très faible vu le nombre suffisant d'apparences d'objets suivis de la séquence Avengers-2 par rapport à la dimension réduite des espaces de descripteurs (10 pour RGB et HSV, et 5 pour I et H). Aussi, fixer un modèle gaussien est logique car le calcul d'une distance de Bhattacharyya par exemple entre deux modèles gaussiens de différent types n'a pas de sens : par exemple un modèle gaussien multivarié de la famille sphérique (variables indépendantes) et un second de la famille générale (variables dépendantes) ne sont pas comparables (voir section 3.6.2.3). Dans ce cas le critère ICL servira seulement à choisir le nombre de composantes gaussiennes le mieux adapté aux données.

Dans l'évaluation de la distance de Kullback par la procédure de Monte Carlo nous générons environ 2000 individus simulés à partir de chaque composante gaussienne référence. Ce nombre est choisi empiriquement mais par contre il est beaucoup plus élevé que la taille de l'échantillon utilisé dans l'estimation des paramètres gaussiens, ce qui veut dire que l'échantillon simulé conduit à une évaluation de la distance stable. Notons que les valeurs de cette distance varient en fonction de la taille des données simulées. En terme de complexité les deux distances implantées se calculent en temps réel sur une machine UltraSparc 256 MHZ mais avec un ordre de grandeur différent; le calcul de la distance de Bhattacharyya est 10 fois plus rapide que celui de la distance de Kullback , car elle est formulée directement sur les descripteurs et ne possède pas une phase de simulation de données.

Maintenant, si les distances calculées ont des valeurs petites entre les objets de la même classe sémantique (distances intra-classes) et des valeurs grandes entre les objets des classes différentes (distance inter-classes) alors on peut s'attendre à une classification automatique parfaite. Sinon, il y aurait certainement des mauvaises classifications. Cela peut s'expliquer comme il suit. La distance est peut-être mal adaptée pour la comparaison des composantes gaussiennes. Ou on est dans le cas où des distances inter-classes sont plus petites que celles intra-classes. Ce dernier cas est dû très probablement à la grande variance dans l'apparence des objets de la même classe sémantique.

La figure 5.9 illustre la matrice de distance de Kullback obtenue durant cette expérimentation pour 9 objets suivis. Parmi ces 9 objets, $1, 2, 3, 4$ et 5 appartiennent à la classe sémantique "Steed" et le reste appartient à la classe "Espion". Sur cette représentation graphique on observe que les distances intra-classes ont des basses fréquences inversement aux distances inter-classes. Selon cette représentation des proximités on peut nettement classifier les objets en deux classes différentes qui correspondent parfaitement aux deux classes sémantiques.

5.4.3 Résultats et procédure d'évaluation

L'application de l'algorithme de CAH sur les matrices de distances calculées produit des hiérarchies de 29 niveaux de détail; la coupure horizontale de la hiérarchie à un certain niveau i génère $(29 - i + 1)$ groupes d'objets suivis.

A un certain niveau de la hiérarchie la qualité des résultats de regroupement est ex-

