Expert Systems with Applications 38 (2011) 10398–10407

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Intelligent computer vision system for segregating recyclable waste papers

Mohammad Osiur Rahman *, Aini Hussain *, Edgar Scavino, Hassan Basri, M.A. Hannan

Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia

ARTICLE INFO

Keywords:
Waste paper sorting
Recyclable waste paper
Paper grade

ABSTRACT

This article explores the application of image processing techniques in recyclable waste paper sorting. In recycling, waste papers are segregated into various grades as they are subjected to different recycling processes. Highly sorted paper streams facilitate high quality end products and save processing chemicals and energy. From 1932 to 2009, different mechanical and optical paper sorting methods have been developed to fill the paper sorting demand. Still, in many countries including Malaysia, waste papers are sorted into different grades using a manual sorting system. Because of inadequate throughput and some major drawbacks of mechanical paper sorting systems, the popularity of optical paper sorting systems has increased. Automated paper sorting systems offer significant advantages over human inspection in terms of worker fatigue, throughput, speed, and accuracy. This research attempts to develop a smart vision sensing system that is able to separate the different grades of paper using first-order features. To construct a template database, a statistical approach with intra-class and inter-class variation techniques are applied to the feature selection process. Finally, the K-nearest neighbor (KNN) algorithm is applied for paper object grade identification. The remarkable achievement obtained with the method is the accurate identification and dynamic sorting of all grades of papers using simple image processing techniques.

1. Introduction

The primary challenge in recycling paper is to obtain raw material with the highest purity. During the recycling process, waste papers are segregated into various grades as they are subjected to different recycling processes. Highly sorted paper streams facilitate a high quality end product and save processing chemicals and energy. The *grade* refers to the quality of a paper or pulp and is based on weight, color, usage, raw material, surface treatment, finish or a combination of these factors (Paper Grades, 2009). In most materials recovery facilities (MRFs), the recovered waste papers are sorted as computer print-out (CPO), white ledger or white paper (WP), colored ledger (CL), old corrugated cardboard (OCC), old newsprint paper (ONP), old magazines (OMG), glossy coated sheet (CS), and mixed office paper. The waste paper sorting systems are classified into manual and automated systems.

In many countries including Malaysia (SPM Paper Recycling Sdn Bhd, 2009), waste papers are sorted into different grades using a manual sorting system. Manual paper sorting faces some major problems: first, it is a labor intensive industry that incurs a relatively high processing cost; moreover, it has an inconsistent end product quality; finally, specific skills and commitment are essential for efficient sorting. In addition, laborers working at the manual sorting facilities are exposed to microorganisms, organic dust, and fungi, which can cause severe infections (Würtz & Breum, 1997). Thus, there is an important need to automate the sorting process for efficient, safe and clean recycling. Automated paper sorting systems offer significant advantages over manual paper sorting systems in terms of human fatigue, throughput, speed, and accuracy.

Automated paper sorting systems are classified into mechanical and optical systems. Mechanical paper sorting cannot achieve commercially viable throughputs and accuracy. The greatest advantages of optical paper sorting systems include the following: consistent and reliable production efficiency with a relatively high hit rate and purity; and low operational cost because of fewer manual workers on the production line.

Manufacturers Standardization Society (MSS) and TiTech are the top two competitors for sensor-based paper sorting. They possess the technology and create partnerships with recyclers. MSS and TiTech utilize Near Infrared (NIR) spectroscopy technology and sell their systems in the United States, European, Australian, and Asian markets (Paper Competitors, 2009). The paper sorting performance reaches 80% for both MSS and TiTech systems (Remade Scotland, 2005). Pellenc, Bollegraaf, Lubo, and RedWave are also competitors in the paper submarket (Paper Competitors, 2009).

The TiTech autosort uses new DUOLINE scanning technology, which conducts a double scan on every pass. TiTech autosort

* Corresponding authors.
E-mail addresses: osiur_rahman@yahoo.com, osi@vlsi.eng.ukm.my (M.O. Rahman), aini@eng.ukm.my (A. Hussain), scavino@eng.ukm.my (E. Scavino), drhb@eng.ukm.my (H. Basri), eehannan@eng.ukm.my (M.A. Hannan).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2011.02.112

combines Visual Imaging Sensor (VIS) and NIR sensors into a universal modular sorting system that meets a wide variety of needs as well as highly specialized ones (TITECH autosort, 2009). The VIS sensor, for example, can be used to recognize print media with Cyan, Magenta, Yellow, and blacK (CMYK) spectral analysis.

The MSS focuses primarily in the United States and United Kingdom markets and is the only U.S.-based company to provide an automated paper sorting machine. They installed the first PaperSort system in Weyerhauser's Baltimore paper recycling facility and another in Denver (MSS Paper Sorting Systems, 2009; Paper Competitors, 2009). The MSS provides different models for optical paper sorting systems (MSS Optical sorting Systems, 2009) that include the MultiWave Sensor (MSS MultiWave™, 2009), FiberSort System (MSS FiberSort™ Module, 2009), or PaperSort System (MSS PaperSort™, 2009). The MSS MultiWave™ is the automated NIR, color, gloss, and lignin sorting system that is able to sort paper and plastics. The FiberSort™ Module is a highly advanced full color spectrum and near-infrared (NIR) sensor that identifies brown OCC from news and mixed paper grades at high speeds. Three types of PAPERSORT™MODELS are designed for three different machine widths, 900 mm (36″), 1200 mm (48″) and 1800 mm (72″), with sorting capacities of 3, 4 and 6 tons/h, respectively.

Pellenc, a French company, recently entered the paper sorting market with two machines: the Zephyr and the Boreas (Pellenc, Paper Sorting Company, 2009). The Boreas incorporates a new technology called Mid-Infrared (MIR) spectroscopy. Pellenc claims that the MIR machine can identify and eject brown, white and grey cardboard and color print (Pellenc, MIR Technology, 2009). RedWave utilizes NIR and color camera technology on all of their machines (RedWave, BT-Wolfgang Binder, 2009). However, they did not develop any multigrade paper sorting systems.

Faibish, Bacakoglu, and Goldenberg (1997) proposed an automated paper recycling system where ultrasonic sound is used to separate different grades of papers. Their proposed system suffered from several problems related to image processing such as non-uniform illumination, segmentation of dark objects with low reflectance (cloths), and detection of subframe bounds. Furthermore, the adaptive threshold mechanism did not work smoothly to automatically tune the image processing parameters. Because of contact manipulation and sensing, the system is too slow (80 ms/sub-frame) for practical industrial applications.

Chandini (2001), Mallapragada (2004) and Ramasubramanian, Venditti, Ammineni, and Mallapragada (2005) developed lignin sensors in their papers. The lignin sensor worked well for separating newsprint samples from others. However, the lignin sensors are influenced by color and sensor distance from the sample. As a result, the lignin sensor alone cannot detect the presence of newsprint.

