International Journal of Innovative Computing, Information and Control
Volume **8**, Number **1(A)**, January **2012**

ICIC International ©2012 ISSN 1349-4198
pp. 471-488

REAL-TIME WASTE PAPER GRADING USING CBR APPROACH

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

Faculty of Engineering and Built Environment
Universiti Kebangsaan Malaysia
Bangi 43600, Selangor DE, Malaysia
osiur_rahman@yahoo.com; { aini; scavino; eehannan; drhb }@eng.ukm.my

Received September 2010; revised February 2011

Abstract. *The popularity of optical paper sorting systems has increased because of inadequate throughput of mechanical paper sorting systems. However, the implementation of existing optical methods is still complex, expensive and only able to segregate two types of paper at one time. Moreover, no image processing or intelligent techniques were used to extract the features from the paper objects. This research attempts to develop a smart vision sensing system that is able to identify waste paper grade using case-based reasoning. In order to construct the reference template database, the mode and energy of the region-of-interest in the paper object image are considered. The paper grade is identified based on the maximum occurrence of a specific reference template in the paper object image. The classification success rates with window size* 3×3 *pixels for white paper, old newsprint paper and old corrugated cardboard are 94%, 92% and 98%, respectively, while the achieved average classification success rate is 95.17%. This remarkable achievement is possible due to the accurate identification and dynamic sorting of all grades of papers. This method is clearly superior to other existing techniques in terms of throughput, performance in identification, adaptability with new paper grades and cost of implementation.*
Keywords: Waste paper sorting, Grades of paper, Template matching

1. **Introduction.** Conservation of resources is a prime concern in ensuring a sustainable environment [1]. Waste paper recycling is one of the key strategies in resource conservation because waste paper makes up the largest percentage of office waste stream [2]. Moreover, by means of effective waste paper recycling, any country will benefit from reduced landfilling, save money, improve image, conserve natural resources, save energy and reduce greenhouse gas emissions [2-4]. It is estimated that making one ton of paper from recycled fiber saves approximately 17 trees, 3.3 cubic yards of landfill space, 360 gallons of water, 100 gallons of gasoline, 60 pounds of air pollutants and 10401 kilowatts of electricity [2-4]. Manufacturing paper from virgin timber uses 60% more energy with respect to using recycled paper [3,4]. Pati et al. [5] revealed that recycled paper was a more economical alternative to wood as a raw material. Laurijssen et al. [6] concluded that recycling was beneficial with regards to CO_2 emissions and (feedstock) energy use over the paper life-cycle. Moreover, they showed that recycling of paper led to increased biomass availability that can significantly be utilized for energy generation. In addition, a highly sorted paper stream facilitates high quality end product, and saves processing chemicals and energy because various grades of papers are subjected to different recycling processes [6].

In this work, the term “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 [7]. Automated paper sorting systems provide more efficient and effective advantages over human inspection from the perspective of worker fatigue, throughput, speed and accuracy. As a result, many automated mechanical [8-14] and optical [15-19] paper sorting methods have been developed to fill the paper sorting demand. The constraints of the aforesaid mechanical and optical methods were mentioned elsewhere in [4,19,20].

From literature review, it was noted that eight sensors, namely (i) ultrasonic [8], (ii) lignin [9,10,15,22-24], (iii) gloss [10,18,25], (iv) stiffness [11,26,27], (v) color sensors [16-19,25], (vi) Near Infrared (NIR), (vii) Mid-Infrared (MIR) and (viii) Visual Imaging Sensor (VIS) have been used in paper grade identification 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.

In the paper recycling industry, the Manufacturers Standardization Society (MSS) and TiTech are the top two competitors for sensor-based sorting. They own the technology and create partnerships with recyclers [28]. TiTech Autosort combines Visual Imaging Sensor (VIS) and NIR sensors into a universal modular sorting system [29]. The VIS sensor is used to recognize print media with Cyan, Magenta, Yellow and blacK (CMYK) spectral analysis. On the other hand, the MSS exploited NIR, color, gloss and lignin sensors in their paper and plastics sorting system. Pellenc, a French company, recently entered the paper sorting market with two machines: the Zephyr and the Boreas [30]. 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 [31]. RedWave utilizes NIR and color camera technology on all of their machines [32]. However, they did not develop any multigrade paper sorting system. Finally, paper sorting performance in industry reaches 80% for both MSS and TiTech systems [33].

Existing automated sorting systems that deploy state-of-the-art technology such as Near Infra-Red, Infra-Red and X-ray, require high investments [34], are still complex, and sometimes offer limited reliability. All these 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.

Three electronic image-based waste paper sorting techniques were proposed by Rahman et al. [4,20,21] to overcome the drawbacks of the previous optical methods. The first technique [4] 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 of the entire paper object, it may provide misleading information regarding the paper grade. The other two methods, template matching [20] and co-occurrence features [21], achieved good paper grade identification success rate. However, for real time applications, both methods are slow because of the significant computing time.

The primary goal of this study is to develop a Smart Vision Sensing (SVS) system capable of separating the different grades of paper using Case-Based Reasoning (CBR) artificial intelligence technique [35]. The case-based reference database is created using first order features [36] i.e., mode and energy. The mode and energy are calculated from the region-of-interest (ROI) in the paper object image with RGB color space [37]. In this system, the paper object image is divided into N-cells based on given window size [36] while the first order features of the N-cells are considered to create respective candidate templates. The algorithm provides robust and fast result because the proposed method avoids the extra computational burden for preprocessing and only two features, namely mode and energy for each RGB component [36,37] are used to identify the dominating color value of the paper object image.

2. **The SVS System for Waste Paper Grade Identification.** Figure 1 illustrates the basic block diagram of the recyclable waste paper grade identification system. The SVS system operates in two phases, i.e., enrollment and identification. Both phases have some common components. The enrollment phase consists of image acquisition, feature extraction, template construction and creation of the case-based reference database. The identification phase consists of image acquisition, window-based subdivision of the entire paper image into N-cells, feature extraction, construction of N-candidate templates for N-cells, matching and decision. The matching and decision components are implemented using CBR steps, i.e., retrieve, use and revise [35,38]. In Subsections 2.1 to 2.4, all processes of both enrollment and identification phases are discussed.

