Sensor systems for high-speed intelligent sorting of waste paper in recycling

MELUR K. RAMASUBRAMANIAN, RICHARD A. VENDITTI, AND PRADEEP K. GILLELLA

ABSTRACT: Sorting paper into compatible grades is a necessary step before recycling. Current methods of manually sorting are tedious, slow, and expensive. High-speed automation of the sorting process will improve the cost efficiency and increase the amount of paper recycled significantly. Correct identification of the sample in real-time before the sample reaches the actuation station still remains a challenge because different types of paper and board samples are mixed in a waste stream, and have a lot of variation in terms of color, chemical composition, coating, and printing in black and white and color to different degrees. This primarily is because of a lack of satisfactory sensors and sensor fusion algorithms for sample identification in real-time. In this paper, we identify key parameters that must be measured, discuss a sensor system design, and describe integration of the output from sensors to interpret the type of sample using a fuzzy inference algorithm. Results show that the sensor system proposed is capable of identifying the samples at 90% accuracy compared to manual identification. The sensor system can be integrated onto a conveyor and actuation system for automated sorting.

Application: The technology described in this report can be used to sort mixed waste paper more efficiently into recyclable grades at high throughput rates.

Recycling of paper and paper related products is very important for sustainable economic growth. It helps save landfill space and costs, reduces the energy requirements for paper manufacturing, and extends the fiber supply. Paper recycling reduces the rate of consumption of precious natural resources (wood, water, minerals, and fossil fuel). One ton of paper fibers from recycled stock saves an estimated 7000 gal (26500 L) of water and 3.3 yd^3 of landfill space, and reduces greenhouse gases by 1 metric ton of carbon equivalent [1]. In the United States, about 71 million tons of paper and board products entered the waste management system in 2010 and 62.5% of all the paper products discarded were recovered for recycling [2]. A recent study found reductions of -1850 to -4400 kg CO_2 –eq. metric ton^{-1} in greenhouse gases and global warming contributions attributable to recycling of paper [3]. That assumes that recycling will help avoid the use of equivalent amounts of virgin fibers and that wood not used for virgin pulp production can be used in the production of energy to offset fossil fuel use. The trees that are not harvested remain available for carbon sequestration of 1.8-2.2 kg CO_2/kg^{-1} of virgin wood growth [3]. These are very large savings in greenhouse gas emissions [3]. Before waste paper can be recycled efficiently, it has to be sorted into grades, typically newsprint, coated sheet, colored free sheet, white free sheet, board, and mixed waste.

At municipal waste sorting plants, sorting is mostly carried out manually. Here, the paper waste is broadly classified as brown paperboard, newspaper, and mixed office waste. But the inability of the human eye to distinguish between coated inferior grades and good grades results in sorted office waste that contains high-grade paper contaminated by similar looking low-grade paper, which reduces the overall fiber quality when they are recycled together. The throughput from these manual sorting facilities is also very low. A typical manual conveyor operates near 75 ft/min (about 0.40 m/s). In addition, recently conducted research has concluded that laborers working at manual sorting facilities are exposed to microorganisms, organic dust, and fungi, which cause severe infections [4]. Hence, the sorting process needs to be automated to provide efficient recycling and a safer sorting process. Sorting at higher speeds also improves the economic viability of an automated process. Target speeds in the industry are about 1200 ft/min (6 m/sec), with a throughput of 15 tons/h.

In this paper, we identify key parameters that must be measured, discuss a sensor system design, and describe integration of the output from sensors to interpret the type of sample using a fuzzy inference algorithm. The sensor system can be integrated onto a conveyor and actuation system for automated sorting.