Hottenstein, Kenny, Friberg, and Jackson (2000) proposed a sensor-based sorting approach in which a brightness sensor (reflected light intensity at 457 nm) is used to sort papers into three categories: white papers containing optical brighteners, white papers without optical brighteners, and others. In order to further distinguish between different colored papers, they described a multigrade sensor in which they used the properties of gloss and color for identification.

Chakravarthi (2006) developed a stiffness sensor that can be used to sort recovered paper into paperboard and others. The stiffness sensor successfully identified paperboard from other recovered papers (Chakravarthi, 2006; Venditti, Ramasubramanian, & Kalyan, 2007). The stiffness sensor cannot distinguish between a stack of newsprint versus a single paperboard (Ramasubramanian, Venditti, & Gillella, 2008).

Sandberg (1932) proposed a sorting device for waste paper in which waste paper is sorted from different types of wastes. The system used a mechanical process to sort waste paper into paper and others. The system did not handle different grades of papers.

Bialski, Gentile, and Sepall (1978, 1980) proposed a mechanical process to separate one paper grade from a waste paper mixture. Ortner, Bahr, and Musselmann (1980) proposed a method to separate paper objects from various foreign matter. In this system, a beater is employed to rough-pulp the waste paper. Holz and Hutzler (1986) and Spencer (1994) proposed waste paper sorting systems that segregated paper objects from others foreign matter and, like other mechanical processes, their systems did not handle different grades of paper. The attempts by Bialski et al. (1978, 1980) and Grubbs, Kenny, and Gaddis (2001) to create a successful system have not been successful because of the inability to distinguish between paper grades (Ramasubramanian et al., 2005).

Khalfan (2002) first introduced an optical paper sorting method in 2002. They used diffuse reflectance to identify a sheet of paper as either white or non-white. After a minor modification, they proposed two additional patented systems (Khalfan & Greenspan, 2006a, 2006b) that use the same methodology. Their proposed paper sorting system segregates papers into white and ground wood paper based on the lignin content.

Eixelberger, Friedl, and Gschweitl (2003) proposed an optical paper sorting method to separate waste paper into two classes based on the radiation reflected from the surface of the papers. Bruner et al. (2003) proposed one optical paper sorting method to separate waste papers into bright white paper and others based on the amount of fluorescence present on the surface of paper objects. Doak et al. (2007a, 2007b) proposed optical paper sorting methods to separate different grades of paper based upon at least one of the following: color, glossiness and the presence of printed matter.

Gschweitl and Heinz (1998) proposed a color-based paper sorting system. The European patent publication proposed the use of visible light, ultraviolet light, X-rays and/or infrared light to illuminate the paper and observing the reflected light with one or more cameras connected to a central processing unit. The disclosure of the European patent office publication is very vague with regard to the manner in which such a process was conducted.

2. Driving philosophy of recyclable waste paper sorting system

From the review, it was noted that five sensors, ultrasonic (Faibish et al., 1997), lignin (Chandini, 2001; Hottenstein et al., 2000; Khalfan, 2002; Khalfan & Greenspan, 2006a, 2006b; Ramasubramanian et al., 2005), gloss (Doak et al., 2007a, 2007b; Hottenstein et al., 2000), stiffness (Chakravarthi, 2006; Ramasubramanian et al., 2008; Venditti et al., 2007) and color sensors (Bruner et al., 2003; Doak et al., 2007a, 2007b; Eixelberger et al., 2003; Gschweitl & Heinz, 1998) have been used in paper sorting systems. Ultrasonic sensors are slow, which make them unsuitable for industry use. The lignin sensor can only be used to separate the newsprint paper from others, and its performance is strictly color dependent. The stiffness sensor is typically used to separate cardboard from other paper grades, and the gloss sensor is used to separate glossy paper from other papers. The color sensor, on the other hand, measures the radiation of the paper surface and is commonly used to identify white papers.

The implementation of the aforesaid methods is still complex, expensive and sometimes offers limited reliability. All the systems can only segregate two types of paper at one time. Moreover, no image processing or intelligent techniques were used to extract the features or characteristics from the paper objects.

Recently, Rahman, Hannan, Scavino, Hussain, and Basri (2009a, 2009b, 2009c) proposed three electronic image-based waste paper sorting techniques. The first technique (Rahman, Hannan, Scavino, Hussain, & Basri, 2009a) focused on the four points in the periphery of the paper object. Then, features surrounding those four points were extracted. Because the method did not consider texture information from the entire paper object, it may provide misleading information regarding the paper grade. The other two methods, template matching (Rahman et al., 2009b) and co-occurrence features (Rahman et al., 2009c), achieved good paper grade identification success rates; however, in real time applications, both methods are slow because of the significant computing time. First-order (Pham & Alcock, 2003) features provide enough texture information for the surface of paper objects and less computation time, which allows the use of these features in real time paper object grade identification. Thus, the main goal of this study is to develop a smart vision sensing system that will be able to separate different grades of paper using first-order features. In this proposed system, texture information from the entire paper object is considered to recognize the paper grade, which overcomes the major drawback of the previous electronic image-based waste paper sorting techniques. Moreover, the algorithm provides robust and fast results because unexpected color regions of the paper object during feature extraction process are filtered.

3. Proposed intelligent computer vision system for waste paper sorting

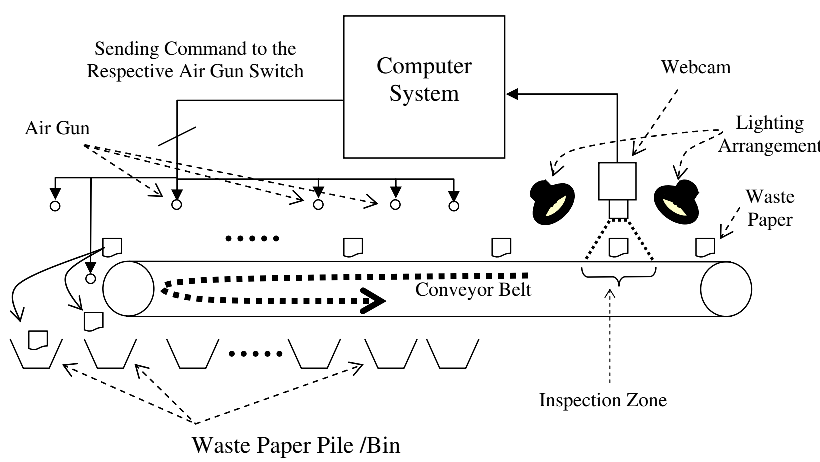
Fig. 1 illustrates the block diagram of the proposed intelligent computer vision system for automatic sorting of recyclable waste paper and a picture of the actual systems. The computer vision process consists of three parts: perception, cognition and action. The perception or image acquisition portion of this vision system consists of a commercially available webcam and a special lighting arrangement. The main responsibility of the action component of the vision system is to segregate waste papers into different grades based on the command of the cognition part of the vision system. Air blasts are used to segregate and to pile different grades of papers according to their respective waste paper bins. In this paper, we emphasize a waste paper grade identification system, which covers the perception and cognition components of the proposed system.

