

Appariement couleur de planches de bois par inférence floue.

E. SCHMITT^{1,2} V. BOMBARDIER¹ P. CHARPENTIER¹

¹Centre de Recherche en Automatique de Nancy, CNRS, UMR 7039
Faculté des Sciences – Bd des Aiguillettes – BP 239 – 54506 Vandoeuvre-lès-Nancy Cedex

²LuxScan Technologies, Z.A.R.E. Ouest L-4384 Ehlerange, Luxembourg

emmanuel.schmitt@cran.uhp-nancy.fr
vincent.bombardier@cran.uhp-nancy.fr
patrick.charpentier@cran.uhp-nancy.fr

Résumé – Cet article présente une méthode de classification de couleurs sur des planches de bois à partir d'un mécanisme à base de systèmes d'inférences floues implémenté à travers l'algorithme d'Ishibuchi-Nozaki-Tanaka. Le raisonnement flou permet de prendre en compte la subjectivité de la perception des couleurs au sein d'un produit « bois ». De plus, les informations extraites des images ne peuvent pas être considérées comme précises et certaines. L'utilisation de notre méthode améliore les taux de reconnaissance en apprentissage et en généralisation dans la classification des couleurs du bois, mais aussi sur le benchmark IRIS.

Abstract – This article presents a method of color classification on wooden boards according to a fuzzy inference system mechanism implemented with Ishibuchi-Nozaki-Tanaka's algorithm. The fuzzy reasoning allows taking into account the subjectivity of color perception in the product « wood ». Moreover, the piece of information extracted from the images cannot be considered as accurate and certain. The use of our method improves the recognition rates in training and operating steps for the wood color classification, and for the IRIS database too.

1. Introduction

L'amélioration de la qualité des produits des industries du bois (scieries, fabricants de meubles, ...) est devenue un enjeu majeur pour rester compétitif sur le marché actuel. En effet, les industriels ne peuvent plus se permettre de fournir des produits uniquement purgés de tous les défauts ; il faut aussi qu'ils soient esthétiquement appareillés.

Ainsi, nous abordons, dans cet article, le contrôle qualité de planches de bois, et plus particulièrement l'appariement couleur de ces planches. Le cadre général de l'étude est d'obtenir un procédé permettant de trier esthétiquement et visuellement des produits finis dans un environnement industriel. Le principal problème provient de l'aspect hautement subjectif de la perception des couleurs qui rend difficile son interprétation. En effet, l'Homme ne dispose pas d'un modèle universel pour définir un ensemble de couleur. Chaque individu possède ses propres caractéristiques visuelles et perceptives, l'œil humain n'étant pas un outil totalement fiable. C'est pourquoi, il est indispensable de rechercher des algorithmes robustes pour traiter des images couleur [13]. Pour tenir compte de la subjectivité de la perception des couleurs, nous avons choisi d'utiliser un système d'inférences basé sur des éléments de la théorie des ensembles flous. Ce système d'inférences floues (S.I.F.) repose en fait sur l'implémentation de règles linguistiques floues pour traiter le vecteur caractérisant la notion de couleur du bois.

Après une description du système de vision et des conditions d'acquisition des images, nous exposerons le principe de la méthode de classification. Enfin, nous évaluerons notre méthode par rapport à un classificateur bayésien, à partir de la base de données IRIS et des échantillons de bois, acquis sur des lignes de production.

Les travaux présentés dans cet article sont réalisés dans le cadre d'une collaboration université-entreprise entre le Centre de Recherche en Automatique de Nancy (CRAN), la société luxembourgeoise LuxScan Technologies.

2. Système de vision

Le premier point traité lors de notre étude consiste à pallier l'imperfection du système d'acquisitions actuellement utilisé (figure 1). En effet, un grand nombre de paramètres externes au système peuvent perturber la reconnaissance des couleurs. Pour exemple, nous pouvons citer la température dont une variation d'environ 10°C peut modifier totalement le spectre d'émission d'un éclairage, et donc modifier complètement la discrimination des couleurs (une pièce initialement reconnue dans la classe « brune » peut se retrouver dans la classe « rouge »). Une étude concernant ces paramètres a permis de fixer les conditions d'éclairage et de choisir le type d'éclairage moins sensible aux variations extérieures et plus stable au niveau de la référence de blanc. Pour des raisons de confidentialité, ces résultats ne peuvent être exposés.