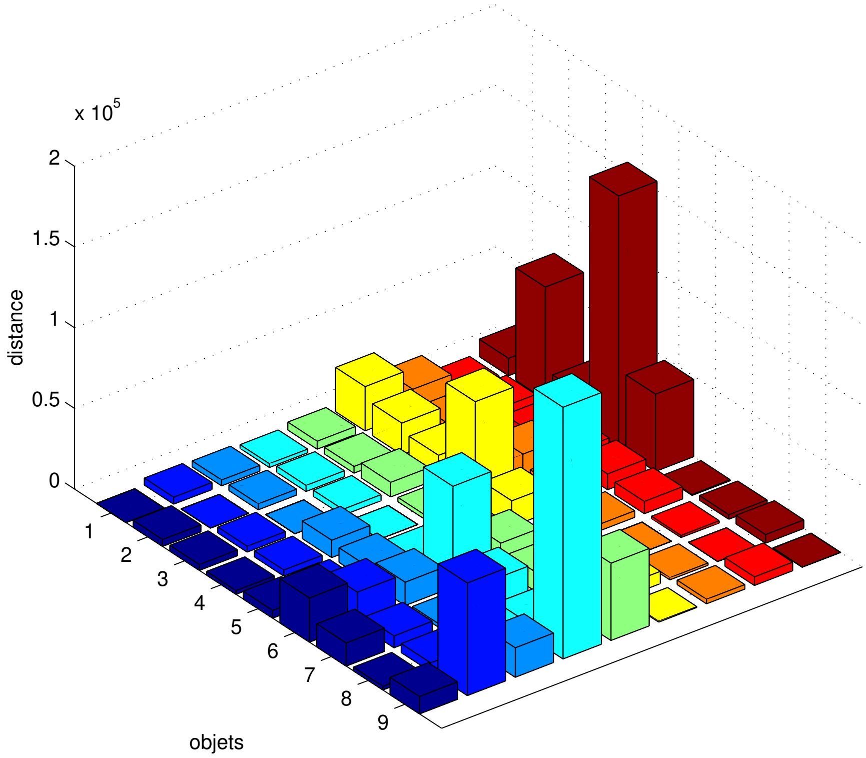


FIG. 5.9: *Illustration de la matrice de distances de Kullback entre* 9 *objets de la séquence Avengers-2; Les objets* $1, 2, 3, 4$ *et* 5 *appartiennent à la classe sémantique "Steed" et le reste appartient à la classe "Espion". Cette matrice de distances décrit bien cette partition.*

primée en fonction de l'homogénéité des individus de chaque classe. Les objets d'une classe sont jugés homogènes ou non selon la vision humaine.

Soit $cc\%$ le pourcentage de classification correcte pour K classes d'un niveau i. L'algorithme d'évaluation proposé ici est le suivant :

0- Initialiser le nombre des objets correctement classifiés dénoté par cc à 0;

1- Compter le nombre des éléments homogènes dans chaque classe k ($1 \leq k \leq K$). Par exemple, la classe numéro 1 regroupe 5 objets suivis, 4 objets de type-1, 1 objet de type-2 et 0 objets de type-[3...L], avec L le nombre correct de classes d'objets fabriquées manuellement ($L = 7$ pour les objets de la séquence Avengers-2). Soit $S = (s_{k\ell})$ la matrice $K \times L$ des scores obtenues pour les K classes; l'élément $s_{k\ell}$ représente le nombre des objets de type-ℓ trouvés dans la classe k.

2- Pour $k = 1$ jusqu'à K faire

   a- calculer $A = \max(s_{k\ell})_{1 \leq \ell \leq L}$; soit j le numéro du colonne de l'élément A dans la matrice S

b- calculer $B = \max(s_{ij})_{k \leq i \leq K}$;

c- **Si** $A \geq B$ **alors** $cc = cc + A$;
affecter les éléments s_{ij} avec $(k+1) \leq i \leq K$ à zéro; (cette opération associe un label unique "type-ℓ" à la classe k);

d- **Sinon**, affecter l'élément $s_{k\ell}$ à zéro, et répéter l'étape a.

3- $cc\% = cc/n \times 100$, où n désigne le nombre total des objets classifiés.

Cet algorithme est utilisé pour évaluer les résultats obtenus au niveau 22 des hiérarchies des différents descripteurs de bas niveaux de cette expérimentation. Ce niveau est choisi car en réalité tous les objets suivis de la séquence Avengers-2 correspondent à 7 classes sémantiques différentes, comme nous l'avons déjà mentionné dans ce chapitre.

Le tableau 5.1 résume les résultats de test. Le meilleur pourcentage de classification correcte est de 79.31%. Les figures 5.10 et 5.11 illustrent les 7 classes d'objets trouvées automatiquement lorsque les paramètres gaussiens sont estimés dans l'espace de descripteurs RGB, la distance de Kullback est calculée et le critère de saut maximum est employé Ces résultats avec les séquences vidéo complètes sont disponible sur le Web à l'adresse suivante :

http://www.inrialpes.fr/movi/people/Hammoud/unsupervisedClassification.htm

Espace de descripteurs	d	d_E	Distance	$cc\%$ Critère du saut		
				minimum	maximum	moyenne
h_{RGB}	64	10	δ_b	44.83	62.07	62.07
h_{RGB}	64	10	δ_k	55.17	79.31	68.97
h_I	32	5	δ_b	44.83	65.52	58.62
h_I	32	5	δ_k	48.28	68.97	58.62
h_{HSV}	64	10	δ_b	44.83	65.52	68.97
h_{HSV}	64	10	δ_k	51.72	68.97	58.62
h_{Hue}	32	5	δ_b	55.17	55.17	58.62
h_{Hue}	32	5	δ_k	58.62	62.07	58.62

TAB. 5.1: *Les pourcentages de classification correcte des d'objets suivis de la séquence Avengers-2, dans le cas où la variabilité intra-plan des objets suivis est modélisée par des densités de mélanges gaussiens dont le nombre maximal de composantes permis est* 3.