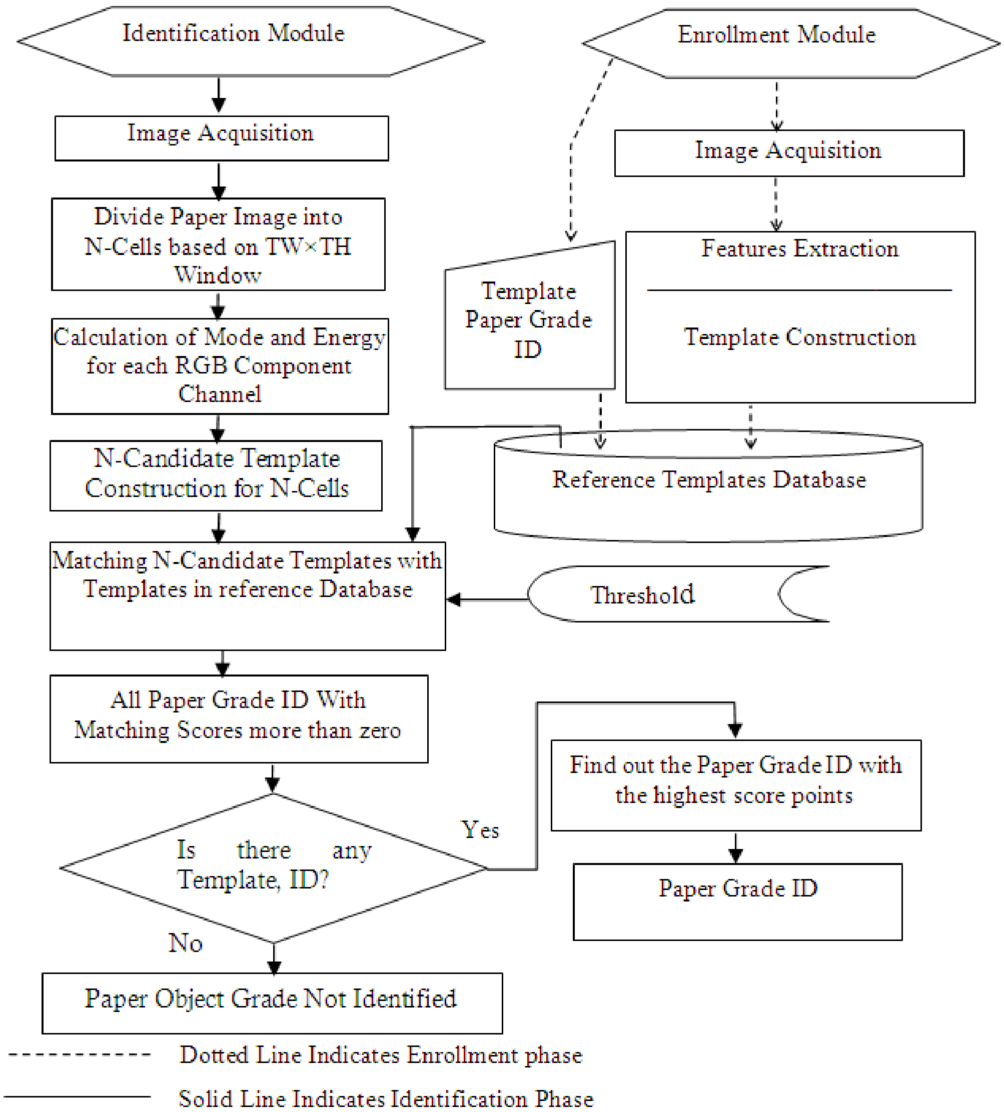


FIGURE 1. Block diagram of the SVS real-time waste paper grade identification system

2.1. **Image acquisition.** In SVS system, 320×240 RGB images are captured from the inspection zone on the conveyor belt using a Logitech QuickCam Pro 4000 Web Camera [39]. To set the webcam properties, the brightness, contrast and saturation were adjusted to 50%, 50% and 100% of their respective scales. Because front lighting-directional-darkfield illumination is widely used in surface scratches or texture analysis [36,40], this illumination technique is adopted for this experiment.

2.2. **Features extraction.** The features extraction process of the enrollment phase is somewhat different from the identification phase. The enrollment feature extraction phase consists of (i) ROI selection, (ii) RGB histogram calculation and (iii) mode and energy calculation for each RGB component of the ROI of the paper object image. In addition to these three processes, the identification phase further consists of (i) window-based paper object image subdivision, (ii) RGB histogram calculation for all N-cells and (iii) mode and energy calculation for each RGB component of all N-cells.

2.2.1. *ROI selection in paper object image for enrollment phase.* The ROI is selected by clicking the mouse twice on the paper object image. The first point (*top, left*) Top-Left and the second point (*bottom, right*) Bottom-Right of the ROI in the paper object image are acquired from the first and second clicks, respectively. The Top-Left and Bottom-Right points of the interested region are shown in Figure 2.

(a) (b)

FIGURE 2. (a) Original captured image and (b) ROI in the old newsprint paper object image

2.2.2. *Window-based paper image subdivision and features extraction.* After obtaining the value of window (or cell) width, *TW*, and window (or cell) height, *TH*, for the cell or window image, the paper image is divided into N-cells shown in Figure 3. The features for all N-cells are obtained using the *WindowBasedSubdivision* and *CandidateTemplateNCell* procedures shown in Tables 1 and 2. The *WindowBasedSubdivision* procedure is used to divide the paper object image into N-cells with width, *TW* and height, *TH*, while the *CandidateTemplateNCell* procedure calculates the mode and energy of RGB components for all N-cells using Equations (1) and (2) as follows:

$$Mode = x \,|h(x) > h(i) : \forall i, 0 \le i, x < Z, i \ne x \tag{1}$$

$$Energy = \frac{\sum_{x=0}^{Z-1} x^2 \cdot h(x)}{pCell} \tag{2}$$

where

$h(x)$: histogram of three RGB channels red, green and blue (one dimensional array of the number of pixels in the image with a color level of x);
x: color level (value range between 0 and $Z-1$);
Z: number of color levels in the image (value range 0 and 255), and
$pCell$: total number of pixels in a cell or window.

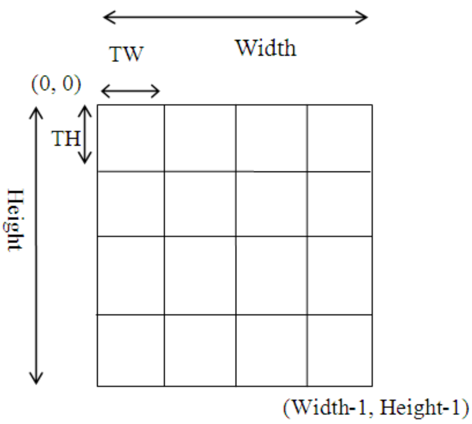


FIGURE 3. Template or cell image (TW × TH) and search image (Width × Height)