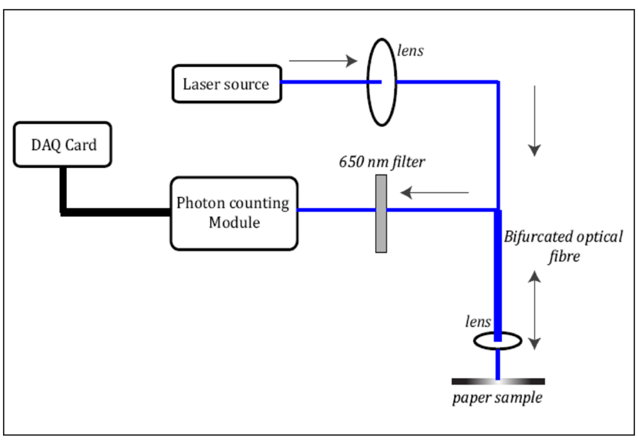
SENSORS FOR SORTING AUTOMATION

An efficient automated sorting system requires robust sensors that can measure various physical and optical properties of paper. The distinguishing physical characteristic of paperboard is its high stiffness compared with other paper grades. The distinguishing characteristic of newsprint is its high lignin content, coated papers have high gloss, and in general, non-office papers are not white and have high color content. To sort papers into various grades, we need to measure the stiffness, chemical content, color, and gloss of specimens and make an inference, all in real-time. This poses many challenges, and further challenges arise when newsprint paper is printed with colors that mask the lignin in the sheet; lignin-free sheets are printed with colors that falsely trip the lignin sensor; or a stack of office papers on the conveyor produce high stiffness mimicking that of paperboard. In this section, we describe sensors developed for the real-time measurement of lignin, color, and stiffness of samples. Gloss data from standard laboratory measurements are integrated in the sensor fusion algorithms described later.

Lignin sensor

An optical sensor that can measure the amount of lignin present in paper based on fluorescence has been developed at North Carolina State University [5-7]. The lignin sensor measures the intensity of fluorescence to predict the actual amount of lignin present in the paper. **Figure 1** shows a schematic of the lignin sensor. This sensor is highly efficient in measuring the amount of lignin in samples and can distinguish between newsprint and other paper grades. The sensor measures lignin fluorescence (near 650 nm) when excited in the visible region (532 nm) with a laser source. Measurement of this peak intensity from paper samples is a measure of lignin content and correlates linearly with the standard chemical method for lignin determination (Klason lignin content). Newsprint samples, typically containing high lignin, produce high intensity fluorescence, while ledger printing and writing grades with low lignin content produce low-intensity fluorescence. Tests with a wider range of samples, including bleached sulfite paper, newsprint, white ledger, and thermomechanical pulp, on a conveyor at high speeds (720 ft/min or about 3.7 m/s) showed clear distinction among samples [5,6].

The lignin sensor results are affected by color present on the surface of the paper and the distance of the sensor from the sample at the time of measurement [7]. We calculated a distance versus intensity calibration curve and used that to correct for the measured intensity. The distance to the top surface of a sample was measured using a standard distance sensor (Sharp GP2D120; Sharp Microelectronics of the Americas, Camas, WA, USA). When a stack of paper passes through the sensing area, it is assumed that the lignin content measured on the surface represents the lignin content of the entire stack. The presence of color on the surface poses a more serious problem. For example, a red patch on the top of an office ledger paper will produce a signal (normalized voltage) that matches that of newsprint, whereas a blue or green patch printed on a newsprint surface will produce signals that match with ledger paper and cause misidentification. Thus, the lignin sensor alone cannot detect the presence of newsprint. Additional considerations from other sensors are necessary to make a determination of the sample type.

