

La couleur en vision par ordinateur : une revue

Color in computer vision : a review



Quang-Tuan LUONG

INRIA Sophia-Antipolis,
2004, route des Lucioles,
06561 Valbonne Cedex, France

Quang-Tuan Luong est né à Paris le 28 mai 1964. Ancien élève de l'École Polytechnique (promotion 1984), puis du Magistère Mathématiques et Informatique de l'École Normale Supérieure, il travaille actuellement à l'Institut National de Recherche en Informatique et Automatique (INRIA), unité de recherche de Sophia-Antipolis, au sein du projet de Robotique et Vision par ordinateur sous la direction d'Olivier Faugeras. Ses recherches portent sur la couleur et l'emploi de méthodes actives.

RÉSUMÉ

L'utilisation de la couleur en vision par ordinateur est un sujet de recherche qui suscite un intérêt croissant. Ce papier fait le point dans ce domaine, en essayant de répondre aux questions : *Qu'est-ce que la couleur ? Quelles en sont les représentations adéquates ? Comment la déterminer ? Que peut-on en faire ?* Pour cela, nous faisons une revue approfondie et très à jour de l'ensemble de la littérature consacrée à ce sujet en cernant les axes de recherche et les problématiques importantes et en tentant de les évaluer.

MOTS CLÉS

Couleur, vision par ordinateur, modélisation, réflectance, colorimétrie, constance des couleurs, détection de contours, segmentation, stéréo, propriétés intrinsèques.

ABSTRACT

The use of colour in computer vision has received growing attention. This paper gives the state-of-art in this subfield, and tries to answer the questions : What is color ? Which are the adequate representations ? How to compute it ? What can be done, using it ? Towards that goal, we make a deep and up-to-date review of the existing literature on this subject, we outline the important research directions and issues, and we attempt to evaluate them.

KEY WORDS

Color, computer vision, modelisation, reflectance, colorimetry, color constancy, edge detection, segmentation, stereo, intrinsic properties.

1. Introduction

L'utilisation de la couleur en vision par ordinateur est une idée naturelle. Notre propre perception visuelle est colorée et a été étudiée en tant que telle depuis très longtemps, ce qui a permis de dégager les lois colorimétriques. La couleur a été employée en vision par ordinateur assez tôt, les premières approches étant issues de la colorimétrie. La disponibilité de caméras couleur, et surtout la récente formalisation des problèmes de réflectance ont donné un vif regain d'intérêt à la recherche sur la couleur en vision par ordinateur.

Cet article se propose de passer en revue l'essentiel de la littérature consacrée à ce sujet, et tente de cerner les axes de recherche et les problématiques qu'il a suscités, en essayant de donner notre avis sur l'intérêt des divers travaux. Il ne traite que des sujets qui intéressent directement la vision, et qui font l'objet d'une formulation (mathématique) suffisamment précise pour pouvoir donner lieu à une utilisation informatique. La couleur a bien sûr aussi été employée pour enrichir des descriptions sémantiques, et son intérêt est évident à ce niveau élevé, mais les questions que nous nous posons portent sur son rôle à « bas niveau », où il n'est pas si évident et où intervient tout le traitement relatif à l'information couleur !

En effet, dans les utilisations⁽¹⁾ à « haut niveau » la couleur n'intervient plus que comme un triplet (voire une paire) de nombres. Néanmoins se posent, pour ce type d'application, des problèmes liés à l'acquisition et à la représentation, la robustesse étant la principale qualité recherchée. Dans cet article centré sur les aspects formels et algorithmiques, nous évoquerons néanmoins brièvement les problèmes à bas niveau (acquisition et prétraitement des images couleur) et quelques résultats psychologiques quant à la vision des couleurs, dans la mesure où ils constituent un préalable indispensable. En revanche on n'y trouvera aucun résultat de recherche nouveau. Les travaux connexionnistes de l'équipe de Grossberg [54, 55, 56] ont mené à une théorie algorithmique de la vision tout à fait intéressante, mais ne sont pas discutés ici à cause de la difficulté que nous avons à évaluer leurs algorithmes.

Plusieurs articles de synthèse, qui couvrent assez partiellement les sujets traités ici, ont été publiés. Il s'agit de ceux de Brown [13] (contient une collection fort précise de définitions et de données de base, surtout sur la photométrie et la colorimétrie), Gershon [48] (un aperçu présenté de manière intéressante des théories et des faits les motivant, avec de bonnes présentations de la colorimétrie, des résultats psychophysiques et neurophysiologiques, et de la recherche en vision par ordinateur en son temps), Connelly [21] (des idées à l'état d'ébauches sur l'utilisation de modèles de réflectance dans les environnements naturels), et Forsyth [41] (sur la constance des couleurs, remarquable). Les livres classiques sur la vision par ordinateur de Ballard et Brown [3], le traitement d'images de Pratt [105], et la synthèse d'images de Foley et Van Dam [36] contiennent aussi une présentation concise des notions de base.

L'organisation des trois parties de cet article est la suivante : la première partie présente d'abord l'état de l'art dans le domaine de la modélisation. Cette partie tente de répondre aux questions : *Qu'est-ce que la couleur ? Quelles en sont les représentations adéquates ?* Elle sert de base pour les deux suivantes, qui sont en quelque sorte complémentaires. L'une traite du problème de déterminer les propriétés spectrales de surface (Constance des couleurs), ce qui revient à calculer la couleur, l'autre de l'emploi de la couleur pour diverses tâches de vision effectuées classiquement avec des images monochromes. Ces parties essaient d'apporter une réponse aux questions : *Comment déterminer la couleur ? Que peut-on en faire ?*

2. Modélisation

Ce chapitre présente les différents outils et modèles disponibles pour décrire le processus de formation des images couleur. La première section expose les notions photométriques et colorimétriques classiques. Le regain

d'intérêt pour la couleur ces dernières années est centré autour de deux nouveaux types de modèles qui ont été à la source d'algorithmes de surbrillance et de constance des couleurs. La seconde section donne une présentation de ces modèles.

2.1. NOTIONS DE BASE

Cette partie expose concisément quelques notions de base, pleinement développées dans les sommes de Wyszecki et Stiles [127] et Judd et Wyszecki [75] qui constituent toujours des ouvrages de référence. Le second, plus récent, traite de sujets un peu plus variés, et sous un angle plus pratique mais certains points sont davantage développés dans le premier, qui contient aussi davantage de données. Pour ce qui concerne la photométrie, d'excellents chapitres peuvent être trouvés dans le livre de Horn [65].

2.1.1. Mesure de la lumière

Cette brève section clarifie, au moyen de quelques définitions la signification du terme « intensité », utilisé aussi bien pour désigner des valeurs relatives à une source de lumière que celles données par les capteurs de vision.

Les mesures de la lumière peuvent être effectuées dans deux types d'unités :

- Les unités radiométriques, relatives aux mesures physiques d'énergie.
- Les unités photométriques, relatives aux mesures psychophysiques de luminance.

Le lien entre les deux est $u_p = K_m V(\lambda) \cdot u_r$, où $K_m = 679,6$ lumens/watt et $V(\lambda)$ est la fonction de sensibilité relative photopique donnant la sensibilité aux différentes longueurs d'ondes des photorécepteurs humains pour la vision diurne (voir plus bas sur la figure 10, la fonction y du système CIE, qui est exactement identique à $V(\lambda)$).

Les trois mesures les plus importantes sont :

- L'énergie W émise, transférée, ou reçue sous forme de radiation

— *flux radiant* :

$$\Phi = dW/dt, \text{ en watt (flux luminant en lumens) .}$$

- L'énergie atteignant une surface en un point donné
- *irradiance* :

$$I = d\Phi/dA, \text{ en watt par mètre carré}$$

(*illuminance* en lumens par mètre carré).

- Le flux radiant Φ émergeant d'une surface dans une certaine direction

— *radiance* :

$$L = d^2\Phi/(d\omega dA \cos \theta), \text{ en watt par stéradian par m}^2$$

(*luminance*, en lux).

Si ces quantités sont spécifiées en fonction de la longueur d'onde, l'adjectif *spectral* précède ces termes et les unités deviennent *par nanomètre*.

La relation entre l'intensité I mesurée par un système de

(1) On pourra consulter divers travaux sur la reconnaissance dans les images aériennes [1], [96], les scènes industrielles [2], [20], [52] ou d'extérieur [87], [106] et, pour le suivi de routes à [25], [120], [116], [82].

formation d'images (*Irradiance d'image*) fondé sur une lentille et la radiance de la scène est :

$$I = \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \alpha L$$

où d est le diamètre de la lentille, f sa distance focale, et α l'angle entre l'axe optique et la ligne d'observation. Il est important de noter que sous l'hypothèse d'une projection orthographique, $\alpha = 0$, donc l'intensité mesurée est proportionnelle à la quantité physiquement significative, la radiance.

2.1.2. L'équation d'intensité

Modèle photométrique. Les radiations que l'on détecte peuvent provenir soit directement de l'objet (sources de lumière), soit d'une transmission (objets transparents ou translucides), soit d'une réflexion (objets opaques). Nous ne nous intéressons qu'au troisième cas, le modèle de réflexion le plus juste est celui utilisant la fonction de distribution bidirectionnelle de la réflectance spectrale (BSRDF), qui indique comment une surface apparaît, vue sous une certaine direction lorsqu'elle est illuminée par une autre direction. Plus précisément cette fonction est donnée par :

$$(1) \quad f_r(\theta_i, \phi_i, \theta_e, \phi_e, \lambda) = dL(\theta_e, \phi_e, \lambda) / dE(\theta_i, \phi_i, \lambda)$$

où dE est l'irradiance spectrale reçue dans la direction d'incidence (θ_i, ϕ_i) et dL est la radiance spectrale dans la direction d'observation (θ_e, ϕ_e) , (θ, ϕ) étant un système polaire-azimut local au point de réflexion (fig. 1). Si la surface n'est pas anisotrope, f_r dépend de $\phi_e - \phi_i$ au lieu de (ϕ_e, ϕ_i) . La radiance spectrale réfléchie sous une

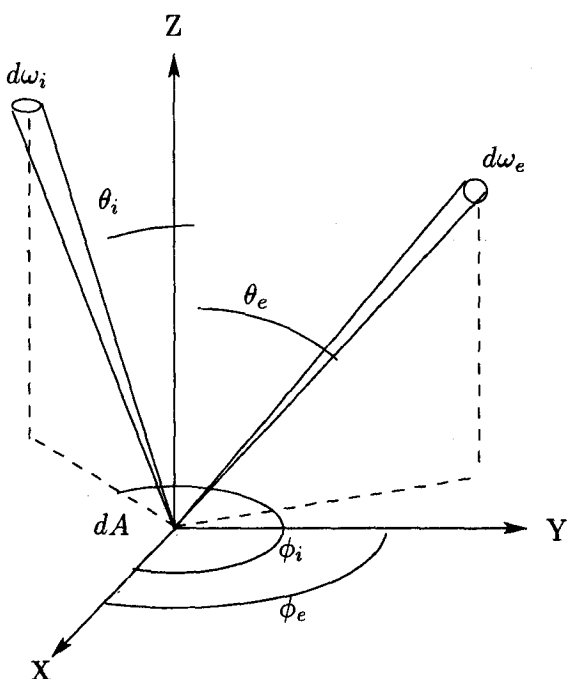


Fig. 1. — Angles photométriques locaux dans le cas général.

illumination de radiance spectrale L_i , d'angle solide du cône de lumière $d\omega_i$ est alors :

$$(2) \quad L_r = \int_{\omega_i} f_r(\theta_i, \phi_i, \theta_e, \lambda) L_i(\theta_i, \phi_i, \lambda) \cos \theta_i d\omega_i.$$

Ce modèle permet de traiter exactement des situations complexes, dont les plus intéressantes sont le cas de sources de lumière étendues et celui de l'interréflexion, mais le nombre élevé de paramètres le rend lourd pour les applications courantes.

L'équation simplifiée. L'équation d'intensité considérée pour pratiquement toutes les applications de vision par ordinateur suppose une source de lumière ponctuelle, et une surface non anisotrope et est de la forme :

$$(3) \quad I(i, e, g, \lambda, \vec{r}) = R(i, e, g, \lambda, \vec{r}) E(\lambda, \vec{r})$$

où (i, e, g) sont les angles d'incidence, de vue, et de phase (fig. 2), et \vec{r} est la position du point observé (il y a dépendance de ce point par l'intermédiaire de la composition matérielle de la surface qui s'y trouve). $E(\lambda)$ est la distribution spectrale d'énergie de l'illuminant et $R(i, e, g, \lambda, \vec{r})$ la fonction de réflectance de l'objet. C'est cette dernière fonction qui est corrélée avec la couleur observée.

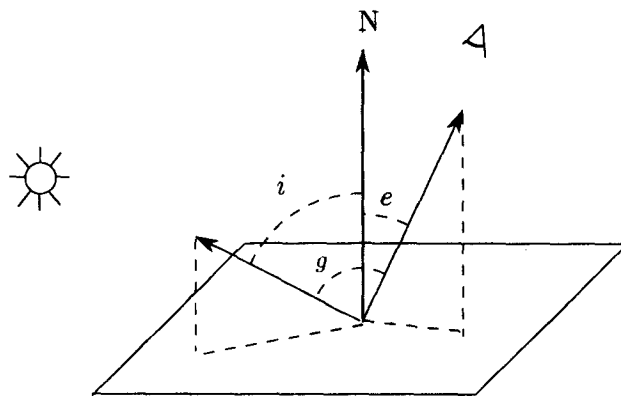


Fig. 2. — Angles photométriques, dans le cas isotrope.

La plupart des travaux ont reposé sur une équation d'intensité dans laquelle facteurs spatiaux et spectraux sont factorisés :

$$(4) \quad I(\lambda, \vec{r}) = \rho(\lambda, \vec{r}) F(i, e, g) E(\lambda, \vec{r})$$

- $\rho(\lambda, \vec{r})$, l'albédo, ne dépend que des propriétés matérielles de la surface au point de réflexion. Si la surface est peinte en blanc, $\rho(\lambda, \vec{r}) = 1$.
- $F(i, e, g)$ est un facteur géométrique tenant compte des directions d'illumination et d'observation. Si la surface est normale à l'illuminant et à la direction d'observation, $F(i, e, g) = 1$.
- $I(\lambda, \vec{r})$ est la distribution spectrale d'énergie de l'illuminant dans l'espace. Si la lumière émise par l'illuminant est blanche, $I(\lambda, \vec{r}) = 1$ aux points sans ombre portée.

L'importance de cette équation me semble grande car la plupart des travaux actuels, bien qu'ils invoquent des modèles plus réalistes (NIR et unichromatique), l'utilisent en dernier ressort, après une série d'hypothèses simplificatrices. Nous verrons plus loin que ceci ne va pas de soi.

2.1.3. Le signal couleur

Dans cette section, nous décrivons le signal couleur en termes physiques faisant intervenir les notions de longueur d'onde et d'énergie, puis nous évoquons, de manière très résumée, les problèmes d'acquisition.

Aspects photométriques : les mesures des senseurs. Le signal perçu par le système visuel, en tant qu'onde électromagnétique, est entièrement caractérisé par sa distribution spectrale d'énergie, une fonction $s(\lambda)$, les longueurs d'onde λ visibles étant celles comprises entre 400 nm et 700 nm. La lumière ordinaire est en général composée d'un continuum d'ondes monochromatiques, mises en évidence par l'expérience du prisme. Pour les ondes monochromatiques, la couleur est une fonction continue de la longueur d'onde.

Le signal couleur $s(\lambda)$ n'est pas représenté par les systèmes de traitement de l'information visuelle. Les senseurs n'ont en effet pas directement accès au signal couleur (il faut des spectrophotomètres coûteux, lents et fragiles), mais seulement à des scalaires du type :

$$(5) \quad s_i = \int_{\lambda_1}^{\lambda_2} s(\lambda) r_i(\lambda) d\lambda$$

où $r_i(\lambda)$ est la réponse (ou sensibilité) spectrale du i -ième senseur, une fonction prenant ses valeurs entre 0 et 1. Aussi bien pour les senseurs humains que pour les senseurs artificiels couramment utilisés, la bande passante de la courbe de réponse est trop large pour que l'on puisse considérer que les s_i soient des valeurs de $s(\lambda)$ à des longueurs d'onde λ_i de réponse maximale des senseurs. En revanche, la réponse d'un senseur de sensibilité $r(\lambda)$ à un signal monochromatique de longueur d'onde λ_0 est simplement $r(\lambda_0)$.

L'existence de trois senseurs permet d'effectuer des discriminations entre des signaux monochromatiques de longueur d'onde distincte. Voyons tout de suite une première justification du nombre de trois (2). Considérons d'abord un système possédant un seul type de senseur, et deux signaux monochromatiques, de longueur d'onde λ_1 et λ_2 . L'effet de l'absorption de photons étant indépendant de leur longueur d'onde, le senseur répondra de manière identique au signal d'intensité 1 et de longueur d'onde λ_1 , et au signal d'intensité $\frac{r(\lambda_1)}{r(\lambda_2)}$ et de longueur d'onde

λ_2 . Dans un système à deux types de senseurs, ce type de confusion ne pourra plus se produire qu'entre des longueurs d'onde λ_1 et λ_2 telles que $\frac{r_1(\lambda_1)}{r_1(\lambda_2)} = \frac{r_2(\lambda_1)}{r_2(\lambda_2)}$. S'il

(2) La justification la plus probante sera exposée plus bas, à partir d'arguments psychophysiques provenant de l'étude de la perception des couleurs.

existe une longueur d'onde λ_0 pour laquelle $r_1(\lambda_0) = r_2(\lambda_0)$, le signal monochromatique de longueur d'onde λ_0 ne peut être distingué du signal achromatique qui excite également les deux senseurs. Ces problèmes ne se posent plus avec trois senseurs indépendants. Remarquons que l'espace des signaux couleur étant de dimension infinie et les senseurs étant en nombre fini, il y aura un nombre infini de signaux généraux (non monochromatiques) aboutissant à un même triplet de mesures donné. La relation (5) entre les mesures des senseurs et le signal couleur permet cependant de contraindre celui-ci. Si on suppose de plus, qu'il peut être décrit par un petit nombre de paramètres, ces relations peuvent permettre de le déterminer, comme nous le verrons plus bas.

Acquisition. Le système visuel humain utilise trois types de cônes, dont la sensibilité spectrale est représentée figure 3, pour la vision diurne (3). Ces cônes sont sensibles respectivement aux longueurs d'onde courte (bleu), moyenne (vert) et longue (rouge), ce qui constitue une décomposition naturelle du spectre, assurant ainsi simplement l'unicité de la décomposition des couleurs pures. En vision par ordinateur, on fait appel à un jeu de filtres colorés, choisis pour décomposer les longueurs d'onde visibles (4) au mieux (au sens où leurs ensemble-supports constituent une partition de l'intervalle des longueurs d'onde visibles), mis en œuvre par différents procédés.

- Utilisation de caméras noir et blanc avec des filtres optiques colorés, de type Kodak Wratten (voir les sensibilités fig. 3). Ce procédé a été très utilisé et permet de bénéficier de la bonne définition et de la bonne dynamique des caméras noir et blanc, en revanche les manipulations sont lourdes et des corrections sont nécessaires du fait de la différence de densité des filtres.

- Caméras couleur mono-CCD, dont le principe est de superposer un masque de filtres sur le CCD, de manière à extraire les trois composantes spectrales. Du fait de l'absence d'éléments mécaniques, ces capteurs sont bon marché, mais chaque capteur CCD étant utilisé pour donner trois valeurs, on perd beaucoup de dynamique, surtout sur les canaux rouge et bleu.

- Caméras couleur tri-CCD, dans lesquelles un prisme permet tout d'abord de dupliquer en trois l'image issue de l'objectif. Chaque image est ensuite filtrée indépendamment avant d'être saisie par un capteur CCD. Les résultats obtenus avec ces caméras sont bons, mais ce matériel est rendu coûteux par la précision mécanique nécessaire pour obtenir une coïncidence pixel à pixel des trois plans couleur (5).

L'acquisition d'une image couleur consiste donc à obtenir

(3) La vision nocturne est assurée par un seul type de récepteur, les bâtonnets, et est donc, d'après la remarque faite plus haut, achromatique.

(4) Il est nécessaire d'utiliser un filtre pour arrêter le rayonnement infrarouge, les récepteurs CCD y sont en effet très sensibles et certains filtres comme les Wratten les transmettent si bien qu'à la fin chaque canal verrait sans cela sa valeur augmenter d'une quantité qui entraînerait une désaturation des couleurs pouvant pratiquement être complète avec un éclairage à incandescence puissant.

(5) Les caméras de télévision sont aujourd'hui des tri-CCD dont un des plans (souvent le vert) est décalé d'un demi pixel, ce qui évite les problèmes d'aliassage.

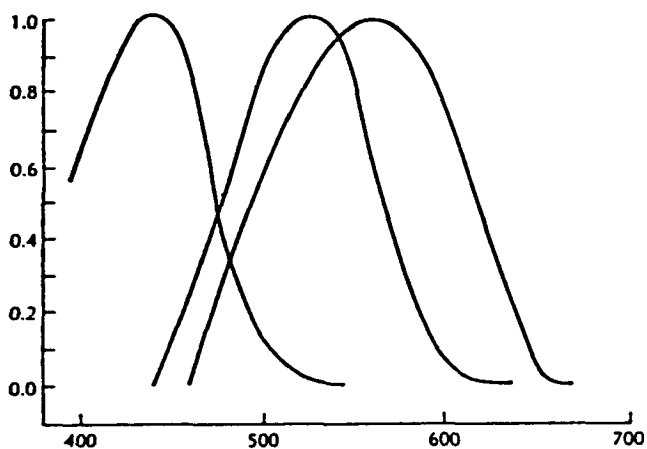
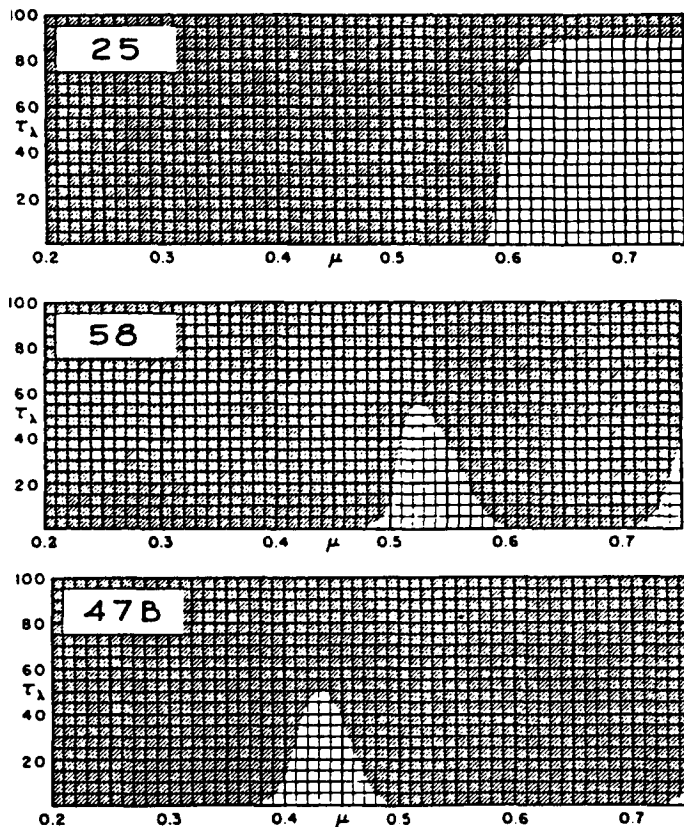


Fig. 3. — Partie gauche : Sensibilités des photorécepteurs de la rétine humaine. Partie droite : Courbes de transmission des filtres Wratten couramment utilisés.



trois images représentant les composantes rouge, verte, et bleue de la scène observée. On a vu que le triplet (r, v, b) permet de reconstituer chaque couleur monochromatique sans ambiguïté, et on peut déjà parler en ce sens de la base (R, V, B) des capteurs, chaque primaire R, V et B étant définie par sa fonction de sensibilité spectrale. La superposition des trois composantes redonne l'image colorée initiale.