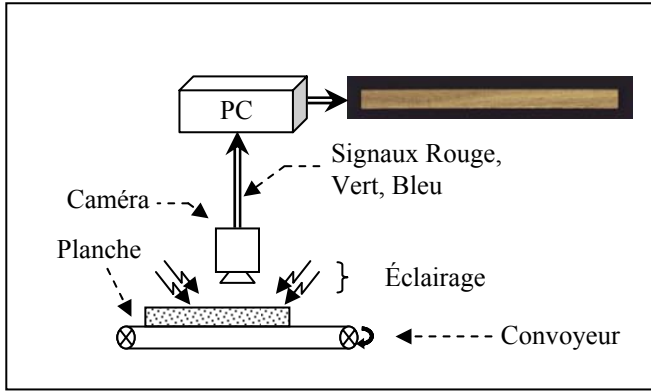


FIG. 1 : Système de vision

La reconnaissance des couleurs doit s'effectuer en temps réel sur des lignes de production pouvant atteindre des vitesses de 180 mètres par minute. Les capteurs utilisés renvoient alors 1500 lignes par seconde, chaque ligne étant composée de 900 pixels. Ainsi, nous pouvons évaluer la quantité de données à traiter, soit environ 1350000 pixels par seconde. Après une conversion des signaux Rouge, Vert et Bleu (RVB) dans le repère colorimétrique Lab conseillé par la Commission Internationale de l'Éclairage, nous pouvons effectuer la classification des données.

La couleur du bois n'étant pas homogène tout au long de la planche, nous travaillons sur des zones dont la taille a été déterminée de façon empirique. Ces Régions d'Intérêts (ROI) ne doivent pas être trop petite pour ne pas rendre la reconnaissance trop sensible aux détails, mais elles ne doivent pas non plus être trop grande pour permettre une discrimination efficace. Les ROIs nous permettent également de mieux traiter les transitions de couleur. Si une planche est composée de plusieurs zones de couleurs différentes, nous devons la classer dans une catégorie spécifique définie au préalable. C'est pourquoi, nous divisons les planches en quatre ROIs (figure 2) sur lesquelles nous calculons le vecteur caractéristique détaillé ci-après.

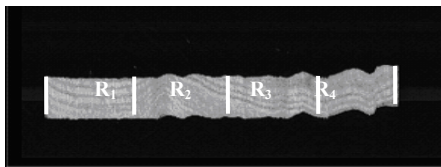


FIG. 2 : Image d'une face composée de 4 ROIs

Pour permettre un traitement rapide des données et pour répondre aux contraintes temps-réel du système industriel, il est nécessaire de caractériser la couleur à partir d'attributs simples. Nous avons choisi le moment statistique d'ordre 0 [12] qui définit le vecteur caractéristique V suivant.

$$V = \begin{pmatrix} m_l \\ m_a \\ m_b \end{pmatrix} \quad (1)$$

où m_i correspond à la valeur moyenne du paramètre i ($i=\{L,a,b\}$).

Nous avons testé d'autres caractéristiques simples comme l'écart type des différentes composantes, ou plus complexes comme des paramètres de texture ou les matrices de cooccurrence. Mais bien souvent le gain obtenu en reconnaissance n'est pas suffisamment significatif par rapport à l'augmentation du temps de calcul qui en découle.

3. Méthodologie de classification couleur

Après avoir défini notre système d'acquisitions, nous devons aborder l'aspect décisionnel pour la classification des planches de bois. A ce niveau, nous avons à résoudre deux problèmes principaux. Tout d'abord, les données délivrées par le capteur peuvent être à la fois imprécises et incertaines ; l'imprécision étant due à la variabilité du capteur lui-même, mais également de l'éclairage bien que nous ayons considérablement amélioré ce point. L'incertitude est due, quant à elle, au mode d'extraction et au choix des caractéristiques. Dans ce cas, nous ne sommes pas certains qu'elles soient totalement représentatives de l'information couleur qui nous intéresse. L'autre difficulté réside dans la subjectivité de la perception de la couleur. En effet, la définition des classes de sortie n'est pas stricte, elle peut varier d'un opérateur à l'autre et présenter des intersections non nulles. Il existe de nombreuses méthodes pour classer les planches de bois selon la couleur [4][7]. Nous avons décidé d'utiliser une méthode basée sur la théorie des sous-ensembles flous [6] afin de faciliter la manipulation de telles données imprécises et incertaines (Soft Computing) et de pouvoir prendre en compte la subjectivité de la classification couleur (Fuzzy Information Engineering).