5.4.4 Analyse comparative et discussion

La classification automatique de données est une tâche difficile en général. De plus, lorsqu'il s'agit de classifier des densités de mélange gaussien qui représentent des objets suivis non-rigides et d'apparences trop variable dans les scènes, on n'attend pas une construction automatique des classes entièrement parfaite. Les tests sur la base réelle d'objets suivis

FIG. 5.10: *Les groupes d'objets suivis de la séquence Avengers-2 fabriquées automatiquement. La suite de ces résultats est illustrée par la figure 5.11. Une seule apparence par objet suivi est illustrée ici; les* 2745 *apparences de ces objets suivis sont disponibles sur le web à l'adresse mentionnée auparavant. Le pourcentage de classification correct obtenu est égal à* 79.31%*.*

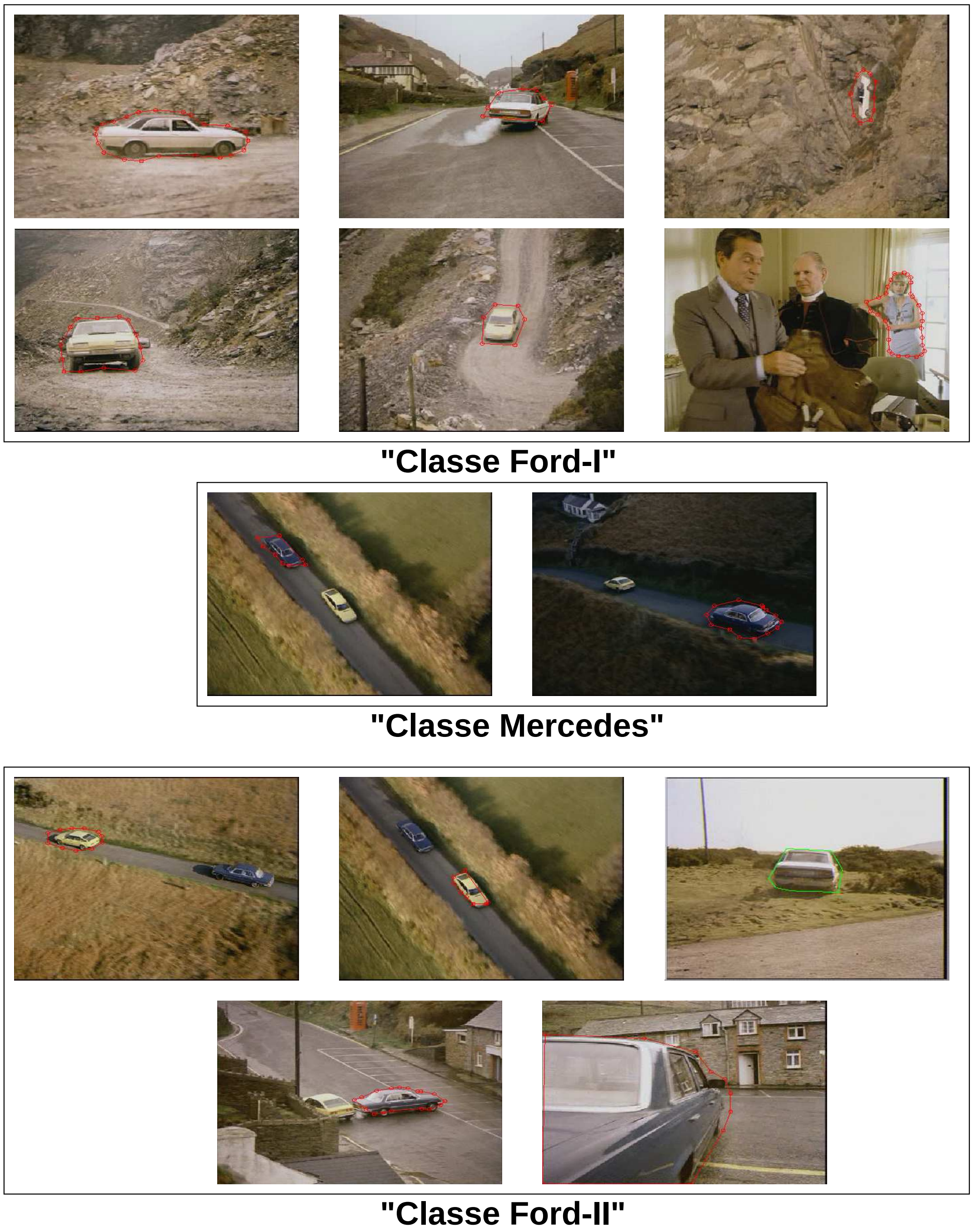


FIG. 5.11: *Suite de la figure 5.10*

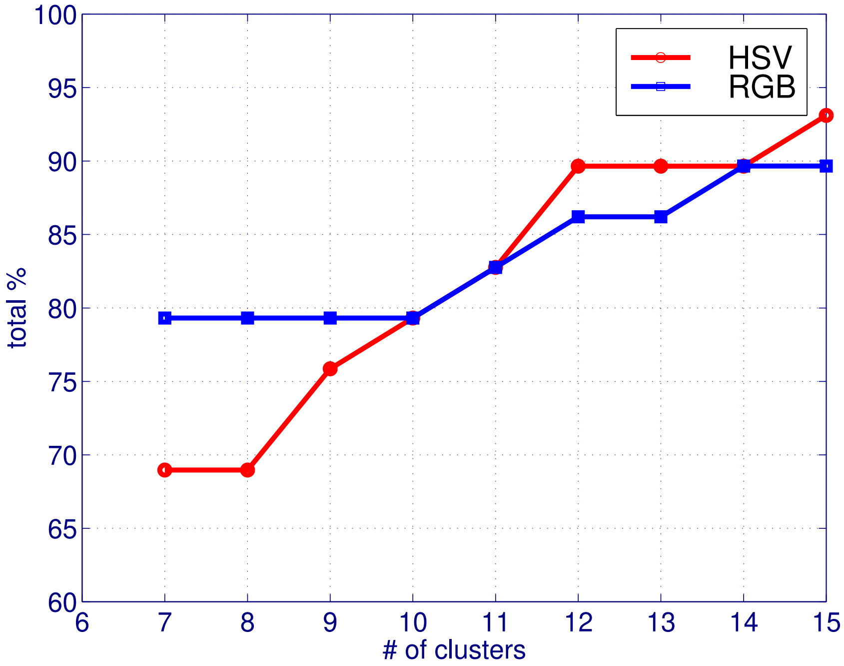


FIG. 5.12: *Illustration du pourcentage de bonne classification des objets suivis par rapport au nombre de classes dans la hiérarchie, dans le cas où le critère du saut maximum et la distance de Kullback sont employés, et les paramètres gaussiens sont estimés dans les espaces de descripteurs RGB et HSV.*

de la séquence vidéo Avengers-2 donnent un pourcentage de bonne classification de 80% environ; ce qui revient à une classification correcte de 23 objets suivis parmi les 29 de la séquence. On peut remarquer que les 6 objets mal classés sont partagés dans les deux dernières classes *Ford-I* et *Ford-II* souvent nommées classes poubelles en classification automatique. Une