Cependant la simple utilisation de la méthode décrite peut conduire, avec certaines caméras, à des couleurs visiblement inexactes et à des données couleur peu exploitables par les algorithmes faisant intervenir des propriétés photométriques de scènes. Les caméras actuelles ne sont en effet pas photométriquement exactes :

- L'intensité mesurée par les caméras dans chaque bande de longueur d'onde n'est pas une fonction linéaire de l'intensité réelle mais une fonction concave de type logarithmique (ceci est dû à la « gamma-correction »). En pratique, l'effet de la non-linéarité est que les couleurs apparaissent moins saturées et que le calcul des coordonnées chromatiques donne des résultats parfois complètement inexacts.
- La densité des filtres n'est pas forcément identique et la réponse de la caméra n'est pas uniforme suivant les longueurs d'onde utilisés et la réponse spectrale de la caméra. Lorsque l'on effectue l'acquisition des trois images dans des conditions identiques, les histogrammes des intensités dans les trois bandes de longueur d'onde n'ont pas la même étendue, ce qui devrait pourtant être le cas

sous un éclairage blanc. Les valeurs mesurées pour le blanc et les gris ne sont pas égales dans les différentes bandes de longueur d'onde. La conséquence de ceci est un déséquilibre chromatique.

Un prétraitement comportant une procédure de linéarisation (consistant par exemple à calculer une table « look-up » attribuant à chaque valeur numérisée une quantité proportionnelle à l'intensité réelle qui a donné cette valeur) et une procédure de balance des blancs (consistant à effectuer un changement d'échelle, pour lequel on peut jouer à différents niveaux, sur les vecteurs couleur pour qu'ils « correspondent » au blanc, c'est-à-dire que les coordonnées finales du blanc soient égales dans chaque canal, et à la valeur maximale) est nécessaire pour les algorithmes faisant intervenir des propriétés photométriques de scènes.

2.1.4. Aspects physiologiques et psychologiques de la perception de la couleur

La base (R, V, B) des senseurs est éloignée de l'expérience humaine des couleurs. Or la notion de couleur n'existe pas dans l'absolu et est étroitement liée à notre perception. Nous décrivons dans cette section deux autres bases qui nous permettront de mieux rendre compte de celle-ci. Toutes utilisent les valeurs $r = \int_{\lambda_1}^{\lambda_2} s(\lambda) R(\lambda) d\lambda$,

$$g = \int_{\lambda_1}^{\lambda_2} s(\lambda) V(\lambda) d\lambda, \quad b = \int_{\lambda_1}^{\lambda_2} s(\lambda) B(\lambda) d\lambda.$$

Couleurs antagonistes. De multiples phénomènes perceptuels n'ont pas trouvé d'explication satisfaisante dans le cadre d'une théorie trichromatique : les faits qu'après avoir fixé un carré jaune, on voit un carré bleu sur une feuille blanche (images induites), qu'il faille quatre noms pour désigner toutes les couleurs, que des mélanges tels que bleu-jaune et rouge-vert, ne soient jamais désignés.

Le modèle des couleurs antagonistes explique les appariements précédents en considérant que les sorties des canaux R, V, B, sont combinées comme dans la figure 4 pour donner un canal achromatique (N-Blanc), et deux canaux chromatiques, (J-B) et (R-V). Une analyse de Gottschalk et Buchsbaum [53] indique que, du point de vue de la théorie de l'information, ce système est optimal, au sens où ses composantes sont orthogonales, et donc ne contiennent pas d'information redondante.

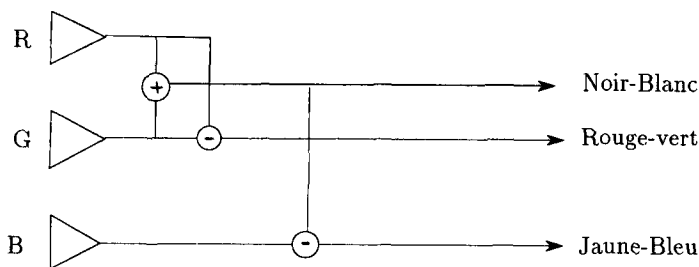


Fig. 4.

Une variante ⁽⁶⁾ de ce modèle a été introduite par Faugeras [34, 33] dans le domaine du traitement d'images. En considérant que la réponse non-linéaire des cônes est proportionnelle au logarithme de l'intensité de la stimulation, il obtient la base :

$$A = a(\alpha \log r + \beta \log v + \gamma \log b),$$

(6)

$$C_1 = u_1 \log (r/v), \quad C_2 = u_2 \log (r/b).$$

Les constantes sont ajustées pour retrouver des données psychophysiques classiques, en particulier on utilise $\alpha = 0,612$, $\beta = 0,369$, $\gamma = 0,019$, pour obtenir une fonction A correspondant à V. Les fonctions (A, C₁, C₂) décrivent les réponses de cellules qui ont été effectivement trouvées dans le cortex visuel ⁽⁷⁾.

L'espace (A, C₁, C₂) a aussi, avec une bonne précision, une interprétation intéressante en termes perceptuels,

⁽⁶⁾ Si l'existence du canal achromatique et des deux canaux opposés est certaine, la question des contributions relatives des différents senseurs n'est pas entièrement clarifiée [90].

⁽⁷⁾ Les coordonnées trichromatiques et le système à couleurs antagonistes peuvent être considérés comme la première et la seconde étape de la perception des couleurs. Le modèle des couleurs antagonistes décrit comment les trois types de récepteurs sont groupés et connectés dans la rétine et le cerveau. Des corrélations importantes entre résultats psychophysiques et propriétés neurophysiologiques ont été mises en évidence d'abord par De Valois [28]. Des références classiques sur la neurophysiologie de la perception colorée sont aussi Daw [27] (contient une discussion des applications neurophysiologiques de la théorie Rétnex), Livingstone et Hubel [90], et Zeki [129].

comme le montre la figure 5 dans laquelle apparaissent des quantités que nous allons à présent caractériser.

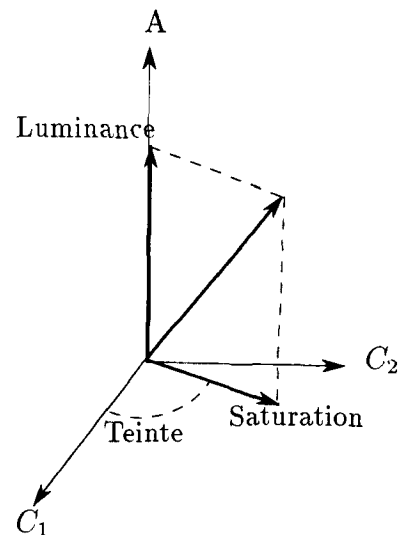


Fig. 5. — Interprétation perceptuelle de l'espace (A, C₁, C₂).

Luminance, teinte (hue), saturation. C'est le triplet de variables le plus proche de la perception consciente humaine, il a une interprétation (imprécise) en termes photométriques, comme le montre la figure 6 qui illustre

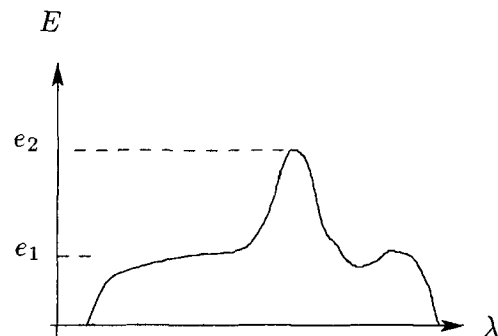


Fig. 6. — Distribution spectrale d'énergie illustrant la longueur d'onde dominante (le pic d'énergie de niveau e₂), la pureté (le rapport entre e₁, l'énergie de la lumière blanche présente dans le signal et e₂), et la luminance (l'intégrale de la courbe).

un signal couleur général (non monochromatique). Le triplet peut aussi être approximé par les valeurs (L, λ, p), définies en termes colorimétriques dans la section suivante, ainsi que par un système de coordonnées pouvant être obtenu simplement dans l'espace (R, V, B) comme un système cylindrique dont l'axe est la bissectrice principale, ce qu'illustre la figure 7, et permettant donc un calcul direct à partir de (r, v, b). Notons que dans toutes ces approximations, comme dans les formules de calcul qui en sont déduites, les trois paramètres sont indépendants (ce qu'illustrent bien les schémas), ce qui n'est pas entièrement exact.

● **Luminance :** Attribut classant les couleurs de « clair » à « sombre ». Distingue les niveaux de gris. L'œil en perçoit

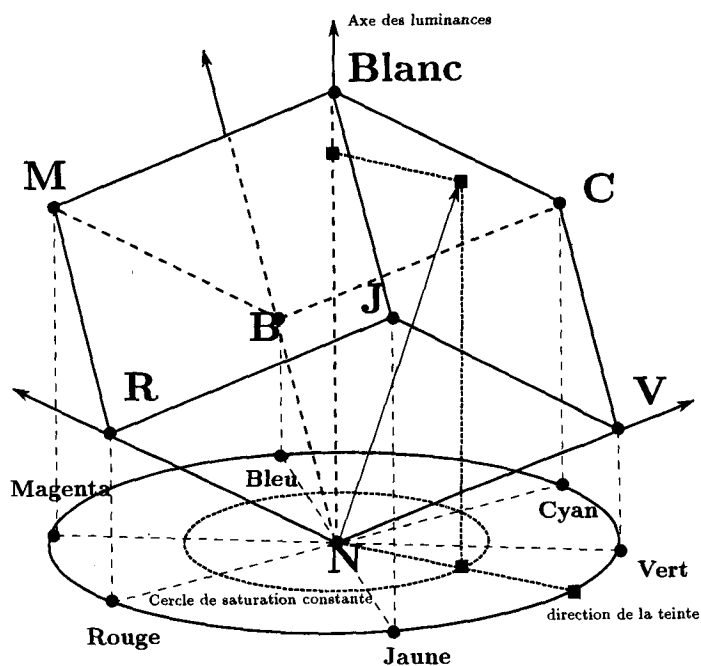


Fig. 7. — Le système de coordonnées (I, H, S), dans le repère (R, V, B) (cube des couleurs).

largement plus d'une centaine. (L : quantité d'énergie du signal.)

- Teinte (hue ⁽⁸⁾) : Psychologiquement, distingue rouge, jaune, vert... Il y a plus d'une centaine de teintes distinctes pour l'œil humain. (Longueur d'onde dominante λ : longueur d'onde de la couleur (pure) observée dans le signal.)
- Saturation : Une couleur pure est saturée à 100 %, le blanc et les gris le sont à 0 %. Distingue rouge et rose, bleu marine et bleu roi... Il y a une vingtaine de niveaux de saturation distincts, suivant la teinte. (Pureté p : proportion de lumière de longueur d'onde dominante et de lumière blanche nécessaire à la reproduction du signal.)

2.1.5. La théorie trichromatique

Nous avons vu que les signaux couleur forment un espace de dimension infinie. L'objet de ce chapitre est de montrer pourquoi on peut utiliser des représentations satisfaisantes, en termes d'espace de dimension 3.

Aspects psychophysiques de la perception des couleurs. La colorimétrie, une théorie psychophysique descriptive bien rodée a établi que dans un nombre significatif de conditions d'observations la plupart des couleurs peuvent être reproduites de manière parfaite par un mélange ⁽⁹⁾ de

⁽⁸⁾ Une traduction plus juste serait « couleur », mot que nous réservons pour les distinctions les plus générales, alors que « teinte » serait alors peut-être plus approprié.

⁽⁹⁾ Notons que par mélange, on entend addition ou soustraction : si les primaires dont on dispose dans un champ sont le vert, le jaune, et le bleu, et la couleur recherchée dans l'autre champ le rouge, on doit utiliser l'égalité Jaune = Rouge + Vert, et donc ajouter une quantité négative de Vert (ce qui est réalisé en l'ajoutant à l'autre champ). Les mélanges sont des opérations algébriques. Si le lecteur se souvient qu'on obtient du vert en mélangeant de la peinture jaune et bleue, qu'il ne soit pas surpris par la relation mentionnée plus haut : il s'agit dans un cas (les sources de lumière) d'une composition additive, dans l'autre (les pigments), d'une composition soustractive !

trois couleurs fixées, et que les proportions de ce mélange sont déterminées de manière unique. Le choix de ces couleurs peut être très large, la seule condition étant leur indépendance, au sens où l'une d'entre elles ne doit pas pouvoir être obtenue par mélange de deux autres. De plus, un résultat plus fort est que ces mélanges satisfont à des lois d'additivité et de proportionnalité (lois de Grassman). Notons que tout ceci repose sur le paradigme expérimental de l'appariement des couleurs, qui définit les concepts colorimétriques : les deux moitiés du champ optique contiennent chacune une zone colorée de faible surface, et le sujet manipule des paramètres jusqu'à ce qu'elles lui paraissent identiques. Ce qui est prévu par la théorie, ce sont les appariements et non la couleur perçue. Par ailleurs, deux distributions spectrales d'énergie différentes peuvent donner naissance à une couleur identique (métamérisme), ce qui est une conséquence de la réduction de la dimension de l'espace des couleurs.

Techniques colorimétriques. Deux des outils importants de la colorimétrie sont :

- L'espace des couleurs : grâce aux lois expérimentales de l'appariement énoncées plus haut, les couleurs peuvent être représentées comme des vecteurs d'un espace de dimension 3. Ceci permet, par exemple d'effectuer des changements de base de manière simple. Étant donné trois couleurs primaires \bar{C}_i et une couleur test \bar{C} , toutes connues par leur distribution d'énergie spectrale, $c_i(\lambda)$ et $c(\lambda)$, les coordonnées C_i sont calculées par la relation suivante, qui fait le lien entre couleur (colorimétrique) et signal couleur :

$$(7) \quad C_i = \int_{\lambda_1}^{\lambda_2} c(\lambda) c_i(\lambda) d\lambda .$$

Ces coordonnées peuvent éventuellement être, rappelons-le, négatives, ce qui signifie physiquement que l'on doit soustraire une primaire.

- Les diagrammes chromatiques : ils permettent de représenter la chrominance, en éliminant la luminance. Des trois coordonnées chromatiques $c_i = C_i / (C_1 + C_2 + C_3)$, on trace les deux premières (la troisième est alors déterminée uniquement) sur un diagramme. La figure 8 montre, sur le diagramme XY de la CIE, les coordonnées des illuminants standard, la courbe des couleurs pures, la ligne des pourpres. Toutes les couleurs physiquement réalisables se trouvent dans le convexe qu'ils définissent. Comme le montre la figure 9, la longueur d'onde dominante et la pureté, quantités psychophysiques correspondant respectivement à la teinte et à la saturation, peuvent être obtenus facilement, ainsi que la couleur pure complémentaire. Une autre propriété, illustrée par cette figure est que l'ensemble des couleurs réalisables par mélange de trois (resp. deux) couleurs données est l'intérieur du triangle (resp. segment) formé par celles-ci. Ceci permet, par exemple, de matérialiser simplement l'ensemble des couleurs pouvant être acquises par un système de filtrage RGB, pour peu que l'on connaisse les coordonnées chromatiques des primaires utilisées.

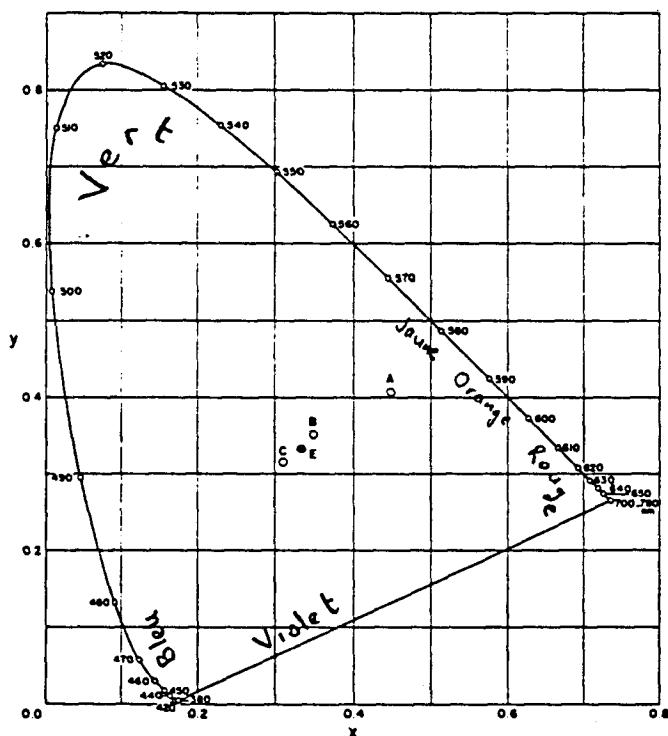


Fig. 8. — Le diagramme XY de la CIE.

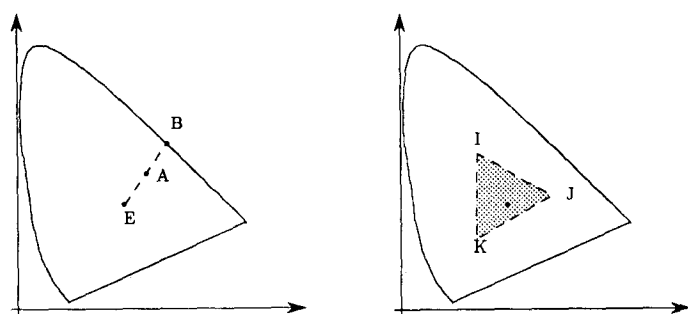


Fig. 9. — Partie gauche : La longueur d'onde dominante de la couleur A est la couleur B, la pureté est $\frac{EA}{EB}$. Partie droite : Couleurs réalisables par mélange de I, J, et K.

La CIE a défini un système de 3 primaires (imaginaires, puisque se situant en dehors du convexe des couleurs réalisables) de référence dit XYZ de 1931 (fig. 10) qui possède plusieurs avantages :

○ Les coordonnées couleurs sont toujours positives (ce qui n'est pas le cas des systèmes fondés sur des primaires RGB).

● La luminance des primaires X et Z est nulle, et donc la composante sur Y représente la luminance du signal. (On a choisi pour cela $Y(\lambda) = V(\lambda)$, la fonction de sensibilité spectrale relative.)

● Le blanc (flux radiant uniforme à toutes les longueurs d'onde) est représenté par le point défini par un mélange en quantité égale des trois primaires XYZ.

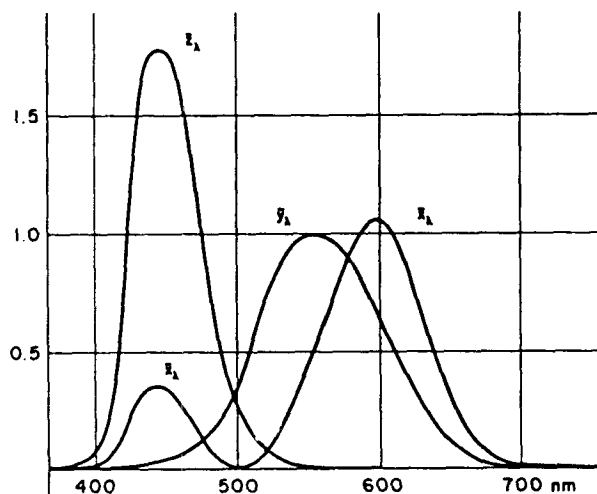


Fig. 10. — Les composantes des couleurs monochromatiques sur la base des fonctions XYZ de la CIE.

Une autre technique, assez souvent employée en pratique pour classer les couleurs, est d'utiliser des atlas de couleur, dont la théorie est donnée par Koenderink [81] et dont le plus connu est celui de Munsell. Par rapport aux techniques colorimétriques, l'emploi d'atlas permet, pour autant que l'on utilise un nombre d'illuminants suffisamment élevé, de comparer directement les fonctions de réflectance, et donc d'obtenir des classements photométriques.

Concluons cette présentation en remarquant que les résultats colorimétriques reposent sur un paradigme expérimental très particulier et qu'ils ne constituent bien sûr pas une théorie de la vision colorée, au sens où celle-ci fait intervenir des scènes (spatialement) complexes.

2.1.6. D'autres espaces couleur

Le système XYZ permet l'échange et la communication d'informations chromatiques de manière très satisfaisante. Il existe par ailleurs beaucoup d'autres systèmes de coordonnées couleur, conçus en vue d'applications spécifiques. Les formules qui suivent, à l'exception de la première série, sont extraites de Pratt ⁽¹⁰⁾ [105].

● I - H - S :

intensité, teinte (*hue*), saturation

$$I = R + V + B$$

$$H = \arccos \left(\frac{\frac{1}{2} ((R - V) + (R - B))}{\sqrt{(R - V)^2 + (R - B)(V - B)}} \right)$$

et si $B > V$, $H = 2\pi - \dots$

$$S = 1 - \min(R, V, B)/I.$$

Ce système possède des propriétés intéressantes, dont l'indépendance de la teinte et de la saturation, et a une interprétation immédiate en termes de perception

⁽¹⁰⁾ On trouve dans la littérature des différences minimales.

humaine, comme nous l'avons vu précédemment (figs. 5 et 7). Cette dernière caractéristique lui donne un grand intérêt « ergonomique », qui peut se manifester en particulier dans les phases de segmentation, dans lesquelles les méthodes, automatiques ou semi automatiques, imposent encore des phases de mise au point très lourdes.

Malheureusement la transformation permettant de passer de R, V, B à I, H, S est non linéaire et souffre par conséquent de deux inconvénients assez graves examinés en détail par Kender [77] :

— il y a des singularités (sans prolongement par continuité possible) de la teinte pour les points achromatiques ($R = V = B$) et de la saturation pour le noir ($R = V = B = 0$), dans le voisinage desquelles une petite perturbation sur R, V, B peut conduire à des variations importantes de H ou S ;

— la distribution des valeurs transformées n'est pas uniforme et la numérisation peut conduire à des pics et trous artéfactuels dans la répartition.

• L - a - b, U* - V* - W* :

systèmes perceptuellement uniformes, dans lesquels la métrique euclidienne est significative des différences de couleur perçues, obtenus en pondérant la chrominance par la luminance

$$L = 25(100 Y/Y_0)^{1/3} - 16,$$

$$a = 500((X/X_0)^{1/3} - (Y/Y_0)^{1/3}),$$

$$b = 200((Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}) \quad U^* = 13 W^* (u - u_0),$$

$$V^* = 13 W^* (v - v_0), \quad W^* = 25(100 Y)^{1/3} - 17$$

où $u_0 = 0,199$, $v_0 = 0,308$, $u = 4X/(X + 15Y + 3Z)$, $v = 6Y/(X + 15Y + 3Z)$, et X_0, Y_0, Z_0 sont les coordonnées du blanc de référence. Ces systèmes comportent les mêmes inconvénients que le précédent, et ne sont pas si utiles pour la vision artificielle, où on s'intéresse aux différences quantifiables et non à celles significatives psychologiquement. En revanche ils permettent d'effectuer simplement des moyennes, fait utilisé par exemple par Westelius [123].