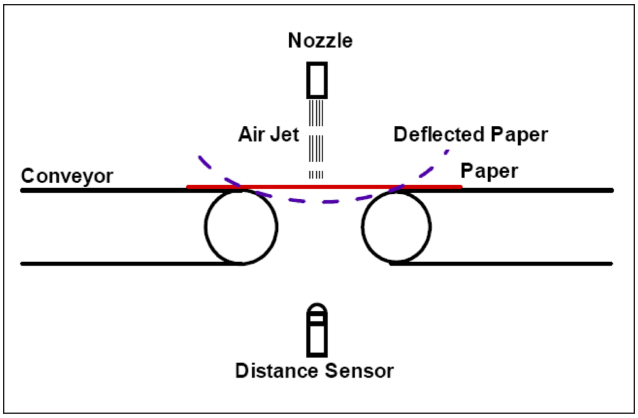


1. Schematic of the noncontact lignin sensor.

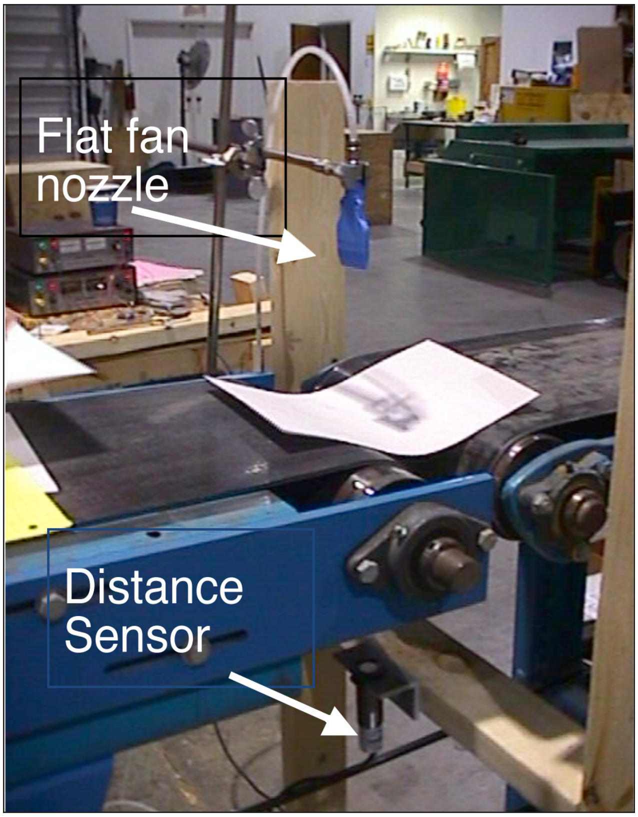
Stiffness sensor

Various types of board are produced, for various uses: brown carton used in packaging, manila folder covers used in offices, thick magazine covers, six-pack can holders, etc. Regardless of type, a board sample should be separated from newsprint and color and glossy papers. All boards have a common physical property, namely, high stiffness. Stiffness sensing on discrete samples randomly oriented on a moving conveyor is challenging. We cannot hold the sample and apply a load to measure the elongation and calculate the stiffness; hence, we developed a quick, non-contact method for measuring the relative stiffness of paper samples moving on a conveyor at high speeds. The method consists of impinging an air jet on a moving paper sample while it is crossing a gap during transfer from one conveyor belt to another. The deflection caused by the impinging air jet can be measured and used to categorize different grades of paper. We carried out finite element simulations of the proposed method to understand the effects of speed, sample orientation, nozzle pressure, and other parameters to make certain that the stiffness estimation has sufficient resolution to distinguish between board and paper grades at high conveyor speeds. **Figure 2** shows a schematic of the sensing concept [8].

The sensor was tested in the static mode in the laboratory and in a dynamic mode on a sorting conveyor moving at 280 ft/min (about 1.42 m/s) [9,10]. The deflection decreased with increasing sample thickness and increasing basis weight; both contribute to bending stiffness. For the case of dynamic measurement, we needed to use a continuous air flow at lower air pressure (34.5 kPa) compared to static tests [9,10]. The sensor was able to identify paperboard samples clearly. When the difference in the thickness of the tested samples is not significant, the difference in deflection resulted from differences in elastic modulus [9,10]. Thus, the stiffness sensor is successful in identifying board from paper on a moving conveyor. However, it cannot distinguish between a stack of



2. Relative bending stiffness sensor.



3. Stiffness sensor in operation on a conveyor moving at 280 ft/min.

newsprint versus a single paperboard. Additional sensor input must be used to resolve such conflicts. Additional sensors to measure other unique properties must be developed for a robust sorting system, followed by a sensor fusion algorithm to realize the full potential of sensor-based sorting. **Figure 3** shows a sample in a recycling research facility crossing a gap of 7 cm between two conveyors instrumented with the stiffness sensor below the gap. The paper sample deflects and clears the gap. Figure 3 also shows the fan air nozzle at the top and the distance sensor below. The conveyor was moving at 280 ft/min. A board sample will have smaller or no measurable deflection.