TABLE 1. Code for window-based sub-division

Procedure	Code
WindowBasedSubdivision	1. INPUT TW, TH 'TW and TH stands for Width and Height of Cell or window 2. MCOL = CInt(picwidth / TW) 'MCOL stands for Number of Columns 3. NROW = CInt(picheight / TH) 'NROW stands for Number of Rows 4. NCELL = MCOL * NROW 'NCELL stands for number of Cells in Paper Image 5. For i = 1 To MCOL 6. For j = 1 To NROW 7. StartCellWidth = (i - 1) * TW 8. EndCellWidth = (i * TW) - 1 9. StartCellHeight = (j - 1) * TH 10. EndCellHeight = (j * TH) - 1 11. Call *CandidateTemplateNCell*(StartCellWidth, EndCellWidth, StartCellHeight, EndCellHeight) 'Store the Mode and Energy of the RGB components for each cell image 12. CellFeatureModeRed(i, j) = CInt(cellModeRed) 13. CellFeatureEnergyRed(i, j) = CLng(cellEnergyRed) 14. CellFeatureModeGreen(i, j) = CInt(cellModeGreen) 15. CellFeatureEnergyGreen(i, j) = CLng(cellEnergyGreen) 16. CellFeatureModeBlue(i, j) = CInt(cellModeBlue) 17. CellFeatureEnergyBlue(i, j) = CLng(cellEnergyBlue) 'End: Store the Mode and Energy of the RGB components for each cell image 18. Next j 19. Next i

Table 2. Code for feature extraction for all N-cell

Procedure	Code
CandidateTemplateNCell	1. Initialize CellHistogramRed(cellRedL), CellHistogramGreen(cellGreenL), CellHistogramBlue(cellBlueL), NoPixelsCell, SumRedEnergy , SumGreenEnergy and SumBlueEnergy with 0 2. For m = StartCellWidth To EndCellWidth 3. For n = StartCellHeight To EndCellHeight 'Code for Red 4. pointRed = red(m, n) 'Calculate the Histogram 5. CellHistogramRed(pointRed) = CellHistogramRed(pointRed) + 1 6. NoPixelsCell = NoPixelsCell + 1 7. Next n 8. Next m 'Calculate the Mode & Energy for Red Compponent 9. cellRedMode = CellHistogramRed(0) 10. For cellRedL = 0 To 255 11. If cellRedMode <= CellHistogramRed(cellRedL) Then 12. cellRedMode = CellHistogramRed(cellRedL) 13. cellModeRed = cellRedL 14. End If 'Calculate Energy for Red component 15. SumRedEnergy = SumRedEnergy + cellRedL * cellRedL * CellHistogramRed(cellRedL) 16. Next cellRedL 17. cellEnergyRed = CLng(SumRedEnergy / NoPixelsCell) '(NoPixelsCell = TW * TH) 18. Repeat Step 2 to Step 17 for Green and Blue Components.

2.3. **Template construction.** The template of the paper grade is defined by the following equation:

$$Template,\ T = \{PID,\ \{Mode,\ Energy\}\ repeat\ for\ red,\ green\ and\ blue\} \tag{3}$$

where

PID: paper grade ID number;
Mode: mode of the RGB components;
Energy: energy of the RGB components.

2.4. **Matching and decision.** The principal function of the matching process is to calculate the score of the case-based reference templates. The matching process consists of *retrieve*, *reuse* and *revised* stages of the CBR protocol [35,38]. After retrieving the case-based reference templates from the case-based reference template database, the matching score of each case-based reference template is calculated while comparing with N-cell candidate templates. Initially, in the matching process, the distances of mode and energy between candidate and reference templates are calculated. A check is made whether their values are less than the threshold value for each RGB channel. For any candidate template, if the difference between candidate template and the reference template is less than the threshold value for all RGB components then the respective reference template scores 1 point. The above process is repeated for the next candidate template. On the other hand, if the reference template does not satisfy all of the three red, green and blue channels, then it will proceed with the next reference template and apply the same conditions again. The processes are continued until the last of the reference templates. In this way, the maximum occurrence of a specific case-based reference template in the paper object image achieves the highest scored points. Finally, the template ID with the highest score points is identified as the candidate paper object grade ID. On the other hand, if the matching score is 0 for all the case-based reference templates then *retain* stage of CBR [35,38] is performed to train the system for a new case. In this experiment, the total number of reference template is 150, and it is divided into three groups, namely 0 to 49 ID for WP, 50 to 99 ID for ONP and 100 to 149 ID for OCC. The candidate paper grade name is identified based on the group of the paper grade ID.

The variables *CellFeatureModeRed*, *CellFeatureEnergyRed*, *CellFeatureModeGreen*, *Cell FeatureEnergyGreen*, *CellFeatureModeBlue* and *CellFeatureEnergyBlue* are used to store the candidate template features mode and energy of the red, green and blue components, respectively. The variables *NoOfTemplate* and *ScoreTemplate (k)* are used for the total number of reference templates and store the score point of the respective reference template respectively. The variables $MCOL$ and $NROW$ are used to indicate the number of cells in the paper image along the column and row respectively. In this experiment, the values of *ThresholdModeRed*, *ThresholdEnergyRed*, *ThresholdModeGreen*, *ThresholdEnergyGreen*, *ThresholdModeBlue* and *ThresholdEnergyBlue* are set at 2, 1500, 2, 1500, 2 and 1500, respectively. The *PaperGradeFlag* is initialized by -1. If any case from the case-based reference database matching with the any cell in the candidate paper object then the *PaperGradeFlag* value turn to positive. After end of the matching process if the *PaperGradeFlag* value remains negative then declared "Unidentified Paper Grade".

The code for both "matching" and "decision" processes are shown in Table 3.

3. Experimental Results and Discussion.

The software tool Microsoft Visual Basic 6.0 is used to develop the SVS system. The experimental results are analyzed using Matlab and MS Excel 2000. The experimental results and discussion section consists of five subsections, namely, (i) case-based reference database creation, (ii) the classification success rate of the SVS system, (iii) comparison with previous results and technique published in literature, (iv) justification of the SVS system from the economic and environmental point of view and (v) the deficiencies of the SVS system.

3.1. Case-based reference database creation.

The success of the CBR approach is significantly depended on the creation of the efficient and perfect reference database. The mode and energy of the ROI from different paper objects are collected using the technique describe in Section 2.2.1 and Equations (1) and (2). The calculated energy values of different paper object samples of the same grade namely WP, are shown in Figure 4. In Figure 4, it is observed that the red, green and blue components energy values are different for different samples of the same paper grade. In order to avoid extra computational time for the unnecessary data retrieval and matching, the case-based reference database size reduced by discarding repeated sample data. For example, in case of WP Samples 12 and 13 shown in Figure 4, the energy values for red, green and blue components are closed each other. Moreover, the mode values are 240 and 238 for red components, 238 and 237 for green components and 255 and 255 for blue components of the Samples 12 and 13, respectively. As a consequence, the sample number 13 is discarded and sample number 12 is considered as first-come-first-serve basis for the enrollment in case-based reference database. The above process was applied for the first time enrollment of the case-based reference database. During running the SVS system, unrecognized candidate