• Y - I - Q, I1 - I2 - I3 :

systèmes obtenus par transformation linéaire des valeurs initiales, obtenue par analyse aux principaux composants de celles-ci

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,273 & -0,322 \\ 0,212 & -0,522 & 0,315 \end{pmatrix} \begin{pmatrix} R \\ V \\ B \end{pmatrix}$$

$$I1 = (R + V + B)/3, \quad I2 = (R - B)/2,$$

$$I3 = (2V - R - B)/4.$$

La linéarité des transformations permet d'éviter les problèmes liés aux singularités, à l'instabilité, et à la non-uniformité. Le premier de ces systèmes, utilisé par la télévision, a été prévu pour minimiser la largeur de bande du signal en conservant la fidélité des couleurs et est fondé sur des observations psychophysiques (achromaticité de

l'œil puis bichromaticité suivant un axe orange-cyan aux faibles angles). Le second est obtenu par Ohta [102] en employant huit images. Sa méthode a été d'effectuer une transformation de Karhunen-Loewe dynamique (utilisée à chaque itération de l'algorithme de segmentation par division de régions avec seuillage récursif de Ohlander [101]), puis d'effectuer une statistique des coordonnées couleur ainsi apparues. Le seul usage des coordonnées I1, I2, I3 apparaissant, dans cet ordre, le plus souvent se révèle aussi efficaces pour la segmentation que la méthode adaptative. Ceci veut dire qu'en général (c'est-à-dire statistiquement), ce système est le meilleur pour rendre compte des différences de couleur, même si pour une image donnée d'autres systèmes de coordonnées pourront être plus efficaces. On peut noter qu'il coïncide assez précisément avec les vecteurs caractéristiques décrivant les réflectances de Cohen, ce qui est, somme toute peu surprenant : la meilleure description des matériaux est donnée par les vecteurs qui permettent de mieux décrire leurs images sous des illuminants relativement moyens.

Notons d'ailleurs que la transformation de Karhunen-Loewe constitue un outil statistique important qu'il est possible d'utiliser pour chaque image afin de trouver le système de coordonnées le plus discriminant pour celle-ci. Les vecteurs X_1, X_2, X_3 calculés ainsi sont :

$$X_i = w_{Ri} R + w_{Vi} V + w_{Bi} B$$

où $W_i = (w_{Ri}, w_{Vi}, w_{Bi})'$ est le vecteur propre normalisé de la matrice de covariance des distributions de R, V, B correspondant à la valeur propre λ_i , avec $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Elles sont non corrélées, X_1 est le vecteur de base optimal (ayant la plus grande variance), X_2 est le meilleur, orthogonal à X_1 , etc...

Vis-à-vis des tâches de type segmentation, la couleur semble être une information bidimensionnelle [102], et tous les systèmes de coordonnées sont à peu près équivalents. Il est intéressant et naturel, si l'on travaille avec un triplet d'images, de conserver l'intensité et un vecteur à deux dimensions pour la chromaticité. Malgré les difficultés évoquées, l'usage de coordonnées normalisées pour cette dernière (par exemple (x, y) , (r, g) , (H, S)) est attirant car il permet de séparer l'information d'intensité de celle de chromaticité. De plus, un avantage important des coordonnées normalisées est de supprimer les petites variations spatiales de couleur dues aux changements d'angle d'observation, dans le cadre du modèle simplifié (sans reflets) de l'équation (4). Ceci est important pour les algorithmes faisant intervenir plusieurs points de vue différents sur la même scène, comme ceux de stéréovision, et pour les algorithmes de segmentation confrontés à des surfaces courbes. Notons cependant que vu l'existence de la singularité au point $I = 0$, les couleurs normalisées correspondant à de faibles intensités ne sont pas significatives, et qu'elles sont généralement très bruitées.

Notons ici une méthode de calcul de la couleur normalisée due à Healey [59], [57], qui est photométriquement fondée et a l'avantage de permettre d'obtenir des résultats précis si l'on utilise beaucoup de filtres colorés différents. L'idée est de chercher à calculer une approximation de $I(\lambda)$ sous

forme d'une combinaison linéaire de n fonctions de base, $\sum_{j=1}^n x_j B_j(\lambda)$ à partir de n mesures $s_i = \int_{\lambda_1}^{\lambda_2} F_i(\lambda) I(\lambda) d\lambda$ obtenues avec des senseurs de sensibilité spectrale $F_i(\lambda)$. Pour cela, on résout le système linéaire :

$$(8) \quad s_i = \sum_{j=1}^n K_{ij} x_j$$

où $K_{ij} = \int_{\lambda_1}^{\lambda_2} F_i(\lambda) B_j(\lambda) d\lambda$. On fait ainsi correspondre à chaque couleur un vecteur (x_i) de \mathbb{R}^n , et il est possible d'utiliser les normes vectorielles classiques, qui, dans ce contexte, ont une interprétation physique.

2.2. MODÈLES RÉCENTS

2.2.1. Modèles physiques de réflectance

Surfaces lambertiennes et spéculaires. Les deux cas extrêmes de propriété géométrique de réflectance sont :

- La surface lambertienne (réflexion parfaitement diffuse, ou isotrope) pour laquelle l'intensité réfléchie est indépendante de l'angle d'observation et ne dépend que de celui d'illumination i et de l'albédo ρ

$$(9) \quad f_r(\theta_i, \phi_i, \theta_e, \phi_e, \lambda) = \frac{\rho}{\pi}$$

Pour une source ponctuelle de radiance E , $L = \frac{\rho}{\pi} E \cos \theta_i$.

- Le miroir (réflexion parfaitement spéculaire) pour lequel l'intensité réfléchie est entièrement concentrée dans une direction, conformément aux lois de Descartes :

$$(10) \quad f_r(\theta_i, \phi_i, \theta_e, \phi_e, \lambda) = \frac{\delta(\theta_i - \theta_r) \delta(\phi_i - \phi_r - \pi)}{\sin \theta_i \cos \theta_i}$$

Dans beaucoup de travaux de vision par ordinateur, les surfaces étaient supposées lambertiennes, mais de nouveaux modèles permettant de prendre en compte les surbrillances et, dans une certaine mesure, la géométrie, sont apparus. Le premier des modèles de cette catégorie est celui de Phong [104], utilisé pour la synthèse, qui est purement empirique. Les meilleurs modèles tenant compte de la couleur utilisés à présent sont au contraire fondés sur la physique de la réflexion.

La plupart des surfaces ont des propriétés intermédiaires et leur réflectance peut être décrite géométriquement comme la somme d'un terme de type lambertien et d'un terme de type (pic) spéculaire. Une telle représentation est justifiée et utilisée par Nayar, Ikeuchi, et Kanade [97] sous le nom de « surface hybride ». Elle leur permet de calculer l'orientation et la proportion lambertien/spéculaire de surfaces par une méthode d'échantillonnage photométrique. Le principe est de fixer la direction d'observation et d'obtenir (en travaillant dans un plan, donc $g = i + e$) la courbe $I(g) = A \cos(i) + B \delta(i - e)$ qui permet alors d'estimer, comme le montre la figure 11, A et B , importance relative des composantes lambertien-

nes et spéculaires, et e , qui donne la direction de la normale. En pratique, ces auteurs utilisent des sources de lumière étendues, ce qui permet d'obtenir à coup sûr le pic spéculaire (qui pourrait échapper autrement à cause de l'échantillonnage) et de mesurer les deux composantes sur une même image, leur intensité étant alors comparable, mais n'apporte que des changements algorithmiques.

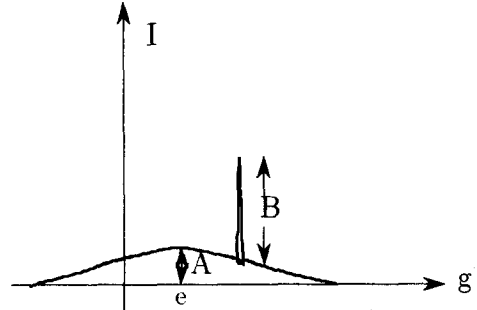


Fig. 11. — Variations d'intensité en fonction de l'angle d'illumination, pour une surface hybride.

Une autre représentation hybride, est utilisée par Ikeuchi et Sato [70]. La différence réside dans la modélisation du terme spéculaire par $B \frac{\exp(-k\alpha^2)}{\cos e}$, qui est la valeur prédite par le modèle de Torrance et Sparrow (voir plus bas). Ces auteurs présentent un algorithme qui extrait A , B , k (paramètre de rugosité) grâce à la connaissance a priori d'une courbe de même nature que la précédente, et ce à partir d'une image d'intensité et d'une image de profondeur.

Réflexion profonde et de surface. Lorsque l'on s'intéresse aux propriétés spectrales, la distinction la plus importante est celle montrée sur la figure 12 entre la réflexion qui se produit en surface de celle qui se produit après un parcours plus profond de la lumière à l'intérieur du matériau (ce n'est pas la même chose que la distinction spéculaire/diffus qui est purement géométrique).

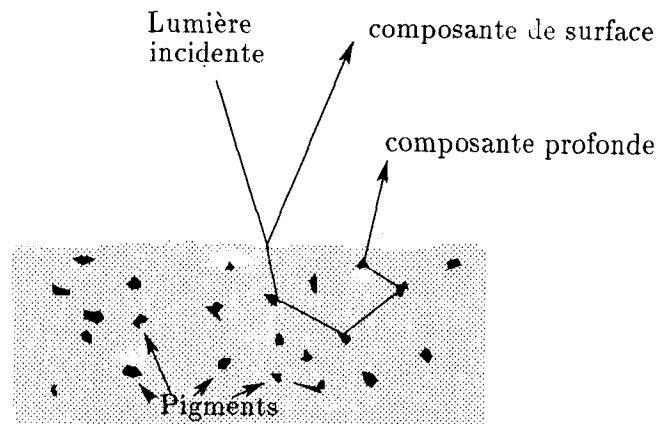


Fig. 12. — réflexion de surface et réflexion profonde.

Les modélisations les plus exactes de la réflexion de surface sont obtenues par des théories d'optique physique⁽¹⁾. Le modèle de Beckmann et Spizzichino [4] (exposé dans [97] que nous suivons) considère les surfaces comme des processus aléatoires stationnaires d'écart type σ_h (représentant la rugosité), de fréquence spatiale T , et l'illumination comme une onde électromagnétique plane de fréquence λ . Ce modèle prédit l'existence de trois composantes réfléchies :

- Un lobe diffus dû aux réflexions multiples. Celui-ci, tout comme la composante diffuse due à la dispersion interne, peut être représenté adéquatement par un modèle lambertien (cosinus).
- Un pic spéculaire, qui est nul en dehors de la direction spéculaire prédite par la loi de Descartes (fonction delta).
- Un lobe spéculaire, éparpillé autour de cette direction (gaussienne).

Les deux dernières composantes forment la réflexion spéculaire. Pour une surface très lisse ($\sigma_h \ll \lambda$), le pic spéculaire domine, pour une surface rugueuse, il est négligeable et le lobe spéculaire domine. La réflexion spéculaire est très concentrée pour une surface lisse ($\sigma_h/\lambda < 1,5$) ou ondulant peu ($\sigma_h/T < 0,02$). Le lobe et le pic ne sont présents simultanément que pour un faible intervalle de rugosités.

Les modèles d'optique géométrique ne permettent pas de rendre compte du pic spéculaire, et ne sont donc pas adéquats pour les surfaces proches de miroirs, en revanche, ils décrivent bien le lobe spéculaire, et donc les surfaces rugueuses. Dans ce cadre, Cook et Torrance [23], qui ont examiné le problème de la couleur, modélisent la surface par une distribution de micro-facettes spéculaires et :

$$(11) \quad R_s = FDA$$

où

- F , quantité de lumière reflétée individuellement par les facettes est donnée par les formules de Fresnel, qui permettent de calculer les composantes réfléchies et transmises à l'interface de deux diélectriques. (Les détails sont par exemple dans le livre de Born et Wolf [8].)
- D , distribution de l'orientation des facettes par rapport à la normale moyenne, permet de prendre en compte divers paramètres tels que la rugosité k , par une relation du type $D(\alpha) = c \exp(-k\alpha^2)$.
- A , facteur d'atténuation géométrique, rend compte du masquage et de l'ombrage mutuel des facettes et est donné par les formules établies par Torrance et Sparrow [118].

F et A varient peu, sauf pour $i \rightarrow 90^\circ$. La couleur de la composante spéculaire dépend à la fois de celle de l'illuminant et de celle du matériau. La dépendance vis-à-vis de l'angle d'incidence i est la suivante :

$$\begin{aligned} &\text{pour } 0 \leq i \leq 70^\circ, R_s(\lambda) \text{ varie peu,} \\ &\text{pour } i \rightarrow 90^\circ, R_s(\lambda) \rightarrow 1. \end{aligned}$$

⁽¹⁾ Les modèles physiques utilisent directement les équations de Maxwell et traitent la lumière comme onde, par opposition aux modèles géométriques, qui la considèrent comme rayon.

Qualitativement la couleur de la spécularité varie donc pour tendre vers celle de l'illuminant, mais cette variation est compliquée à calculer précisément. Sous une forme achromatique, le modèle a été utilisé avec succès par Healey et Binford [60] pour calculer les rayons de cylindres dans des conditions expérimentales bien contrôlées.

Healey et Binford [62] proposent un modèle de réflexion d'une grande généralité combinant d'une part le modèle de Cook et Torrance précédemment décrit pour la réflexion spéculaire et d'autre part le modèle de Kubelka-Munk étendu par Reichman [107] (également bien décrit dans [75]) pour la réflexion profonde. Deux grandes classes de matériaux sont distinguées :

- Les matériaux homogènes (métaux), pour lesquels $R = R_s$.
- Les matériaux inhomogènes (diélectriques), pour lesquels $R = R_s + R_p$. Pour ceux-ci, R_s dépend peu de λ , R_p dépend peu de i , les surbrillances sont localisées (et d'irradiance élevée), si bien que dans les zones mates, $R = R_p$ et dans les zones surbrillantes $R_p = Cte$. L'existence de R_p fournit une base pour la distinction des diélectriques et des métaux.

Le modèle dichromatique, valable pour les matériaux inhomogènes, est :

$$(12) \quad R(i, e, g, \lambda) = R_s(i, e, g, \lambda) + R_p(i, e, g, \lambda) = m_s(i, e, g) C_s(\lambda) + m_p(i, e, g) C_p(\lambda).$$

L'affirmation supplémentaire est empirique et bien sûr contenue dans la seconde équation qui affirme qu'il y a une factorisation indépendante de la géométrie et des facteurs spectraux, les autres hypothèses étant conservées. En pratique $m_s(i, e, g)$ est une fonction prenant des valeurs quasi nulles sauf sur un pic, et $m_p(i, e, g)$ varie doucement. Le modèle NIR (réflexion de surface neutre) consiste à supposer en plus que $C_s(\lambda) = Cte$, c'est-à-dire que la couleur spéculaire est celle de l'illuminant et la couleur profonde celle de l'objet. Ce modèle est de loin le plus utilisé. Le modèle unichromatique, valable pour les matériaux homogènes, est identique au modèle dichromatique, mais sans le terme profond. On peut noter que dans le cadre du modèle dichromatique, si la composante de surface est négligeable, ou si un algorithme a pu permettre de l'isoler, une équation d'intensité du type de l'équation (4) peut être retrouvée.

L'adéquation des modèles, y compris ceux mis au point par des physiciens, n'est pas simple à montrer, eu égard au grand nombre de causes physiques de la couleur. Par ailleurs les équations exactes sont toujours très compliquées. Les trois derniers modèles cités ont été justifiés et testés expérimentalement : Healey [61] trouve des résultats satisfaisants pour le modèle unichromatique sur 11 métaux et pour le modèle dichromatique sur 25 Munsell. Lee, Breeneman et Schulte [89] testent le modèle NIR sur 8 matériaux différents. Les résultats sont bons pour les plastiques, feuilles de plantes, surfaces peintes, peaux d'orange, mais faibles pour le papier coloré et la céramique. Ceci est en partie expliqué par un argument physique, l'existence d'une interface grasse pour les matériaux vivants.

2.2.2. Modélisation linéaire

Les modèles linéaires consistent à décomposer sur des bases de dimension réduite, dont les vecteurs sont des fonctions de l'intervalle 400 nm-700 nm,

- la distribution spectrale d'énergie de l'illuminant $\vec{E}(\lambda)$
- la fonction de réflectance des surfaces considérées $\vec{R}(\lambda)$
- les fonctions de sensibilité des senseurs $\vec{S}(\lambda)$.

L'intérêt de ces modèles est de représenter toute l'information spectrale par un faible nombre de paramètres, en séparant la dépendance spatiale (contenue dans les coefficients) de la dépendance spectrale (contenue dans les fonctions de base). Cette séparation est un point commun avec le modèle dichromatique, mais une différence importante est que si un plus grand nombre de fonctions spectrales sont utilisées, celles-ci sont spécifiées a priori et sont communes à l'ensemble des surfaces considérées. La représentation du signal couleur (c'est-à-dire de l'information spectrale) est nécessaire pour modéliser la réflectance. Des représentations trichromatiques sont insuffisantes.

Un choix de trois fonctions de base compatibles avec le modèle NIR est utilisé par DZmura et Lennie [32] et également discuté par Wandell [122]. Ces trois fonctions sont les trois premiers termes de base d'une décomposition de Fourier (fig. 13) : $\vec{R}_1(\lambda)$ est constante, $\vec{R}_2(\lambda)$ est une fonction Rouge-Vert, $\vec{R}_3(\lambda)$ est une fonction Jaune-Bleu. Il ne semble cependant pas qu'un tel choix conserve l'information spectrale précisément.

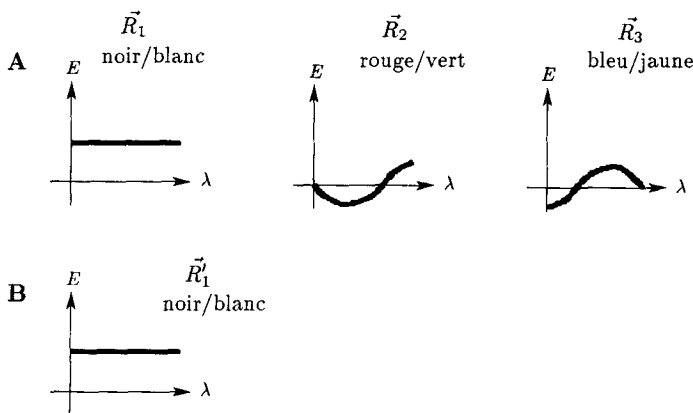


Fig. 13. — Fonctions de base compatibles avec le modèle dichromatique, pour la réflexion profonde (A) et de surface (B).

Il est plus aisé de construire de tels modèles que de démontrer leur validité, tout modèle obtenu par des mesures étant a priori de dimension finie. Le point intéressant est d'obtenir une dimensionalité faible. Or l'un des problèmes de ces modèles est que le nombre de paramètres requis pour une représentation correcte d'un grand nombre d'objets semble en fait plus grand que le nombre raisonnable de 3. D'autre part la détermination de fonctions de base pour une représentation fidèle de l'information spectrale est purement empirique, et donc

dépend très largement de l'échantillon sélectionné pour les calculs statistiques.

Des données utilisées par de nombreux auteurs sont les spécimens de l'atlas de couleur de Munsell et les résultats de Krinov qui a mesuré la réflectance spectrale de 370 matériaux naturels. Les premières analyses sur le sujet sont celles de Judd pour l'illuminant ([74], tables aussi dans [127]) et de Cohen pour la réflectance [18], qui utilisent une méthode statistique de vecteurs caractéristiques. La figure 14 présente ces résultats classiques. La contribution récente est celle de Maloney [91] qui reprend ces analyses et les étend à 462 spécimens de Munsell et à 337 de Krinov pour conclure que, de manière générale, 5 à 7 paramètres sont nécessaires; mais que si l'on tient compte des courbes de sensibilité des photorécepteurs humains et de leurs propriétés de filtrage, un modèle à 3 ou 4 paramètres est suffisant. Il explique la faible dimensionalité requise par un argument de physique nucléaire assez curieux: les lignes d'absorption spectrale sont élargies par des effets d'états solides.

Jaaskelainen, Parkkinen, Kuittinen et Oja [71], [103] déterminent statistiquement des fonctions de base pour représenter la couleur et parviennent à des conclusions à peu près identiques.

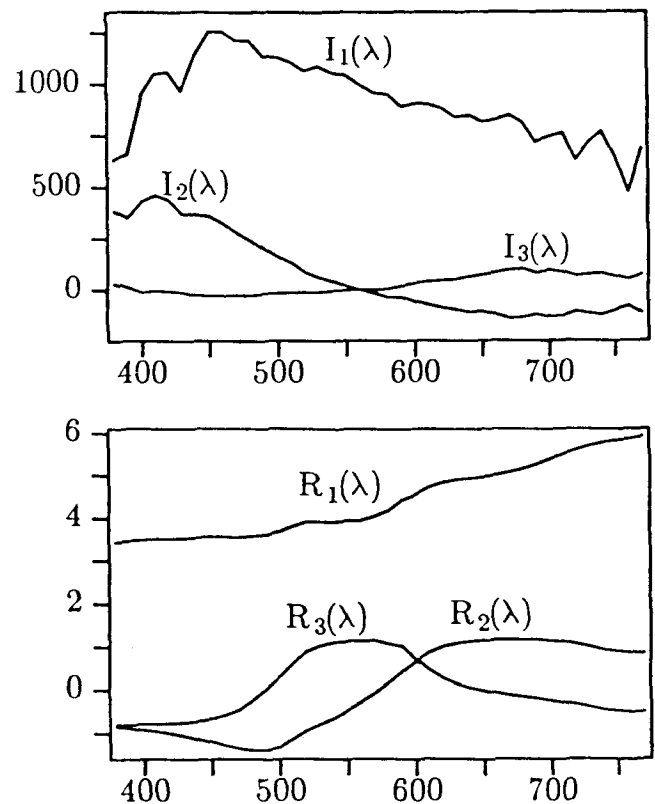


Fig. 14. — Fonctions de base pour la représentation des illuminants (d'après Judd) et des réflectances (d'après Cohen).

2.2.3. Modélisation de l'illumination et synthèse d'images

Tous ces modèles supposent que l'on est en présence d'un illuminant unique. Pour des illuminants multiples, il faut

faire la somme des termes correspondant à chaque illuminant (aucune application n'a été réalisée dans ce contexte et cela semble compliqué). Tenir compte de l'interréflexion (illumination mutuelle) est encore plus difficile, car chaque point d'un objet de la scène doit alors être considéré comme une source de lumière. En pratique, l'illumination mutuelle représente 30 % de l'illumination totale dans les scènes d'intérieur d'après Nishita [100] et on peut s'attendre à ce que sous des illuminants relativement neutres chromatiquement, ils introduisent des couleurs d'illumination non négligeables, comme dans une image produite par Gershon dans [48]. Tant que l'on n'aura pas trouvé le moyen de prendre en compte de tels effets d'illumination, il semble difficile d'utiliser des modèles très fins de réflectance.

Ces problèmes constituent le plus grand obstacle à leur utilisation dans un contexte naturel, ce qui explique que les seules réalisations les exploitant pleinement utilisent un dispositif expérimental très contrôlé. Sauf dans le cas de surfaces presque noires ou d'une surface convexe isolée, les phénomènes d'interréflexion ont un impact qualitatif important pour les images monochromes comme le montrent Forsyth et Zissermann [38], [39] en comparant les profils d'intensité de scènes de géométrie identique mais d'albédo différents : seules les discontinuités de la radiance sont physiquement significatives⁽¹²⁾, les valeurs elles-mêmes résultent d'une interaction entre les surfaces qui est globale et non linéaire. L'équation de radiosité :

$$(13) \quad I(M) = I_0(M) + \rho(M) \int K(M, M') I(M') dM'$$

dans laquelle $I_0(M)$ est la composante de la radiance en M due à la seule source et $K(M, M')$ (le *noyau*, qui dans le cas lambertien vaut 0 si les points M et M' ne sont pas visibles mutuellement, et $\frac{1}{\pi} (\vec{n}_{M'} \cdot \vec{u}_{MM'}) (\vec{n}_M \cdot \vec{u}_{M'M})$ sinon) représente le facteur de gain pour la composante de gain en M due à M' , permet à partir d'un modèle des surfaces de calculer explicitement la radiance par une technique d'éléments finis. Pour m facettes, l'équation (13) devient

$$(14) \quad I = I_0 + RKI$$

où les m composantes de I et I_0 sont respectivement les radiances et composantes de radiance dues à la seule source, et R et K sont des matrices très simples d'albedos et noyaux. Cette technique a été introduite dans le domaine de la synthèse d'images par Cohen et Greenberg [19] et vérifiée expérimentalement par Forsyth et Zissermann [38]. Dans le cas lambertien achromatique, Nayar, Ikeuchi, et Kanade [98] montrent que l'équation (14) permet, en partant d'estimées de la forme et de l'albédo données (inexactement) par un algorithme de « shape from shading » de converger par des itérations successives vers la forme et l'albédo corrects. Leur travail est important, car il montre donc que dans un cas simplifié, le problème de l'analyse peut être résolu au moyen d'un

⁽¹²⁾ Elles ne peuvent, montrent ces auteurs dans un cadre de modèle lambertien, être provoquées que par des discontinuités de surface, des ombres, ou des changements de réflectance.

modèle qui a été introduit pour la synthèse, et qui n'est pas inversible de manière explicite. Pour ce qui est de la couleur, ou des surfaces non lambertiennes, il n'y a pas encore de travaux.