Color sensor

Sensing color is very important in paper sorting for recycling. The color in the surface affects the determination of lignin content in paper. Color prints with high gloss may suggest a coated magazine paper. Hence, rapid measurement of color is an important component in the determination of the type of paper being examined.

Some sophisticated digital single lens reflex (SLR) cameras provide the option of capturing RAW file format data from the image sensor. Digitized values of the analog signals generated from each pixel in the sensor are logged into a file. This essentially converts the digital camera into a highly flexible color sensor. These cameras provide the ability to alter the aperture size and exposure time of the sensor. To obtain color information in real-time, we used a commercially available digital SLR camera (Nikon D50). The RAW data from the camera's sensor is captured on a digital memory card. The logged data is then processed using proprietary software from the camera manufacturer (Nikon; Tokyo, Japan) to convert the images into standard red, green, blue (RGB) format. This RGB data is further processed using MATLAB software (MathWorks, Natick, MA, USA) to extract color parameters for each of the samples.

Although color can be quantified by RGB components, the components themselves do not have a physical meaning. We transformed the data into a more practical system, L* a* b* color space, through the standard coordinate transformations [11]. While interpreting RGB primary colors is obvious, L* a* b* color space needs some explanation. L* a* b* color space is designed to approximate human vision and perception of color. The L* component closely matches the human perception of brightness. The a* component represents the position of the color on the green-magenta axis, and the b* component represents the position of the color on the yellow-blue axis. The reason for choosing to work with this color space is because it closely resembles the way we perceive color (i.e., light/dark, warm/cool descriptors). After color measurements, the mean and standard deviations of L* a* b* components for each sample were calculated and used as input parameters for developing the sensing algorithm.

SENSOR FUSION

A sample moving on a conveyor should be assessed for all necessary attributes and the information from the sensors should be collectively interpreted to make a decision on the paper's grade. In this study, we have used lignin, stiffness, and gloss sensors, and color sensor parameters to explore the use of artificial intelligence algorithms for sensor fusion. Because of their recent success in the fields of pattern recognition and classification, we first evaluated an algorithm based on artificial neural networks [12]. The possibility of integrating human-like reasoning and expert knowledge into algorithms based on fuzzy logic was a primary reason for choosing a fuzzy inference system (FIS) as another alternative [13].

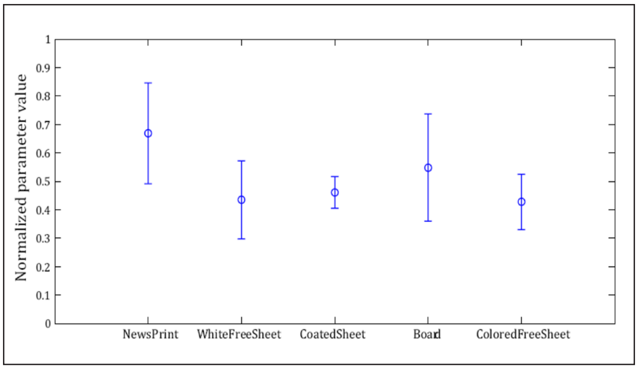
Samples tests and data acquisition

Different types form a typical unsorted waste paper stream. We identified 224 different regions in 63 different samples, each representing one of five categories of paper: newsprint, white lignin-free sheet, coated sheet, board, and colored lignin free sheet. We then determined six parameters to measure for each of the 224 samples: average lignin value across the sample, gloss meter reading from the center of the sample, deflection in the downward direction from the stiffness sensor, color variance parameters LAB(I) and LAB(II) calculated by the sum of deviations of a* and b*, and the sum of average values of a* and b*, respectively.

