Engineering Applications of Artificial Intelligence 23 (2010) 978–988

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Fuzzy rule classifier: Capability for generalization in wood color recognition

Vincent Bombardier*, Emmanuel Schmitt

Research Centre for Automatic Control, CRAN, CNRS UMR 7039, Université Henri Poincaré, Campus scientifique, BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France

A R T I C L E I N F O

Article history:
Received 21 April 2009
Received in revised form
21 April 2010
Accepted 4 May 2010
Available online 3 June 2010

Keywords:
Classification
Fuzzy logic
Image processing
Fuzzy rules
Color recognition

A B S T R A C T

In this paper, a classification method based on fuzzy linguistic rules is exposed. It is applied for the recognition of the gradual color of wood in an industrial context. The wood, which is a natural material, implies uncertainty in the definition of its color. Moreover, the timber context leads obtaining imprecise data. Several factors can have an impact on the sensors (ageing of the acquisition system, variation of the ambient temperature, etc.). Finally, the data sets are often small and incomplete. Thus the proposed method must work within these constraints, and must be compatible with the time-constraint of the system. This generally imposes a weak complexity of the recognition system. The Fuzzy Rule Classifier is split in two main parts, the fuzzification step and the rule generation step. To improve the tuning of this classifier, a specific fuzzification method is presented and compared with more classical ones. Several comparisons have been made with other classification method such as neural network or support vector machine. This experimental study showed the suitability of the proposed approach essentially in term of generalization capabilities from small data sets, and recognition rate improvement.

1. Introduction

The works presented in this article deals with color classification on wooden boards in an industrial environment. Color recognition is an important step for matching wood pieces which affects many activities of the timber industry (veneer, paneling, manufacture of squares, etc.). The wood pieces have to be assembled according to their perceived color. The aim is mainly to provide a wood piece that seems to be homogeneous and massive. This kind of problems belongs to the aesthetic aspect of wood characterization. The aspect control issue can be divided into two categories:

- detection, localization and identification of singularities and
- classification and identification of color and/or grain (textured aspect of wood).

In the context of wood classification (color, texture and singularities), different methods can be found in the literature. For the defect recognition, there are essentially "compilation" methods which "compile" an important training data set to obtain the output classes' model: Neural Networks (NN) (Cho et al., 1991; Schmoldt et al., 1996; Kauppinen et al., 1999) and Genetic Algorithms (GA) (Estevez et al., 2003). For color or texture identification, the used methods are principally based on Neural Networks (NN) (Bai and Wang, 2007; Pölzleitner and Schwingshakl, 2004; Wang and Bai, 2007) but also on k-Nearest Neighbor algorithm (k-NN) (Kline et al., 1999; Maenpaa et al., 2003). Distance Minimization algorithms (Lu, 1997; Daul et al., 2000; Hanbury, 2002; Srikanteswara et al., 1997) or Genetic Algorithms (Sathyanath and Sahin, 2001) are used too.

The aim of the global vision system is to identify the wood color in a continuous mode during the production. Such a system involves lots of constraints.

Firstly, the classification method must be able to work with small training data sets. Indeed, some classes which are rare in nature are defined with few samples. Moreover, providing a training data set requires the industrialist to label each sample. This task is really painful and highly time-consuming. To adapt the system to the changes in the system and/or products, they should be renewed if necessary.

Then, colors which must be identified are subjective. Because of the impact of the wooden fiber to the perceived color of the wood, human operators can have different perceptions of the wood color. The output classes are gradual and non-disjointed too. For example, there are no strict bounds between a "red" wood and a "light red" wood.

Finally, there are other specific constraints to respect such as the real-time aspect of the production system or how easy the method is to set up.

So, the used classification method must take into account these constraints and present a low complexity of the recognition model.

The existing methods do not answer exactly to these specific constraints. Indeed, Neural Networks are popular machine

* Corresponding author. Tel.: +33 383 68 44 52.
E-mail addresses: vincent.bombardier@cran.uhp-nancy.fr (V. Bombardier), emmanuel.schmitt@cran.uhp-nancy.fr (E. Schmitt).

0952-1976/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2010.05.001

learning algorithms that remain to be widely used for wood recognition problems. Their main advantages are their ease of use and their good behavior in term of classification rates. But their main drawback is that they need lots of training samples (Tou et al., 2009). Other compilation methods such as Genetic Algorithm have the same drawback. The other most popular method in wood color recognition is k-Nearest Neighbor. The k-NN is easily implemented as it does not require a training process. It is useful especially when there is a small dataset available. However, the major drawback of the Nearest Neighbor algorithms is that the computing time will increase according to the k number of neighbors used (Tou et al., 2009). Moreover, the classification is very dependent on the choice of these neighbors. So, the setting must be done by an expert in image processing.