classe poubelle est le fait d'une coupure de la hiérarchie à une échelle unique; elle regroupe généralement des éléments non-homogènes forcés à être ensemble. Une coupure unique de la hiérarchie à un niveau plus bas que 22 ou bien une coupure à plusieurs échelles permet probablement de garder les 5 premières classes et de couper les deux dernières classes en quatre sous-classes plus homogènes. Dans ce cas, une surestimation du nombre de classes d'objets est effectuée mais par contre le contenu des classes est plus homogène. La figure 5.12 illustre le pourcentage de bonne classification par rapport aux nombres de classes. On peut voir que lorsque le nombre de classe est double du vrai nombre, un pourcentage de 95% de bonne classification est réalisé. Pratiquement, le pourcentage obtenu n'a pas atteint le 100%. Cela est expliqué par le fait que certains objets mal classifiés sont certainement regroupés au début de la classification hiérarchique ascendante. Malgré ces cas inévitables la construction des classes à des niveaux prédécesseurs de 22 améliore considérablement le pourcentage de classification correcte. La généralisation de cette technique "fournir à l'utilisateur le double du nombre de classes d'objets qui existent en réalité" permet de simplifier considérablement le coût d'interaction manuelle effectuée par l'auteur de la vidéo hyperliée qui corrigera les résultats de classification. Il s'agit d'éditer par exemple $2m$ classes au lieu de $n \times m$ classes, si on suppose qu'il s'agit de classifier m objets suivis qui apparaissent n fois dans la séquence vidéo traitée.