4. Waste paper grade identification method

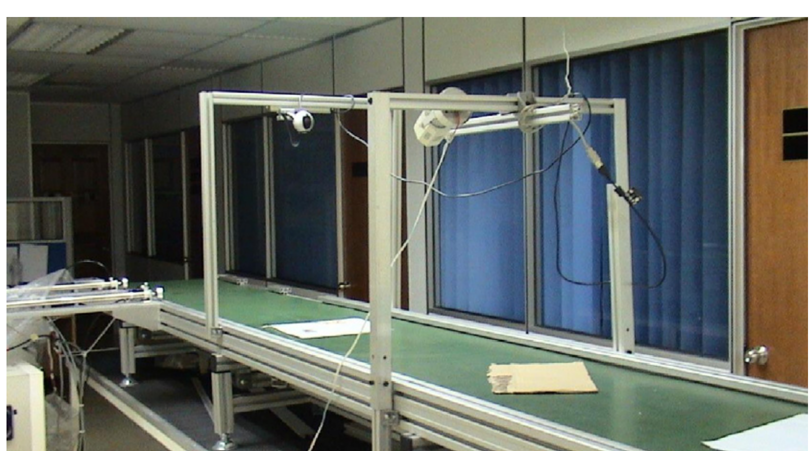
Fig. 2 illustrates the basic block diagram of the automated waste paper grade identification system. The proposed system operates in two phases: enrollment and identification. In Sections 4.1 through 4.4, the processes of both the enrollment and identification phases are discussed.

4.1. Image acquisition

In this proposed system, 640×480 RGB images are captured from the inspection zone on the conveyor belt using a Logitech QuickCam Pro 4000 Web Camera (Logitech QuickCam Pro 4000 Web Camera Specification, 2009). To set the webcam properties,



(a)



(b)

Fig. 1. (a) Block diagram of the intelligent computer vision system for automatic sorting of recyclable waste paper, and (b) a picture of the actual system.

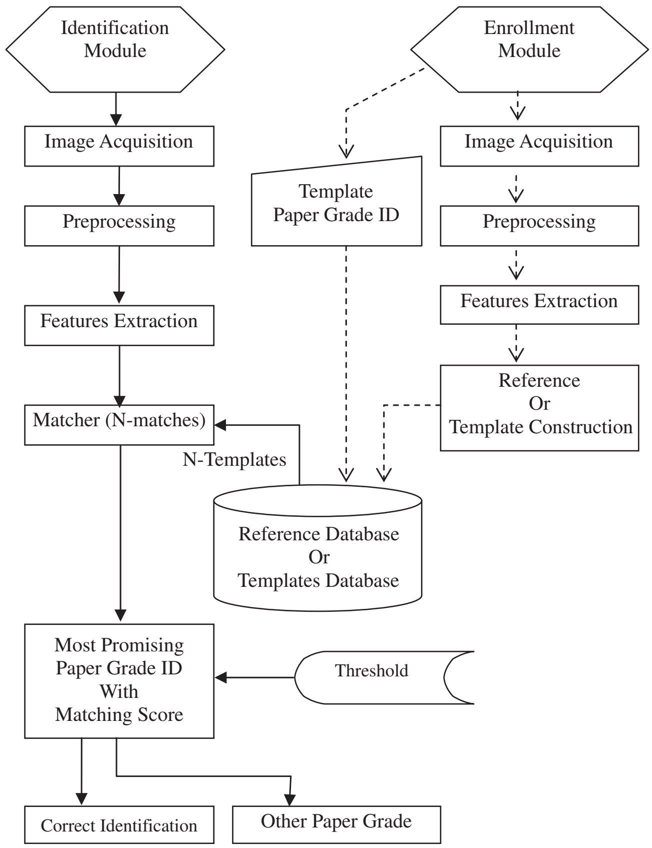


Fig. 2. Block diagram of the paper grade identification system.

the brightness, contrast and saturation were adjusted to 50%, 50% and 100% of their respective scales. In the experiment, it is observed that the performance of the vision system is extremely influenced by the lighting arrangement. Because front lighting-directional-darkfield illumination is widely used in surface scratches or texture analysis (Burke, 1996; Pham & Alcock, 2003), this illumination technique is adopted for this experiment.

The conveyor belt speed is 14 feet per minute. The system processes eight images per second. The real time scanning process produces two types of signals: presence of an object (*PObj*) and absence of an object (*AObj*). The system always performs a scanning operation to detect the presence of an object. The system captures the images from the inspection zone based on the two signals. If the *AObj* signal is detected after the *PObj* signal, then the system captures the image of the paper object from the inspection zone. This technique separates individual paper objects from the sequence of paper objects.

4.2. Preprocessing

The first step in the preprocessing block is to capture the image within the inspection zone after trimming the unnecessary boundary portion of the image. Next, the background noise is eliminated from the image using the combined operation of threshold and morphological operation erosion (Pham & Alcock, 2003) with a 3×3 minimum convolution filter.

4.3. Features extraction and template construction

In the feature extraction phase, both color and gray scale images are considered. Brunner, Maristany, Butler, Leeuwen, and Funck (1992) converted the usual RGB color space into other potentially more useful color spaces, but they found that none provided any improvement over RGB. Thus, the RGB color space is considered in this research. For color images, each of the three color components – red, green and blue – are considered separately. For gray scale image, standard grayscale transformation is obtained from the original RGB image. Grade identification is primarily based on the dominating color level of the paper objects. In the feature selection process, special emphasis is placed on those features, which provides significant information regarding the dominant color level. Initially, seventeen first-order features, such as mean, standard deviation, skewness, kurtosis, dispersion, lowest color level, highest color level, mode of the color level, entropy, energy, lower quartile, upper quartile, histogram tail length on dark side, histogram tail length on light side, median color level, range of the color level, and inter-quartile range, are extracted from the image using Eqs. (1)–(17) to determine the significant features in paper grade identification.