La vision par ordinateur et la synthèse d'images ont des problématiques duales : l'utilisation de techniques d'analyse permet à la vision par ordinateur de produire des descriptions intrinsèques particulières (qui sont des modèles particuliers) à partir d'images numériques obtenues par les caméras, alors que la synthèse part de modèles de scènes pour produire les images numériques affichées sur les écrans. Étant donné ses objectifs, la vision par ordinateur opère dans les scènes (plus ou moins) naturelles, mais ne peut alors que produire des descriptions simples et incomplètes, en particulier au niveau des caractéristiques photométriques des scènes et de leur illumination. La synthèse doit nécessairement partir de descriptions complètes, et peut donc utiliser des modèles très précis. Ses primitives ne correspondent pas toujours à celles utilisées en vision (telles que les segments). Depuis de nombreuses années les modèles de réflexion et d'illumination utilisés pour la synthèse sont bien plus élaborés que ceux utilisés pour la vision. Ils s'attachent en effet essentiellement au rendu de surfaces particulières (transparentes, texturées), à l'interréflexion et à l'illumination ambiante, tous phénomènes qui sont très loin d'être pris en compte pour la vision. Cependant la modélisation demeure assez heuristique, l'aspect réaliste des images synthétisées servant à valider les modèles, si bien que les scènes naturelles complexes ne peuvent encore être reproduites. Il n'y a eu que peu d'interactions entre les deux disciplines⁽¹³⁾, le transfert ayant lieu presque exclusivement de la synthèse vers l'analyse, qui reprend, pour ses propres besoins, les modèles mis au point par la première. Cependant ces modèles ne sont pas d'une utilisation immédiate pour la vision, d'une part à cause de la complexité élevée (il faut des heures de calcul déjà pour la synthèse), d'autre part à cause de la nature même des méthodes employées à présent. En effet les techniques de base (qui ont été combinées : voir Wallace, Cohen et Greenberg [121]) sont :

- Le lancer de rayons [22], aboutissement logique de l'emploi de l'optique géométrique, qui est dépendant du point de vue et que son principe rend non inversible.
- La méthode de radiosité [51], [19], [100] fondée sur la conservation de l'énergie, qui est physiquement exacte et impose de disposer a priori d'un découpage de l'espace pour obtenir les nombreuses équations d'équilibre.

Ce sont les seuls modèles d'illumination fins que nous connaissons, qui ne font pas l'hypothèse de la source de lumière unique. Le calcul de l'illumination résultante est effectué à partir de la connaissance complète de la réflectance et de l'orientation des surfaces et de l'émittance des sources principales. L'utilisation de modèles de ce type

⁽¹³⁾ Une tentative de faire collaborer les deux disciplines est présentée par Gagalowicz [46], qui avait déjà mis cette approche en pratique pour la texture. L'objectif est d'obtenir par l'analyse d'une paire stéréo un modèle synthétisable, de produire une image à partir de celui-ci, et de comparer l'image synthétique à l'image initiale. Cependant les modèles utilisés dans la partie analyse sont assez simplifiés.

pour l'analyse est rendue délicate par l'interdépendance de l'illumination (totale) et des caractéristiques des surfaces. De même, en vision, se pose aussi le problème de l'interdépendance du découpage de l'espace servant au calcul et de celui physiquement réalisé.

2.3. CONCLUSION

Les transformations colorimétriques peuvent présenter des propriétés intéressantes, mais d'une portée limitée. Beaucoup d'espaces couleur sont à peu près équivalents, et il ne faut pas en attendre des miracles !

Les modèles récents sont fondés sur une formalisation du processus de réflexion et du signal couleur, ils permettent donc de représenter des propriétés physiques intrinsèques des scènes. Un des grands intérêts des modèles récents (linéaires et dichromatiques) est de permettre, grâce à une factorisation, de prendre en compte simultanément en les séparant facteurs spatiaux et spectraux, ce qui devrait permettre théoriquement de mener calcul de la couleur et calcul des caractéristiques géométriques de pair. Il nous semble que dans un cadre de *vision* c'est une question de première importance. Une analyse locale des interactions entre ces deux facteurs, au niveau des opérateurs de traitement de type biologique, a été entamée par Gershon [49]. Hurlbert [67] donne aussi une discussion des fonctions algorithmiques possibles des différentes cellules à rôle chromatique qui ont été trouvées dans le cortex. Étant donné que ces analyses font largement appel à des connaissances dépassant le cadre de ce papier, nous préférons envoyer le lecteur intéressé à ces deux travaux, dans lequel il trouvera de nombreuses références physiologiques.

Il nous semble cependant que l'utilisation de modèles de réflectance très fins se révèle délicate dans des environnements qui ne sont pas contrôlés (artificiellement) avec une grande minutie. Aussi nous pensons que les limitations à présent viendront moins de l'insuffisance propre des modèles que de la difficulté de les appliquer dans les cas complexes. Plus précisément, ce qui manque, c'est un modèle de l'illumination suffisamment riche pour dépasser l'hypothèse de la source unique en tenant compte des phénomènes de diffusion et d'interréflexion, et suffisamment simple pour être exploitable algorithmiquement. Il est peut-être plus important de chercher un tel modèle que d'essayer de perfectionner ce qui existe déjà pour la réflectance.

3. Constance des couleurs

Ce terme désigne le fait que les couleurs perçues dans des scènes réelles demeurent à peu près stables dans des conditions d'illumination variant assez largement et dans des scènes de composition variée. Un tel phénomène montre que la *vision* des couleurs est un processus complexe, et en tous cas qu'il requiert un traitement dépassant la mesure des quantités physiques décrites dans le chapitre précédent. On peut dire, sans rentrer dans les détails, que les expériences de psychophysique mettent en

évidence des effets spatiaux puissants. D'un point de vue calculatoire, le problème de la constance des couleurs est celui, sous-déterminé, de déterminer la réflectance spectrale à partir des mesures de senseurs.

C'est une question qui a son importance car le simple signal couleur mesuré par les senseurs ne nous indique pas nécessairement grand chose de très fiable à propos du monde. En revanche, en déterminant la réflectance, on peut ainsi obtenir une propriété intrinsèque des objets, qui est particulièrement utile pour la génération des descriptions sémantiques, et pour l'identification d'indices physiques dans une scène. Il faut cependant reconnaître que vis-à-vis de tâches ne nécessitant qu'une connaissance relative des attributs (la détection des contours et l'étiquetage pour l'appariement), la constance des couleurs n'est pas toujours indispensable.

Ce problème a suscité une grande partie des travaux traitant de la couleur. Cependant, si l'on considère ceux qui ont été effectués dans un cadre algorithmique, force est de constater que :

- tous ces travaux reposent sur des hypothèses assez restrictives et peut-être pas très réalistes ;
- les rares réalisations convaincantes ne fonctionnent que dans des mondes très simplifiés.

Forsyth [41] fournit une revue très complète et claire de la question, avec, entre autres, une analyse fouillée du Retinex. Ce papier, un peu plus récent, prend en compte les derniers développements.

3.1. RETINEX

Land a publié un nombre important de papiers à la fois très influents et très controversés, et en général algorithmiquement assez vagues. En revanche, plusieurs séries d'expériences sont décrites précisément, qui mettent bien en évidence la sophistication du mécanisme de perception des couleurs dans des scènes complexes. La première [83] consistait à superposer deux diapositives achromatiques d'une scène, l'une prise avec un filtre rouge et projetée avec un projecteur à lumière colorée⁽¹⁴⁾, l'autre prise avec un filtre vert, et projetée en lumière blanche. Bien que l'application des méthodes colorimétriques prédit que seules des nuances de rouge et rose seraient observées, les sujets perçoivent une bonne partie des couleurs de la scène originale. Dans la seconde série d'expériences [86], trois diapositives, obtenues avec des filtres rouge, vert, et bleu, sont projetées avec des projecteurs équipés des mêmes filtres colorés (choisis pour correspondre aux sensibilités des cônes rétinien). Les scènes utilisés sont des plans formés de juxtapositions de 100 zones de couleur uniforme disposés aléatoirement (appelés « Mondrians » du fait de leur ressemblance avec certaines peintures de ce peintre). Une zone est choisie, puis sa radiance est mesurée dans chacune des trois bandes de longueur d'onde, ce qui donne les informations atteignant l'œil. Puis une seconde zone est choisie, et les intensités des trois projecteurs sont ajustées pour que les radiances issues de celles-ci soient identiques

⁽¹⁴⁾ On constate, de manière surprenante, que les couleurs finalement perçues ne varient que peu, lorsque la couleur de lumière de ce projecteur est changée.

à celles issues de la précédente. On constate qu'alors, bien que les valeurs mesurées photométriquement par l'œil soient exactement identiques, les couleurs perçues sont différentes. Dans une troisième série d'expériences [85], les sujets doivent apparier les 17 zones d'un Mondrian avec des échantillons de l'atlas de Munsell. Alors que ces derniers sont présentés sous un illuminant constant et sur un fond gris, l'illuminant du Mondrian consiste en trois projecteurs qui sont successivement réglés pour que cinq zones différentes envoient tour à tour la même radiance. On trouve, en examinant les cinq séries de 17 appariements, que ceux-ci sont effectués sur la base, non des radiances, mais des réflectances.

Plusieurs modèles comprennent des équations qui permettent de prédire les phénomènes précédents (par exemple, ceux de Helson et Judd [63], [73], ou Richards et Parks [108], plus simple), ce qui a donné lieu à une critique des travaux de Land par Judd. Notons cependant la différence de nature des deux types de théories : celle de Judd est essentiellement empirique et descriptive (et a son efficacité en tant que telle), alors que celle de Land a un propos algorithmique. Land a en effet proposé plusieurs formulations d'un algorithme appelé Retinex (il était conjecturé qu'il est implémenté par la rétine et le cortex) qui a à son tour inspiré un certain nombre de travaux. Ces algorithmes reposent sur trois principes :

- La couleur en un point ne dépend pas uniquement de son signal couleur.
- La couleur en un point ne dépend que de la combinaison de trois « désignateurs ».
- Les « désignateurs » sont calculés *indépendamment* dans trois bandes de longueur d'onde.

Ils s'attachent donc essentiellement à un calcul de clarté dans différentes bandes spectrales, et ont pour but d'éliminer l'effet de l'illuminant. De plus les « désignateurs » ne devraient pas dépendre de la composition de la scène, mais seulement des propriétés de réflectance au point considéré. Un problème rencontré par cette famille de méthodes est le fait que les valeurs qui servent de données de base pour les algorithmes sont en fait des intégrales comme le montre l'équation (5). Le quotient des intensités de part et d'autre d'un contour n'est donc pas égal au quotient des réflectances.

L'équation d'intensité sous-jacente est de la forme :

$$(15) \quad I(\lambda, \vec{r}) = \rho(\lambda, \vec{r}) E'(\lambda, \vec{r}).$$

C'est l'équation d'intensité (4) où on a regroupé les facteurs pour séparer les propriétés de surface du reste en définissant l'*irradiance effective* E' , et où on a supposé l'absence de diverses dépendances. Elle n'est utilisée que pour trois valeurs de λ fixées. On suppose de plus qu'il y a une asymétrie entre ρ et E' , permettant de résoudre l'équation (15) : ρ consiste en zones uniformes changeant abruptement, alors que E' varie lentement sur la scène : c'est le (*micro-*)*monde de Mondrian*.

En conséquence, les algorithmes sont prévus pour fonctionner dans les conditions suivantes :

- Hypothèse 1 (Monde de Mondrian) : La scène est un

Mondrian plat, l'irradiance effective y varie lentement et continûment et est indépendante de la position de l'observateur.

- Hypothèse 2 (Monde gris) : La moyenne des réflectances des scènes dans chacune des bandes de longueur d'onde est la même.

La première hypothèse permet d'effectuer, en chaque location spatiale, la décomposition de l'intensité en ses deux composantes. La deuxième hypothèse garantit que la normalisation spectrale (rendue nécessaire par le fait que le processus précédent ne fournit des résultats qu'à une constante multiplicative près) donne un triplet correspondant bien à la couleur de la surface. Ces hypothèses ne sont clairement pas vérifiées telles quelles dans les scènes réelles et il semble difficile d'envisager des techniques comme la segmentation de l'image en Mondrians, puisqu'elles demandent de résoudre les problèmes que les Mondrians éliminent.

3.1.1. Les algorithmes originaux

Land a décrit plusieurs algorithmes différents (mais très proches), fondés sur les idées de quotient, chemin, seuillage et moyennage. La simple utilisation d'un quotient entre deux points adjacents permet de détecter un contour, et d'éliminer les effets d'une illumination non uniforme, par seuillage de ce quotient. Cette opération peut être effectuée par des cellules chromatiques, qui sont sensibles uniquement aux différences de couleur. Par transitivité des quotients, on obtient le rapport des réflectances en deux points quelconques, même distants, de l'image, ce qui rend compte de la globalité de la perception. Les chemins utilisés à cet effet formalisent les balayages effectués par l'œil. Les algorithmes sont compliqués à analyser et il n'a pas été publié de résultat satisfaisant quant à leurs performances, contrairement à ce qui est affirmé par plusieurs. En particulier, dans l'article de McCann, McKee et Taylor [94], le seul contenant des simulations, il a été utilisé comme illuminant une combinaison de trois sources de lumière à bande étroite, et des récepteurs ayant chacun une sensibilité voisine de celle de la courbe de distribution d'énergie d'une de ces sources, ce qui fait qu'il n'était sensible qu'à la lumière de celui-ci. Le problème de la séparation des différentes bandes de longueur d'onde ne se posait pas trop !⁽¹⁵⁾ De plus la composition de la scène n'a pas été variée. Il n'est donc pas réellement possible d'évaluer les algorithmes Retinex sur la base de la littérature existante.

⁽¹⁵⁾ Plus précisément, l'illuminant $E(\lambda)$ est de la forme $\sum_{i=1}^3 \alpha_i \delta(\lambda - \lambda_i)$, et les longueurs d'onde choisies, $\lambda_1 = 450$ nm, $\lambda_2 = 530$ nm et $\lambda_3 = 630$ nm sont telles que pour $i \neq k$, $S_k(\lambda_i) = 0$, si bien que les valeurs obtenues par les senseurs sont : $s_k = \int S_k(\lambda) E(\lambda) \rho(\lambda) d\lambda = \alpha_k S_k(\lambda_k) \rho(\lambda_k)$. Il est clair qu'il suffit de disposer de deux points pour calculer dans ces conditions une couleur $(\rho(\lambda_1), \rho(\lambda_2), \rho(\lambda_3))$, et que si l'on change les proportions des lumières monochromatiques de l'illuminant, cette couleur reste invariante. Par contre, le changement de l'une des longueurs d'onde, ne permet plus d'obtenir ce résultat, sans parler de ce qui se passerait pour des lumières quelconques.

L'ALGORITHME RETINEX ([84], le plus récent). — Le désignateur à la zone i est la moyenne des relations à cette zone :

$$\bar{R}^{\Lambda}(i) = \frac{\sum_{j=1}^N R^{\Lambda}(i, j)}{N}$$

la relation de la zone i à la zone j :

$$R^{\Lambda}(i, j) = \sum_k \sigma \left(\log \frac{I_{k+1}}{I_k} \right)$$

où :

$$\sigma(x) = \begin{cases} x & \text{si } x > \text{seuil} \\ 0 & \text{sinon} \end{cases}$$

les I_k étant les intensités, dans la bande de longueur d'onde Λ , des zones d'un chemin reliant i et j .

Cet algorithme est stochastique et dépend de choix qui ne sont pas précisés. Une analyse de Brainard et Wandell [9] utilisant des chemins markoviens donne les résultats suivants :

- De manière formelle, ces calculs compliqués sont équivalents à une simple normalisation par rapport à une moyenne.

- Une simulation montre que l'algorithme se comporte de manière acceptable vis-à-vis du changement d'illuminant, mais non de celui de la composition de la scène (ce qui n'est pas surprenant au vu de l'hypothèse du monde gris). Parmi les variantes de l'algorithme décrit plus haut, citons un « Retinex avec remise à zéro » [95] dans lequel, de plus, les quotients supérieurs à 1 sont remis à 1. Celui-ci se comporte de manière assez semblable, tant formellement (normalisation par rapport à un maximum) que pour une simulation.

La dernière proposition en date de Land contient un opérateur fort différent, de type centre/bords, qui est inspiré par les découvertes neurophysiologiques récentes, et ressemble à celui synthétisé par Hurlbert et Poggio que nous décrivons plus bas.

3.1.2. Algorithmes formels

Le principe. De multiples propositions ont été faites pour des algorithmes effectuant la même tâche que le Retinex (c'est-à-dire le calcul monochromatique de la clarté) mais calculatoirement plus simples et plus clairs. Ces algorithmes sont assez satisfaisants : contrairement à l'algorithme de Land qui utilise des chemins, ils fonctionnent directement dans le cas bidimensionnel et peuvent être formulés de manière locale et parallèle. Leur parenté avec le Retinex est certaine, mais d'un point de vue quantitatif elle n'est toutefois aussi claire que l'affirment plusieurs auteurs (y compris Marr [93]), tout ce que l'on peut dire, c'est qu'il y a les ressemblances suivantes :

- Le calcul de la relation entre les zones donne un nombre qui devrait être proche du quotient de leurs réflectances, et fait intervenir une sorte de différentiation (le long d'un chemin) et un seuillage.

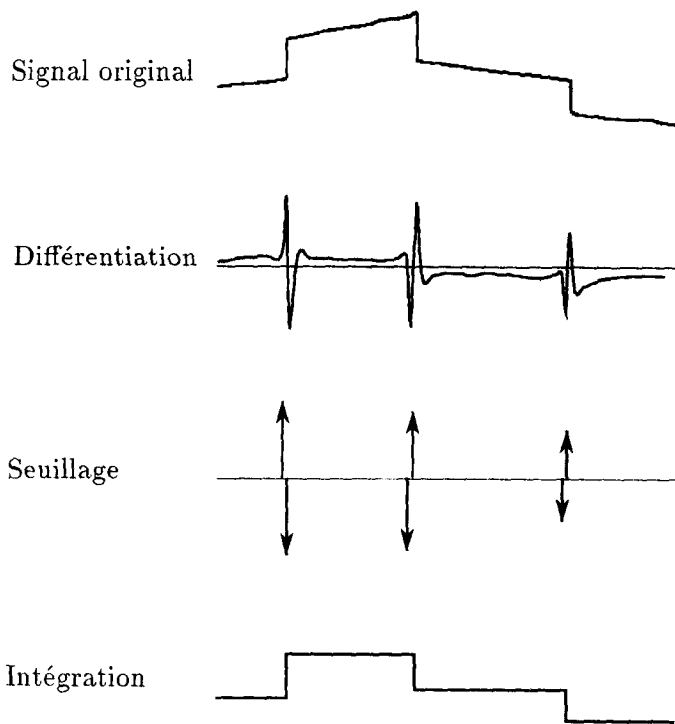


Fig. 15. — Calcul de la clarté (lightness).

- Le calcul du désignateur fait intervenir une opération de moyennage proche d'une intégration. Ils utilisent la méthode suivante (fig. 15) :

1. Différencier spatialement l'intensité.
2. Seuiller pour éliminer les petites valeurs (dues aux petites variations de l'illuminant) en gardant les grandes valeurs (dues aux changements abrupts de réflectance).
3. Intégrer pour récupérer la réflectance.

Les opérations impliquées sont de deux types bien différents :

- la différentiation est un processus *local* effectuant la décomposition spatiale ;
- l'intégration est un processus *global* effectuant la normalisation spectrale en utilisant une large portion du champ visuel pour obtenir une moyenne.

Ces opérations, qui consistent à éliminer les faibles fréquences spatiales, sont réalisées physiologiquement par le phénomène d'inhibition latérale, dans lequel une cellule inhibe l'activité des cellules voisines d'une quantité proportionnelle à son excitation, et décroissant avec la distance.

Le problème du logarithme. Les algorithmes ayant été prévus pour fonctionner de manière achromatique (la longueur d'onde n'est pas un paramètre), la supposition supplémentaire importante dans un cadre de constance des couleurs est que le signal renvoyé par les senseurs est décomposé en une *somme* de deux composants représentant la réflectance et l'irradiance effective. Pour un calcul achromatique (clarté), ceci est obtenu par une simple application de logarithme : en posant $I = \log I'$, $\rho = \log \rho'$, $E = \log E'$, les valeurs originales étant I' ,

ρ' et E' , on a bien $I = \rho + E$ et les techniques qui suivent permettent de calculer alors ρ . L'intérêt de cette application du logarithme avait déjà été mise en évidence par le modèle homomorphique de Stockham [113]. Le caractère homomorphique signifie ici que la structure du processus de formation d'images (produit de I et de R) est envoyée sur celle, additive, du système visual humain, par les sensibilités des récepteurs rétinaux, qui sont logarithmiques. Le modèle A, C_1, C_2 de Faugeras déjà mentionné plus haut, est une extension, pour la couleur du modèle de Stockham. Il permet d'expliquer comment des changements d'illumination pourraient être éliminés : si l'illuminant change uniformément sur l'image, de sorte que les valeurs des réponses des trois cônes $R, V,$ et B soient multipliées respectivement par k_R, k_V, k_B , les canaux A, C_1 et C_2 vont renvoyer une valeur augmentée uniformément des constantes respectives

$$a(\alpha \log(k_R) + \beta \log(k_V) + \gamma \log(k_B)),$$

$u_1 \log(k_R/k_G)$, et $u_2 \log(k_R/k_B)$. Ces dernières peuvent être éliminées par un filtrage spatial (inhibition latérale). Cependant, si les changements d'illuminants sont ainsi pris en compte par le système au niveau perceptuel, la séparation de la réflectance et de l'illumination ne peut être effectuée complètement par cette méthode. Il y a effectivement un problème dans le cas du signal couleur, puisque prendre son logarithme n'entraîne aucune séparation (ce serait toutefois le cas si les senseurs avaient accès à des valeurs du signal couleur, mais malheureusement ils ont accès à des intégrales, comme nous l'avons remarqué, et on a donc affaire à une intégrale de produits), en conséquence les algorithmes ne calculent pas la réflectance ρ . Même dans le cas favorable où l'illuminant est séparable en espace et longueur d'onde, on n'obtient, en appliquant un logarithme, qu'une séparation des facteurs spatiaux.

Les algorithmes de calcul de clarté.

L'ALGORITHME DE HORN [64]. — Calculer :

$$r = (\nabla^2)^{-1} \sigma_1 \nabla^2 I$$

en résolvant par relaxation de Gauss-Seidel le système d'équations aux différences obtenu en discrétisant l'équation (16) sur une grille hexagonale.