Table 3. Code for "Matching" and "Decision" processes

Process	Pseudocode

Matching	For *i* = 1 To *MCOL*
	For *j* = 1 To *NROW*
	For *k* = 0 To (*NoOfTemplate* - 1)
	DisModeR = Abs(TemplateModeRed(k) - CellFeatureModeRed(i, j))
	DisEgR = Abs(TemplateEnergyRed(k) - CellFeatureEnergyRed(i, j))
	DisModeG = Abs(TemplateModeGreen(k) - CellFeatureModeGreen(i, j))
	DisEgG =Abs(TemplateEnergyGreen(k) - CellFeatureEnergyGreen(i, j))
	DisModeB = Abs(TemplateModeBlue(k) - CellFeatureModeBlue(i, j))
	DisEgB = Abs(TemplateEnergyBlue(k) - CellFeatureEnergyBlue(i, j))
	If (DisModeR <= ThresholdModeRed)AND (DisEgR <= ThresholdEnergyRed)Then
	RedPass = 1
	End If
	If(DisModeG <= ThresholdModeGreen)AND(DisEgG <= ThresholdEnergyGreen)Then
	GreenPass = 1
	End If
	If (DisModeB <= ThresholdModeBlue)AND(DisEgB <= ThresholdEnergyBlue)Then
	BluePass = 1
	End If
	Pass = RedPass + GreenPass + BluePass
	If Pass = 3 Then
	If PaperGradeFlag = -1 Then PaperGradeFlag = 1 End if
	ScoreTemplate(k) = ScoreTemplate(k) + 1
	GoTo nextcell
	End If
	Next *k*
	nextcell:
	Next *j*
	Next *i*

Decision	If PaperGradeFlag = -1 Then Print "Unidentified Paper Grade" Else
	MaxScoreTemplate = *ScoreTemplate*(0)
	For *k* = 0 To (*NoOfTemplate* - 1)
	If *MaxScoreTemplate* < *ScoreTemplate(k)* Then
	MaxScoreTemplate = *ScoreTemplate(k)*
	PaperGradeID = *k*
	End If
	Next *k*
	Print the Paper grade name of the *PaperGradeID* , k
	End If

paper samples also enrolled in the case-based reference database giving new paper grade ID.

3.2. **Classification success rate of the SVS system.** In this section, a comparison is made between the outcomes of the SVS system for three types of waste papers, i.e., old corrugated cardboard (OCC), old newspaper (ONP) and white paper (WP). These three types of waste papers had been taken in this experiment because these three types

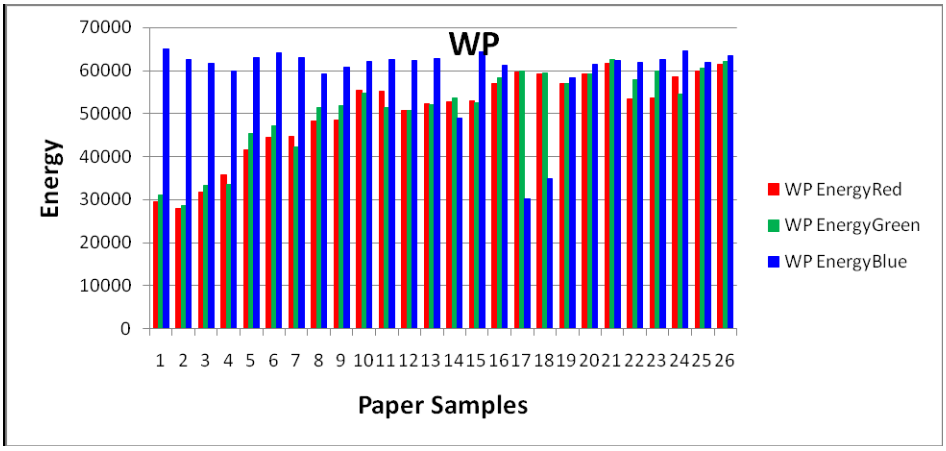


FIGURE 4. The energy values of the red, green and blue components for different samples of the WP paper

of waste papers constitute the main contents of waste papers. In this experiment, 485 WP samples, 375 ONP samples and 485 OCC samples (in total 1345) were considered.

Figure 5 illustrates the images of ONP in both original and 15×15 pixels window-based subdivided forms. In the creation of case-based reference database, only the ROI from the paper object images are considered. Hence, during identification phase, when the paper object images are divided into N-cell then at least one or more than one cell exist without printed material on the surface of the paper objects shown in Figure 5(b). As a consequence, the paper grade identification rates increased significantly.

(a) (b)

FIGURE 5. Paper object image: (a) original and (b) subdivided 15×15

Figure 6 illustrates the visualization of the waste paper objects using mode and energy of the RGB components. It is observed that the different grades of papers, namely WP, ONP and OCC, are separable from each other in the mode and energy spaces. In Figure 6(a), the mode values of the WP, ONP and OCC constitute the different clusters in the mode space. Moreover, the different clusters are constituted for the different paper grades in the energy space shown in Figure 6(b). In both Figures 6(a) and 6(b), few paper objects are located in their foreign clustered such as few WP objects located in the ONP objects clustered and few ONP objects located in the OCC objects clustered. In this situation, most of the paper objects were classified among their respective paper grades either using mode or energy space but for the robust identification the mode and energy are used together in the matching process.

Another important throughput from the observation of the visualization of the paper objects in Figure 6 is to select the appropriate paper objects identification technique. In Figure 6, it is observed that it is essential to select non-linear classifier to identify the grades of different paper objects. In this problem, case-based reasoning is tuned perfectly to recognize the grades of different paper objects.

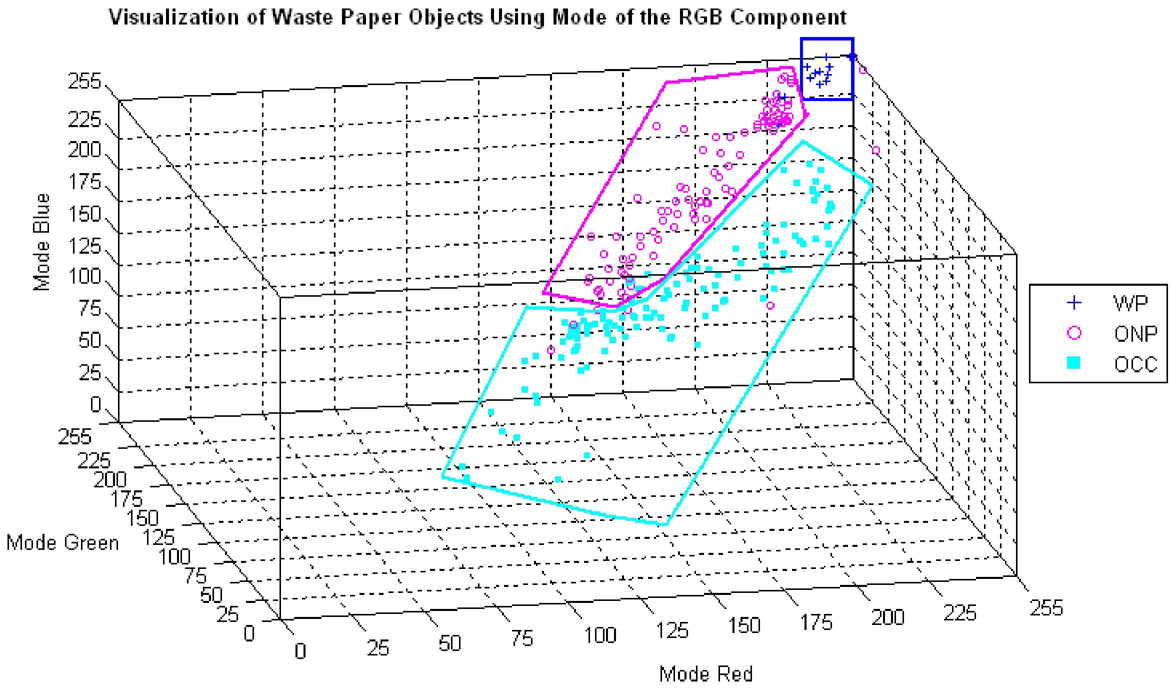
Table 4. Classification success rate for different window sizes