To calculate the first-order features, the gray level histogram of the image is calculated first. The histogram, $h(x)$, is a one dimensional array that represents the number of pixels in the image with a gray level of x. The x parameter can take any value between 0 and $Z-1$, where Z is the number of gray levels in the image. For color images, three histograms are calculated for the three color components: red, green and blue.

$$\text{Mean, } \mu = \frac{\sum_{x=0}^{Z-1} h(x)}{Z} \tag{1}$$

$$\text{Standard Deviation, } \sigma = \sqrt{\frac{\sum_{x=0}^{Z-1} (h(x)-\mu)^2}{Z}} \tag{2}$$

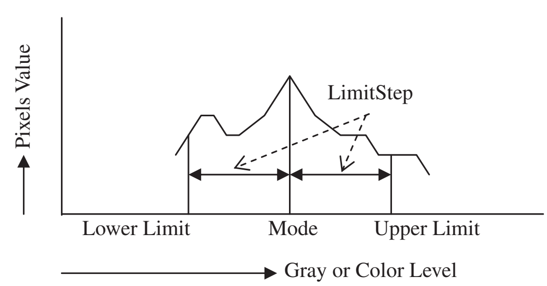


Fig. 3. Determine the lower and upper limits to calculate energy and entropy.

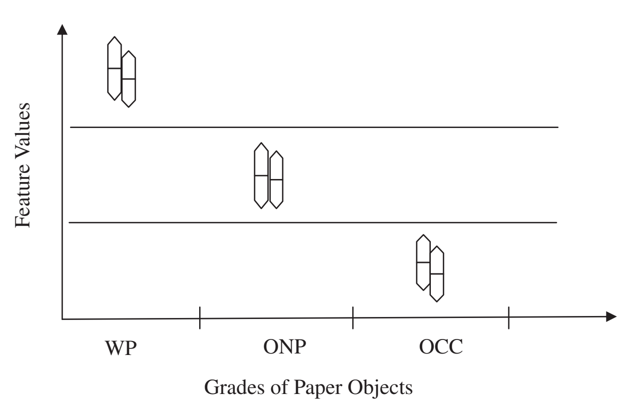


Fig. 4. Values of one feature for different grades of paper objects.

$$\text{Skewness} = \frac{\sum_{x=0}^{Z-1}(h(x)-\mu)^3}{Z\sigma^3} \tag{3}$$

$$\text{Kurtosis} = \frac{\sum_{x=0}^{Z-1}(h(x)-\mu)^4}{Z\sigma^4} \tag{4}$$

$$\text{Dispersion} = \sum_{x=0}^{Z-1}|h(x)-\mu| \tag{5}$$

$$\text{lowest color level, } c = x|h(x)\neq 0 \quad \text{where } 0 \leqslant x < Z \text{ and } h(i) = 0\ \forall i : 0 \leqslant i < x \tag{6}$$

$$\text{highest color level, } d = x|h(x)\neq 0 \quad \text{where } 0 \leqslant x < Z \text{ and } h(i) = 0\ \forall i : x < i < Z \tag{7}$$

$$\text{Mode} = x|h(x) > h(i) : \forall i,\ 0 \leqslant i,\ x < Z,\ i \neq x \tag{8}$$

$$\text{Entropy} = \frac{\sum_{x=LowerLimit}^{UpperLimit} h(x).\log_2(h(x))}{tpixelsp} \tag{9}$$

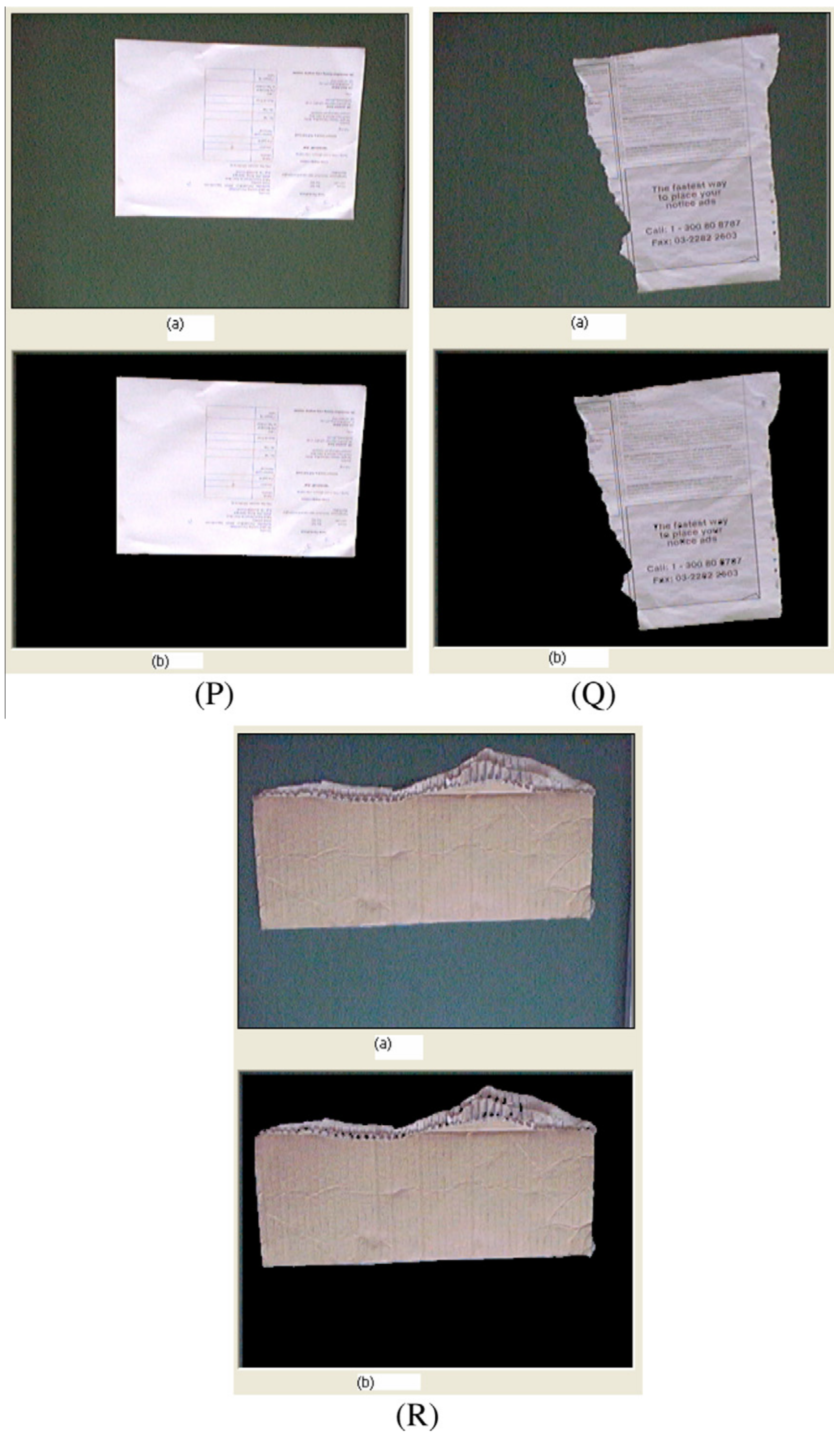


Fig. 5. Original and preprocessed paper image: (P) white paper, (Q) old newsprint paper, and (R) old corrugated cardboard.

