

Classification par analyse des degrés de connexité d'ensembles couleur

Michèle Gouiffès , Bertrand Zavidovique

Institut d'Electronique Fondamentale, UMR 8622, Paris 11, Bat 220, 91405 Orsay cedex, France

michele.gouiffes@ief.u-psud.fr

Résumé – Cet article propose une méthode de classification basée sur l'analyse du degré de connexité couleur (DCC) dans un espace chrominance/luminance. Ce paramètre informe sur la probabilité qu'une couleur donnée soit connectée à un ensemble de couleurs similaires. De plus, les couleurs, après quantification, sont représentées dans un espace sphérique permettant une bonne distinction des couleurs intrinsèques du modèle dichromatique. La chrominance est analysée en priorité, puis la luminance est étudiée seulement si nécessaire. Finalement, la méthode dépend principalement d'un paramètre : le pas de quantification λ . Cette approche est évaluée en terme de qualité ($PSNR$ and CIE_{76}) et compacité (nombre de couleurs et régions finales) sur la base d'images Kodak. La représentation couleur proposée est comparée aux espaces Ohta et HSV et l'impact de λ est étudié.

Abstract – The paper presents an adaptive method for spatio-color classification based on the analysis of the *color connectedness degrees* (CCD) in a chrominance-luminance space. The CCD embeds information about the probability that a given color is connected (in the image) to a set of similar colors. The colors, after quantization, are converted to a spherical space, well-suited to reveal the body color vectors. The chrominance is analyzed first while the luminance is studied only when necessary. Eventually, the method depends mainly on one parameter : the quantization step λ . The method is evaluated quantitatively in terms of quality ($PSNR$ and CIE_{76}) and in terms of compactness (number of finals colors and regions) on the Kodak image database. Our color representation is compared to Ohta and HSV spaces, and the impact of the parameter λ is studied.

1 Introduction

En vision par ordinateur, nombreuses sont les applications requérant des prétraitements spécifiques tels que de la classification ou de la segmentation afin de réduire la quantité de données à traiter : en détection d'objets, en indexation [15], en suivi d'objets [12]. Partant généralement d'une distribution dans un histogramme, la classification consiste à déterminer les k classes de pixels les plus représentatives de l'image.

Les performances de la classification dépendent largement des choix effectués quant à l'espace couleur utilisé et le type d'histogramme. Les histogrammes couleur ont pour avantage d'être robustes à de faibles changements de point de vue ou d'apparence. Néanmoins, ils ne préservent pas les informations topologiques (connexité) ou spatiales, et peuvent s'avérer peu discriminants ou informatifs. Plusieurs méthodes élégantes incorporent les informations spatiales et colorimétriques dans une même structure de données [1, 15, 10], la façon la plus classique étant d'utiliser des caractéristiques statistiques du 2nd ordre [7]. Le Degré de Connexité Couleur (DCC) [5] a récemment été utilisé dans une procédure de segmentation sécurisant une bijection entre *cluster* couleur et région [9]. Cette méthode suppose k classes et détermine les intervalles RGB cubiques de DCC et d'homogénéité couleur maximales, c'est-à-dire les intervalles couleur les plus susceptibles de correspondre à de réelles régions dans l'image.

De façon similaire à [9], notre procédure de *clustering* repose sur le DCC. Nos contributions se portent à la fois sur la représentation couleur et la stratégie de classification :

1) Un espace couleur luminance-chrominance est proposé

pour séparer au mieux les couleurs intrinsèques des matériaux. En outre, il permet une bonne robustesse vis-à-vis des variations d'illumination tout en évitant des problèmes d'apparition de bruit comme dans l'espace HSV par exemple.

2) Chrominance et luminance sont analysées séparément dans le but de simplifier les structures de données mises en œuvre et réduire les temps d'exécution. La classification en chrominance est privilégiée et la luminance est traitée uniquement si nécessaire.

3) La forme des *clusters* couleur est adaptative, et n'est donc pas restreinte à une forme parallépipédique (3D) ou rectangulaire (2D).

4) La méthode n'est pas supervisée, *i.e.* le nombre de classes k n'est pas connu *a priori*.