Window Size	WP		ONP		OCC		Average Success Rate %
	No of Successes	Success Rate %	No of Successes	Success Rate %	No of Successes	Success Rate %	
3×3	458	94	345	92	477	98	95.17
5×5	455	94	351	94	465	93	94.50
8×8	443	91	308	82	435	87	88.18
10×10	438	90	286	76	417	83	84.83
12×12	428	88	271	72	394	79	81.26
15×15	418	86	233	62	361	72	75.24
18×18	411	85	189	50	334	67	69.44
20×20	406	84	197	53	311	62	67.96
25×25	399	82	128	34	268	54	59.11
30×30	383	79	100	27	234	47	53.31
35×35	370	76	71	19	195	39	47.29
40×40	363	75	57	15	175	35	44.24
45×45	357	74	42	11	162	32	41.71
50×50	346	71	29	8	142	28	38.44

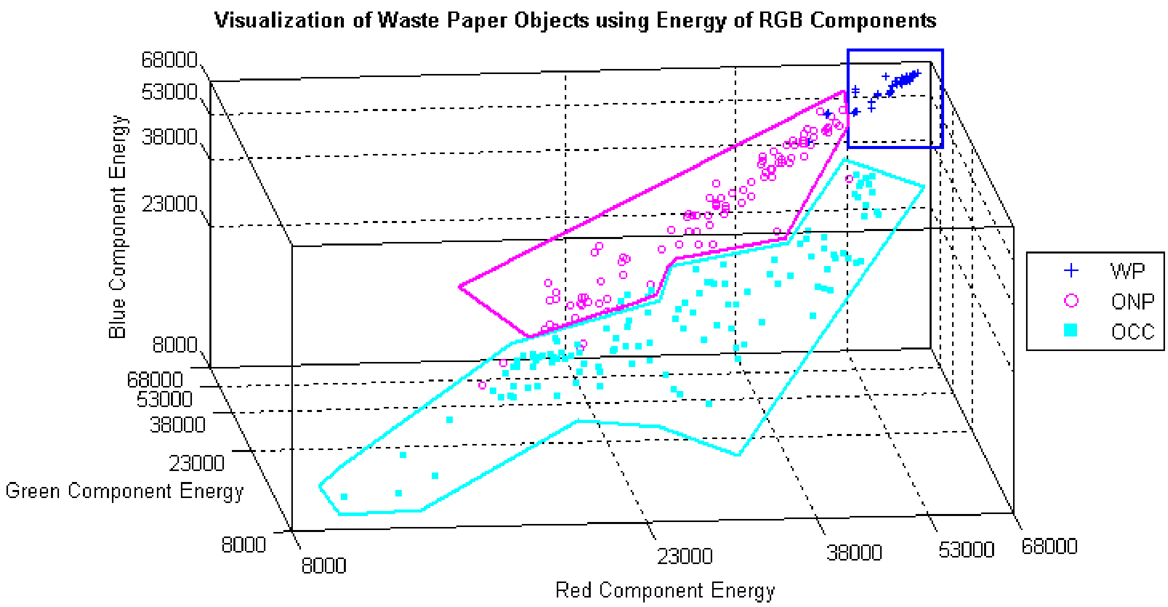
The success rates of the paper grade identification process for the three major waste papers types are shown in Table 4 and the graphical representation of the result is illustrated in Figure 7. In Table 4, the success rate and average success rate are obtained using Equations (4) and (5), respectively. The correct identification rate or success rate is calculated based on the percentage of the number of paper objects that are classified into their respective paper grades.

In Table 4, when the window size is 3×3, the achieved classification success rates are 94%, 92% and 98% for WP, ONP and OCC respectively. For ONP a better classification rate of 94% is achieved with window size 5×5. When the template size is 3×3 the achieved average classification success rate is 95.17%. In Figure 7, it is observed that when the window size is less than or equal to 8×8 pixels then average success rates stayed within 88.18 to 95.17. For WP, the success rate gradually decreased with the window size but for ONP, the success rate dramatically decreased with the window size. The success rate of OCC for window size is stayed in between WP and ONP.

From Table 4 and Figure 7, it is concluded that the performance of this experiment is greatly influenced by the window size. If the window size was decreased, the success rate of paper grade identification increased. However, the computational time increased due to the number of cells in paper object image. In order to get the optimum performance in terms of classification success rate and computation time, the suggested window size



(a)



(b)

FIGURE 6. Visualization of paper objects: (a) mode of RGB components and (b) energy of RGB components

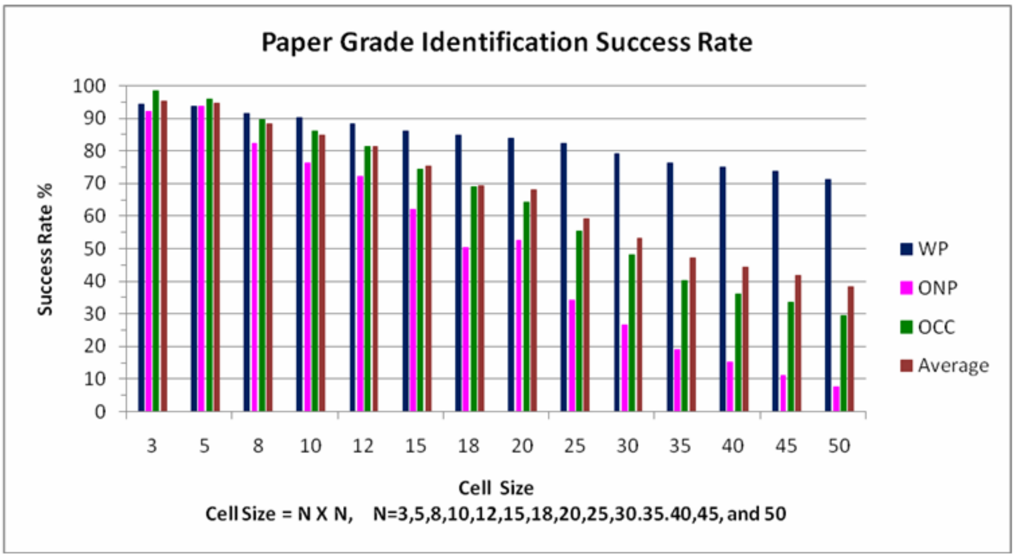


FIGURE 7. Paper grade identification success rates for different window sizes

is less than or equal to 8×8 pixels.