where *tpixelsp* is the total number of pixels used to calculate entropy

$$\text{Energy} = \frac{\sum_{x=LowerLimit}^{UpperLimit} x^2.h(x)}{tpixelsg} \tag{10}$$

where *tpixelsg* is the total number of pixels used to calculate energy

The Lower Limit and Upper Limit are used to fix the number of gray levels for the entropy and energy calculation shown in Fig. 3. The terms are determined using the following pseudocode:

```
If (Mode-Lowest)<LimitStep Then
   LowerLimit = Lowest Color Level Value
Else LowerLimit = Mode –LimitStep
If (Highest-Mode)<LimitStep Then
   UpperLimit = Highest Color Level Value
Else UpperLimit = Mode + LimitStep
```

The *LimitStep* stands for the number of gray levels or color levels considered above and below the *Mode* value.

$$\text{lower quartile, } a = y\left|\sum_{x=0}^{x=y} h(x) \geqslant 0.25s \text{ and } \sum_{x=0}^{x=y-1} h(x) < 0.25s\right. \tag{11}$$

$$\text{where } s = \sum_{x=0}^{x=Z-1} h(x)$$

$$\text{upper quartile, } b = y\left|\sum_{x=0}^{x=y} h(x) \geqslant 0.75s \text{ and } \sum_{x=0}^{x=y-1} h(x) < 0.75s\right. \tag{12}$$

$$\text{where } s = \sum_{x=0}^{x=Z-1} h(x)$$

$$\text{histogram tail length on dark side, } e = a - c \tag{13}$$

$$\text{histogram tail length on Light side, } f = d - b \tag{14}$$

$$\text{median color level, } g = o\left(\frac{s}{2}\right) \tag{15}$$

Paper Grade: WP PG Comments: White Office Paper
Pg ID: 50

	BLUE	GREEN	RED	GRAY
1. Mean Value	108.8125	108.8125	108.8125	108.8125
2. Standard Deviation	457.53760	250.58589	280.3676	269.1531
3. Skewness	10.986167	3.9906475	3.404556	2.778388
4. Kurtosis	147.08243	6.1452799	14.62905	6.948726
5. Dispersion	44764.75	41651.25	43738.62	43492.62
6. Entropy	10	9.1830284	9.497977	9.414771
7. Energy	59658	53443	57465	56815
8. Lower quartile (a)	236	223	231	230
9. Upper Quartile (b)	254	239	248	247
10. Lowest Color Level (c)	66	84	82	66
11. Highest Color Level (d)	255	255	255	255
12. Histogram tail length on dark side (e)	170	139	149	164
13. Histogram tail length on light side (f)	1	16	7	8
14. Median Color Level (g)	8	12	9	9
15. Range of Color Level (h)	189	171	173	189
16. Inter-Quartile Range (i)	18	16	17	17
17. Mode Gray or Color Level (j)	255	234	255	249

(a)

Paper Grade: ONP PG Comments: News paper with text
Pg ID: 150

	BLUE	GREEN	RED	GRAY
1. Mean Value	104.58984	104.58984	104.5898	104.5898
2. Standard Deviation	215.02861	239.11215	237.2471	233.7258
3. Skewness	2.3865506	4.1821378	2.735702	2.681843
4. Kurtosis	4.7053834	7.0256573	6.659980	6.304882
5. Dispersion	37202.656	38238.554	38397.73	38189.73
6. Entropy	9.0240659	9.0334707	9.024065	9.009449
7. Energy	35437	30925	33908	33410
8. Lower quartile (a)	178	168	177	175
9. Upper Quartile (b)	197	183	192	190
10. Lowest Color Level (c)	68	70	67	68
11. Highest Color Level (d)	253	246	247	249
12. Histogram tail length on dark side (e)	110	98	110	107
13. Histogram tail length on light side (f)	56	63	55	59
14. Median Color Level (g)	33	32	32	35
15. Range of Color Level (h)	185	176	180	181
16. Inter-Quartile Range (i)	15	15	15	15
17. Mode Gray or Color Level (j)	188	176	184	184

(b)

Paper Grade: OCC PG Comments: Old Corrugated Cardbc
Pg ID: 250

	BLUE	GREEN	RED	GRAY
1. Mean Value	128.01562	128.01562	128.0156	128.0156
2. Standard Deviation	254.58609	271.19164	279.5012	273.8559
3. Skewness	2.2392509	3.6874292	2.600184	2.502700
4. Kurtosis	3.9454888	5.4079625	5.870732	5.322826
5. Dispersion	45238.375	45904.406	46420.56	46262.5
6. Entropy	9	9.2530451	9.270051	9.215610
7. Energy	23936	25905	31730	27000
8. Lower quartile (a)	144	101	76	95
9. Upper Quartile (b)	164	162	179	165
10. Lowest Color Level (c)	71	65	47	71
11. Highest Color Level (d)	249	231	251	244
12. Histogram tail length on dark side (e)	73	36	29	24
13. Histogram tail length on light side (f)	85	69	72	79
14. Median Color Level (g)	50	50	32	50
15. Range of Color Level (h)	178	166	204	173
16. Inter-Quartile Range (i)	20	61	103	70
17. Mode Gray or Color Level (j)	155	170	181	167

(c)

Fig. 6. First order feature values: (a) white paper (b) newsprint paper, and (c) old corrugated cardboard.