Le choix interactif du nombre de classes ou du niveau de coupure de la hiérarchie se fait par l'auteur de la vidéo hyperliée, soit directement par précision du nombre désiré, soit par sélection d'un point de courbure sur la courbe des indices hiérarchiques (distance minimale entre groupes) aux itérations du CAH. La figure 5.13 illustre la courbe des indices hiérarchiques par rapport au nombre de classes, dans le cas où le critère du saut maximum est employé, et les paramètres sont gaussiens estimés dans les espaces de descripteurs RGB et HSV. En filtrant les points de courbures les plus significatifs pour ces courbes (par exemple 9, 7, etc), la tâche devient plus simple à l'utilisateur pour prendre sa décision. Dans le système développé pour ce travail, l'utilisateur pourra visualiser les classes d'objets à un niveau désiré de la hiérarchie, ensuite servir des outils de correction manuelle de résultats obtenus (voir section 5.3.3.4).

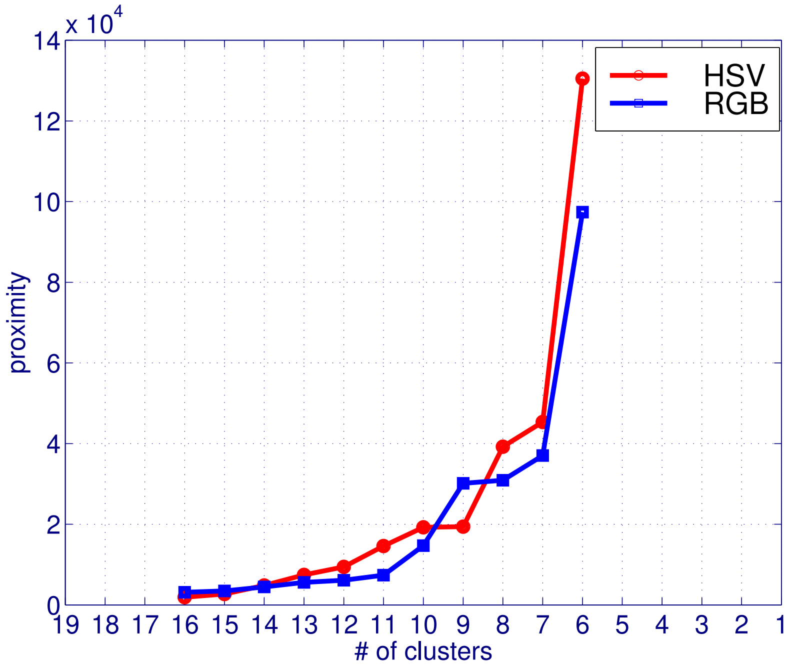


FIG. 5.13: *Illustration de l'indice de la hiérarchie (distance minimale entre groupes) par rapport au nombre de classes, dans le cas où le critère du saut maximum est employé, et les paramètres sont gaussiens estimés dans les espaces de descripteurs RGB et HSV.*

Comme nous avons mentionné dans la section 5.3.3.3, un problème du choix de la distance entre groupes se présente à l'application de la classification hiérarchique. Les expériences ci-dessus montrent que le critère de saut maximum fournit les meilleurs résultats. Par contre, les mauvais résultats sont obtenus quand le critère de saut minimum est employé. Un peu de chevauchement entre les groupes non homogènes à une itération i du CAH conduit à un fusionnement par le critère de saut minimum lors de la prochaine itération. Par contre, le critère du saut maximum retarde ce fusionnement le plus possible. Le critère de la distance moyenne un est compromis entre les deux sauts minimum et maximum et devra fournir des résultats meilleurs que les deux, mais expérimentalement ceci n'est pas vérifié. Une augmentation des pourcentages de bonne classification allant de 10% jusqu'à 20% est obtenue par l'adoption du critère du saut maximum.