L'article est organisé comme suit. La Section 2 explique les pré-requis sur lesquels reposent ces travaux, c'est-à-dire l'espace couleur puis la définition du degré de connexité d'un ensemble de couleurs. La méthode proposée fait l'objet de la Section 3. Enfin, la Section 4 montre la pertinence de l'approche par le biais de résultats sur des images réelles.

2 Quelques pré-requis

2.1 Le modèle dichromatique

Représentation couleur. Selon le modèle dichromatique [13] pour les objets lambertiens, les couleurs de la plupart des images se concentrent le long d'un nombre fini de segments dans l'es-

pace *RGB*, i.e le long de chaque *body vector*, allant de l'origine de l'espace à la couleur intrinsèque du matériau c_b . Dans ces travaux, le lien entre les coordonnées *RGB* notées c et le modèle dichromatique est facilité par une conversion dans un repère sphérique (ρ, θ, ϕ) , où ρ exprime la distance de l'origine $O = (0, 0, 0)$ à la couleur c (alors $\rho \simeq$ intensité). (θ, ϕ) exprime la chrominance, indépendamment de la luminance. Partant de cette représentation notée *TPR*, la méthode utilise le Degré de Connexité Couleur (DCC) pour former les *clusters*.

Degré de connexité couleur. Soit une image couleur $c = (c^1, c^2, c^3)$ et $c_i = (c_i^1, c_i^2, c_i^3)$ les composantes couleurs d'un pixel i au point p_i . Supposons que la chrominance est représentée par les composantes (c_i^1, c_i^2) , la luminance étant donnée par c^3 . Un intervalle de chrominance de taille λ^2 , d'origine c_i , est défini comme l'ensemble de *bins* couleur $I_i^\lambda = [c_i^1, c_i^1 + \lambda][c_i^2, c_i^2 + \lambda]$. La probabilité du 1^{er} ordre $\mathcal{P}_1(I_i^\lambda)$ est la probabilité qu'un pixel de couleur c_a appartienne à I_i^λ . Elle est calculée comme la somme de toutes les probabilités du er ordre \mathcal{P}_1 des composantes $c \in I_i^\lambda$. Ensuite, on définit la probabilité de co-occurrence de deux couleurs $\mathcal{P}_{cc}(c_a, c_b)$ comme la somme des probabilités d'occurrence de c_b dans un voisinage 8-connexe \mathcal{N} d'un pixel de couleur c_a . La probabilité du 2nd ordre $\mathcal{P}_2(I_i^\lambda)$ d'un intervalle I_i^λ est alors la somme des probabilités de co-occurrence de tout couple de couleurs $(c_a, c_b) \in I_i^\lambda$. Ainsi, le degré de connexité d'un intervalle couleur $\mathcal{D}(I_i^\lambda)$ est donné par [5]:

$$\mathcal{D}(I_i^\lambda) = \frac{\mathcal{P}_2(I_i^\lambda)}{\mathcal{P}_1(I_i^\lambda)} = \frac{\sum_{c_a \in I_i^\lambda} \sum_{c_b \in I_i^\lambda} \mathcal{P}_{cc}(c_a, c_b)}{\sum_{c_a \in I_i^\lambda} \mathcal{P}_1(c_a)} \quad (1)$$

Le DCC $\mathcal{D}(I_i^\lambda)$ est maximum lorsque I_i^λ correspond à une unique composante connexe dans l'image. Plus il y a de régions, plus \mathcal{D} est faible. Ainsi, contrairement au corrélogramme ou à l'histogramme couleur, une petite (mais éventuellement pertinente) région homogène peut obtenir un \mathcal{D} élevé malgré une faible probabilité d'occurrence \mathcal{P}_1 .

La procédure basée sur l'analyse du DCC dans *TPR* est détaillée dans la section qui suit.

3 Méthode de classification

Tout d'abord, l'image *RGB* est convertie vers *TPR*, et la saturation S de image est calculée. Ensuite, une quantification uniforme de la dynamique 2^M (M bits de codage, 8 ici) est effectuée sur N niveaux (intervalles de taille $\lambda = 2^M/N$). Chrominance et luminance sont ensuite analysées séparément et les résultats sont fusionnés.