$$\text{Success rate} = \frac{\text{Number of paper object correctly identified}}{\text{Total number of paper object}} \times 100\% \tag{4}$$

$$\text{Average Success rate} = \frac{\text{WP}_\text{S} + \text{ONP}_\text{S} + \text{OCC}_\text{S}}{\text{WP}_\text{T} + \text{ONP}_\text{T} + \text{OCC}_\text{T}} \times 100\% \tag{5}$$

where

WP_S: number of WP objects correctly identified;
ONP_S: number of ONP objects correctly identified;
OCC_S: number of OCC objects correctly identified;
WP_T: total number of WP objects are examined;
ONP_T: total number of ONP objects are examined;
OCC_T: total number of OCC objects are examined.

3.3. **Comparison with previous results.** In this section, the best result of the SVS system is compared with the results published in literature using other methods. The detailed comparison among the SVS system and other methods is shown in Table 5. It is observed that the performance of the SVS system is the best among all systems. The template matching [20] method showed the closest performance. The average maximum classification success rate of the template matching system is 94.67% [20], while the SVS system offered 95.17% classification success rate. In real time implementation, the SVS system is more effective and convenient than the template matching technique with regards to computational time and lighting consistency. For instance, in template matching, significant time is allocated for preprocessing, while in the SVS system preprocessing is not required. The performance of the template matching method also depends on lighting consistency during the enrollment and identification phases. With the SVS system, the lighting dependency has been alleviated because the system uses different reference templates for the same paper grade which are taken in different lighting conditions. For template matching, a 5×5 template consists of 25 pixels; and for each pixel the RGB string length is 4 to 16. The RGB string length for 5×5 template is thus 100 to 400. As a consequence, there are 100 to 400 comparisons between one reference template and one cell image template. Additionally, the template matching method is inconvenient for real time implementation. For the SVS system, the template consists of only two values, namely mode and energy of the RGB components, which greatly improves the speed of the matching process.

TABLE 5. The result of the SVS system compared with results published in literature

	Techniques Applied for Identification	Types of Sensor	Features	Classification Success Rate
Template Matching [20]	Template Matching	Logitech QuickCam Pro 4000 Web Camera	RGB String	94.67%
Window Features [44]	Window-based sub-division, Distance, and voting	Logitech QuickCam Pro 4000 Web Camera	Mode of Hue, Mean of the Hue and Saturation	91.07%
Co-occurrence Features [21]	Rule based Classifier	Logitech QuickCam Pro 4000 Web Camera	Energy for the Co-occurrence matrices	90.67%
TiTech Systems [33]	Not Mentioned	NIR, CMYK sensor and color camera	Materials, shape, color, texture and four color printing	80%
MSS Systems [33]	Not Mentioned	NIR, Color sensor, Gloss and Lignin	The sensor measures the intensity of the material's fluorescence at a specific wavelength in the ultra-violet light.	80%
Mechatronic Design of a Waste Paper Sorting System for Efficient Recycling [41]	Artificial Neural Network	Four Sensors: Lignin, Gloss, Stiffness and Nikon D50 Digital SLR camera as a Color.	(i) Average Lignin value, (ii) Gloss meter reading, (iii) Deflection in the up-ward derection, (iv) Def-lection in the downward direction, (v) color varia-nce parameter I, (vi) color variance parameter II.	36.6%
	Fuzzy Inference System Algorithm			90.4%
Dominated Color [45]	KNN, Absolute Distance Metric	Logitech QuickCam Pro 4000 Web Camera	First Order Features: energy, mode, histogram tail length on the dark side, histogram tail length on the light side, lower quartile and upper quartile	93%
The SVS System	Case Based Reasoning	Logitech QuickCam Pro 4000 Web Camera	Mode and Energy of the RGB Components	95.17%

The window features [44] method showed the 2$^{\text{nd}}$ closest performance to the SVS system. Since the HSI color space was used in the window features method, the extra computation time is needed for RGB color space to HSI color space conversion. On the other hand, the SVS system used RGB color space, which overcomes the problem of extra color space conversion time. Moreover, the SVS system is superior to the window features method in terms of throughput. Figure 8 illustrated the performance comparison between both methods. It is observed that for ONP, the window features method showed the better performance than the SVS system. On the other hand, for WP, the SVS system showed the better performance than the window features method. The average maximum classification success rate of the window features system is 91.07% [44], while the SVS system offered 95.17%.

The Co-occurrence Features [21] method showed the 3$^{\text{rd}}$ closest performance to the SVS system. The Co-occurrence Features method used the rule based classifier which is much faster than the SVS system. However, the computational time to create co-occurrence matrix is excessive large and it is the main constraint to implement the system for real time application. Moreover, the experimental results lead the SVS system to the superior position over the co-occurrence features method shown in Table 5.

The 4th closed performance to the SVS system found in [41]. Kumar [41] did his experiment in off line and integrated the outcomes of four sensors namely Lignin, Gloss, Stiffness, and Nikon D50 Digital SLR camera. In addition, for color information L∗a∗b color space was used, which involved two color space conversion such as RGB color space to XYZ color space and then XYZ color space to L∗a∗b color space. Moreover, the color variance parameters were used as partial features of the feature vector. On the other hand, the SVS system used RGB color space and only two features mode and energy that make the SVS system viable to implement in real time application with providing low-cost solution.