L'opération délicate est le calcul de l'inverse de ∇^2 , laplacien en deux dimensions qui est solution de l'équation de Poisson

$$(16) \quad \nabla^2 r(x, y) = \sigma(\nabla^2 I(x, y))$$

où $\sigma_1(x)$ est un opérateur de seuillage scalaire. Si le domaine de définition de I est infini, ou si l'image est entièrement contenue dans celui-ci et est entourée d'une bordure de réflectance constante pour laquelle $\sigma(\nabla^2 I(x, y)) = 0$, le laplacien peut être inversé de manière unique par convolution avec la fonction de Green $g = \frac{1}{2\pi} \log \vec{r}$. Dans le cas discret, l'inversion se ramène à la résolution d'un système d'équations aux différences de

la forme :

$$(17) \quad r_{i,j} - \sum \alpha_{k-i,l-j} r_{k,l} = \sigma(\nabla^2 I_{i,j})$$

où les $\alpha_{k,l}$ donnent une approximation discrète du laplacien (pour une grille hexagonale obtenue en prenant pour éléments de base des triangles équilatéraux, les $r_{k,l}$ sont les 6 plus proches voisins et $\alpha_{k,l} = 1/6$). L'algorithme a fait l'objet d'un traitement en multirésolution par Terzopoulos [115] qui en augmente notablement l'efficacité en apportant une solution au problème de la difficulté de propager des contraintes dans de larges représentations, typique aux algorithmes locaux et itératifs issus d'équations aux dérivées partielles.

Comme l'équation de Poisson n'impose que $\nabla^2 r(x, y) = 0$ à l'intérieur des zones, la solution pourrait contenir un terme harmonique arbitraire, ce qui a imposé une condition aux limites restrictive (clarté constante en bordure de l'image, ce qui suppose une réflectance constante sur cette bordure). Si celle-ci n'est pas satisfaite, ce qui est souvent le cas, l'algorithme peut conduire à des résultats inexacts. Ce point a été corrigé par Blake [5], qui démontre que l'on obtient une solution unique par :

L'ALGORITHME DE BLAKE. — Calculer :

$$r = (\nabla^2)^{-1} \nabla \sigma_2 \nabla I$$

en résolvant par relaxation de Gauss-Seidel le système d'équations aux différences obtenu en discrétisant l'équation (18) sur une grille hexagonale.

Il est clair que l'emploi du gradient ⁽¹⁶⁾ préserve davantage d'informations que celui du laplacien. Il est montré en utilisant pour σ_2 l'opérateur défini par $\sigma_2(\vec{a}) = H(|\vec{a}|^2 - k^2) \vec{a}$ ($H(x)$ étant une fonction de type « marche » régularisée et k un seuil) que dans le cas des Mondrians, $\nabla \times \sigma \nabla I = 0$, ce qui entraîne que l'équation $\nabla r = \sigma \nabla I$ admet une solution unique, et qu'elle est équivalente à l'équation de Poisson :

$$(18) \quad \nabla^2 r = \nabla(\sigma \nabla I)$$

avec de bonnes conditions aux limites ⁽¹⁷⁾. Si cela n'est pas le cas (par exemple à cause de la discrétisation), l'algorithme minimise une mesure d'erreur, et est donc robuste. La discrétisation de Blake est analogue à celle de Horn, en plus complexe. L'algorithme fonctionne avec des données synthétiques.

Une analogie est faite avec le Retinex avec reset : l'équation précédente peut en effet être résolue en calculant :

$$r(x, y) = \int \sigma \nabla I dl \text{ le long d'un chemin reliant } (x_0, y_0) \text{ où}$$

la valeur de r est définie arbitrairement, et (x, y) . Cependant il n'est pas expliqué ce qui équivaut à l'opération de reset.

Cette méthode a été à son tour améliorée par Brelstaff et

⁽¹⁶⁾ Ce n'est pas un opérateur isotrope, mais il est suffisant qu'il soit covariant, afin que le calcul ne dépende pas du choix du système de coordonnées.

⁽¹⁷⁾ Conditions de Neuman : $\vec{n} \cdot \nabla r = \vec{n} \cdot (\sigma \nabla I)$ sur la frontière.

Blake [10] de manière à fonctionner avec des images réelles. Les deux modifications sont :

- La discrétisation sur une grille carrée (à l'aide de triangles rectangles isocèles) qui est celle utilisée par les dispositifs de digitalisation habituels.
- L'intégration d'un détecteur de contours de type Canny pour tenir compte du bruit mieux que la simple étape de seuillage (qui avait justement pour but de distinguer les changements brusques de réflectance).

Les résultats sont satisfaisants sur les images (monochromes) de Mondrians, même bruitées, mais dans le cas de scènes comportant des contours dus à d'autres facteurs que le changement de réflectance, des valeurs incorrectes de celle-ci sont obtenues.

3.1.3. Méthode d'apprentissage

Mentionnons pour finir une approche très différente proposée par Hurlbert et Poggio [69], qui fait intervenir un « apprentissage » sur des exemples où sont fournis en même temps les images de mondrians monochromatiques et les réflectances recherchées. Malgré la formulation linéaire, par son principe monochromatique cette technique est étroitement apparentée aux algorithmes de la famille Retinex dont elle partage nécessairement les hypothèses méthodiques. L'idée est d'estimer l'opérateur L tel que $LS = R$, à partir de matrices S et R dont les colonnes sont respectivement des profils d'intensité de l'image et les profils de réflectance correspondants. La technique, l'estimation linéaire optimale⁽¹⁸⁾, suppose l'hypothèse de la linéarité de l'opérateur L , et conduit à un résultat dont la partie spatialement invariante est un opérateur de type centre/bords soustrayant à la valeur en chaque pixel la valeur moyenne des logarithmes des voisins. Cet opérateur ressemble beaucoup au dernier opérateur Retinex proposé par Land⁽¹⁹⁾. La spécificité de ces opérateurs réside dans leur champ réceptif, qui comporte une petite zone centrale activatrice, et une très large bordure inhibitrice. Outre le fait qu'on ne sache pas dans quel canal chromatique ils opéreraient, ils n'ont pas encore été trouvés de manière aussi systématique que d'autres opérateurs « classiques » (voir [67] pour des références). Les cellules chromatiques bien connues à présent comprennent en effet essentiellement les cellules à double opposition ($L+M-/L-M+$), et celles à simple opposition ($L+/M-$), opposition spatiale (B/B), opposition chromatique ($L+M-$), qui en sont des versions plus simples. Une cellule de type $L+M-/L-M+$ est, au centre de son champ réceptif, excitée par les longueurs d'onde L et inhibée par les longueurs d'onde M , et inversement au bord de son champ réceptif. Le comportement d'une telle cellule vis-à-vis d'un changement d'illumination est intéressant : si par exemple l'intensité des longueurs d'onde L de la source de lumière augmente, l'excitation au centre de la cellule va augmenter, mais l'inhibition sur les bords également, deux tendances qui peuvent se compenser.

⁽¹⁸⁾ Les réseaux connexionistes linéaires et à « back-propagation » ont aussi été essayés et ne se révèlent pas meilleurs.

⁽¹⁹⁾ A la différence que ce dernier utilise le logarithme de la moyenne.

Cependant les phénomènes impliqués dans la constance des couleurs sont très globaux, ce qui ne s'explique pas suffisamment si on ne considère que ces cellules, qui génèrent des effets purement locaux (et expliquent ainsi les phénomènes de contraste simultané, par exemple).

3.2. MODÈLES LINÉAIRES

Le principe de ces algorithmes est d'exprimer toutes les réflectances et illuminants comme combinaison linéaire d'un nombre limité de vecteurs. Par rapport aux techniques de Retinex, celles-ci présentent l'avantage d'effectuer les calculs sur toutes les longueurs d'onde à la fois, et donc d'éviter les problèmes liés au calcul du logarithme d'une intégrale et à la normalisation séparée. Ces algorithmes supposent cependant la même équation d'intensité que les précédents, et donc ne fonctionnent que sous l'hypothèse du monde de Mondrian. Dans ce cadre, cette ligne de recherche me semble la plus avancée, au sens où il y a à présent plusieurs algorithmes reposant sur des bases théoriques claires et donnant des résultats plutôt satisfaisants. La problématique commune à ces algorithmes est de résoudre un système d'équations sous-déterminé au moyen de contraintes de différents types. Le nombre d'inconnues souhaitable est au minimum de 6 (3 paramètres pour l'illuminant, 3 paramètres pour la réflectance), alors que celui des senseurs est de 3, si l'on veut obtenir un modèle de la vision humaine.

3.2.1. Les méthodes à invariants

Les algorithmes de Buchsbaum et de Gershon font l'hypothèse supplémentaire (analogue à celle du monde gris) qu'une certaine quantité moyenne est un invariant d'une scène à l'autre. Ils estiment l'illuminant (ce qui nécessite de plus l'hypothèse qu'il y a une seule distribution spectrale d'illuminant dans la scène) en mettant en correspondance :

- \vec{U} représentant un invariant des réflectances ;
- \vec{V} une moyenne des intensités dans l'image.

La détermination de ces deux quantités est la principale difficulté de ces algorithmes.

L'ALGORITHME DE BUCHSBAUM [14]. — Soient \vec{U} un gris de référence et T le tenseur défini par $(T)_{i,j,k} = \sum_{\lambda} E_{\lambda} R_{j,\lambda} S_{k,\lambda}$,

E_{λ} , $R_{j,\lambda}$, $S_{k,\lambda}$ étant les fonctions de base pour l'illuminant, la réflectance et les senseurs.

1. Déterminer \vec{V} par moyennage spatial.
2. Calculer l'illuminant $\vec{e} = \Lambda_{\vec{U}}^{-1} \cdot \vec{V}$ où $\Lambda_{\vec{U}} = T \cdot \vec{U}$ est une représentation des réflectances.
3. Calculer les couleurs intrinsèques $\vec{C}_{\text{constant}} = \Lambda_{\vec{e}}^{-1} \cdot \vec{C}_{\text{image}}$ où $\Lambda_{\vec{e}} = T \cdot \vec{e}$ est une représentation de l'illuminant.

Cet algorithme repose sur l'hypothèse selon laquelle une moyenne spatiale pondérée des réflectances spectrales est constante et est un gris. Le mode de calcul choisi pour la moyenne nécessite de plus que les couleurs occupent des surfaces égales. Gershon [49] améliore un peu l'algorithme :

- En prenant pour invariant une moyenne prise sur les données de Krinov.
- En segmentant l'image et en prenant donc une moyenne sur les régions.

D'Zmura et Lennie [32] décrivent un modèle linéaire tenant aussi compte du modèle NIR, mais l'estimation de la chromaticité de l'illuminant est effectuée par une simple moyenne spatiale sur la scène.

3.2.2. Les méthodes à structures linéaires limitées

Une autre famille d'algorithmes a été proposée, qui ne nécessite que des hypothèses générales sur la structure des illuminants et réflectances, et non plus sur la composition des scènes. L'équation de base relie par une relation linéaire, en chaque point \vec{r} , les réponses des p senseurs \vec{S} et les n coordonnées de la réflectance \vec{R} :

$$(19) \quad \vec{S}(\vec{r}) = \Lambda_{\vec{E}} \cdot \vec{R}(\vec{r})$$

avec $\Lambda_{\vec{E}} = T \cdot \vec{E}$, opérateur ne dépendant que de l'illuminant d'une manière connue. L'idée, illustrée par la figure 16, est que sous certaines conditions, discutées plus bas, les réponses observées dans l'espace des récepteurs sont restreintes à un sous-espace propre de celui-ci. Ce sous-espace ne dépend que de l'illuminant, les positions des réponses, elles, ne dépendent que des réflectances.

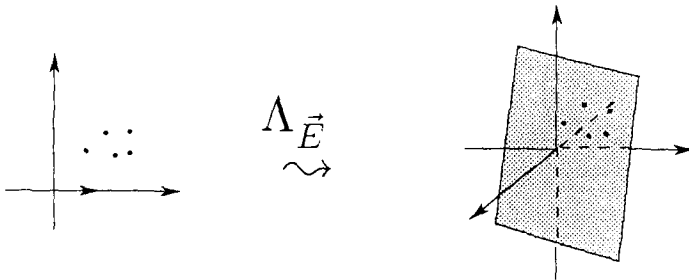


Fig. 16. — La méthode de Maloney et Wandell.

L'ALGORITHME DE MALONEY ET WANDELL [92].

1. Identifier le sous-espace $\text{Vect}(\Lambda_{\vec{E}} \cdot \vec{R}(\vec{r}))_{\vec{r}}$ en trouvant le meilleur vecteur \vec{n} normal à ce sous-espace.
2. Calculer le vecteur illuminant à un scalaire près en résolvant le système $\vec{n}^t \cdot \Lambda_{\vec{R}(\vec{r})} \cdot \vec{E} = 0$ (chaque surface donne une équation) avec $(\Lambda_{\vec{R}(\vec{r})}) = T \cdot \vec{R}$.
3. Calculer les réflectances en chaque point \vec{r} par $\vec{R}(\vec{r}) = \Lambda_{\vec{E}}^{-1} \vec{S}(\vec{r})$ (pseudo-inversion).

Le modèle des illuminants et des réflectances impliqué est restrictif dans la mesure où les hypothèses suivantes sont faites :

- Le nombre des senseurs est strictement supérieur à la dimensionnalité de l'espace des réflectances, pour obtenir un sous-espace propre. Même dans le cas le plus favorable du sous-espace de codimension 1, avec le nombre classique de trois senseurs, on ne peut donc calculer que deux

paramètres de réflectance, ce qui, nous l'avons vu plus haut, est insuffisant.

- Deux surfaces de réflectance différentes donneront toujours des réponses de senseurs distincts, quel que soit l'illuminant.

Wandell [122] étend le précédent au cas d'un illuminant variant spatialement : la méthode consiste à appliquer cet algorithme à un certain nombre de zones, puis à effectuer une interpolation sur l'illuminant. Les problèmes viennent de la difficulté du découpage (il faut des zones assez grandes pour que plusieurs réflectances soient présentes, mais assez petites pour que l'illuminant puisse y être supposé constant) et de l'interpolation (il y aura peu d'échantillons).

Un algorithme tout à fait similaire est décrit par Yuille [128].

3.2.3. Les derniers développements

Régularisation. Rubner et Schulten [110] se placent dans le cas d'un gradient d'illumination linéaire ($\vec{E}(x, y) = \vec{E} + x \cdot \vec{E}_x + y \cdot \vec{E}_y$) et traitent la sous-détermination⁽²⁰⁾ des équations (19) par une méthode de régularisation : ils choisissent des fonctionnelles stabilisantes P_k telles que le minimum de la fonction de coût $\left\| \sum_{\vec{r}} (T \cdot \vec{E}) \cdot \vec{R}(\vec{r}) - \vec{S}(\vec{r}) \right\|^2 + \sum_k \lambda_k \|P_k(\vec{E}, \vec{R})\|^2$ soit uni-

que. Ces fonctionnelles correspondent à des contraintes physiques plausibles :

- le spectre de l'illumination est aussi proche que possible du « blanc » de la lumière du jour D_{65} : $P_1 = \vec{E} - \vec{D}_{65}$;
- la moyenne des réflectances est aussi proche que possible d'un gris G : $P_2 = \sum_{\vec{r}} \frac{\vec{R}(\vec{r})}{N} - \vec{G}$;
- les gradients d'illumination sont les plus faibles possibles : $P_3 = \vec{E}_x$, $P_4 = \vec{E}_y$.

Les coefficients λ_i doivent être ajustés. Il faut noter que le cas $\lambda_1 = 0$ correspond à celui de la constance des couleurs parfaite (ne marche qu'avec une moyenne exactement grise), et celui $\lambda_2 = 0$ à celui d'une perception « brute ». L'algorithme détecte les contours couleur, y calcule les valeurs de $\vec{R}(\vec{r})$ en utilisant une méthode de descente de gradient, et effectue une interpolation. Des simulations sont présentées, elles semblent montrer que les résultats sont bons.

Utilisation de Rétinex. Funt et Drew [44] montrent comment appliquer dans un cadre de modèle linéaire les idées de Blake et Horn. Leur méthode exige un modèle linéaire où les dimensions de l'espace des récepteurs, celui des illuminants, et celui des réflectances sont égales. L'idée est d'appliquer à l'équation (19) les opérateurs gradient et seuillage, puis gradient. Grâce à l'hypothèse du

⁽²⁰⁾ Si il y a N zones, puisqu'ils prennent tous les espaces de dimension 3, il leur faut calculer $3N$ coefficients $R_i(\vec{r})$ et 9 coefficients d'illumination à partir des $3N$ valeurs mesurées $S_k(\vec{r})$.

monde de Mondrian, des termes s'annulent pour donner :

$$(20) \quad \Lambda_{\vec{E}} \nabla^2 \vec{R} = \nabla(\sigma(\nabla \vec{S}))$$

Il est alors possible de déterminer \vec{R} à une constante multiplicative près par une méthode itérative. On part d'un illuminant arbitraire \vec{E}_{init} , on calcule \vec{R} avec (20), puis \vec{E} avec (19)... Seule une exécution (au demeurant très satisfaisante) sur des données achromatiques synthétiques est présentée par les auteurs.

Comme les approches précédentes, celles-ci ne fonctionnent que dans le cadre du monde de Mondrian, mais par rapport aux autres méthodes linéaires vues plus haut, elles présentent l'avantage de ne pas faire d'hypothèse de « Monde gris » (ou de la faire seulement sous une forme affaiblie), de gagner un degré de liberté pour les réflectances, et prennent directement en compte le cas de l'illuminant variant spatialement. Leurs désavantages sont liés à l'aspect itératif et non déterministe des calculs : la complexité algorithmique est grande, et contrairement au cas des méthodes précédentes, on ne peut savoir a priori si celles-ci fonctionnent ou non. Cette difficulté d'analyse devrait inciter à la prudence, car on doit s'en remettre à des simulations sur quelques images...

Images multiples. Tsukada et Ohta [119] proposent d'obtenir des contraintes supplémentaires par une approche originale : en utilisant plusieurs images d'une même scène avec des illuminants différents (ce qui revient à faire intervenir la mémoire). L'idée est qu'en écrivant les équations (19), pour k objets, ℓ illuminants, et en prenant 3 degrés de liberté pour les réflectances et les illuminants, on a $3k$ équations avec $3(k + \ell)$ inconnues. Si on suppose que l'on a identifié q objets entre deux images, alors on peut disposer d'un nombre réduit d'inconnues de réflectance : $3(k - q)$. Le cas traité est celui où $\ell = 2$, alors pour $q \geq 2$, il est possible de résoudre le système (non-linéaire) d'équations. L'algorithme a été testé sur des scènes réelles de 12 Munsell. Deux appariements sont donnés entre les deux scènes au programme, qui détermine alors avec une bonne précision les réflectances des autres échantillons. Les illuminants variables ne sont pas pris en compte, cependant dans le cadre du monde de Mondrian, l'algorithme semble efficace.

L'approche de Forsyth. Forsyth [42], [40], [37] a utilisé une approche nouvelle et fort intéressante d'un point de vue théorique, exploitant de manière très précise la structure mathématique du problème, en cherchant, non pas à calculer la réflectance des surfaces, mais l'apparence qu'elles ont lorsqu'elles sont éclairées par un illuminant canonique. L'idée généralise la remarque que si, par exemple, on observe une image ne comportant que des nuances de vert, l'illuminant ne pouvait être une lumière rouge vif. Pour cela, on utilise la contrainte physique que toute fonction de réflectance prend ses valeurs entre 0 et 1, ce qui permet de circonscrire l'ensemble des couleurs observables sous un illuminant donné. Ceci conduit à rejeter, pour une image donnée, les illuminants ne pouvant produire toutes les couleurs observées dans cette image.

Techniquement, le problème de la constance des couleurs

est traité comme celui de construire une fonction Ψ faisant correspondre aux mesures des p senseurs en lumière canonique \vec{i}^c et à l'illuminant dont le paramètre est \vec{i} , les mesures des senseurs en lumière \vec{i} , puis d'estimer \vec{i} pour une image, et enfin d'appliquer $\Psi^{-1}(., \vec{i})$.

La fonction Ψ recherchée vérifie donc en chaque point, et pour chacune des p composantes l'équation :

$$(21) \quad \Psi_k \left(\int \Phi_1(\lambda, \vec{i}^c) \rho(\lambda) d\lambda, \dots, \int \Phi_p(\lambda, \vec{i}^c) \rho(\lambda) d\lambda, \vec{i} \right) = \int \Phi_k(\lambda, \vec{i}) \rho(\lambda) d\lambda$$

où on a posé : $\Phi_k(\lambda, \vec{i}) = s_k(\lambda) e(\lambda, \vec{i})$ produit de la sensibilité du récepteur k par la distribution d'énergie de l'illuminant \vec{i} et où $\rho(\lambda)$ est la réflectance de la surface. Ceci peut se réécrire :

$$(22) \quad \Psi_k \left(\left(\sum_{j=1}^p a_{ij} \rho_j \right)_{1 \leq i \leq p}, \vec{i} \right) = \sum_{s=1}^p \sum_{j=1}^p r_{ks}(\vec{i}) a_{sj} \rho_j + \int F_k(\lambda, \vec{i}) \rho^*(\lambda) d\lambda$$

après les décompositions

$$\Phi_k(\lambda, \vec{i}) = \sum_{i=1}^p r_{ki}(\vec{i}) \Phi_i(\lambda, \vec{i}^c) + F_k(\lambda, \vec{i}),$$

$$\rho(\lambda) = \sum_{j=1}^p \rho_j \phi_j(\lambda) + \rho^*(\lambda)$$

où les $\phi(\lambda)$ sont définis comme une base orthogonale dans laquelle : $\Phi_i(\lambda, \vec{i}^c) = \sum_{j=1}^p a_{ij} \phi_j(\lambda)$. Le problème de la

constance des couleurs ne peut être alors résolu que si le dernier terme de l'équation précédente est constant (on peut alors supposer qu'il est nul), ce qui peut être dû à trois types d'hypothèses différentes quant à la structure des illuminants et des réflectances (lesquelles se retrouvent simplement à partir des définitions indiquées) :

1. Les illuminants satisfont une contrainte telle que ce terme est toujours nul, comme $F_k(\lambda, \vec{i}) = 0$.
2. Les réflectances sont de dimensionnalité inférieure à p , et appartiennent à Vect $(\phi_i(\lambda))_{1 \leq i \leq p}$.
3. Les réflectances sont de dimensionnalité finie, mais supérieure à p , et les illuminants satisfont à certaines contraintes assurant que seulement un sous-espace de dimension p des réflectances intervient dans les réponses.

Le premier cas est le seul qui soit simple et pour lequel un algorithme ait été implanté, il est montré que celui-ci est une forme de renormalisation de Von Kries (on multiplie les réponses de chacun des récepteurs par un coefficient indépendant). Les deux autres semblent mener à des algorithmes très compliqués avec des schémas de renorma-

lisation plus généraux dans lesquels un nombre élevé de paramètres intervient, mais permettent néanmoins une discussion intéressante sur les modèles à structures linéaires finies.

Dans les conditions où le problème de la constance des couleurs peut être résolu, c'est-à-dire sous l'une de ces hypothèses, l'équation (22) montre que les réponses des récepteurs (la couleur observée) sont les images par une application linéaire I (l'illuminant), d'une certaine projection sur p vecteurs, des fonctions de réflectances (la couleur propre de la surface).

L'estimation de l'illuminant I est facilitée par plusieurs remarques : la première est que l'ensemble des illuminants plausibles est convexe, ainsi que $\Delta = \{\bar{p}, \forall \lambda, 0 \leq \bar{p}(\lambda) \leq 1\}$, l'ensemble des réflectances physiquement réalisables, et $\bar{\Delta}$, sa projection sur les p vecteurs. La seconde est qu'on peut éviter de déterminer $\bar{\Delta}$, en travaillant à I^c (l'illuminant canonique) près, qui est linéaire et bijective. L'algorithme qui suit permet alors d'exploiter l'idée de base que si l'on observe une réponse n'appartenant pas à $I(\bar{\Delta})$, alors l'illuminant I n'a pas pu être celui de l'image.

L'ALGORITHME DE FORSYTH.

1. Construire $D = I^c(\bar{\Delta})$, par observation d'une gamme aussi étendue de couleurs que possible, sous un illuminant fixe.