3.1 Analyse de la chrominance

La classification de la chrominance est réalisée suivant quatre étapes :

1) Calcul des statistiques couleur et spatiales de chaque *bin* couleur \mathbf{b}_k de largeur λ (au total $N \times N$ *bins*) : $\mathcal{P}_1(\mathbf{b}_k)$, $\mathcal{P}_2(\mathbf{b}_k)$,

$\mathcal{D}(\mathbf{b}_k)$ et les coordonnées $x_c(\mathbf{b}_k)$ et $y_c(\mathbf{b}_k)$ du centre d'inertie p_c de l'ensemble des coordonnées des pixels de couleur \mathbf{b}_k .

2) Détection des *maxima* locaux de valeurs \mathcal{D}_λ . Un maximum local \mathcal{D}_m est détecté en un bin (\mathbf{b}_m) lorsque les deux critères \mathcal{C}_1 and \mathcal{C}_2 sont respectés :

$$\mathcal{C}_1 = \mathcal{D}(\mathbf{b}_m) > \mathcal{D}(\mathbf{b}_v) \quad \forall \mathbf{b}_v \in \mathcal{N}(\mathbf{b}_m) \quad (2)$$

$$\mathcal{C}_2 = \mathcal{D}(\mathbf{b}_m) > \mathcal{T}_1 \quad (3)$$

\mathbf{b}_v est un bin couleur appartenant au voisinage 8-connexe de \mathbf{b}_m (noté $\mathcal{N}(\mathbf{b}_m)$) dans l'espace des DCC. \mathcal{T}_1 est un seuil défini afin d'éviter des maxima locaux non significatifs : $\mathcal{T}_1 = 0.01 \times N_{pix}/(N \times N)$. Il est suffisamment faible pour ne pas être critique. N_{pix} est le nombre de pixels dans l'image. De plus, lorsque deux *maxima* locaux sont considérés trop proches (séparés par un *bin*), le maximum local de plus faible valeur est éliminé.

3) Définition des *clusters* couleur, c'est-à-dire des ensembles de *bins* couleur. À partir des *maxima* locaux, la procédure consiste à faire croître chaque *cluster* de manière à aboutir aux *clusters* de DCC maximum. Chaque *bin* est affecté à un *cluster* en se basant sur un critère dépendant conjointement de : la distance colorimétrique dans (θ, ϕ) ; le centre d'inertie dans l'image; la modification du DCC après affectation. Un bin \mathbf{b}_i est ajouté au mode qui maximise une mesure de similitude impliquant trois mesures différentes :

1. sa distance chromatique (distance euclidienne) au maximum local \mathbf{b}_m : $d^c(\mathbf{b}_i, \mathbf{b}_m)$
2. la distance spatiale (euclidienne) entre son centre d'inertie $p_c(\mathbf{b}_i)$ et celui du maximum local $p_c(\mathbf{b}_m)$: $d^s(\mathbf{b}_i, \mathbf{b}_m)$
3. le DCC de l'ensemble des couleurs incluses dans l'union de deux *bins* \mathbf{b}_i et \mathbf{b}_m : $d^D(\mathbf{b}_i, \mathbf{b}_m) = \mathcal{D}(\mathbf{b}_i \cup \mathbf{b}_m)$

Finalement la mesure de similarité s'écrit :

$$S(\mathbf{b}_i, \mathbf{b}_m) = \frac{d_{max}^c}{\max(d^c(\mathbf{b}_i, \mathbf{b}_m), \epsilon)} + \frac{d_{max}^s}{\max(d^s(\mathbf{b}_i, \mathbf{b}_m), \epsilon)} + \frac{d^D(\mathbf{b}_i \cup \mathbf{b}_m)}{d_{max}^D} \quad (4)$$

Notons que chaque distance est normalisée entre 0 et 1 par division par sa valeur maximale. ϵ est la plus petite valeur du critère, utile pour éviter un dénominateur nul.

4) Nettoyage des *clusters* couleur, de manière à éliminer les éventuels défauts de classification. En quelques mots, un *bin* est reclassé lorsque la majorité de ses voisins appartient à une classe différente de la sienne.

En guise d'exemple, la figure 1(a) montre l'image PERROQUETS de Kodak. La figure Fig.1(b) affiche les DCC de chrominance tandis que la figure 1(c) montre les maxima locaux. Finalement, les *clusters* obtenus sont représentés sur la figure 1(d).