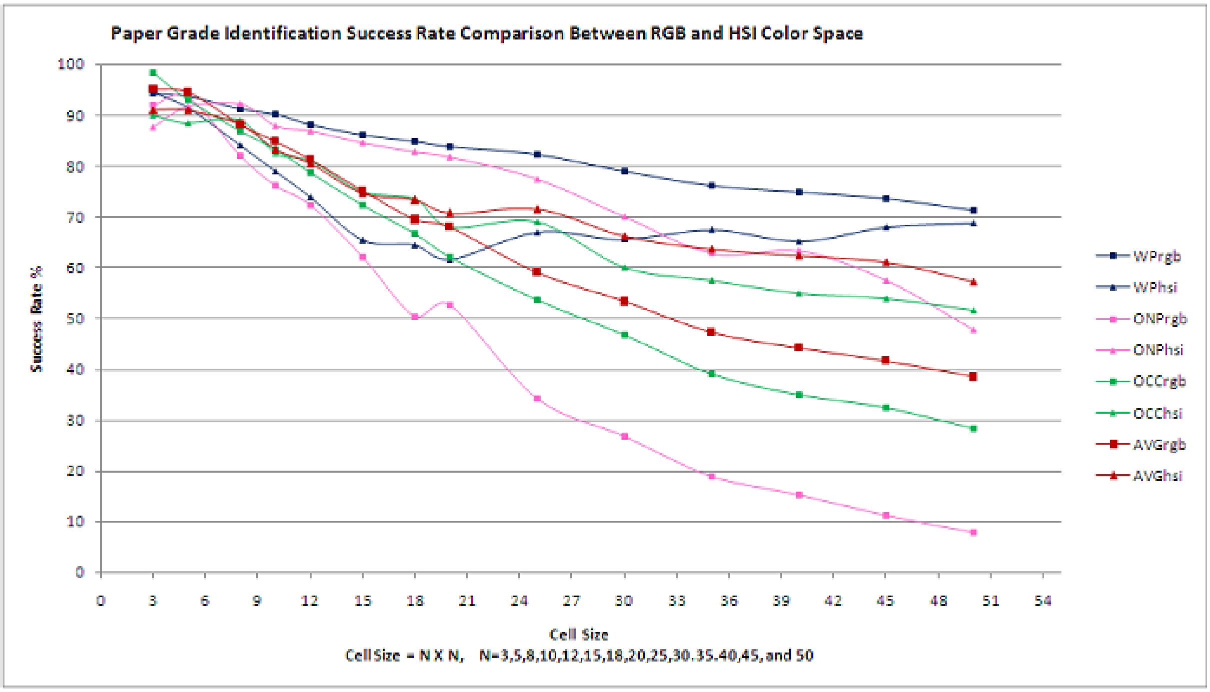


FIGURE 8. Paper grade identification success rate comparison between RGB and HSI color space

3.4. **Economic and environmental impact.** The reverse distribution network [42] for wastepaper recycling consists of five entities, i.e., vendor-customer (initial source of waste paper), dealer, godown owner, supplier and manufacturer. Waste papers are segregated into different grades on the godown stage of the reverse distribution network [42]. However, society gains at all the stages from the point of economic and environmental impact [4,6,42,43]. The waste paper purchase rates of Alam Flora Sdn Bhd, one of the leading dealers collecting waste papers in Malaysia, are shown in Table 6. Another company, SPM Paper Recycling Sdn Bhd, sells different grades of paper to the paper manufacturing company at different rates as shown in Table 7. In the waste paper market, the price of mixed waste paper is always lower than segregated waste papers. In addition, the paper manufacturer will benefit more from the segregated waste paper because they can produce high quality paper and paper products using less processing chemicals and energy.

High capital costs are required for implementing the existing automated sorting systems that deploy state-of-the-art technology, namely NIR, IR and X-ray. While these technologies may be feasible for high volumes of throughput [34], it is generally uneconomic for small scale operators. The recycling volume of the Malaysian local paper recycling facilities can be categorised as low to medium. As a consequence, an electronic image based approach is cheaper, more efficient and therefore justifiable for Malaysian recycling facilities.

Table 6. Waste paper purchase rate of Alam Flora Sdn Bhd

Paper Grade	[cents/kg]	1000Kg-RM
newspaper	24	240
black & white paper	35	350
carton	18	180
magazines	20	200
mix paper	16	160

Table 7. Waste paper sales rate of SPM paper recycling Sdn Bhd

Paper Grade	[cents/kg]	1000Kg-RM
old corrugated cardboard (OCC)	32	320
old newsprint (ONP)	35	350
black white (BW)	50	500
computer form (CF)	70	700
mixed white (MW)	36	360
mixed color (MC)	20	200
pure white (PW)	95	950

3.5. **Deficiencies of the SVS system.** In Figure 5, the subdivision of the paper object image into N-cells is shown. All the cells are considered for feature calculation in the SVS system. However, considering all the background cells is an unnecessary overburden in the SVS system for the paper grade identification. Thus, the efficiency of the SVS system will be improved in terms of computational time if the unnecessary background cells are discarded.

4. **Conclusion.** The primary emphasis of this work is on the development of a new waste paper grade identification method for automated paper sorting systems known as the SVS system. The SVS system is implemented based on CBR approach. The CBR is one of the learning based classifier and working using the principle of template matching [46,47]. The method described involved window-based subdivision of the paper image into N-cells, construction of N-candidate template for N-cells, calculation of matching scores of reference templates for the N-cells paper image, and application of matching score to identify the grade of the paper object. The SVS system performance for correct paper grade identification is 95.17% with estimated throughput of 21,600 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 is the adaptability to new subcategories of the primary paper grades. The wide range of subcategories of paper grades is used to train the system to recognize new subcategories, and as a result the system is scalable and able to provide robust decisions for paper grade identification tasks. Besides, the method was trained with many reference templates using different lighting conditions, which overcame the need to maintain lighting consistency during enrollment and identification phases.

The most important point addressed in this work is that the method, which uses computer vision, can be implemented easily to sort multiple grades of paper. Moreover, the algorithm provides robust and fast results because the proposed method avoids the extra computational burden for preprocessing since only two features, mode and energy, of the RGB components are used to identify the dominating color value of the paper object image. The proposed method can identify three major paper grades, WP, ONP and OCC, using the CBR approach with coalescing window-based subdivision technique and first-order window-features. 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-36-184/2011. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] J. Petek and P. Glavic, An integral approach to waste minimization in process industries, *Resource, Conservation and Recycling*, vol.17, no.3, pp.169-188, 1996.

[2] *Paper Recycling*, http://www.newsprint.com.my/, 2010.

[3] *WasteCap*, http://www.wastecap.org/wastecap/commodities/paper/paper.htm, 2008.

[4] M. O. Rahman, M. A. Hannan, E. Scavino, A. Hussain and H. Basri, An efficient paper grade identification method for automatic recyclable waste paper sorting, *European Journal of Scientific Research*, vol.25, no.1, pp.96-103, 2009.

[5] R. K. Pati, P. Vrat and P. Kumar, Economic analysis of paper recycling vis-a-vis wood as raw material, *Int. J. Prod. Econ.*, vol.103, no.2, pp.489-508, 2006.

[6] J. Laurijssen, M. Marsidi, A. Westenbroek, E. Worrell and A. Faaij, Paper and biomass for energy?: The impact of paper recycling on energy and CO_2 emissions, *Resources, Conservation and Recycling*, vol.54, no.12, 2010.

[7] *Paper Grades*, http://www.paperonweb.com/ppmanf.htm, 2009.

[8] S. Faibish, H. Bacakoglu and A. A. Goldenberg, An eye-hand system for automated paper recycling, *Proc. of the IEEE International Conference on Robotics and Automation*, Albuquerque, New Mexico, pp.9-14, 1997.

[9] M. K. Ramasubramanian, R. A. Venditti, C. M. Ammineni and M. Mallapragada, Optical sensor for noncontact measurement of lignin content in high-speed moving paper surfaces, *IEEE Sensors Journal*, vol.5, no.5, pp.1132-1139, 2005.