3.1 Système d'inférences flous (S.I.F.)

Il existe deux types de raisonnement utilisant des règles floues : le raisonnement abductif qui permet d'obtenir des informations sur une entrée X à partir d'une variable de sortie Y ; et le raisonnement déductif qui permet de déduire la sortie Y à partir des valeurs de l'entrée X . Dans notre cas, nous utilisons le Modus Ponens [14], qui correspond au deuxième cas. Les règles générées peuvent être classées en deux catégories : les règles conjonctives et implicatives. Dans le cas des règles implicatives, le raisonnement est dirigé par la connaissance. Plus nous aurons d'informations sur le produit, plus nous préciserons le résultat. Les informations, que nous devons manipuler dans le cadre de la classification couleur, ne sont pas assez « riches ». C'est pourquoi nous préférons utiliser les règles conjonctives qui sont mieux adaptés à ce type de problème d'analyse de données [2]. L'intérêt d'un tel principe réside notamment dans la rapidité de calcul et dans la certitude d'obtenir une base de règles cohérente. En effet, pour notre application, nous devons aboutir à une méthode simple à mettre en place pour rester en accord avec l'aspect « temps réel » des systèmes de production visés.

De nombreuses méthodes permettent d'obtenir automatiquement les règles conjonctives floues à partir d'un lot de données [1]. Il existe deux principaux modèles qui utilisent ce type de règles : celui de Larsen et celui de Mamdani [12]. La différence entre ces deux modèles est faite

à partir du choix de la T-norme pour l'opérateur d'implication. Finalement, notre choix s'est posé sur une méthode basée sur l'algorithme d'Ishibuchi-Nozaki-Tanaka [5][10] qui correspond à l'implémentation d'un modèle de Larsen. Nous avons choisi ce modèle car le **Produit** est plus adapté que le **Minimum** pour la T-norme. En fait, en utilisant le **Produit**, nous autorisons un découpage de l'espace des variables d'entrée de façon non linéaire [11]. Après le traitement des prémisses, nous combinons les différents résultats partiels avec l'opérateur de disjonction (T-conorme), **Maximum**. Ainsi, la sortie de notre classificateur à bases de règles linguistiques floues est un vecteur flou fournissant les différents degrés d'appartenance de l'échantillon traité (ROI) à chaque classe de sorties envisageable. Pour finir, la décision est prise à partir du maximum de ces degrés d'appartenance.

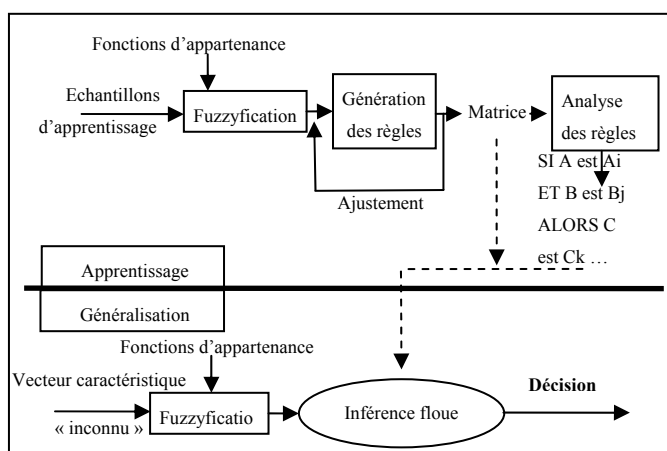


FIG. 3 : Module de reconnaissance des couleurs

Les fonctions d'appartenance (figure 3) sont construites suivant la connaissance des experts. En fait, nous divisons l'espace de représentation des différents attributs (moyennes e L, a et b) grâce à la répartition des couleurs.

3.2 Description de la méthode

Deux phases sont nécessaires pour exploiter en production notre méthode. Tout d'abord, la phase d'apprentissage des moteurs d'inférences, qui se décompose en trois étapes :

- La fuzzification des paramètres d'entrée, qui représente la traduction des données quantitatives en variables linguistiques en fonction d'un nombre de termes adéquats. Dans notre cas, nous cherchons à représenter une caractéristique par des termes répartis sur son univers de discours en fonction de sa plage de variation utile.
- La génération des règles floues du type « Si ... Alors ... » qui décrit ainsi la perception que le système a des couleurs qui sont à identifier.
- L'ajustement du modèle suivant le lot d'échantillons utilisés pour la phase d'apprentissage des couleurs. Cette dernière étape ajuste le découpage de l'espace de représentation des paramètres par itérations successives.