3.2 Analyse de la luminance ρ

Pareillement à [2], nous faisons l'hypothèse que l'information de luminance est utile uniquement dans le cas où les couleurs ne sont pas suffisamment saturées. Ainsi, lorsque la saturation moyenne d'un *cluster* couleur a une valeur inférieure à

un certain seuil, le *cluster* est déclarée non-saturée et doit être divisée. Le seuil est choisi à 20% comme dans [14] si bien que les *clusters* faiblement saturés sont également fractionnés.

Ensuite, la classe est divisée uniquement lorsque sa population est suffisamment répartie dans l’histogramme. L’énergie est choisie comme critère de division, elle est en effet maximale pour une région parfaitement homogène. Ainsi, lorsque l’énergie $E(\rho)$ dans le *cluster* couleur est inférieure à un seuil (0.8), il est réparti en *clusters gris* suivant la même procédure que précédemment excepté que: 1) les données sont monodimensionnelles; 2) seuls les pixels peu saturés ($S < 20\%$) sont traités.

Fig.1(e) montre la classification finale chrominance/luminance avec $N = 64$. Notons que certains *clusters* couleur n’ont pas été fractionnés en *clusters gris*: ceux de forte saturation et ceux dont la luminance est homogène (haute énergie).

Dans le but d’une amélioration des résultats, une classification Markovienne¹ (Fig.1 (f)) peut être préférée à la précédente classification.

4 Expérimentations

Les premières expérimentations se basent sur 24 images de la base de données Kodak² qui est un ensemble varié d’images, composées de scènes naturelles en extérieur, de personnes, d’objets manufacturés, de bâtiments.

La classification est supposée satisfaisante lorsqu’elle fournit un bon compromis entre une forte réduction de couleurs et/ou régions dans l’image (compacité) tout en préservant l’information (qualité objective et subjective). Partant de ce postulat, quatre critères de comparaison sont utilisés: le *PSNR* (Peak Signal to Noise Ratio) calculé entre l’image résultat et l’image initiale; la distance *CIE76* calculée dans l’espace perceptuellement uniforme *CIELAB*; le nombre de classes N_c et de régions N_r (en 8-connexité). Tout d’abord, l’espace couleur pris en compte est *TPR*, le pas de quantification est $\lambda = 4$ ($N = 64$), les couleurs sont classées selon le critère (4) (noté $D + C + S$, pour **D**egré, **C**ouleur et **S**patiale). Le tableau 1 regroupe les résultats de chaque expérience (moyenne du critère calculée sur la base d’images Kodak) pour trois études différentes visant à évaluer la qualité et la compacité des résultats. De plus, chaque couple de résultats (*PSNR*, Nb. classes) pour les 3 expériences sont représentées sur les figures 2(a) à 2(c). Ces représentations 2D sont utiles pour visualiser et comparer les résultats en un coup d’œil, pour chaque paramétrisation de la méthode. Étant donné que la meilleure classification est associée à un faible nombre de classes et un fort *PSNR*, les résultats jugés bons sont situés en bas à droite de la figure. Ces résultats nous renseignent sur:

(1) l’impact de la quantification (tableau 1 et figure 2(a)). Finalement, $N = 64$ (1^{ère} entrée, tableau 1) est un paramètre adéquat, permettant un bon compromis entre qualité et compa-

TAB. 1: Résultats des critères d’évaluation. (1) Critère $C + S + D$ dans *TPR*. (2) Critère $C + S + D$ avec $N = 64$. (3) *TPR*, $N = 64$.

		<i>PSNR</i>	<i>CIE76</i>	N_c	N_r
(1)	$N = 64$	20,19	10,48	22,38	3550,08
(1)	$N = 32$	18,02	14,31	5,83	801,17
(1)	$N = 96$	20,68	9,43	51,79	7789,58
(2)	<i>HSV</i>	19,59	10,47	92,83	13754,21
(2)	$I_1 I_2 I_3$	18,69	13,67	16,13	2423,13
(3)	$C + S$	20,09	10,74	32,31	4819

cité. L’utilisation d’un N plus élevé implique une forte consommation en mémoire et en temps.

