Expert Systems with Applications 36 (2009) 8807–8815

Contents lists available at ScienceDirect

Expert Systems with Applications

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

# Recycled paper visual indexing for quality control

Jose Orlando Maldonado, Manuel Graña *

Dpto. EEST, Faculty of Engineering and Architecture, University of Pamplona, University City, Pamplona, Colombia, USA

ARTICLE INFO

*Keywords:*
Gabor filter banks
Texture features
Discrete wavelet transform
Quality control
Classification

ABSTRACT

In this paper, we describe the development of a system for evaluating an specific quality characteristic of recycled paper sheets using techniques of image analysis and pattern recognition. We call Bumpiness the phenomenon of interest, which is new in the literature on paper quality. This phenomenon is characterized by the appearance of macroscopic undulations on the paper sheet surface that may emerge shortly or some time after its production. We explore the detection and measurement of this defect by means of computer vision and statistical pattern recognition techniques that may allow early detection at the production site. Our goal is to give an scalar continuous measure of Bumpiness. We propose features computed from Gabor filter banks (GFB) and discrete wavelet transforms (DWT) for the characterization of paper sheet surface Bumpiness in recycled paper images. The starting point is to state the problem as a classification of the paper sheet images into two classes: low and high Bumpiness. In this setting we obtain, with both proposed texture modelling approaches (GFB and DWT), classification accuracies comparable to the agreement between human observers. The best performance is obtained using DWT features. Finally, we propose as the scalar index of Bumpines the fisher discriminant analysis (FDA) function defined on the space of the best features for the classification task. We perform an innovative validation process of this Bumpiness index, based on the ordering of random pairs of images, obtaining a very high agreement with the human observers.



## 1. Introduction

The works reported in this paper arise from the demand of a company devoted to recycled paper production to provide a way to measure a defect in the paper sheets. The Pulp and Paper Industry and the Printing Industry have set a number of quality standards along with their measurement processes. The major standardization organizations, ISO and Tappi, have defined standards for diverse characteristics of the paper:

- *Physical properties:* grammage, moisture content, thickness, water absorption and roughness.
- *Strength properties:* bending resistance/stiffness, bursting strength, ply bond/scott bond, stiffness, tearing resistance, tensile strength, folding endurance, internal bonding strength and *Z*-directional Tensile.
- *Optical properties:* color, brightness, gloss, opacity and whiteness.
- *Printing properties:* such as mottle, graininess, darkness and contrast.

None of these measurements is relevant to the paper defect the company was interested in, and that we have called Bumpiness for lack of a name in the literature. This Bumpiness is the deformation of the paper sheet showing undulations even before any printing or coating process has taken place. This paper defect can be worsened by printing processes i.e., laser printers. We do not know the microscopic causes for this paper defect, which can be intuitively associated with inhomogeneous spatial fiber distribution. The spatial range of these undulations is in the order of centimeters or inches, so it seems difficult that a microscopic study of paper sheet samples using tools like phase-contrast X-ray microtomography (Holmstad, Antoine, Silvy, Costa, & Antoine, 2001), or scanning electron microphotographs (Klungness, Ahmed, Ross-Sutherland, & AbuBakr, 2000), could lead to a characterization of the paper Bumpiness. Besides, we do not have access to these imaging facilities. Texture features developed for computer vision applications could provide the visual spatial information needed to characterize this paper sheet defect. We have found some, but not many, precedents of the application of visual pattern recognition techniques in the paper and printing industries: to determine the distribution of local strain during a tensile test (Considine, Scott, Gleisner, & Zhu, 2005), for defect detection in paper pulp images, Calderon-Martinez and Campoy-Cervera (2003), for mottling assessment (Sadovnikov, Salmela, Lensu, Kamarainen, & Kalviainen, 1995). Instances of works that report other industrial quality inspection

* Corresponding author.
*E-mail address:* ccpgrrom@gmail.com (M. Graña).
*URL:* http://www.ehu.es/ccwintco (M. Graña).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.11.023

using computer vision and texture features are Abouelelaa, Abbasb, Eldeeba, Wahdanb, and Nassara (2004), Anagnostopoulos et al. (2002), Chi-Ho and Pang (2000), Kumar and Pang (2002), Ngana, Panga, Yungb, and Ngb (2005), Sari-Sarraf and Goddard (1999), Scharcanski (2006) for textile defect detection, Martinez-Alajarin, Luis-Delgado, and Tomas-Balibrea (2005) for the classification of marble slabs and Funck, Zhong, Butler, Brunner, and Forrer (2003) for classification of wood surfaces. There are other paper defects that can be modeled using spatial information extracted from images. Computer vision techniques have been used to characterize mottling (Sadovnikov et al., 1995). The mottling defect is due to the uneven distribution of ink, and it is a different phenomenon from the Bumpiness. After some initial research, we found that the techniques applied to the mottling visual characterization and measure were not applicable to the Bumpiness phenomenon.