2. Estimer l'ensemble des illuminants plausibles :

- déterminer les points extrémaux C_i de l'enveloppe convexe des réponses observées ;
- déterminer les points extrémaux correspondants de l'enveloppe convexe des illuminants plausibles (les $I \in GL(p)$ qui sont diagonalisés⁽²¹⁾ tels que $C_i \in I(D)$) pour chacun de ces points ;
- intersecter ces dernières enveloppes (utilise un algorithme d'approximation par partition de l'espace en cubes).

3. Choisir à l'aide d'un estimateur l'illuminant, par exemple celui donnant naissance à la gamme de couleurs la plus étendue (c'est celui ayant la trace la plus grande).

4. Appliquer l'inverse de I pour obtenir la couleur constante.

L'algorithme a été testé avec succès sur plusieurs images, dans la version décrite. Il devrait se généraliser au cas d'un illuminant variant spatialement par une technique de propagation des contraintes utilisant l'hypothèse du monde de Mondrian, mais ceci n'a pas été encore implanté.

Au-delà du Monde de Mondrian ? Ces derniers travaux me semblent un peu épuiser les possibilités du monde de Mondrian, dans lequel les problèmes sont à présent à peu près résolus. Il faut noter qu'alors que les trois autres algorithmes sont difficiles à analyser, la machinerie théorique de Forsyth a abouti à une simple normalisation de Von Kries ! Cependant, pour ces algorithmes, comme pour tous les précédents, la généralisation à des cas réels

⁽²¹⁾ Ceci met à profit le fait que l'on ait fait une hypothèse de type 1 (contrainte sur les illuminants), et qu'alors on montre que l'ensemble des illuminants est inclus dans un sous-espace de dimension p .

ne sera pas du tout immédiate car beaucoup de facteurs nouveaux seront à prendre en compte, lesquels facteurs étaient supposés, en hypothèses de base, distinguables a priori des changements de réflectance. Les principales difficultés viendront de :

- La dépendance vis-à-vis de la géométrie : on s'attend à ce que des algorithmes fonctionnant avec un illuminant variant spatialement puissent traiter le cas de faibles dépendances continues⁽²²⁾, en revanche les effets dus aux changements d'orientation et aux reflets ne peuvent être pris en compte par ces méthodes.

- La structure de l'illumination, qui peut :

- Varier brusquement (ombres portées⁽²³⁾ introduisant de nouvelles discontinuités).

- Avoir une distribution spectrale variable suivant la position, à cause des reflets, mais surtout des inter-réflexions, qui ont commencées à être étudiées par Forsyth et Zissermann. Ce dernier phénomène fait des valeurs de radiance spectrales mesurées une fonction globale complexe de la géométrie (voir plus haut la fin de la section « modélisation »). Il est bon de noter ici qu'aucun algorithme ne fonctionne avec des illuminants de composition spectrale multiple.

Il semble donc capital de pouvoir distinguer les contours dus à ces nouveaux facteurs de ceux de réflectance.

3.3. APPROCHES PHYSIQUES

Ces algorithmes exploitent un modèle physique du processus de formation des images. Ils présentent l'avantage de ne pas faire l'hypothèse du « monde de Mondrian » mais nécessitent des conditions d'illumination assez contrôlées.

3.3.1. Méthodes fondées sur les surbrillances

Observation directe de la couleur de l'illuminant. D'après plusieurs (voir par exemple Flock [35]), la couleur de l'illuminant serait observée directement dans les spécularités plus ou moins prononcées d'une scène. Une première méthode consiste donc à identifier les surbrillances, dont la couleur, d'après des modèles du type NIR, donne celle de l'illuminant. Elle est illustrée par les calculs suivants, qui reposent sur le modèle des mêmes auteurs :

L'ALGORITHME DE HEALEY ET BINFORD [62].

1. Trouver une frontière spéculaire/diffus sur un objet inhomogène.

2. Mesurer, dans la zone diffuse, $I'(\lambda) = R_p(\lambda) E(\lambda)$.

3. Mesurer, dans la zone spéculaire, $I(\lambda) = (R_s(\lambda) + R_p(\lambda)) E(\lambda)$.

4. Obtenir l'illuminant par $E(\lambda) R_s(\lambda) = I(\lambda) - I'(\lambda)$ en supposant une variation spectrale de R_s faible et une variation spatiale lente de R_p .

Utilisation de plusieurs couleurs d'objets. Une seconde méthode consiste à faire intervenir plusieurs objets (deux

⁽²²⁾ Encore qu'alors l'information géométrique serait perdue par l'opération « d'élimination de l'illuminant ».

⁽²³⁾ A noter aussi les problèmes liés à la dynamique en présence de zones sombres : celles-ci risquent d'être pratiquement noires !

sont nécessaires et suffisants en principe) et à raisonner de manière purement spectrale (et non plus spatialement comme précédemment) grâce à un modèle séparant facteurs spatiaux et spectraux. Klinker, Shafer et Kanade [80], font remarquer qu'en présence d'un illuminant commun, les plans dichromatiques de différents objets, contenant le même vecteur de réflexion de surface, sont sécants le long d'une ligne qui est la couleur de l'illuminant. Ceci nécessite de segmenter la scène pour savoir quelles lumières parviennent de quels objets, ce qui ne va pas de soi puisqu'il est présumé que la couleur joue un rôle important dans une telle segmentation comme le suggère Healey [61]. Une solution par prédiction et vérification d'hypothèses a été donnée dans [78] par ces mêmes auteurs mais elle ne résout cependant pas tous les problèmes de segmentation comme nous le verrons plus bas.

Par un modèle NIR, Lee [88] exploite directement la même idée : dans un diagramme chromatique, les couleurs d'une même surface définissent une droite passant par la couleur de l'illuminant et celle propre de la surface (fig. 17).

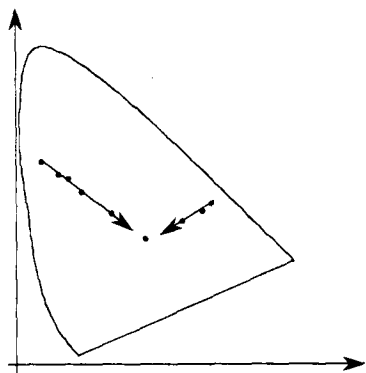


Fig. 17. — La méthode de Lee.

Étant donné deux surfaces de réflectance différente, il y aura deux droites dont l'intersection représentera la couleur de l'illuminant. Des résultats sont présentés uniquement pour une image synthétique assez simple.

L'ALGORITHME DE LEE.

1. Détecter des contours dans trois images G/R, G, B/R.
2. A chaque point de couleur correspondant à un contour, stocker les valeurs G/R et B/R correspondant aux directions de variations d'amplitude maximale (positive et négative) de G.
3. Effectuer une régression linéaire sur ces deux ensembles et conserver une droite si ils sont alignés (ceci est destiné à ne garder que les droites définies par deux parties d'un même objet).
4. Déterminer l'intersection des droites ainsi obtenues (par une méthode d'accumulateur par ex.).

On voit que des hypothèses contenues dans le modèle NIR doivent être faites pour ces deux classes d'algorithmes qui présentent les difficultés suivantes :

- Ils ne peuvent évidemment fonctionner qu'en présence

de surbrillances sur des objets inhomogènes, or il est clair que celles-ci sont absentes des mondes de Mondrian, par exemple.

- Ils doivent distinguer les frontières de specularités des autres contours (voir plus bas une discussion de cette tâche).
- Les mesures de surbrillances sont délicates, entre autres à cause des problèmes posés par la dynamique des caméras.
- Il n'est ni nécessaire ni suffisant de calculer la couleur de l'illuminant pour la constance des couleurs (néanmoins la connaissance de la distribution spectrale d'énergie est suffisante).
- Un contrôle de l'illuminant rigoureux est nécessaire (le modèle dichromatique devrait pouvoir s'accommoder des illuminants multiples, de l'inter-réflexion et des ombres projetées mais aucune réalisation n'a été effectuée avec ces éléments).

3.3.2. Autres méthodes

Aberration chromatique. Funt et Ho [45] utilisent une technique bien particulière : l'aberration chromatique. L'idée est que l'index de réfraction des milieux varie en fait avec la longueur d'onde, si bien que par un effet de prisme, l'image d'un point couleur par une lentille n'est pas exactement un point mais une superposition d'une infinité de disques chacun correspondant à une certaine longueur d'onde. Ce processus de formation donne des motifs complexes, dont on montre, dans deux cas simples (spot sur fond noir, deux régions uniformément colorées adjacentes) qu'ils peuvent être analysés pour obtenir le signal couleur. L'utilisation d'un modèle linéaire (sans limitation), permet alors de calculer les paramètres d'illumination et de réflectance. Cette approche paraît intéressante et ouvre de nouvelles perspectives (cependant l'aberration chromatique des caméras habituelles est faible et les trois canaux ont la même résolution spatiale, contrairement à l'œil), mais il est à craindre que les calculs ne soient pas simples pour des images complexes, de plus un modèle très précis de capteur est nécessaire.

Inter-réflexion. Une autre méthode, due à Drew et Funt [31] consiste à utiliser l'inter-réflexion. L'idée est d'observer sur deux zones voisines, outre les valeurs de radiance provenant de la réflexion directe de l'illuminant sur chacune des zones, celles provenant de l'inter-réflexion au voisinage de la jonction des zones, et d'obtenir ainsi des équations supplémentaires. Pour les réflexions directes, on obtient en utilisant un modèle linéaire comme dans l'équation (19) (les notations sont celles des modèles linéaires) :

$$(23) \quad \vec{S}_1 = \Lambda_{\vec{E}} \cdot \vec{R}_1 \quad \vec{S}_2 = \Lambda_{\vec{E}} \cdot \vec{R}_2$$

Les réflexions multiples sont approximées par une réflexion double⁽²⁴⁾ entre les deux zones, donnant dans la zone d'inter-réflexion les valeurs du signal couleur

⁽²⁴⁾ Ceci signifie que la lumière est réfléchi de la zone 1 (resp. 2) vers la zone 2 (resp. 1), puis directement vers l'observateur.

$I_{mi}(\lambda)$ à partir des signaux couleur des zones sans inter-réflexion $I_i(\lambda)$, des facteurs géométriques (inconnus) α_{ij} , fraction de $I_i(\lambda)$ atteignant la surface j , et des réflectances $R_i(\lambda)$:

$$(24) \quad \begin{aligned} I_{m1}(\lambda) &= I_1(\lambda) + \alpha_{12} I_2(\lambda) R_1(\lambda) I_{m2}(\lambda) \\ &= I_2(\lambda) + \alpha_{21} I_1(\lambda) R_2(\lambda) \end{aligned}$$

En utilisant les fonctions de base, on obtient l'équation :

$$(25) \quad \vec{S}_{m2} = \vec{S}_2 + \alpha_{12} \mathbf{H} \cdot \vec{E} \cdot \vec{R}_1 \cdot \vec{R}_2$$

où $(\mathbf{H})_{i,j,k,\ell} = \sum_{\lambda} E_i R_j R_k S_{\ell}$. Compte tenu du facteur multiplicatif entre \vec{E} et \vec{R}_i , le compte des inconnues et des équations donne la condition $p \geq (m + 2n)/3$, avec p nombre de senseurs, m degrés de liberté de \vec{E} , n degrés de liberté de \vec{R}_i . Elle est satisfaite pour $p = n = m = 3$, on obtient alors facilement α_{21} . L'algorithme consiste à résoudre itérativement le système non linéaire (23) (25). Des tests ont été effectués sur des données synthétiques, mais générées par un modèle (radiosité) prenant en compte les réflexions multiples et un découpage fin de chaque zone. D'un point de vue théorique, cette nouvelle méthode est intéressante : elle permet de conserver 3 degrés de liberté pour l'illuminant et la réflectance, ne nécessite pas de couleurs multiples (l'algorithme marche même si $R_1(\lambda) = R_2(\lambda)$), calcule un facteur géométrique α significatif, et permet de supprimer l'inter-réflexion. Il n'y a pas eu de tests sur images réelles. Dans celles-ci, il faudra détecter les occurrences de la situation physique précise exploitée par la méthode.

3.4. CONCLUSION

Après des perfectionnements successifs, il y a à présent un algorithme complètement opérationnel de calcul de la clarté (*lightness*) dans les scènes monochromes [10]. Cet algorithme ne permet pas d'obtenir l'albédo dans des scènes complexes, car il est induit en erreur par toutes les discontinuités physiques autres que le changement d'albédo. De même, dans le cadre du monde de Mondrian, le problème du calcul de la réflectance semble avoir reçu plusieurs solutions assez satisfaisantes, encore que le cas de l'illumination variable n'ait été implanté que récemment et dans un petit nombre de travaux.

Cependant, on a l'impression que ces deux lignes de recherche se sont un peu éloignées des vrais problèmes de la vision, en se concentrant sur leurs micro-mondes simplifiés. Les algorithmes effectuant directement le calcul de la couleur semblent donc peu utiles dans un contexte réel. Les approches physiques, quant à elles, nécessitent des conditions d'illumination parfaitement contrôlées. Adapter tous ces algorithmes au monde réel nécessitera d'identifier l'origine physique des discontinuités pour segmenter l'image en zones dans lesquelles leurs hypothèses s'appliquent, tâche qui, nous le voyons plus bas, est

ardue. En attendant, en présence des difficultés que nous avons mentionnées plus haut (facteurs géométriques, inter-réflexions), nous craignons que ces algorithmes n'apportent pas grand chose de plus aux étapes ultérieures de *vision* par rapport à l'image brute !

Une remarque générale est que la robustesse des algorithmes n'a pas été évaluée de manière entièrement satisfaisante, même pour ceux fonctionnant sur des images réelles, ce qui limite considérablement leur utilité pratique. En effet, par robustesse, il faut entendre deux choses :

- Exactitude en présence de bruit, d'imperfections diverses des senseurs.
- Exactitude en présence de scènes où les hypothèses simplificatrices des modèles sous-jacents ne sont pas vérifiées exactement.

Seule la première a (parfois) fait l'objet d'attention.

Une seconde remarque est qu'un changement d'illuminant peut affecter sérieusement la qualité des mesures des récepteurs. Forsyth [40] remarque judicieusement que si l'on opère avec des paramètres d'acquisition constants et réglés pour la lumière blanche, sous une lumière rouge, par exemple, la dynamique du signal vert n'est plus que de trois bits, et le système devient à peu près ignorant de ce qui peut être mesuré dans ce canal. Un système de constance des couleurs devrait contrôler les paramètres de ses transducteurs en fonction de son estimation de la couleur de la lumière. C'est ce que fait le système humain, avec le phénomène de l'adaptation.

4. Utilisation de l'information couleur

Cette section présente essentiellement les utilisations qui ont été faites de la couleur dans des tâches de vision *universelles* (et donc à bas niveau). Celles-ci comprennent tout ce qui peut être calculé sans connaissances a priori autres que celles générales au monde physique, et donc valables pour tout type de scène. Nous avons distingué d'une part des tâches qui ont déjà largement été étudiées dans les images monochromes et pour lesquelles l'apport de la couleur consiste en une généralisation *quantitative*, d'autre part celles où elle intervient de manière *qualitative*. On attend alors de son utilisation qu'elle permette de déterminer des propriétés physiques intrinsèques des scènes, ce qui suppose l'utilisation de modèles généraux a priori. Dans les deux cas, les problèmes de segmentation et de discrimination sont au premier plan. Les premiers travaux sur la couleur ont en effet eu pour but d'améliorer la segmentation et se concentraient sur la recherche de coordonnées chromatiques efficaces pour des algorithmes dérivés de techniques de reconnaissance des formes, conjointement à l'utilisation de modèles sémantiques. Parmi les travaux les plus récents, le problème de la segmentation en parties physiquement significatives est abordé en utilisant une formalisation du processus de réflectance et traité de manière assez satisfaisante sans faire intervenir d'autre connaissance a priori.

4.1. TROIS VALEURS AU LIEU D'UNE SEULE

4.1.1. Détection de contours d'images couleur

Plusieurs méthodes ont été essayées pour améliorer les techniques fondées uniquement sur l'intensité. Elles peuvent se répartir en trois catégories, suivant le type de représentation adopté :

- L'utilisation d'une métrique particulière sur l'espace des couleurs (correspondant aux résultats de psychophysique, en ce que l'on veut quantifier la notion de différence de couleur perceptible) et la détection des contours par rapport à cette métrique. Cette technique est illustrée par Claxton et Kwok [17], et ramène donc le problème à de la détection en une dimension, moyennant quoi on ne voit pas pourquoi ceci donnerait des meilleurs résultats que la détection sur l'intensité. Une approche utilisant une métrique globale (déformation d'un espace des couleurs de type intensité-teinte-saturation) poserait le problème de la définition d'une dérivée seconde.

- La détection indépendante dans plusieurs images correspondant à des codages de bandes de longueur d'onde différentes par un détecteur classique. Par exemple, Forsyth [43] propose d'utiliser deux images à codage opposé (Bleu/Jaune et Rouge/Vert), et donne quelques arguments intéressants en faveur d'une telle représentation. Il demeure toutefois le problème de la fusion. Des techniques de fusion multisensorielle peuvent être utilisées et devraient conduire à davantage de stabilité et de robustesse.

- L'utilisation combinée de plusieurs composantes. Les alternatives sont de mener tous les calculs de manière vectorielle, ou de faire une combinaison (moyenne, moindres carrés) de différences calculées sur chaque composante. Novak et Shafer sont évoqués par Kanade [76] pour des travaux portant sur les deux techniques et proposent une méthodologie pour construire différents opérateurs couleur n/m en combinant à divers stades :

- effectuer de manière séparée sur chaque composante les n premières étapes,

- combiner avec la norme m ,

- poursuivre à partir de l'étape $n + 1$ avec les valeurs combinées.

Nevatia [99] généralise Huckel par une contrainte de cohérence d'orientation du contour entre les trois composantes, les paramètres d'ajustement de la « marche » pouvant varier indépendamment et le meilleur ajustement étant déterminé par une minimisation aux moindres carrés. Une approche purement vectorielle est utilisée par Coutance, Baron et Briot [24] et Shafer et Novak [76] (le premier donnant peu de détails en dehors du calcul du gradient détaillé plus bas, le second parlant d'une méthode à la Canny). La clé est le calcul du gradient vectoriel (présenté aussi par DiZenno [30]) qui se fait par l'intermédiaire de la matrice jacobienne : soit la fonction image couleur : $f : (x, y) \mapsto \vec{C} = (C_1, C_2, C_3)$, et m la plus grande valeur propre de $J'(f) J(f)$, avec $J(f) = [\vec{C}_x, \vec{C}_y]$. Alors le vecteur propre associé à m donne la direction

produisant la plus grande variation et \sqrt{m} est la valeur de cette variation. Il vient :

$$\theta = \frac{1}{2} \arctan \frac{2 \vec{C}_x \cdot \vec{C}_y}{\|\vec{C}_x\|^2 - \|\vec{C}_y\|^2}$$

$$m = (\vec{C}_x \cos(\theta) + \vec{C}_y \sin(\theta))^2$$

Cependant Novak et Shafer trouvent que l'opérateur $2/\infty$, consistant à (1) calculer les dérivées partielles dans chaque image, (2) calculer la direction et l'amplitude des contours dans chaque image, combiner avec L_∞ , (3) supprimer les non-maxima par seuillage, donne pratiquement les mêmes contours que l'opérateur vectoriel.

La conclusion typique de tous ces auteurs est que ceci n'apporte que peu de nouveaux contours, mais que ceux obtenus sont un peu meilleurs, car correspondant plus souvent à des changements significatifs dans l'image. La plupart des contours détectés en couleur seraient compris dans les contours d'intensité, et la couleur ne serait utile qu'en second ressort, pour affiner la détection. Il est cependant difficile d'évaluer quantitativement ce type de résultats (les auteurs se contentent de présenter côte-à-côte les images avec les segments extraits), ce qui explique les divergences. Et surtout il me semble toutefois que l'intérêt de la couleur n'est pas tant d'améliorer l'extraction de contours que d'enrichir leur description. En fait, les propriétés photométriques des surfaces étant liées aux surfaces, et non aux contours, la couleur est plus naturellement attachée aux régions, si bien que c'est le problème dual qui a surtout attiré l'attention.

4.1.2. Segmentation d'images couleur

La couleur étant une information que l'on attache naturellement aux régions, on a pensé assez tôt à l'utiliser pour segmenter des images, avec l'espoir que les régions obtenues soient davantage significatives : en effet la couleur est un attribut plus stable⁽²⁵⁾ que la seule luminance et constitue aussi une information plus riche.

La couleur comme attribut vectoriel. Une première famille de méthodes consiste à utiliser la couleur comme un attribut vectoriel. Il s'agit alors d'une simple généralisation de méthodes monochromatiques, qui ne posent que le problème du passage à l'espace tridimensionnel. Le problème supplémentaire posé par la couleur consiste surtout dans la recherche d'une métrique adéquate, ce qui passe par le choix d'axes de coordonnées couleur. Plusieurs systèmes sont décrits par Ohlander [101], mais c'est Ohta [102] qui systématise, par une méthode que nous avons déjà décrite, la recherche des coordonnées les plus efficaces. Les algorithmes, que nous ne décrivons pas car le concept de couleur n'y joue pas un rôle important, se répartissent en deux catégories :

- Algorithmes globaux, utilisant les histogrammes : division par seuillage [101], détection de groupements (*clusters*) dans l'espace des couleurs par projection de l'histo-

⁽²⁵⁾ En un sens qui est ici assez vague, et que nous précisons lors de l'examen des méthodes de segmentation « physique ».

gramme sur des plans [112] ou sur des droites [16], [15], extraction dans trois histogrammes de domaines de validité de classes [117].

- Algorithme fondé sur les relations géométriques, utilisant un critère d'uniformité : séparation et fusion (*split and merge*) de régions [49], [24].

Il faut remarquer que ces algorithmes ne donneront pas toujours des segmentations significatives : des petites différences causées uniquement par la géométrie peuvent provoquer la séparation de régions uniformes, comme l'illustre le cylindre présenté par Ohta [102].

Utilisation de champs markoviens. Une seconde famille, plus récente, considère une image couleur comme un ensemble de trois images monochromes couplées par l'intermédiaire de champs markoviens. Le formalisme des champs markoviens permet pour une image d'exploiter les propriétés de continuité par morceaux de la couleur en utilisant seulement des relations locales, et, par l'emploi de probabilités bayésiennes, de traiter explicitement le bruit en détectant des discontinuités. De plus, un triplet peut être corrélé au niveau du pixel, sans nécessiter de corrélation directe entre les différentes valeurs de ceux-ci (ce qui permet l'utilisation de représentations diverses de la couleur, testées par Dailly [26] qui discute aussi de l'influence du type de treillis utilisé, ou même de données de senseurs différents comme les 4 modules utilisés par l'équipe du M.I.T. [47]), et on peut pour cela imposer aux discontinuités de coïncider plus ou moins fortement d'une image à l'autre, comme le montre Wright [126] qui illustre les effets de diverses valeurs de couplage. Au passif de ces approches, il faut signaler qu'il y a encore une dépendance vis-à-vis de paramètres (discutés [26], ou choisis empiriquement [47]) et surtout la nécessité de techniques d'optimisation globale (recuit simulé, techniques de Monte-Carlo) qui sont vraiment très coûteuses⁽²⁶⁾.

4.1.3. Stéréo

L'utilisation de la couleur permet d'obtenir des attributs supplémentaires qui doivent permettre de lever des ambiguïtés lors de la procédure d'appariement. Jordan et Bovik [72] rendent compte d'une expérience de dénombrement d'appariements (sur 3 paires d'images) où ils comparent ceux obtenus sur des segments (zéros de Laplacien de gaussiennes) avec les seules contraintes classiques de proximité, signe de contraste, et d'orientation, et ceux obtenus en rajoutant la contrainte de signe de gradient chromatique dans des images R-G, G-B, B-R normalisées, et rapportent d'une part une réduction significative aussi bien du nombre des segments intervenant dans au moins une hypothèse d'appariement (- 40 %) que du nombre total d'hypothèses (- 60 %) et d'autre part une augmentation du nombre d'hypothèses sans ambiguïté (+ 70 %) et de celui des appariements corrects (+ 1 %). Coutance, Baron et Briot [24] ont effectué des travaux similaires en utilisant toutefois des segments détectés dans les images

couleur, et notent une réduction de 10 % du nombre d'hypothèses, mais des résultats finaux à peu près identiques. Auparavant une utilisation de la couleur dans un système stéréo spécialisé pour le traitement d'objets filiformes colorés par Demazeau et Gordillo [29] avait donné des résultats satisfaisants, mais il s'agissait d'une application très particulière.