(2) l’impact de l’espace couleur (tableau 1 et figure 2(b)). Notre approche peut s’appliquer à d’autres espaces chrominance/luminance, bien que tous les espaces ne permettent pas de révéler aisément les couleurs intrinsèques. À titre de comparaison, les espaces *HSV* et *Otha* (composantes indépendantes $I_1 I_2 I_3$) sont testés. Avec *HSV*, le nombre de classes est élevé pour un faible *PSNR* et un fort *CIE76*. Cet espace couleur est générateur de bruit, la teinte étant mal définie pour une saturation faible. L’utilisation de $I_1 I_2 I_3$ produit une sous-segmentation (faible nombre de régions mais grande réduction du *PSNR*). En comparaison, *TPR* offre une bonne qualité avec un nombre raisonnable de classes.

(3) l’impact du DCC (tableau 1 et figure 2(c)). Nous comparons l’utilisation du critère $D + C + S$ pour la création des *clusters* (1^{ère} entrée et eq.(4)) avec ce même critère sans utilisation du DCC ($C + S$) et en détectant les maxima locaux dans les histogrammes classiques. Le *PSNR* moyen obtenu avec $C + S$ n’est pas significativement plus élevé qu’avec $D + C + S$ mais le nombre de classes est beaucoup plus important. En effet, le DCC contribue à croître la taille des *clusters* couleur. Il favorise les ensembles de couleurs larges et significatifs et isole les ensembles peu pertinents associés à des pixels éparpillés dans l’image. La similarité basée sur une simple distance couleur donne une importance équivalente à tous les *clusters*, indépendamment de leur pertinence en termes de connexité dans l’image.

Pour finir, la figure 3 montre quelques résultats supplémentaires.

5 Conclusion

Partant du degré de connexité d’ensembles de couleurs, une nouvelle méthode de classification spatio-colorimétrique a été proposée. La représentation couleur utilisée est inspirée du modèle dichromatique afin d’exhiber aisément les couleurs intrinsèques. L’analyse est réalisée successivement sur la chrominance puis sur la luminance. En effet, seuls les ensembles de couleurs peu saturés sont divisés en *clusters gris* si nécessaire. Comparé aux approches précédentes basées sur le DCC, la forme des ensembles couleur est adaptative et la procédure est non-supervisée. Le paramètre le plus critique de l’approche est le pas de quantification λ . Toutefois la valeur $\lambda = 4$ abou-

1. Cette étape n’est pas expliquée ici par souci de concision, ce n’est pas le cœur de la contribution.

2. <http://r0k.us/graphics/kodak/>

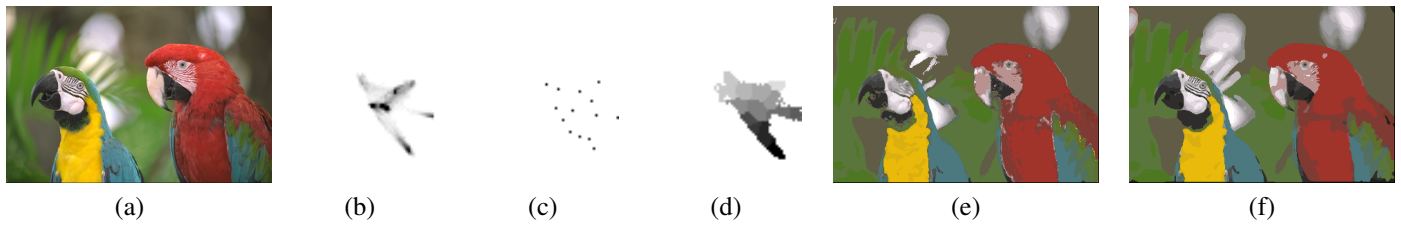


FIG. 1: (a) Image PERROQUETS (Kodak). (b) Espace des DCC \mathcal{D} . (c) Maxima locaux. (d) Intervalles couleur. (e) Classification finale (PSNR = 18,98 dB, CIE76=11,76, 27 classes, 4597 régions). (f) Classification Markovienne.

tit à des résultats satisfaisants sur l'ensemble des images traitées. Finalement, les résultats montrent la pertinence de l'espace TPR comparé aux espaces classiques tels que HSV et $I_1 I_2 I_3$.