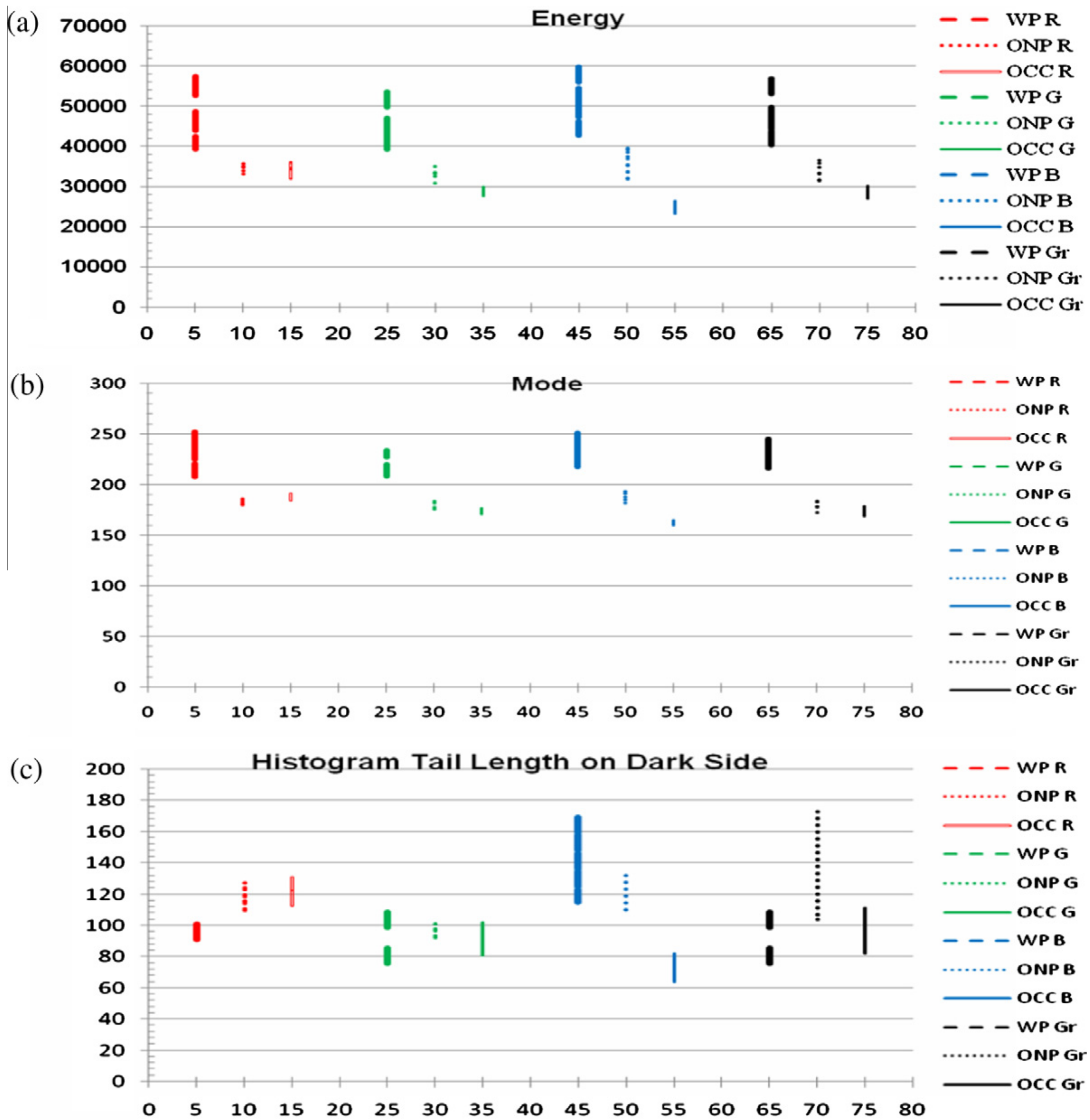


Fig. 7. (a) Energy, (b) mode, (c) histogram tail length on the dark side, (d) histogram tail length on the light side, (e) lower quartile, and (f) upper quartile.

where $o(j)$ is the color or gray level of the jth pixel in the ordered window

$$\text{range of the color or grey levels, } h = d - c \tag{16}$$

$$\text{inter quartile range, } i = b - a \tag{17}$$

To determine the significance of one feature, the different paper objects are plotted along the x-axis, and the feature values are plotted along the y-axis as shown in Fig. 4. Using the plotted values the inter-class and intra-class variation are determined. When the feature is good at discriminating certain classes from each other based on inter-class variation and intra-class variation then the feature is considered in the template construction.

Finally, the template of the paper grade is

$$\text{Template, } T = \{\text{PG}, \text{PID}, \{E_g, Q_L, Q_U, T_D, T_L, M\} \text{ repeat for red, green, blue, and gray}\} \tag{18}$$

where PG = paper grade; PID = paper grade ID number, E_g = energy; Q_L = lower quartile; Q_U = upper quartile; M = mode; T_D = histogram tail length on dark side; T_L = histogram tail length on light side.

4.4. Technique applied for identification

The K-nearest neighbor (KNN) (Pham & Alcock, 2003; Teknomo, 2009) algorithm is applied for paper object grade identification. The KNN algorithm is described in Procedure 1.

Procedure 1: KNN algorithm

1. Determine parameter K, which is the number of nearest neighbors.
2. Calculate the distances between the candidate paper object and all the training samples.
3. Sort the distances and determine the K-nearest neighbors based on minimum distances.