Le travail le plus exhaustif sur ce sujet est celui de Brockelbank et Yang [12], qui utilisent une approche consistant à combiner différents opérateurs à différents stades. Ceux-ci sont l'extraction de contours, la création d'ensembles d'attributs, et la mise en correspondance (méthode assez compliquée, par relaxation et étiquetage). Les modèles examinés, en plus du modèle achromatique sont :

- *Trois canaux* : le traitement du modèle achromatique est appliqué à chacun des canaux R, G, B.
- *Trichromatique* : la détection des contours est effectuée dans les images R, G, B, puis ceux-ci sont combinés pour créer un ensemble de contours avec attributs chromatiques pris en compte lors de l'étape de mise en correspondance.
- *Couleurs antagonistes* : les images R, G, B sont combinées en images à couleurs antagonistes. Le traitement se poursuit comme dans le modèle trichromatique.

Les tests ont été effectués sur trois images : un stéréogramme aléatoire avec couleurs décorréliées, un stéréogramme aléatoire à contours corrélés entre les couleurs, et une scène naturelle. A cause de la nature différente de ces images, les résultats ne sont pas très réguliers (le rapport des contours appariés correctement et du nombre de contours détectés est, pour l'image 1, 48 % pour le modèle achromatique et 73,5 % pour les modèles chromatiques, pour l'image 2 respectivement 90 %, 55 %, 75 %, 91 %, quant à l'image 3, le pourcentage d'appariements est toujours de l'ordre de 47,5 %). Les auteurs fournissent cependant d'autres décomptes et concluent que :

- Le modèle achromatique détecte beaucoup moins de contours (c'est la principale différence aux yeux des auteurs, mais ceci me semble un artéfact dû aux images choisies : il reste à mener une étude détaillée avec un nombre important d'images réelles) que les autres. La mise en correspondance semble à peu près identique (ce qui n'est pas très encourageant).
- Le modèle à trois canaux donne des résultats qui peuvent être peu cohérents entre ceux-ci.
- La représentation des attributs est meilleure en couleurs antagonistes qu'en couleurs trichromatiques : les composantes R, G, B sont à bande plus étroite que les composantes antagonistes, et donc peuvent « manquer » des contours que les autres auraient permis de détecter⁽²⁷⁾. Ce point confirme l'intérêt d'un système de représentation « Intensité/Attribut-couleur1/Attribut-couleur2 ».

Le type d'approche utilisé pour la stéréo devrait aussi se révéler utile pour tous les algorithmes fondés sur des mises

⁽²⁶⁾ Daily [26] mentionne 60 heures de calcul avec 8 machines Lisp travaillant en parallèle, Wright [126] n'utilise que des images 32 x 32, quant au M.I.T. [47], ils précisent que les algorithmes tournent beaucoup plus vite sur leur *connexion machine*, mais ne donnent pas les temps !

⁽²⁷⁾ Soit par exemple une image avec une discontinuité rouge détectable et une discontinuité verte tout juste en dessous du seuil de détection. Le modèle trichromatique ne détecte que la composante rouge du contour, alors que le modèle à couleurs opposées donne, dans sa représentation, la discontinuité de luminance et celle du canal R/G.

en correspondance entre plusieurs images, en particulier ceux de l'analyse du mouvement et ceux de fusion de plusieurs vues. On peut s'attendre avec certitude à une réduction de la combinatoire de ces algorithmes, mais la question qui reste ouverte est de savoir si elle justifie le prétraitement supplémentaire.

4.2. UNE PISTE VERS LES PROPRIÉTÉS INTRINSÈQUES DES SCÈNES ?

4.2.1. Détection des surbrillances

L'intérêt de détecter les surbrillances est de pouvoir disposer séparément de celles-ci et de la couleur mate, intrinsèque, de l'objet.

- Les surbrillances peuvent se révéler utiles pour le calcul de la géométrie : le calcul de leur disparité permet de fortes contraintes sur les surfaces [6], [7], et l'utilisation du modèle de Cook et Torrance donne accès à la courbure locale, dans le cas d'un cylindre [60]. Cependant nous pensons que dans les scènes naturelles les surbrillances ne sont pas suffisamment denses si bien que les techniques utilisant la vision active pour en tirer le maximum d'information [111], [130] nous semblent plus conséquentes. D'un autre côté, les surbrillances, de par leur dépendances géométriques, posent des problèmes aux algorithmes classiques de mise en correspondance d'images (stéréo et mouvement), si bien qu'il peut être intéressant de les éliminer.

- — La couleur mate est beaucoup plus utile pour le calcul de la géométrie par les nombreuses techniques de reconstruction de la forme à partir des dégradés (*shape from shading*, voir Horn [65] pour une présentation détaillée) qui supposent qu'on a affaire à des surfaces lambertiennes. Ces techniques, d'ailleurs, pour autant que nous ne nous soyons pas trompés, ne retirent pas de bénéfice particulier de l'emploi de la couleur.

— Dans le cadre de modèles de type dichromatique, une représentation normalisée de la couleur du type de celle donnée par Healey et Binford [59], [57] fait disparaître les dépendances dues aux facteurs géométriques pour préserver uniquement celles dues à l'illuminant et au matériel, permettant une sorte de segmentation physique.

Les surbrillances diffèrent des parties mates par deux faits :

- Leur *chromaticité* se rapproche de celle de l'illuminant.
- Leur *intensité* est notablement plus grande.

Récemment un certain nombre de techniques de détection fondées sur la couleur ont été proposées avec un certain succès. Il faut remarquer au passage leur lien avec des techniques de constance des couleurs.

Exploitant parfaitement le modèle dichromatique, Klinker, Shafer et Kanade [80], [79] remarquent que si l'on effectue un histogramme des lumières provenant d'un objet constitué d'un matériau unique, les points correspondant à la réflexion de surface et ceux correspondant à la réflexion profonde se situent sur deux segments de droite différents formant un T comme le montre la figure 18 et proposent un algorithme purement spectral pour identifier ces deux segments.

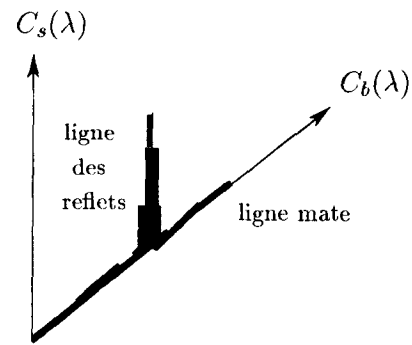


Fig. 18. — La répartition des points dans le plan dichromatique (d'après Klinker).

L'ALGORITHME DE KLINKER.

1. Effectuer l'histogramme en projetant tous les pixels dans le cube RGB.
2. Déterminer le plan dichromatique et y projeter les pixels.
3. Déterminer le segment des surbrillances et celui des points de couleur mate (techniques de reconnaissance de formes assez délicates) ; ceci donne les vecteurs \vec{C}_p et \vec{C}_s .
4. Calculer les coordonnées des pixels dans cette base, qui sont la composante mate et surbrillante.

Cet algorithme nécessitait la segmentation préalable de l'image en zones de composition matérielle constante, mais une extension présentée dans [78] utilisant un algorithme assez compliqué de prédiction et vérification effective celle-ci automatiquement et donne de jolis résultats sur des images bien contrôlées. Néanmoins les limitations conceptuelles de cette approche sont celles du modèle dichromatique. La méthode de segmentation ne pourra distinguer des objets aux lignes mates colinéaires ni analyser des zones sans surbrillances. De plus, généralement, les histogrammes ne seront pas des T, vu les changements de couleur d'illumination et de composition matérielle.

L'approche de Gershon, Jepson et Tsotsos [50] repose sur la simple observation que la transition d'une région diffuse à une région spéculaire s'accompagne d'un changement de couleur vers celle de l'illuminant résultant en une sorte d'inflexion dans le cube des couleurs illustrée par la figure 19. Les hypothèses quant au modèle de réflexion sont moins contraignantes, mais l'algorithme nécessite la constance des couleurs, ce qui, nous l'avons vu, ne va pas de soi dans les scènes naturelles. Pourtant de bons résultats sont présentés sur quelques images naturelles.

L'ALGORITHME DE GERSHON.

1. Segmenter en régions de couleur uniforme.
2. Appliquer un algorithme de constance des couleurs.
3. Ajuster les points de chaque région à une droite.
4. Repérer entre chaque paire de régions adjacentes si l'« inflexion » se produit, et si elle a bien lieu dans la direction du miroir parfait.

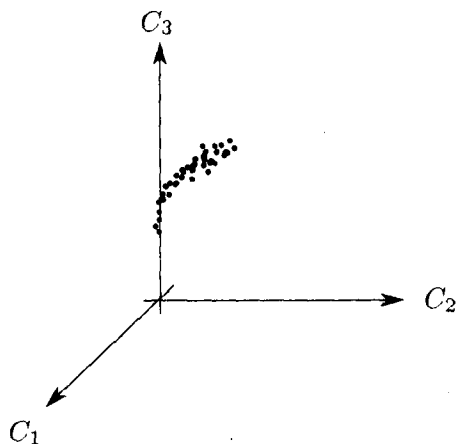


Fig. 19. — Inflexion dans le cube des couleurs (d'après Gershon).

Toutes ces méthodes sont malheureusement rendues coûteuses par le recours aux histogrammes et dépendent de techniques de segmentation qui leur ôtent beaucoup de caractère local.

Il semble qu'il soit suffisant de travailler de manière achromatique, sans évoquer de modèles explicites de spécularité. Brelstaff et Blake [11], argumentent que les surfaces maintiennent un caractère à peu près lambertien en dehors des spécularités, c'est-à-dire que les déviations par rapport à une fonction BSRDF f_r , constante sont au plus de 3, pour toutes directions d'illumination et d'observation. Deux types de tests sont alors proposés : les premiers détectent les régions trop brillantes pour être lambertiennes (nécessite un calcul de la clarté par une technique de type Horn-Blake), les seconds les pics d'irradiance trop aigus pour être lambertiens.

Notons, pour clore cette section, une approche [125] utilisant les propriétés de polarisation de la lumière réfléchie, qui fonctionne quelle que soit la ou les couleurs des objets et illuminant, et aussi bien pour les diélectriques que les métaux.

4.2.2. Segmentation matérielle

Par opposition aux algorithmes de segmentation présentés précédemment, ce qui est recherché ici est un découpage physiquement significatif, en zones de composition matérielle constante (en fait de réflectance constante). La dernière version de l'algorithme de Klinker, Shafer et Kanade, que nous avons déjà évoquée remplit aussi cette fonction. Mentionnons en outre deux autres travaux qui sont tous les deux fondés sur l'utilisation conjointe de la couleur normalisée et de l'équation d'intensité 4.

Le but des travaux de Healey est de classifier les surfaces suivant leur composition matérielle. L'idée de base est que si on fait l'hypothèse d'une équation d'intensité du type 4, la normalisation

$$\frac{\rho(\lambda, \vec{r}) F(i, e, g) E(\lambda)}{\|\rho(\lambda, \vec{r}) F(i, e, g) E(\lambda)\|} = \frac{\rho(\lambda, \vec{r}) E(\lambda)}{\|\rho(\lambda, \vec{r}) E(\lambda)\|}$$

permet d'éliminer les dépendances géométriques et donc d'obtenir sous l'hypothèse d'un illuminant unique, $\rho(\lambda, \vec{r})$ caractéristique de la surface. On note que cette propriété est conservée par intégration, ce qui permet de tenir compte du modèle de senseur 5. Après l'utilisation d'une simple technique de « clustering » dans l'espace des senseurs sur des régions déjà segmentées [61], un algorithme plus sophistiqué est décrit (sans résultats) [58] : après une détection de contours, il examine par division des régions de plus en plus petites de l'image, jusqu'à trouver des régions ne contenant pas de contour, qui sont supposées de composition matérielle constante. Une technique de classification statistique permet alors soit de les rattacher à une catégorie matérielle déjà trouvée, soit d'en rajouter une à la liste. L'algorithme n'est pas prévu pour fonctionner en présence de reflets et d'ombres.

Les travaux de Hurlbert et Poggio [68] utilisent aussi les contours d'intensité pour guider la segmentation qui est également fondée sur le modèle de l'équation 4. La couleur est représentée par les coordonnées normalisées

$$u = \frac{R}{R + G + B} \quad \text{et} \quad v = \frac{G}{R + G + B}$$

qui sous l'hypothèse de surfaces lambertiennes soumises *en tout point* à une source de lumière de spectre constant⁽²⁸⁾, sont caractéristiques des surfaces. Le but est donc de trouver les discontinuités de u ou v en s'aidant des discontinuités d'intensité. Il faut lutter contre le bruit (important avec la couleur normalisée, rappelons-le), grâce à la régularité par morceaux. Des techniques fondées sur les champs markoviens ont été examinées, mais compte tenu de la complexité, il a été préféré finalement un algorithme à réseau. Le réseau fonctionne de manière itérative dans l'image des valeurs de u (ou v) en remplaçant chaque valeur de pixel par la moyenne de ses voisins à condition qu'ils ne soient pas séparés par un contour et que leurs valeurs soient à peu près similaires. Il en résulte un moyennage qui supprime le bruit pour conduire à des régions de couleur constante, par un processus de diffusion bloquée par les contours assez proche de la perception humaine. Des bons résultats sont présentés sur des images simples, mais il est clair que les limitations de cet algorithme sont celles du modèle sous-jacent.

4.2.3. Origine physique des contours

Il s'agit du problème de classifier les contours détectés comme discontinuités de l'intensité, suivant le type de changement physique qui leur a donné naissance. L'intérêt de cette tâche est bien sûr de donner accès à des propriétés intrinsèques des scènes, l'idée de partir des contours se justifiant par le fait qu'ils correspondent effectivement à des discontinuités physiques (ils sont stables par les problèmes liés aux senseurs et à l'interréflexion, qui sont des distorsions continues). Notons aussi qu'elle nous semble indispensable pour résoudre le problème de la constance des couleurs dans les scènes naturelles, étant

⁽²⁸⁾ Si cette hypothèse était exacte, l'algorithme fonctionnerait bien en présence d'ombres, mais il semble qu'en général la composition spectrale de l'illumination directe ne soit pas la même que celle de l'illumination indirecte reçue par une surface dans l'ombre, si bien que la couleur normalisée changera de part et d'autre.

donné les limitations des algorithmes actuels que nous avons largement discutées plus haut.

Considérons l'équation d'intensité résultant du modèle NIR⁽²⁹⁾ :

$$(26) \quad I(\lambda, \vec{r}) = (m_s(i, e, *g) + m_p(i, e, g) C_p(\lambda, \vec{r})) E(\lambda, \vec{r})$$

Chacun des termes de cette équation, peut varier brusquement⁽³⁰⁾, donnant les discontinuités des types suivants :

- Spéculaire : discontinuité de $m_s(i, e, g)$ entre une zone dans laquelle il est négligeable et une zone où il ne l'est pas.
- Orientation : discontinuité de $m_p(i, e, g)$.
- Réflectance (marques de couleur) : discontinuité de $C_p(\lambda, \vec{r})$.
- Illuminant (ombres) : discontinuité de $E(\lambda, \vec{r})$.

Horn [66] donne une discussion générale du problème en indiquant quelles peuvent être les causes de changement dans les images, mais ne fournit pas de méthode de classification. Très peu de travaux ont été effectués sur ce problème. Il est à peu près connu que :

- de part et d'autre d'un contour de réflectance, la direction du gradient d'intensité est continue, et le quotient des intensités est identique à celui des gradients ;
- de part et d'autre d'une ombre, la couleur propre⁽³¹⁾ (celle que donnerait un algorithme de constance des couleurs⁽³²⁾) est constante, la texture continue, et le contraste élevé.

L'idée que la couleur peut servir pour donner des informations sur la nature physique des contours a d'abord été mise en avant par Rubin et Richards [109]. La méthode repose sur l'utilisation de l'équation d'intensité 4 et du signal couleur $I(\lambda)$ et $I'(\lambda)$ de part et d'autre du contour. Chaque type de contour provoque une variation dans l'un des termes de 4. Pour un contour :

- dû à une ombre, du côté ombré, l'illuminant est une composante diffuse alors que du côté éclairé s'ajoute en plus la composante directe. On donne une relation additive $I(\lambda) = I'(\lambda) + k(\lambda)$ où $k \geq 0$;
- dû à un changement d'orientation, seul le terme F , qui ne dépend pas de λ varie, d'où une relation multiplicative entre $I(\lambda)$ et $I'(\lambda)$;
- dû à un changement de pigmentation (sur un même matériau), les deux albédos $\rho(\lambda)$ et $\rho'(\lambda)$ sont tous deux une puissance d'une même fonction (loi de Beer).

La remarque importante (fig. 20) est que ces variations sont caractérisées par le fait que l'on a à toute longueur d'onde λ , $I(\lambda) \geq I'(\lambda)$ (ou vice-versa), et que l'on peut donc distinguer ainsi les contours dûs aux trois causes

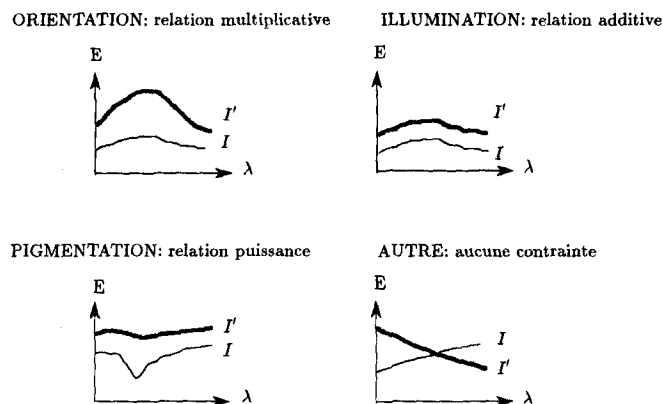


Fig. 20. — Relations entre les distributions spectrales d'énergie selon les types de contour (d'après Rubin et Richards).

mentionnées des contours dûs aux changements matériels par l'application d'un opérateur très simple (analogue aux cellules doublement antagonistes) testant si $I(\lambda_1) > I'(\lambda_1)$ et $I(\lambda_2) < I'(\lambda_2)$.

Cette idée est intéressante par son utilisation directe de l'équation d'intensité, et la technique de comparaison de signes, qui devrait permettre de remédier robustement au fait que les comparaisons exactes sont rendues difficiles par tous les problèmes attachés aux techniques devant utiliser des régions. Néanmoins elle n'a pas été testée par une simulation, qui permettrait de voir dans quelle mesure les difficultés qui suivent peuvent être surmontées :

- Les senseurs n'ont pas directement accès à la valeur du signal couleur en une longueur d'onde.
- Le modèle physique utilisé est assez peu réaliste, et l'hypothèse sur l'illumination est fautive en présence d'inter-réflexion, comme le montre Gershon [49].

Gershon [49] a proposé une solution qui ne présente pas ces inconvénients mais l'algorithme dépend du choix d'un seuil qui doit lui être donné a priori. La difficulté qu'il met en évidence⁽³³⁾ illustre bien d'une part la raison pour laquelle il serait souhaitable de disposer de constance des couleurs, et d'autre part la difficulté de trouver un algorithme opérant dans des scènes de ce type.

Un travail d'un autre type sur ce sujet est celui de Witkin [124], qui n'utilise que l'intensité. Son idée est que les propriétés physiques varient en général continûment, et que celles de parties distantes de scènes sont en général indépendantes. La méthode consiste à construire une famille de courbes parallèles au contour, et à effectuer une mesure de corrélation entre elles. Une faible corrélation indique un contour de profondeur, et une bonne corrélation avec changement des paramètres linéaires indique un changement d'illuminant (fig. 21). Cette technique est difficile à évaluer (elle n'est pas entièrement implantée) et présente l'inconvénient d'utiliser des histogrammes.

Récemment, Gamble, Geiger, Poggio et Weinshall [47] ont proposé une nouvelle méthodologie. A l'aide d'un

⁽²⁹⁾ On utilise souvent par simplicité l'équation (4), ce qui revient simplement à ignorer le terme $m_s(i, e, g)$.

⁽³⁰⁾ Les discontinuités de profondeur se combinent toujours au moins avec une des trois dernières discontinuités que nous mentionnons, et, dans le cas le plus général, avec les trois.

⁽³¹⁾ Assimilée incorrectement parfois à la teinte (Hue).

⁽³²⁾ Ceci montre pourquoi constance des couleurs et étiquetage doivent être effectués de pair, étant interdépendants.

⁽³³⁾ En exhibant une image où une bonne partie de l'illumination diffuse est due à un objet fortement coloré, si bien que sa couleur est bien distincte de celle de l'illumination directe.

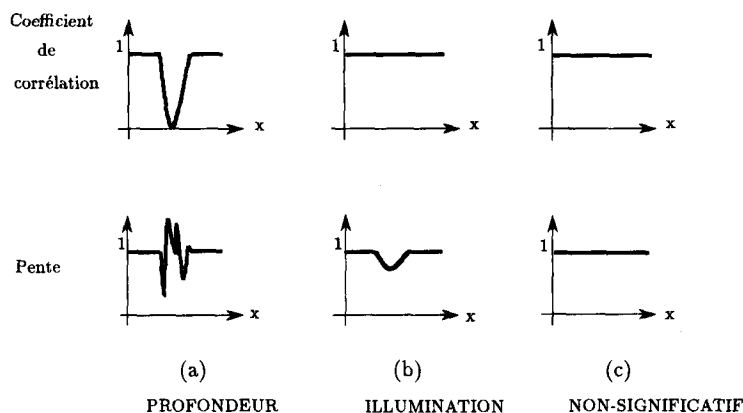


Fig. 21. — Coefficient de corrélation et paramètres de régression selon le type de contour (d'après Witkin).

modèle fondé sur les champs de Markov, la sortie de chacun des modules de stéréo, mouvement, couleur et texture est traitée pour restituer les données bruitées ou manquantes, puis pour trouver les discontinuités du module en les alignant avec celles obtenues par le détecteur de contours. On obtient ainsi une représentation des discontinuités, étiquetées en termes de leur effets observés. Plutôt que d'employer a priori des connaissances dérivées de l'équation d'intensité (et du modèle sous-jacent) du type de celles que nous mentionnons plus haut pour obtenir l'étiquetage en fonction de l'origine physique, une approche heuristique est adoptée. Ces auteurs utilisent un classificateur linéaire dont les entrées sont, outre les 4 unités signalant présence, absence, ou indétermination quant aux discontinuités détectées par les modules, des unités représentant l'intensité moyenne de part et d'autre du segment considéré et la longueur de celui-ci, et les sorties des unités représentant chacun des 5 types de discontinuité physique. Ils ont effectué un apprentissage sur une image de nature morte, ce qui correspond essentiellement à l'ajustement de seuils, et montré qu'alors l'étiquetage réalisé par le réseau était bon pour des images de composition très voisine. En conséquence, les limites de cette approche sont celles, bien connues, des réseaux à apprentissage. Il nous semble que les bons résultats présentés sont davantage dûs à la présence au départ d'une information (relativement) dense et redondante obtenue par les 4 modules de bas niveau, qu'à la technique d'étiquetage elle-même.

4.3. CONCLUSION

Voyons à présent d'un point de vue global l'apport de la couleur à la vision par ordinateur. On peut considérer que cette discipline a pour but de calculer des descriptions intrinsèques de l'environnement, à partir de capteurs visuels. L'essentiel de la recherche actuelle s'est concentrée sur les représentations de bas niveau, et porte donc sur les thèmes suivants :

- extraction d'attributs
- segmentation
- reconstruction 3D

- analyse du mouvement
- représentation des formes
- reconnaissance d'objets.