La distance entre deux objets suivis est considérée comme la distance minimale entre leurs classes d'apparences intra-plan. Deux distances de Kullback et de Bhattacharyya ont été calculées entre ces classes d'apparences; chaque classe d'apparences est modélisée par une gaussienne. Ainsi, les tests montrent que la distance de Kullback fournit toujours des résultats de classification meilleurs que la distance de Bhattacharyya , allant de 3% jusqu'à 17%.

Mieux modéliser la variabilité intra-plan des objets suivis induit une meilleure classification. Cette règle est aussi validée dans la classification automatique quand la modélisation est effectuée dans les espaces de descripteurs RGB et I. Le tableau 5.2 résume les résultats de la classification automatique lorsque la variabilité des objets suivis est modélisée par une seule gaussienne. La comparaison de ces résultats avec les précédents (tableau 5.1) prouve une amélioration de 10% à 16% quand le mélange gaussien est adapté dans les espaces RGB et I. Par contre, les résultats sont presque similaires pour les autres espaces de descripteurs. Cette dernière situation reste pour nous inexpliquées.

Espace de descripteurs	d	d_E	Distance	$cc\%$ Critère du saut minimum	maximum	moyenne
h_{RGB}	64	10	d_B	44.83	55.17	55.17
h_{RGB}	64	10	d_K	44.83	62.07	62.07
h_I	32	5	d_B	48.28	55.17	55.17
h_I	32	5	d_K	41.38	55.17	51.72
h_{HSV}	64	10	d_B	51.72	62.07	65.52
h_{HSV}	64	10	d_K	51.72	62.07	62.07
h_{Hue}	32	5	d_B	51.72	58.02	62.07
h_{Hue}	32	5	d_K	55.17	62.07	62.07

TAB. 5.2: *Les pourcentages de classification correcte des d'objets suivis de la séquence Avengers-2, dans le cas où la distribution de chaque objet suivi est modélisée par une seule composante gaussienne.*

5.5 Conclusion

Dans ce chapitre nous avons décrit une approche de classification automatique des objets suivis. Le but d'une telle approche est d'aider au maximum l'auteur de la vidéo hyperliée à construire la structure de "groupes d'objets". Dans cette approche la classification ascendante hiérarchique est employée pour fabriquer des partitions imbriquées. D'un point de vue technique le choix de cette méthode est motivé par le fait que les classes d'objets fabriquées peuvent être examinées par l'auteur de la vidéo hyperliée à différents niveaux sans aucun coût de calcul.

Ici, les données à classifier ne sont pas des descripteurs classiques mais des densités de mélanges gaussiens. Les classes d'apparences intra-plan de chaque objet suivi sont d'abord recherchées dans les espaces de descripteurs de couleurs. Ensuite, les distances de Kullback et de Bhattacharyya sont calculées localement.

Les expérimentations ont été menées sur la séquence vidéo Avengers-2 dont le contenu est très complexe en termes de variation de l'apparence intra- et inter-plan des objets suivis. Sur une telle base d'objets bruités et pour un processus de classification automatique les résultats obtenus sont vraiment intéressants. Un pourcentage de bonne classification de 80% est atteint. Ces résultats sont jugés acceptables pour une initialisation qui serait reprise par l'auteur de la vidéo hyperliée. Le choix du nombre de classes est effectué par l'utilisateur d'une manière interactive. Aussi, des outils interactifs sont fournis par le système développé pour ce travail afin de permettre une édition des résultats. Ces expérimentations permettent de suggérer de procéder ainsi :

- Modéliser la variabilité intra-plan des objets suivis par un mélange gaussien;
- Utiliser la distance de Kullback qui est mieux adaptée que celle de Bhattacharyya à la comparaison des densités gaussiennes;
- Employer la distance de "saut maximum" dans l'application de la classification ascendante hiérarchique.

Les données classifiées ici sont très bruitées. On ne peut donc pas s'attendre à une classification automatique parfaite. Une approche de "relevance feedback" adoptée lors de la construction de la hiérarchie des classes d'objets, surtout les premiers niveaux, permettrait vraisemblablement d'améliorer la qualité de la classification.