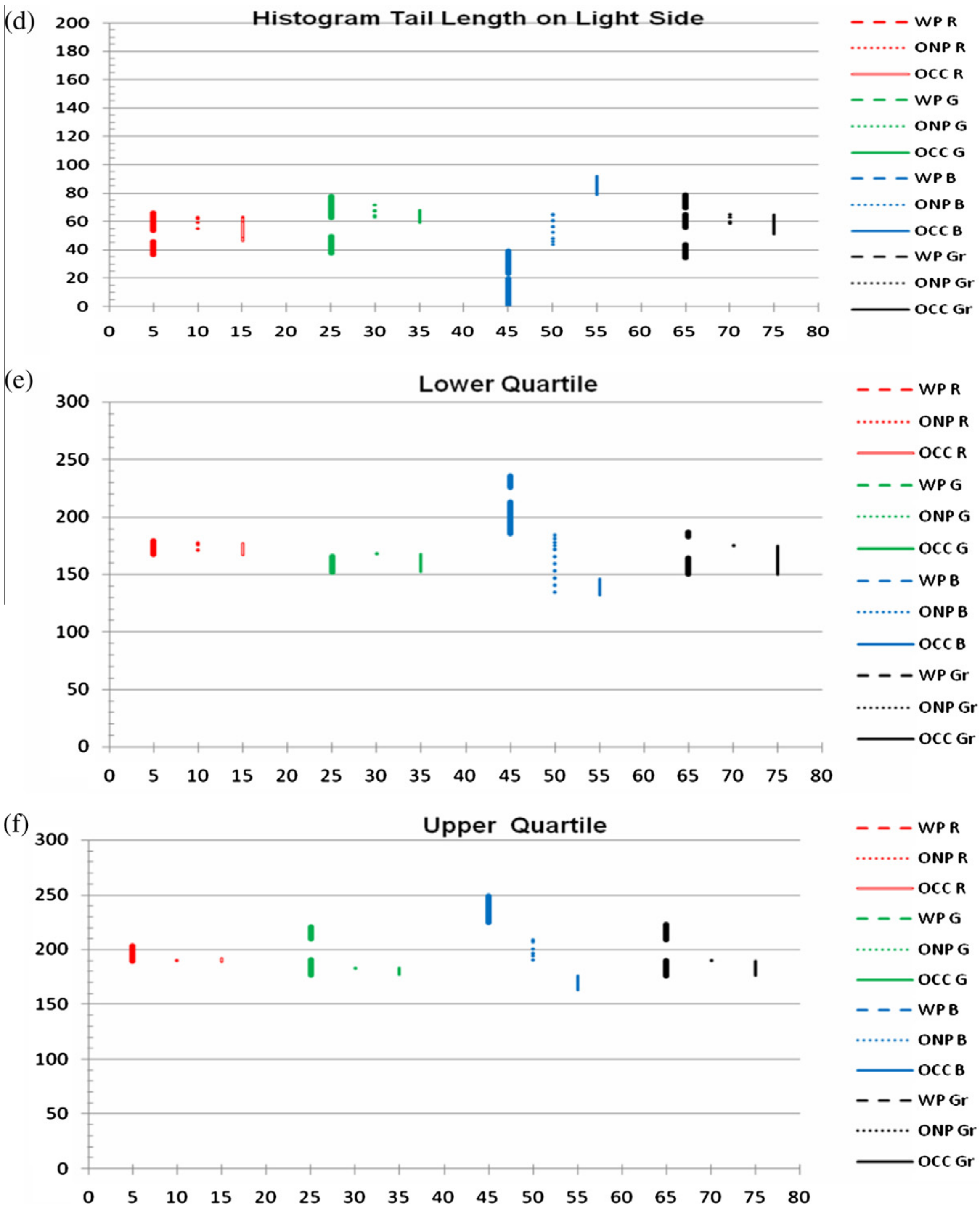


Fig. 7 (*continued*)

Table 1
Identification success rate for the two distance metrics at different values of *K*.

Method	K-value	Name of the distance metrics	Correct identification rate (%)
K-nearest neighbor (KNN)	3	Absolute distance	90
	5	Absolute distance	93

4. Gather the paper grade category of the nearest neighbors.
5. Use the simple majority of the paper grade category of the nearest neighbors as the paper grade of the candidate paper object.

The distances between candidate paper objects and the training samples are calculated using the absolute distance metric (Rahman, 2005) shown in Eq. (19). Let us suppose $F = (f_1, f_2, \ldots, f_n)$ represents the n-dimensional template of the paper object category in the reference database and $Y = (y_1, y_2, \ldots, y_n)$ represents the template of the candidate paper object, then the absolute distance, D is

$$D = \sum_{j=1}^{n} |y_j - f_j| < \varepsilon_a \qquad (19)$$

where ε_a is the threshold value for the distance metric.

5. Experimental results and discussion

In order to develop the proposed system, the software tools Microsoft Visual Basic 6.0 for front-end application, Microsoft Access 2000 for backend database support, and MS Excel 2000 for data sheets and experimental results analysis are used.

The three types of waste papers, white paper (WP), old newsprint paper (ONP) and old corrugated cardboard (OCC), were considered in this experiment because of their dominating role in waste papers with 1500 paper samples. Different templates were created for the same grade of papers. Ten samples were considered to create an accurate feature vector for the reference template paper grade.

In this section, a relative comparison is made based on the outcomes of the proposed method for WP, ONP and OCC. In Fig. 5, the images P(a), Q(a) and R(a) represent the original images of WP, ONP and OCC with background noise; in addition, the images P(b), Q(b) and R(b) represent the preprocessed images of WP, ONP and OCC, respectively. The calculated first-order features of the WP, ONP and OCC papers are illustrated in Fig. 6. The discriminating capabilities of the significant feature energy, mode, histogram tail length on the dark side, histogram tail length on the light side, lower quartile, and upper quartile are illustrated in Fig. 7. Fig. 7(a–f) are drawn considering Fig. 4. The feature energy is the most significant for red, green, blue and gray scales for discriminating certain classes from each other based on inter-class variation and intra-class variation. The feature mode separates white paper from others well using red, green and gray scales, but when the blue scale is included, then the feature mode also shows discrimination ability among different grades of papers. The histogram tail length on the dark side and the histogram tail length on light side, lower quartile and upper quartile features provided good performance in paper grade identification in the blue scale. When all RGB components are considered, then the gray scale features did not play an important role in paper grade identification. Thus, the red, green and blue component features are sufficient to separate different grades of papers. Moreover, blue scale features have shown the most significant power to identify different paper grades.

The success rates of the paper grade identification process for absolute distance metrics at different values of K are tabulated in Table 1. The correct identification rate is calculated based on the percentage of the number of paper objects that are classified into their respective paper grades. Using the absolute distance metric with KNN, the results are 90% and 93% for $k = 3$ and $k = 5$, respectively.

6. Conclusion

This work emphasizes the development of a new method for an automated paper sorting system that utilizes an image processing technique with the K-NN classifier. The proposed system performance for correct paper grade identification is more than 90% with an estimated throughput of 28,800 paper objects per hour with a conveyor belt width of 18″. The weight of the throughput depends on the size and grade of the paper objects.

Another important idea that has been implemented in this proposed system is adaptability to new subcategories of the primary paper grades. The wide ranges of subcategories of paper grades are used to train the system to recognize new subcategories, and thus the system is scalable and able to provide robust decisions in paper grade identification.

In the experiment, it is observed that the performance of the vision system is influenced by the lighting arrangement. Thus, in order to achieve the best performance from this method, the lighting consistency must be maintained during the enrollment and identification phases of this system.

The most important point addressed in the proposed method is that the method, which uses computer vision, can be implemented easily to sort multiple grades of paper. Moreover, the computational burden for paper identification using the proposed technique is simple. The proposed method can identify three major paper grades, WP, ONP and OCC, using simple image processing techniques. Further work should focus on all paper grades and extend the method to other solid waste sorting, such as plastic, metal, and glass.

Acknowledgements

The project is sponsored by the Universiti Kebangsaan Malaysia under the OUP Grant of UKM-OUP-ICT-36-184/2010.

References

Bialski, A., Gentile, C., & Sepall, O. (1978). *Paper sorting method*. US Patent No. 4124,168.

Bialski, A., Gentile, C., & Sepall, O. (1980). *Paper sorting apparatus*. US Patent No. 4, 236, 676.

Bruner, R. S., Morgan, D. R., Kenny, G. R., Gaddis, P. G., Lee, D., & Roggow, J. M. (2003). *System and method for sensing white paper*. US Patent No. 6570,653.

Brunner, C. C., Maristany, A. G., Butler, D.A., Leeuwen, D. V., & Funck, J.W. (1992). An evaluation of colour spaces for detecting defects in Douglas-fir veneer. *Industrial Metrology*, Part 2, pp. 169–184.

Burke, M. W. (1996). *Image acquisition: Handbook of machine vision engineering* (Vol. 1). London: Chapman and Hall.

Chakravarthi, K. K. (2006). *Development of online stiffness sensor for high speed sorting of recovered paper*. Master's thesis, North Carolina State University, Raleigh, NC, USA.

Chandini, C. M. (2001). *Design of lignin sensor for identification of waste paper grades for an automatic waste paper sorting system*. Master's thesis, North Carolina State University, Raleigh, NC, USA.