Neural network

We used a neural network algorithm based on radial basis functions [12,13] for this study. A set of 203 sample data were used to train the neural network. The remaining 21 sample data sets were used for model validation. Five sample categories were used, as indicated above. The predictions of the neural network were rounded to the nearest integer between 1 and 5 and associated with the sample identification. As shown in **Table I**, the accuracy of prediction (correct identification of sample) was poor; we attributed this to data inconsistencies inherent to waste paper streams. **Figure 4** shows the lignin sensor data for various samples when tested repeatedly and randomly. The average values show the expected trend. However, the error bars overlap significantly, indicating that using a lignin sensor alone to identify newsprint will be misleading and can lead to the misidentification of board as newsprint or of a board or red color printed writing paper as newsprint. Similarly, newsprint could register a low lignin sensor value because of the presence of blue or green shades on the surface, which could quench the amount of fluorescence given off by the sample. Some newsprint samples made from recycled paper also register low values of lignin if, during recycling, they were mixed with fibers from higher grades to increase the strength of the paper. These variations are not controllable in a waste paper stream.

The neural network builds a model that is purely based on the training set, without the possibility of incorporating rule-based decision making. For instance, if the stiffness of a sample is very high, it is almost always a board sample, and if the lignin content is very high and the stiffness value is low, then the sample is newsprint. Newsprint also can be identified by another rule, which states that the sample is newsprint if the stiffness value is very low and the gloss value also is very low. This can be implemented easily into a fuzzy logic scheme.