The goal of our work is to obtain some quantitative evaluation of paper sheet Bumpiness using computer vision and statistical pattern recognition techniques. The long term research goal would be to construct predictive models of paper quality from production parameters aiming to identify quality loss causes and introducing correcting control strategies in the production process. For this task, we need a quantitative measure of paper quality to define and train these predictive models, starting from the human labelling of the images obtained from quality control paper sheet samples. This quantitative scalar value can be also used as an index for the search in quality control image databases, easing and automating the search for similar images. This paper is devoted to describe our works to obtain such an image index. Fig. 1 shows a general scheme of the system proposed to compute the continuous index measuring the Bumpiness level. To build this system we went through the following steps:

1. We state the Bumpiness measurement problem as a classification problem into two degrees of Bumpiness.
2. We have gathered validation information from the human classification of paper sheet images. We studied the agreement between human classifications.
3. We have tested texture features based alternatively on Gabor filter banks (GFB) and on discrete wavelet transforms (DWT) to characterize the paper images. We made independent classification experiments for each definition of the feature vector with a collection of supervised classification methods available from the Weka suite (Witten & Frank, 2005).
4. Based on the classification results, we selected a feature vector as the best characterization of the defect, and we define a continuous Bumpiness index as an scalar transformation of this feature vector. This scalar transformation is the fisher discriminant analysis (FDA).
5. We validate the continuous Bumpiness index with human observers introducing a new validation methodology, that compares the responses of the human observers with the ordering induced by the continuous Bumpiness index over a set of randomly selected image pairs.

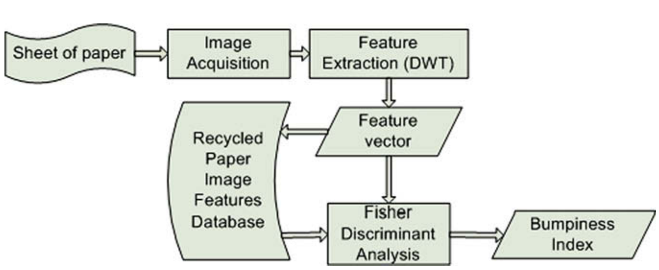


**Fig. 1.** General scheme of system proposed to compute the Bumpiness index.

Fig. 2 shows several images of recycled paper sheets. Image acquisition was performed with a conventional flatbed office scanner at an optical resolution of 1200 dpi, which generated spatial high-resolution images. Images of recycled paper were pre-processed applying a contrast enhancement that consists on the selection of the upper grey levels that encompass 90% of the image accumulative histogram and renormalization of the image intensity range, to emphasize the visual features of the Bumpiness phenomenon, because scanning paper sheets produces white images with very little contrast. The enhanced images show the presence of certain textures that do not keep any regular pattern. They show also some strong impulsive noise due to the contrast enhancement procedure. Impulsive noise is dealt with by Gaussian smoothing previous to any feature extraction process. Original size of each paper sheet is standard A4: $24 \times 29.7$ cm, the size of the scan images was $850 \times 1170$ pixels.

The paper is organized as follows: Section 2 reviews some definitions of GFB and their design for texture characterization. Section 3 reviews the DWT and its application to define visual texture features. Section 4 presents the results of the manual labelling of the recycled paper images that define the ground truth for the construction of supervised classifiers. Section 5 gives the results on the classification of the recycled paper images based on GFB and DWT texture features. Section 6 is devoted to describing the scalar Bumpiness index and its validation process. Section 7 gives some conclusions and directions for further work.

## 2. Gabor filter bank design

Gabor elementary functions (Gabor, 1946) have the property of being highly selective both in the space domain and in the frequency domain. Although Gabor original works focused in 1D representation, their principles were extended later by Daugman (1985) to 2D domain. Gabor filter banks (GFB) are a well-established texture feature extraction method for image segmentation (Clausi & Jernigan, 2000) and image data retrieval (Manjunath & Ma, 1996). Filters based on Gabor functions have been applied to texture characterization and image segmentation since the early 90s (Dunn & Higgins, 1995; Dunn, Higgins, & Wakeley, 1995; Teuner, Pichler, & Hosticka, 1995), but also recent works demonstrate the renewed interest in Gabor filters for texture analysis, Clausi and Deng (2005). A recent review of Gabor filters in image processing can be found in Kamarainen, Kyrki, and Kalviainen (2006). Each GFB is a collection of band-pass filters that define a particular sampling of the Fourier transform space. Varying the parameters of the filters allows defining diverse sampling strategies with different properties.

A two-dimensional Gabor function can be written as

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2}+\frac{y'^2}{\sigma_y^2}\right)\right] \exp[2\pi i(Ux+Vy)], \quad (1)$$

where $(x', y')$, are Euclidean coordinates $(x, y)$ rotated in the space domain:

$$x' = x\cos(\theta) + y\ \text{sen}(\theta), \quad (2)$$

$$y' = -x\,\text{sen}(\theta) + y\ \text{sen}(\theta). \quad (3)$$

Thus, a Gabor function is a Gaussian function modulated by a complex sinusoid. Parameters $\sigma_x$, $\sigma_y$ characterize the spatial support and bandwidth of the filter. Filters are usually assumed to be non-isotropic, thus $\sigma_x \neq \sigma_y$. The Gaussian's major axis is rotated by an angle $\theta$ around the positive $z$-axis. If we denote $(u, v)$ the Euclidean coordinates in frequency domain, the point $(U, V)$ represents a particular 2D frequency. The complex exponential is a 2D