Doak, A. G., Roe, M. G., & Kenny, G. R. (2007). *Multi-grade object sorting system and method*. US Patent No. US2007/0002326.

Doak, A. G., Roe, M. G., & Kenny, G. R. (2007). *Multi-grade object sorting system and method*. US Patent No. 7173709.

Eixelberger, R., Friedl, P., & Gschweitl, K. (2003). *Method and apparatus for sorting waste paper of different grades and conditions*. US Patent No. 6506,991.

Faibish, S., Bacakoglu, H., & Goldenberg, A. A. (1997). An eye-hand system for automated paper recycling. In *Proceedings of the IEEE international conference on robotics and automation* (pp. 9–14). New Mexico: Albuquerque.

Grubbs, M., Kenny, G. R., & Gaddis, P. G. (2001). *Paper sorting system*. US Patent No.6, 250, 472.

Gschweitl, H., & Heinz, K. (1998). *Method for sorting waste paper*. European Patent, EP0873797.

Holz, E., & Hutzler, H. (1986). *Apparatus for Dissolving and Sorting Waste paper*. US Patent No. 4632,320.

Hottenstein, F. A., Kenny, G. R., Friberg, T., & Jackson, M. (2000). High-speed automated optical sorting of recovered paper. *Proc. TAPPI recycling symposium*, Vol. 1 (pp. 149–158). Atlanta, GA.

Khalfan, Z. (2002). *Optical paper sorter*. US Patent No. 6335,501.

Khalfan, Z., & Greenspan, S. (2006). *Optical paper sorting method device and apparatus*. US Patent No. 2006/0124511.

Khalfan, Z., & Greenspan, S. (2006). *Optical paper sorting method device and apparatus*. US Patent No. 7081,594.

Logitech QuickCam Pro 4000 Web Camera Specification, (2009). Available at <http://www.reviews.cnet.com/webcams/quickcam-pro-4000-web/4507-6502_7-20276742.html?tag=rnav/> Accessed 09.09.

Mallapragada, V. (2004). *Online lignin sensor for high speed sorting of newsprint from mixed waste*. Master's thesis, North Carolina State University, Raleigh, NC, USA.

MSS FiberSort™ Module (2009). Available at <http://www.magsep.com/node/19/> Accessed 09.09.

MSS MultiWave™ (2009). Available at <http://www.magsep.com/node/20/> Accessed 09.09.

MSS Optical sorting Systems (2009). Available at <http://www.magsep.com/> Accessed 09.09.

MSS Paper Sorting Systems (2009). Available at <http://www.recyclingtoday.com/news/news.asp?ID=742/> Accessed 09.09.

MSS PaperSort™ (2009). Available at <http://www.magsep.com/node/21/> Accessed 09.09.

Ortner, H., Bahr, T., & Musselmann, W. (1980). *Process and apparatus for treating waste paper*. US Patent No. 4231,526.

Paper Competitors (2009). Available at <http://separation.wikispaces.com/Paper+Competitors/> Accessed 09.09.

Paper Grades (2009). Available at <http://www.paperonweb.com/ppmanf.htm/> Accessed 02.09.

Pellenc, MIR Technology (2009). Available at <http://www.pellencst.com/en/21/54/mir-technology/>, accessed September 2009.

Pellenc, Paper sorting Company (2009). Available at <http://www.pellencst.com/en/1/products/> Accessed 09.09.

Pham, D. T., & Alcock, R. J. (2003). *Image Acquisition and Enhancement, Feature Extraction and Selection, and Classification Chap. 2, 4, and 5 in Smart Inspection System -Techniques and Applications of Intelligent Vision*. Great Britain: Academic Press. pp. 38-47, 55-59, 99-102, 115-116, 133-134.

Rahman, M. O. (2005). *Methodology of the proposed system chap. 4', in an efficient method of person verification using hand geometry*. M.Sc. Engg Thesis, BUET, Dhaka, Bangladesh, pp. 52–53.

Rahman, M. O., Hussain, A., Scavino, E., Hannan, M. A., & Basri, H. (2009). *Recyclable waste paper sorting using template matching*. LNCS 5857, Springer-Verlag, pp. 467–478.

Rahman, M. O., Hussain, A., Scavino, E., Hannan, M. A., & Basri, H. (2009). Segregating recyclable waste papers using co-occurrence features. *The 9th WSEAS international conference on applied computer science (ACS'09)* (pp. 187–191). Italy: Genova.

Rahman, M. O., Hannan, M. A., Scavino, E., Hussain, A., & Basri, H. (2009). An efficient paper grade identification method for automatic recyclable waste paper sorting. *European Journal of Scientific Research, 25*(1), 96–103.

Ramasubramanian, M. K., Venditti, R. A., & Gillella, P. K. (2008). *Sensor systems for high speed intelligent sorting of waste paper in recycling*. Available at <http://www.osti.gov/bridge/servlets/purl/919471-VGwxA0/919471.PDF/> Accessed 11.08.

Ramasubramanian, M. K., Venditti, R. A., Ammineni, C. M., & Mallapragada, M. (2005). Optical sensor for noncontact measurement of lignin content in high-speed moving paper surfaces. *IEEE Sensors Journal, 5*(5), 1132–1139.

RedWave, BT-Wolfgang Binder (2009). Available at <http://www.redwave.at/altpapierrecycling/> Accessed 11.09.

Remade Scotland, (2005). Initial business case for utilisation of automated optical paper sorting technology. Caledonian Environment Centre, Glasgow Caledonian University. Web address: <http://www.remade.org.uk/filesInitialBusinessCaseforUtilisationofAutomatedOpticalPaperSortingTechnology_16122222124.pdf/> Accessed 11.09.

Sandberg, N. H. (1932). *Sorting device for waste paper*. US Patent No. 1847,265.

Spencer, M. W. (1994). *Waste paper disintegration, classification and pulping system*. US Patent No. 5324,389.

SPM Paper Recycling Sdn Bhd (2009). Kuala Lumpur, Malaysia. Available at <http://spmholdings.com.my/index.php?page=staticpages/paper/> Accessed 03.09.

Teknomo, K. (2009). Available at <http://people.revoledu.com/kardi/tutorial/KNN/index.html/> Accessed 09.09.

TITECH autosort (2009). Available at <http://www.titech.com/recycling-equipment/titech-autosort-10715/> Accessed 09.09.

Venditti, R. A., Ramasubramanian, M. K., & Kalyan, C. K. (2007). A noncontact sensor for the identification of paper and board samples on a high speed sorting conveyor. *Appita Journal: Journal of the Technical Association of the Australian and New Zealand Pulp and Paper Industry, 60*(5), 366–371.

Würtz, H., & Breum, N. O. (1997). Exposure to microorganisms during manual sorting of recyclable paper of different quality. *Annals of Agricultural and Environmental Medicine, 4*, 129–135.