We propose to use some tools provided by the Fuzzy Sets Theory (Zadeh, 1965), which seems to be well adapted for taking into account the above detailed constraints. Fuzzy Logic is able to provide non-separated output classes. While in the past fuzzy rule-based systems have been mainly applied to control problems, they have also been used recently in pattern classification tasks (Roubos et al., 2003; Nakashima et al., 2007) and they proved their ability to work with few learning data sets (Wang et al., 2008; Schmitt et al., 2007). That is why a classifier based on fuzzy linguistic rules seems to be more appropriate. Another advantage is the interpretability of such Rule Systems (Roubos et al., 2003; Nauck and Kruse, 1998, 1999). On the other hand, for instance, Neural Network are used as "black box" and there is no explicit link between the vocabularies used to define the output classes and the image features characterizing the color, it is very hard to interpret the decision taken by the classification method. The aim of this study is to obtain a vision sensor which delivers an answer in the "wood" vocabulary. This problem is often referred to as the "semantic gap", defined as "the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation" (Smeulders et al., 2000) Indeed, the wood expert will use words like "Red" or "Brown", whereas the vision expert will define colors with numerical values like {90, 35, 12} in RGB space for instance.

Moreover, the setting parameters must be comprehensible for non-specialists too. In this way, it is judicious to be interested in the classification methods based on fuzzy set theory, because it allows the integration of the information expressed under linguistic form (Dubois et al., 1996). Finally, in highly specialized field like the timber industry, the experts make decisions whatever the work conditions. Thus, the vision system, which must carry out the colorimetric control of the wooden boards, can be hatched as a decision system reproducing the human expert reasoning. D'Acquila recommends using inference engines which allows not only work from a representation space with n dimensions, but also takes into account all forms of uncertainty and imprecision (D'Acquila et al., 2002).

This paper is organized as follows: Section 2 introduces the problems of the wood color recognition and the vision system used to do that. Section 3 details the proposed Fuzzy Rule Classifier by explaining the different steps of the method and its settings are explained in Section 4. Finally, Section 5 presents several comparisons with other classifiers and analyses the experimental results obtained from several University of California Irvine learning databases (Blake and Merz, 1998) and the industrial database corresponding to the application described in Section 2.

2. Industrial vision process

For about fifty years, the timber industry has been placed on a competitive market. In the order to get away from the industries which provide "bottom-of-the-range" products, some companies have given each other consequent means, like the use of vision systems in order to control and to enhance the production performances. As mentioned in the introduction, the context of this study is the classification of wooden colors.

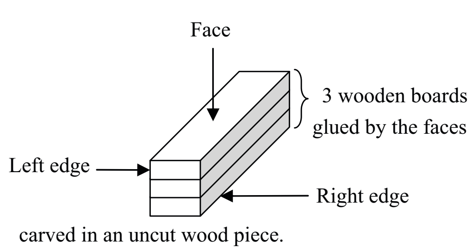
2.1. Acquisition step

This recognition is carried out in real time on the industrial production line. These lines may reach speeds up to 400 m of board length per minute. After the color identification step, done by the vision system, color information is sent to an optimization step. Then each board is sent to a sorting line or to a cutting line. The cutting line aims to split the boards into uniformly colored piece of wood. The sorting line aims to group pieces of wood into specific classes, whose number and definition are given by the final customer. The boundary classes are very subjective in both cases.

The originality of the process concerns the color sorting which is only realized on the wooden board edges (board thickness). Indeed, the machining of handrails requires a uniform color in a large thickness (Fig. 1). To obtain this large thickness, three boards are glued by their face. So, the final product makes illusion of a product carved in an uncut wood piece.

The entire vision system is detailed in Fig. 2. The acquired images are processed to obtain color descriptors (see Section 2.2). These features are then used by the classification stage (see Section 3) to provide the color label of the wooden board or a part of it if the color changes along the board.

Fig. 3 shows an image obtained with the acquisition system which is made up of one type of linear sensors: CCD color cameras. This CCD sensor provides the red, green and blue components of the signal. The signals are sampled at the rate of 1500 lines per second along the y-axis (Fig. 3). Each line is composed of 900 pixels (x-axis in Fig. 3). In the industrial case presented, the wooden boards are around 3 m tall. With a longitudinal resolution equal to 1.5 mm per pixel, the images are made up of around 2000 lines. Thus, working in real-time, the data processing must be carried out under time constraints of around 1.5 s.



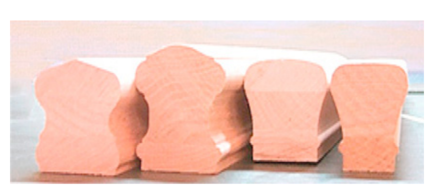


Fig. 1. Schematic representation of the final products and Handrail picture.