La base de règles floues est sauvegardée sous la forme d'une matrice numérique (figure 3) composée des attributs, des classes de sortie à reconnaître et des degrés de confiance. Ce degré associé à chaque règle peut être assimilé à un degré de pertinence de la règle en accord avec les données d'entrée. Les résultats, que nous obtenons en sortie de notre module de reconnaissance, sont décrits sous la forme d'un degré de similarité à la classe de sortie reconnue [3]. En d'autres termes, nous nous intéressons à la ressemblance de la couleur inconnue avec la couleur la plus proche.

Nous proposons d'enchaîner plusieurs S.I.F. afin d'améliorer les temps de calcul. En effet, au lieu de travailler avec un S.I.F. unique où toutes les caractéristiques sont entrées en même temps, nous préférons utiliser une structure arborescente, enchaînant plusieurs SIF plus simples. Le nombre total de règles se trouve ainsi réduit. Nous exposons dans la suite de l'article l'apport de notre méthode dans la classification des couleurs ainsi que son positionnement par rapport à d'autres méthodes classiquement utilisées.

4. Résultats

Le choix du nombre de termes de fuzzification pour définir une variable est souvent empirique mais peut être très important pour améliorer les taux de reconnaissance de notre méthode. Pour valider cette partie, nous avons utilisé la base de données IRIS afin de quantifier l'impact du découpage de l'espace de définition des paramètres d'entrées. Dans le cas de cette base composée de 150 échantillons, nous utilisons les 4 paramètres de la base de données (longueur du sépale, largeur du sépale, longueur du pétale et largeur du pétale) afin d'identifier trois classes en sorties (Setosa, Versicolor et Virginica). Le nombre de termes choisi permet de renseigner le nombre de règles linguistiques floues qui seront générées. Pour la base IRIS, nous obtiendrons par exemple 2^4 règles pour une fuzzification en 2 termes et 7^4 règles pour une fuzzification en 7 termes. La fuzzification peut être réalisée de deux manières différentes : suivant une équirépartition des termes ou suivant une répartition adaptée aux classes de sorties.

TAB. 1 : Taux de reconnaissance en généralisation sur la base IRIS pour différents types de fuzzification des entrées

	Classificateur Bayésien	Algorithme de Nozaki	
		2 termes équirépartis	3 termes équirépartis
Taux de reconnaissance	96.83%	52.67%	93.33%
e	Algorithme de Nozaki		
	5 termes équirépartis	7 termes équirépartis	3 termes ajustés
Taux de reconnaissance	94.66%	98.00%	97.33%

Les résultats (tableau 1) montrent l'amélioration du taux de reconnaissance pour différentes fuzzifications dans le cas de la généralisation, ainsi qu'une comparaison par rapport à un classificateur bayésien. Dans le cas de la base IRIS, nous constatons que le choix du nombre de termes de fuzzification

est très important pour améliorer les résultats. Cependant, il faut faire attention à ne pas trop décomposer l'univers du discours des attributs car le nombre de règles générées augmente de manière rapide (150 fois plus de règles pour une décomposition en 7 termes par rapport à une décomposition en 2 termes). De plus, en comparaison aux méthodes classiquement utilisées, nous pouvons remarquer que notre méthode semble être justifiée.

Dans un deuxième temps, nous avons utilisé un lot de données plus complexes que le benchmark IRIS afin de tester une structure arborescente. Nous avons sélectionné 500 échantillons de couleur séparés en 6 classes (Brun foncé, Brun, Brun clair, Rouge foncé, Rouge et Rouge clair).

TAB. 2 : Comparaison des taux de reconnaissance et des temps de traitement pour deux types de SIF

	Classificateur bayésien	Algorithme de Nozaki	
	Inférence arborescente	Inférence unique 5 termes	Inférence arborescente 5 termes
Apprentissage	77.20%	91.07%	91.10%
Généralisation	71.50%	83.33%	84.70%
Temps de traitement	0.12 ms	0.10 ms	0.05 ms

Pour notre étude de la couleur de planches de bois, nous avons comparé notre méthode avec celle utilisée par l'industriel, qui prend essentiellement en compte des notions bayésiennes [4][7]. Les résultats exposés dans le tableau 2 montrent 3 atouts de notre méthode :

- une amélioration du taux de reconnaissance en utilisant une structure arborescente ;
- une amélioration du taux de reconnaissance par rapport à la méthode actuellement utilisée ;
- une amélioration du temps de traitement des données.