Références

- [1] Birchfield, S. and Rangarajan, S. (2005). Spatiograms versus histograms for region-based tracking. In *CVPR*, pages 1158–1163.
- [2] Carron, T. and Lambert, P. (1994). Color edge detector using jointly hue, saturation and intensity. In *IEEE ICIP*, volume 3, pages 977–981.
- [3] Cheng, H., Chen, C., Chiu, H., and Xu, H. J. (1998). Fuzzy homogeneity approach to multilevel thresholding. *IEEE Trans. on Image Processing*, 7(7):1–5.
- [4] Davis, L., Johns, S., and Aggarwal, J. K. (1979). Texture analysis using generalized co-occurrence matrices. *IEEE Trans. PAMI*, 1:251–259.
- [5] Fontaine, M., Macaire, L., and Postaire, J.-G. (2000). Unsupervised segmentation based on connectivity analysis. In *ICPR*, volume 1, pages 600–603, Spain.
- [6] Gevers, T. and Smeulders, A. W. M. (1999). Colour based object recognition. *Pattern Recognition*, 32(3):453–464.
- [7] Haralick, R., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans. on Systems, Man & Cybernetics*, 3(6):610–621.
- [8] Huang, J., Kumar, S., Mitra, N., Zhu, W., and Zabih, R. (1997). Image indexing using color correlograms. In *CVPR*, pages 762–768, Puerto Rico.
- [9] Macaire, L., Vandenbroucke, N., and Postaire, J.-G. (2006). Color image segmentation by analysis of subset connectedness and color homogeneity properties. *CVIU*, 102(2006):105–116.
- [10] Nezamabadi-Pour, H. and Kabir, E. (2004). Image retrieval using histograms of uni-color and bi-color blocks and directional changes in intensity gradient. *Pattern Recognition Letters*, 25(14):1547–1557.
- [11] Ohta, Y., Kanade, T., and Sakai, T. (1980). Color information for region segmentation. *Comp. Graph. and Im. Proc.*, 13(2):222–241.
- [12] Peihua, L. (2006). A clustering-based color model and integral images for fast object tracking. *Signal Processing: Image Communication*, 21(8):676–687.
- [13] Shafer, S. (1985). Using color to separate reflection components. *Color Research and Applications*, 10(4):210–218.
- [14] Sural, S., Qian, G., and Pramanik, S. Segmentation and histogram generation using the hsv color space for image retrieval. In *IEEE ICIP*.
- [15] Wada, N., Kaneko, S., and Takeguchi, T. (2006). Using color reach histogram for object search in colour and/or depth scene. *Pattern recognition*, 39(5):881–888.

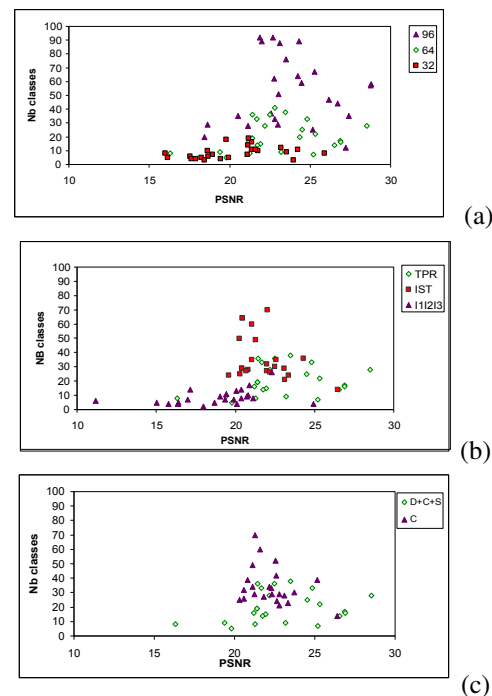


FIG. 2: Comparaison des résultats obtenus sur la base de données Kodak. (a) Impact de N . (b) Impact de la représentation couleur. (c) Impact du DCC.

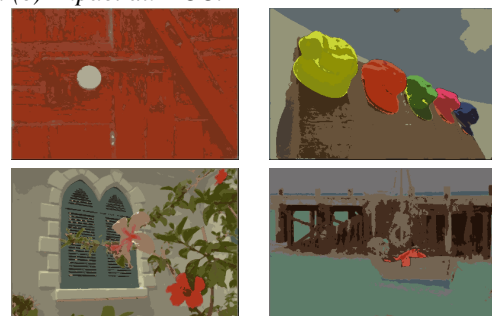


FIG. 3: Exemples de classification avec $N = 64$ dans TPR , avec utilisation du DCC. De gauche à droite: Kodim 2, 3, 7 and 11.