[10] F. A. Hottenstein, G. R. Kenny, T. Friberg and M. Jackson, High-speed automated optical sorting of recovered paper, *Proc. of the TAPPI Recycling Symposium*, Atlanta, GA, USA, pp.149-158, 2000.

[11] R. A. Venditti, M. K. Ramasubramanian and C. K. Kalyan, A noncontact sensor for the identification of paper and board samples on a high speed sorting conveyor, *Journal of the Technical Association of the Australian and New Zealand Pulp and Paper Industry*, vol.60, no.5, pp.366-371, 2007.

[12] N. H. Sandberg, Sorting Device for Waste Paper, *US Patent No. 1,847,265*, 1932.

[13] A. Bialski, C. Gentile and O. Sepall, Paper Sorting Apparatus, *US Patent No. 4,236,676*, 1980.

[14] M. Grubbs, G. R. Kenny and P. G. Gaddis, Paper Sorting System, *US Patent No. 6,250,472*, 2001.

[15] Z. Khalfan and S. Greenspan, Optical Paper Sorting Method Device and Apparatus, *US Patent No. 7,081,594*, 2006.

[16] R. Eixelberger, P. Friedl and K. Gschweitl, Method and Apparatus for Sorting Waste Paper of Different Grades and Conditions, *US Patent No. 6,506,991*, 2003.

[17] R. S. Bruner, D. R. Morgan, G. R. Kenny, P. G. Gaddis, D. Lee and J. M. Roggow, System and Method for Sensing White Paper, *US Patent No. 6,570,653*, 2003.

[18] A. G. Doak, M. G. Roe and G. R. Kenny, Multi-Grade Object Sorting System and Method, *US Patent No. 7173709*, 2007.
[19] Gschweitl and K. Heinz, Method for Sorting Waste Paper, *European Patent, EP0873797*, 1998.
[20] M. O. Rahman, A. Hussain, E. Scavino, M. A. Hannan and H. Basri, *Recyclable Waste Paper Sorting Using Template Matching*, Springer-Verlag, Germany, 2009.
[21] M. O. Rahman, A. Hussain, E. Scavino, M. A. Hannan and H. Basri, Segregating recyclable waste papers using co-occurrence features, *Proc. of the 9th WSEAS International Conference on Applied Computer Science*, Genova, Italy, pp.187-191, 2009.
[22] C. M. Chandini, *Design of Lignin Sensor for Identification of Waste Paper Grades for an Automatic Waste Paper Sorting System*, Master Thesis, North Carolina State University, 2001.
[23] Z. Khalfan, Optical Paper Sorter, *US Patent No. 6,335,501*, 2002.
[24] Z. Khalfan and S. Greenspan, Optical Paper Sorting Method Device and Apparatus, *US Patent No. 2006/0124511*, 2006.
[25] A. G. Doak, M. G. Roe and G. R. Kenny, Multi-Grade Object Sorting System and Method, *US Patent No. US2007/0002326*, 2007.
[26] K. K. Chakravarthi, *Development of Online Stiffness Sensor for High Speed Sorting of Recovered Paper*, Master Thesis, North Carolina State University, 2006.
[27] M. K. Ramasubramanian, R. A. Venditti and P. K. Gillella, *Sensor Systems for High Speed Intelligent Sorting of Waste Paper in Recycling*, http://www.osti.gov/bridge/servlets/purl/919471-VGwxA0/919471.PDF, 2008.
[28] *Paper Competitors*, http://separation.wikispaces.com/Paper+Competitors, 2009.
[29] *TITECH Autosort*, http://www.titech.com/recycling-equipment/titech-autosort-10715, 2009.
[30] Pellenc, *Paper Sorting Company*, http://www.pellencst.com/en/1/products, 2009.
[31] Pellenc, *MIR Technology*, http://www.pellencst.com/en/21/54/mir-technology, 2009.
[32] RedWave, *BT – Wolfgang Binder*, http://www.redwave.at/altpapierrecycling/, 2009.
[33] Remade Scotland, *Initial Business Case for Utilisation of Automated Optical Paper Sorting Technology*, http://www.remade.org.uk/files/InitialBusinessCaseforUtilisationofAutomatedOpticalPaperSortingTechnology_16122222124.pdf, 2009.
[34] D. A. Wahab, A. Hussain, E. Scavino, M. M. Mustafa and H. Basri, Development of a prototype automated sorting system for plastic recycling, *American Journal of Applied Sciences*, vol.3, no.7, pp.1924-1928, 2006.
[35] S. H. Chen, A. J. Jakeman and J. P. Norton, Artificial intelligence techniques: An introduction to their use for modelling environmental systems, *Mathematics and Computers in Simulation*, vol.78, no.2-3, pp.379-400, 2008.
[36] D. T. Pham and R. J. Alcock, *Smart Inspection System – Techniques and Applications of Intelligent Vision*, Academic Press, Great Britain, 2003.
[37] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd Edition, Prentice Hall, USA, 2008.
[38] A. Aamodt and E. Plaza, Case-based reasoning: Fundamental issues, methodological variations, and system approaches, *AI Communication*, vol.7, no.1, pp.35-59, 1994.
[39] *Logitech QuickCam Pro 4000 Web Camera Specification*, http://reviews.cnet.com/webcams/quick-cam-pro-4000-web/4507-6502_7-20276742.html?tag=rnav, 2009.
[40] M. W. Burke, , *Image Acquisition – Handbook of Machine Vision Engineering: Volume 1*, Chapman and Hall, London, 1996.
[41] R. G. P. Kumar, *Mechatronic Design of a Waste Paper Sorting System for Efficient Recycling*, Master Thesis, North Carolina State University, 2007.
[42] R. K. Patia, P. Vrat and P. Kumar, A goal programming model for paper recycling system, *Omega the International Journal of Management Science*, vol.36, no.3, pp.405-417, 2008.
[43] T. Ekvall, Key methodological issues for life cycle inventory analysis of paper recycling, *Journal of Cleaner Production*, vol.7, no.4, pp.281-294, 1999.
[44] M. O. Rahman, A. Hussain, E. Scavino, N. E. A. Basri, H. Basri and M. A. Hannan, Waste paper grade identification system using window features, *Journal of Computational Information Systems*, vol.6, no.7, pp.2077-2091, 2010.
[45] M. O. Rahman, A. Hussain, M. A. Hannan, E. Scavino and H. Basri, Intelligent computer vision system for segregating recyclable waste papers, *Expert Systems with Applications*, vol.38, no.8, pp.10398-10407, 2011.
[46] S. Karungaru, M. Fukumi, N. Akamatsu and T. Akashi, Detection and recognition of vehicle license plates using template matching, genetic algorithms and neural networks, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.1975-1985, 2009.

[47] H. Ryu, V. Dinh and M. Kim, Real-time multi-view face tracking combining learning based classifier and template matching, *ICIC Express Letters*, vol.1, no.2, pp.185-189, 2007.