Tous ces atouts répondent entièrement aux critères imposés par l'environnement de travail sur une ligne de production devant réaliser le tri des planches de bois en temps-réel.

5. Conclusion

La perception de la couleur est une notion très subjective. Cet article présente une méthodologie permettant de tenir compte la perception humaine pour la classification des couleurs. Notre modèle est constitué de règles conjonctives floues, qui sont activées en parallèle et fusionnées par un opérateur de disjonction. Ce modèle est généré à partir de l'algorithme d'Ishibuchi-Nozaki-Tanaka, qui répond totalement aux contraintes liées à nos travaux.

Cette étude démontre l'avantage et l'utilité de la logique floue pour notre problème d'appariement de planches de bois selon la couleur. Outre l'amélioration des taux de reconnaissance, notre méthode permet de diminuer les temps de calculs. Nous notons une diminution de 50% du temps de

traitement par rapport à une méthode bayésienne classique, ce qui s'avère non négligeable dans notre cadre d'application temps-réel. Néanmoins, nous pouvons encore améliorer notre méthode en modifiant le vecteur caractéristique d'entrée du S.I.F.. En effet, actuellement seules les moyennes des composantes couleurs calculées sont utilisées. Ce type de données ne permet pas de prendre en compte la grande variabilité du veinage au sein d'une planche de bois, celle-ci pouvant perturber la perception humaine de la couleur. Certaines personnes seront plus sensibles que d'autres aux fortes densités de veinage au sein d'un bois clair. Il serait donc intéressant de travailler sur des aspects de texture [8]. Ce projet a pour but de construire un capteur flou délivrant une information d'aspect du bois en fonction de sa texture et sa couleur [9].

Références

- [1] O. Cordon, M.J. Del Jesus, F. Herrera. « A proposal on reasoning methods in fuzzy rule-based classification systems », *Int. Jour. of Approximate reasoning*, 20, 1999, pp. 21-45.
- [2] D. Dubois, H. Prade. « Fuzzy rules in knowledge-based systems – Modelling gradedness, uncertainty and preference », *An introduction to fuzzy logic application in intelligent systems*, Kluwer, Dordrecht, 1992, pp. 45-68.
- [3] D. Dubois, H. Prade. « The semantics of fuzzy sets », *Fuzzy Sets and Systems*, Vol. 90, 1997, pp.141-150.
- [4] A. Hanbury, « Morphologie Mathématique sur le Cercle Unité avec applications aux teintes et aux textures orientées », *Thèse doctorale, Ecole Nationale Supérieure des Mines de Paris*, 2002.
- [5] H. Ishibuchi, K. Nozaki, H. Tanaka, « Distributed representation of fuzzy rules and its application to pattern classification », *Fuzzy Sets and Systems*, Vol. 52, n° 1, 1992, pp. 21-32.
- [6] A. Kaufmann, « Introduction à la théorie des sous-ensembles flous », *Édition Masson*, 1975.
- [7] D.E. Kline, R.W. Conners, Q. Lu, P.A. Araman, « Automatic color sorting of hardwood edge-glued panel parts », *Hardwood Symposium Proceedings*, 1997.
- [8] Topi Mäenpää, Matti Pietikäinen, « Classification with color and texture: jointly or separately? », *Pattern Recognition*, 2003.
- [9] G. Maurice, E. Benoit, L. Foulloy, « Fuzzy sensors: another view », *Information Engineering*, 1997.
- [10] K. Nozaki, H. Ishibuchi, H. Tanaka, « A Simple but powerful heuristic method for generating fuzzy rules from numerical data », *Fuzzy sets and systems*, Vol. 86, 1997, pp. 251-270.
- [11] E. Schmitt, V. Bombardier, P. Charpentier, R. Vogrig, « Classificateur flou pour l'appariement couleur de pièces de bois », *LFA*, Nantes, 2004.
- [12] E. Schmitt, C. Mazaud, V. Bombardier, P. Lhoste, P. Charpentier, R. Vogrig, « Fuzzy linguistic rules identification system for wood quality », *AISTA*, Luxembourg, 2004.
- [13] G. Sharma, H.J. Trussel, « Digital color imaging », *IEEE Transactions on Image Processing*, 6(7), juillet 1997, pp. 901-932.
- [14] L.A. Zadeh, « Outline of a new approach to analysis of complex systems and decision processes », *IEEE Trans. On Systems, Man and Cybernetics*, Vol 3, 1973, pp. 28-44.