La couleur constitue un attribut important, pour peu que l'on puisse accéder à sa valeur intrinsèque. C'est le problème de la constance des couleurs, dont nous avons vu qu'il est très difficile dans les environnements réels. C'est aussi la question de l'extraction de composantes de réflectance, et des reflets. Pour ce qui est des attributs et segmentations pouvant être obtenus par des procédés monochromatiques, il semble que l'ajout de la couleur permette d'obtenir des améliorations réduites mais néanmoins sensibles, que l'on peut attribuer à une simple augmentation d'information, sans qu'un progrès conceptuel notable soit effectué, nous semble-t-il. Les travaux reposent surtout sur la colorimétrie, des choix judicieux de représentations chromatiques s'avèrent utiles, dans la mesure où ils déterminent la qualité des attributs ou la segmentation. Comme les prétraitements peuvent être effectués en parallèle de manière évidente, dans une perspective d'application en temps réel, ce n'est pas sans intérêt. De manière générale, l'augmentation de la complexité des algorithmes est faible (linéaire). Cependant le coût du matériel est multiplié au moins par trois, sans compter les problèmes supplémentaires liés à la moins bonne qualité des capteurs ou à leur difficulté de mise en œuvre. En revanche, l'emploi de formalisations de la réflectance peut permettre des segmentations qui sont physiquement fondées, et donc beaucoup plus utiles. Nous avons vu qu'il commence à y avoir des algorithmes opérationnels dans ce domaine.

En ce qui concerne des problèmes purement géométriques, tels que la reconstruction 3D et l'analyse du mouvement, la couleur peut être d'une part traitée comme un simple attribut, servant à guider les processus (géométriques) d'appariements. Dans ce cas, son rôle peut se révéler utile grâce à sa stabilité et à son indépendance relative vis-à-vis de la géométrie, mais est d'un intérêt relativement secondaire puisque la couleur n'intervient plus qu'en tant que valeurs numériques sans significations physiques directes. On peut, d'autre part, considérer qu'il y a une interaction entre les processus de forme et ceux de couleur. Là alors, les progrès *qualitatifs* ne résulteront, nous semble-t-il, que d'algorithmes fondés sur des modèles et destinés à inférer des propriétés physiques, et en particulier géométriques, intrinsèques. L'idée de calculer conjointement forme (orientation) et réflectance (albedo) pour des objets lambertiens achromatiques a donné lieu déjà à une recherche abondante (ce sont les algorithmes de « shape form shading » : voir Horn [65] pour une présentation détaillée). Ces algorithmes ont toujours nécessité des conditions expérimentales bien contrôlées, et ce n'est que récemment [98] que l'inter-réflexion a pu être prise en compte. Comment généraliser au cas d'objets colorés ? La première difficulté importante qui vient à l'esprit est de taille, c'est la modélisation des environnements réels. Une seconde difficulté majeure que nous voyons est que pour ce type de tâche, contrairement au cas des familles précédentes d'algorithmes, une certaine forme de constance des couleurs est nécessaire. Or les discontinuités que

l'on cherche à identifier (en utilisant entre autres la couleur) sont celles qui posent problème aux algorithmes actuels de constance des couleurs... Ce n'est qu'en surmontant ces difficultés que l'on rattrapera à la question : à quoi peu vraiment servir la couleur, de manière universelle ?

Assez paradoxalement, alors que le dernier thème, la reconnaissance d'objets, suggère une application intuitive de la couleur, depuis les premiers travaux allant dans cette direction [101], peu de progrès ont été effectués. Récemment, Swain et Ballard [114] ont montré qu'un objet multicolore peut être représenté de manière très stable, vis-à-vis des changements de points de vue, par l'histogramme de ses couleurs, et qu'une technique très simple d'intersection d'histogrammes permet d'effectuer une recherche efficace dans une grande base de modèles, à partir de l'histogramme d'un objet observé dans la scène. Une opération dérivée permet aussi la localisation rapide d'un objet donné. La couleur pourrait donc être un attribut dont l'intérêt serait de premier ordre, pour ces tâches importantes. Cependant, une telle approche nécessite un algorithme de constance des couleurs, ce qui expliquerait les faibles progrès réalisés dans cette ligne de recherche. La couleur pourrait jouer un rôle très important en vision par ordinateur, mais le problème de son calcul doit être auparavant résolu.

Remerciements

Ce travail a été rendu possible grâce au service de documentation de l'INRIA-Sophia. Olivier Faugeras, Xavier Lebegue et Régis Vaillant ont relu entièrement les versions préliminaires.

Manuscrit reçu le 17 mai 1990

RÉFÉRENCES

- [1] M. ALI, W. MARTIN and J. AGGARWAL, Color-based computer analysis of aerial photographs. *CVGIP*, 9:282-293, 1979.
- [2] J. BAJON, M. CATTOEN and L. LLANG, Identification of multicoloured objects using a vision module. In *Proc. Sixth int. conf. on robot vision and sensory controls*, pages 21-30, 1986.
- [3] D. H. BALLARD and C. M. BROWN, *Computer vision*. Prentice-Hall, 1982.
- [4] P. BECKMANN and A. SPIZZICHINO, *The scattering of electromagnetic waves from rough surfaces*. The Macmillan Company, 1963.
- [5] A. BLAKE, Boundary conditions for lightness computation in mondrian world. *CVGIP*, 32:314-327, 1985.
- [6] A. BLAKE, Inferring surface shape by specular stereo. In *Proc. IJCAI*, pages 973-976, 1985.
- [7] A. BLAKE and G. BRELSTAFF, Geometry from specularities. In *Proc. ICCV*, pages 394-403, 1988.
- [8] M. BORN and E. WOLF, *Principles of optics*. Pergamon Press, 1959.
- [9] D. H. BRAINARD and B. A. WANDELL. Analysis of the retinex theory of color vision. *JOSA-A*, 3(10):1651-1661, 1986.
- [10] G. BRELSTAFF and A. BLAKE, Computing lightness. *Pattern Recognition Letters*, 5:129-138, 1987.
- [11] G. BRELSTAFF and A. BLAKE, Detecting specular reflection using lambertian constraints. In *Proc. ICCV*, pages 297-302, 1988.
- [12] D. BROCKELBANK and Y. H. YANG, An experimental investigation in the use of color in computational stereopsis. *IEEE-PAMI*, 19(6):1365-1383, 1989.
- [13] C. M. BROWN, *Color vision and computer vision*. Technical Report TR 108, University of Rochester, 1982.
- [14] G. BUCHSBAUM, A spatial processor model for object colour perception. *J. Franklin Inst.*, 310:1-26, 1980.
- [15] M. CELENK, A color clustering technique for image segmentation. *CVGIP*, 52:145-170, 1990.
- [16] M. CELENK, Color image segmentation by clustering and parametric-histogramming technique. In *Proc. IEEE ICPR*, pages 883-886, 1986.
- [17] P. R. CLAXTON and E. K. Y. KWOK, The use of colour to segment and label images. In *Proc. Alvey vision conf.*, pages 295-302, 1987.
- [18] J. COHEN, Dependency of the spectral reflectance curves of the munsell color chips. *Psychon. Sci.*, 1:369-370, 1964.
- [19] M. F. COHEN and D. P. GREENBERG. The hemi-cube: a radiosity solution for complex environments. In *Proc. SIGGRAPH, Comp. Graph 19(3)*, pages 31-40, 1985.
- [20] D. CONNAH and C. FISHBOURNE, The use of color information in industrial scene analysis. In *Proc. 1st ROVISEC*, pages 340-347, 1981.
- [21] S. CONNELLY, *Qualitative approaches to image analysis using color, shading, and highlights*. Technical Report CAR-TR-292, University of Maryland, 1987.
- [22] R. COOK, T. PORTER and L. CARPENTER, Distributed ray tracing. In *Proc. SIGGRAPH, Comp. Graph 18(3)*, pages 137-145, 1984.
- [23] R. L. COOK and K. E. TORRANCE, A reflectance model for computer graphics. In *Proc. SIGGRAPH, Comp. Graph 15(3)*, pages 307-316, 1981.
- [24] R. L. COOK and K. E. TORRANCE, A reflectance model for computer graphics. In *Proc. SIGGRAPH, Comp. Graph 15(3)*, pages 307-316, 1981.
- [24] V. COUTANCE, T. BARON and M. BRIOT, Segmentation of color images in robotics. In *Proc. Reconnaissance des formes et intelligence artificielle*, pages 1115-1122, 1989.
- [25] J. CRISMAN and C. THORPE, Color vision for road following. In *Proc. SPIE Conference on mobile robots*, 1988.
- [26] M. DAILY, Color image segmentation using markov random fields. In *Proc. IEEE CVPR*, pages 304-312, 1989.
- [27] N. W. DAW, The psychology and physiology of colour vision. *Trends in Neurosci.*, 7:330-335, 1984.
- [28] R. L. DE VALOIS, Central mechanisms of color vision. In *Handbook of sensory physiology*, pages 209-253, 1973.
- [29] Y. DEMAZEAU and J. L. GORDILLO, A color stereo system applied to wire identification and localisation. In *Proc. 4th Canadian CAD/CAM and robotics conference*, pages 12.1-12.8, 1985.
- [30] S. DiZENZO, A note on the gradient of a multi-image. *CVGIP*, 33:116-125, 1986.
- [31] M. DREW and B. FUNT, Calculating surface reflectance using a single-bounce model of mutual reflection. In *proc. ICCV*, pages 394-399, 1990.
- [32] M. D'ZMURA and P. LENNIE, Mechanisms of color constancy. *JOSA-A*, 3(10):1662-1272, 1986.
- [33] O. D. FAUGERAS, *Digital color image processing and psychophysics within the framework of a human visual model*. PhD thesis, Univ. Utah, Salt Lake City, June 1976.
- [34] O. D. FAUGERAS, Digital color image processing within the framework of a human visual model. *IEEE ASSP*, 27(4):380-393, 1979.
- [35] H. R. FLOCK, Illumination : inferred or observed ? *Perception and Psychophysics*, 35(3):293, 1984.

- [36] J. D. FOLEY and A. VANDAM, *Fundamentals of interactive computer graphics*. Addison-Wesley, 1983.
- [37] D. FORSYTH, A novel algorithm for color constancy. *IJCV*, 5(1):5-36, 1990.
- [38] D. FORSYTH and A. ZISSERMANN, Mutual illumination. In *Proc. CVPR*, pages 466-473, 1989.
- [39] D. FORSYTH and A. ZISSERMANN, Shape from shading in the light of mutual illumination. *Image and vision computing*, 8(1):42-49, 1990.
- [40] D. A. FORSYTH, Color constancy and its application in machine vision PhD thesis Oxford university, 1988.
- [41] D. A. FORSYTH, Color constancy : a review, 1988, Oxford university.
- [42] D. A. FORSYTH, A novel approach to colour constancy. In *Proc. ICCV*, pages 9-18, 1988.
- [43] D. A. FORSYTH, A system for finding changes in colour. In *Proc. Alvey vision conf.*, pages 285-294, 1987.
- [44] B. FUNT and M. S. DREW, Color constancy computation in near-mondrian scenes using a finite dimensional linear model. In *Proc. IEEE CVPR*, pages 544-549, 1988.
- [45] B. FUNT and J. Ho, Color from black and white. In *Proc. ICCV*, pages 2-8, 1988.
- [46] A. GAGALOWICZ, Collaboration between computer graphics and computer vision. In *Proc. ICCV*, pages 733-739, 1990.
- [47] E. GAMBLE, D. GEIGER, T. POGGIO and D. WEINSHALL, Integration of vision modules and labeling of surface discontinuities. *IEEE SMC*, 19:1576-1581, 1989.
- [48] R. GERSON, *Survey on color : Aspects of perception and computation*. Technical Report RCBV-TR-84-4, University of Toronto, 1984.
- [49] R. GERSON, *The use of color in computational vision*. PhD Thesis, Dept of Comp. Sc.-University of Toronto, 1987.
- [50] R. GERSON, A. D. JEPSON and J. K. TSOSTOSOS, Highlight identification using chromatic information. In *Proc. ICCV*, pages 161-170, 1987.
- [51] C. GORAL, K. TORRANCE, D. GREENBERG and B. BATAILLE, Modeling the interaction of light between diffuse surfaces. In *Proc. SIGGRAPH, Comp. Graph 18(3)*, pages 213-222, 1984.
- [52] J. L. GORDILLO, Caicou : l'analyse d'image couleur. *Traitement du signal*, 4(3):217-228, 1987.
- [53] A. GOTTSALK and G. BUCHBAUM, Information theoretic aspects of color signal processing in the visual system. *IEEE SMC*, 13(5):864-873, 1983.
- [54] S. GROSSBERG, Cortical dynamics of three-dimensional form, color, and brightness perception : 1. monocular theory. *Perception and psychophysics*, 41(2):87-116, 1987.
- [55] S. GROSSBERG, Cortical dynamics of three-dimensional form, color, and brightness perception : 2. binocular theory. *Perception and psychophysics*, 41 (2):117-158, 1987.
- [56] S. GROSSBERG and D. TODOROVIC, A neural network architecture for brightness perception under constant and variable illumination conditions. In *Proc. IEEE CNN*, pages IV 185-192, 1987.
- [57] G. HEALEY, Color discrimination by computer. *IEEE PAMI*, 19(6):1613-1617, 1989.
- [58] G. HEALEY, A parallel color algorithm for segmenting images of 3-d scenes. In *Proc. ARPA image understanding Workshop*, pages 1038-1041, 1989.
- [59] G. HEALEY and T. O. BINFORD, A color metric for computer vision. In *Proc. IEEE CVPR*, pages 10-17, 1988.
- [60] G. HEALEY and T. O. BINFORD, Local shape from specularity. *CVGIP*, 42:62-86, 1988.
- [61] G. HEALEY and T. O. BINFORD, Predicting material classes. In *Proc. ARPA image understanding Workshop*, pages 1140-1146, 1988.
- [62] G. HEALEY and T. O. BINFORD, The role and use of color in a general vision system. In *Proc. ARPA image understanding Workshop*, 1987.
- [63] H. HELSON, Fundamental principles in color vision. 1. the principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. *J. Exp. Psychol.*, 23:439-471, 1938.
- [64] B. K. P. HORN, Determining lightness from an image. *CGIP*, 3:277-299, 1974.
- [65] B. K. P. HORN, *Robot vision*. MIT Press, 1986.
- [66] B. K. P. HORN, Understanding image intensities. *AI*, 2(8):201-231, 1977.
- [67] A. HURLBERT, *The computation of color*. PhD thesis, MIT-AI Lab, 1989.
- [68] A. HURLBERT and T. POGGIO, A network for image segmentation using color. In *Proc. Denver conf. on neural networks*, 1988.
- [69] A. HURLBERT and T. POGGIO, Synthesizing a color algorithm from examples. *Science*, 239:482-485, 1988.
- [70] K. IKEUCHI and K. SATO, Determining reflectance parameters using range and brightness images. In *Proc. ICCV*, 1990.
- [71] T. JAASKELAINEN, J. PARKKINEN and E. OJA, Color discrimination by optical pattern recognition. In *Proc. ICPR*, pages 766-768, 1986.
- [72] J. JORDAN and A. BOVIK, Computational stereo vision using color. *IEEE Control systems magazine*, 31-36, June 1988.
- [73] D. JUDD, Hue, saturation, and lightness of surface colors with chromatic illumination. *JOSA*, 30:2-32, 1940.
- [74] D. JUDD, D. MACADAM and G. WYSZECKI, Spectral distribution of typical daylight as a function of correlated color temperature. *JOSA*, 54:1031-1040, 1964.
- [75] D. JUDD and G. WYSZECKI, *Color in business, science and industry*. John Wiley and Sons, 1975.
- [76] T. KANADE, Image understanding research at cmu. In *Proc. ARPA image understanding Workshop*, pages 35-37, 1987.
- [77] J. KENDER, *Saturation, hue, and normalized color : calculation, digitization effects, and use*. Master's thesis, Dept of CS. Carnegie-Mellon university, 1976.
- [78] G. J. KLINKER, S. A. SHAFER and T. KANADE, Color image analysis with an intrinsic reflection model. In *Proc. ICCV*, pages 292-296, 1988.
- [79] G. J. KLINKER, S. A. SHAFER and T. KANADE, The measurement of highlights in color images. *IJCV*, 2(1):7-32, 1988.
- [80] G. J. KLINKER, S. A. SHAFER and T. KANADE, Using a color reflection model to separate highlights from object color. In *Proc. ICCV*, pages 145-150, 1987.
- [81] J. KOENDERINK, Color atlas theory. *JOSA-A*, 4(7):1314-1321, 1987.
- [82] D. KUAN, G. PHIPPS and A. HSUEH, Autonomous robotic vehicle road following. *IEEE PAMI*, 10(5) : 648-658, 1988.
- [83] E. LAND, Color vision and the natural image. In *Proc. Natl. Acad. Sci.*, pages 115-129, 636-645, 1959.
- [84] E. H. LAND, Recent advances in retinex theory. *Vision Res.*, 26(1):7-21, 1986.
- [85] E. H. LAND, The retinex theory of color vision. *Scient. Am.*, 237:108-128, 1977.
- [86] E. H. LAND and J. J. McCANN, Lightness and retinex theory. *JOSA*, 61:1-11, 1971.
- [87] X. LEBEGUE, D. BAKER and J. AGGARWAL, 2-d and 3-d model-based recognition of man-made objects in outdoor scenes. In *Proc. SCIA*, pages 423-439, 1989.
- [88] H.-C. LEE, Method for computing the scene-illuminant chromaticity from specular highlights. *JOSA-A*, 3(10):1694-1699, 1986.

- [89] H.-C. LEE, E. J. BRENEMAN and C. P. SCHULTE, Modeling light reflection for computer color vision. *IEEE PAMI*, 12(4): 402-409, 1988.
- [90] M. S. LIVINGSTONE and D. H. HUBEL, Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.*, 4:309-356, 1984.
- [91] L. T. MALONEY, Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA-A*, 3(10):1673-1683, 1986.
- [92] L. T. MALONEY and B. A. WANDELL, Color constancy: a method for recovering surface spectral reflectance. *JOSA-A*, 3(1):29-33, 1986.
- [93] D. MARR, *Vision*. Freeman, 1982.
- [94] McCANN, MCKEE and TAYLOR, Quantitative studies in retinex theory. *Bision Res.*, 16:445-458, 1976.
- [95] J. J. McCANN and K. L. HOUSTON, Calculating color sensations from arrays of physical stimuli. *IEEE SMC*, 13(5):1000-1007, 1983.
- [96] M. NAGAO, T. MATSUYAMA and Y. IKEDA, Region extraction and shape analysis in aerial photographs. *CGIP*, 10:195-223, 1979.
- [97] S. NAYAR, K. IKEUCHI and T. KANADE, Extracting shape and reflectance of hybrid surfaces by photometric sampling. In *Proc. ARPA image understanding Workshop*, pages 563-583, 1989.
- [98] S. NAYAR, K. IKEUCHI and T. KANADE, Shape from inter-reflections. In *Proc. ICCV*, pages 1-11, 1990.
- [99] R. NEVATIA, A color edge detector and its use in scene segmentation. *IEEE SMC*, 7(11):820-826, 1977.
- [100] T. NISHITA and E. NAKAMAE, Continuous tone representation of three-dimensional objects taking account of shadows and inter-reflection. In *Proc. SIGGRAPH, Comp. Graph 19(3)*, pages 23-30, 1985.
- [101] R. OHLANDER, K. PRICE and D. RAJ REDDY, Picture segmentation using a recursive region splitting method. *CGIP*, 8:313-333, 1978.
- [102] Y.-I. OHTA, T. KANADE and T. SAKAI, Color information for region segmentation. *CGIP*, 13:222-241, 1980.
- [103] J. PARKKINEN, T. JAASKELAINEN and M. KUITTINEN, Spectral representation of color images. In *Proc. ICPR*, pages 933-935, 1988.
- [104] B. T. PHONG, Illumination for computer generated pictures. *Comm. ACM*, 18, 1975.
- [105] W. PRATT, *Digital Image processing*. John Wiley and Sons, 1978.
- [106] G. RABATEL, *Localisation de fruits en environnement naturel par segmentation d'images polychromes*. PhD thesis, Université de Bordeaux I, 1988.
- [107] J. REICHMAN, Determination of absorption and scattering coefficients for nonhomogeneous media. 1: theory. *Applied Optics*, 12(8):1811-1815, 1973.
- [108] W. RICHARDS and E. A. PARKS, Model for color constancy. *JOSA*, 61:971-976, 1971.
- [109] J. M. RUBIN and W. A. RICHARDS, *Color vision and images intensities: When are changes material?* Technical Report AI Memo 631, Ai Lab-MIT, 1981.
- [110] J. RUBNER and K. SCHULTEN, A regularized approach to color constancy. *Biological cybernetics*, 61:29-36, 1989.
- [111] A. C. SANDERSON, L. E. WEISS and K. N. NAYAR, Structured highlight inspection of specular surfaces. *IEEE PAMI*, 10(1):44-55, 1988.
- [112] A. SARABI and J. K. AGGARWAL, Segmentation of chromatic images. *Pattern recognition*, 13:417-427, 1981.
- [113] T. STOCKHAM, Image processing in the context of a visual model. *IEEE (special issue on picture processing)*, 60:828-842, 1972.
- [114] M. SWAIN and D. BALLARD, Indexing via color histograms. In *proc. ICCV*, pages 390-393, 1990.
- [115] D. TERZOPOULOS, Efficient multiresolution algorithms for computing lightness, shape-from-shading, and optical flow. In *Proc. AAAI*, pages 314-317, 1984.
- [116] C. THORPE, M. HEBERT, T. KANADE and S. SHAFER, Vision and navigation for the carnegie mellon navlab. *IEEE PAMI*, 10(3):362-372, 1988.
- [117] TOMINAGA, Color image segmentation using three perceptual attributes. In *Proc. CVPR*, pages 628-630, 1986.
- [118] K. TORRANCE and E. SPARROW, Theory for off-specular reflection from roughened surfaces. *JOSA*, 57:1105-1114, 1967.
- [119] M. TSUKADA and Y. OHTA, An approach to color constancy using multiple images. In *Proc. ICCV*, pages 385-389, 1990.
- [120] M. TURK, K. MORGENTHAUER, K. GREMBAN and M. MARRA, Vits a vision system for autonomous land vehicle navigation. *IEEE PAMI*, 10(3):342-361, 1988.
- [121] J. R. WALLACE, M. F. COHEN and D. P. GREENBERG, A two-pass solution to the rendering equation: a synthesis of ray tracing and radiosity methods. In *Proc. SIGGRAPH, Comp. Graph 21(4)*, pages 311-320, 1987.
- [122] B. A. WANDEL, The synthesis and analysis of color images. *IEEE PAMI*, 9(1):2-13, 1987.
- [123] C. WESTELIUS and C. WESTIN, A colour representation for scale-spaces. In *Proc. SCIA*, pages 890-893, 1989.
- [124] A. P. WITKIN, Intensity-based edge classification. In *Proc. AAAI*, pages 36-41, 1982.
- [125] L. WOLFF, Using polarisation to separate reflection components. In *Proc. IEEE CVPR*, pages 363-369, 1989.
- [126] W. A. WRIGHT, A markov random field approach to data fusion and colour segmentation. *Image and vision computing*, 7(2):144-150, 1989.
- [127] G. WYSZECKI and W. S. STILES, *Color science, concepts and methods, quantitative data and formulas*. John Wiley and Sons, 1967.
- [128] A. YUILLE, A method for computing spectral reflectance. *Biological cybernetics*, 56:195-201, 1987.
- [129] S. ZEKI, The representation of colours in the cerebral cortex. *Nature*, 284:412-418, 1980.
- [130] A. ZISSERMAN, P. GIBLIN and A. BLAKE, The information available to a moving observer from specularities. *Image and vision comput.*, 7(1):38-42, 1989.