4. Lignin sensor data variability due to print, color, and fiber content variations in waste paper samples

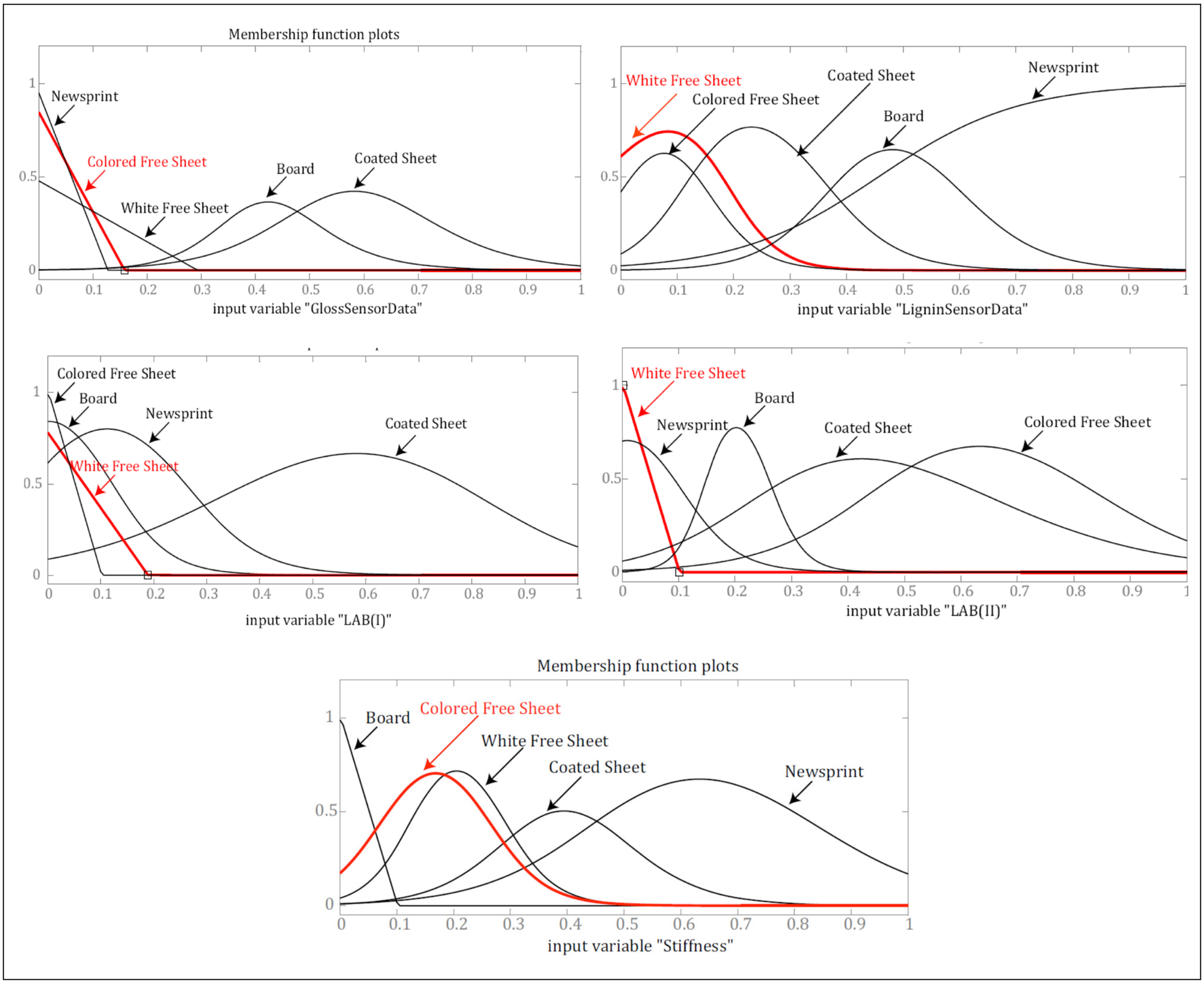
Fuzzy inference system

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic [13]. **Figure 5** shows the membership functions for the sensor inputs. The experimental data were organized for each sensor and paired with the 224 samples, each sample consisting of a known paper category as determined by a human judge. If the lignin sensor value is high for a given sample, the probability that the sample is newsprint is nearly 1 (Fig. 5). On the other hand, if the lignin sensor value is midrange, the probability that the sample is paperboard or newsprint is nearly equal, but the probability is small that the sample is lignin-free. Other membership functions are interpreted in the same manner. Next, the fuzzy rules need to be developed. The following is a listing of conditions currently implemented as the fuzzy rules for the FIS algorithm for this application.

- If lignin is high and stiffness is low, then the sample type is newsprint.
- If lignin is medium and stiffness is low and gloss is low, then the sample type is newsprint. (This rule assists when a newsprint sample may have significant recycled fiber, causing low lignin reading.)
- If lignin, stiffness, and gloss are moderate then the sam-

Training Data (sample range)	Test Data (sample range)	Prediction Accuracy
1 - 199; 221 - 224	200 - 220	30%
1 - 159; 181 - 224	160 - 180	20%
1 - 139; 161 - 224	140 - 160	15%
1 - 119; 141 - 224	120 - 140	20%
1 - 99; 121 - 224	100 - 120	30%
1 - 39; 61 - 224	40 - 60	20%

I. Prediction accuracy of neural network model.



5. Membership functions for sensor outputs based on experimental observations. Note that the input variable "stiffness" is based on sheet deflection in an air stream, and therefore varies oppositely with mechanical stiffness as used in the text.

ple is a free sheet (depending on the color variance parameter, it can either be a colored or white free sheet).

- If stiffness is low and gloss is high, then sample type is magazine (color variance is optional; if it is high, confidence increases; if low, it can be ignored).
- Any sample with very high stiffness is generally considered to be paperboard.

The membership functions and the fuzzy rules that have been derived are input into the MATLAB's FIS toolbox. This toolbox is a design time tool used to quickly visualize the inputs, outputs, and various fuzzy parameters. Upon successfully entering all the data into the FIS toolbox, we tested it using the rule viewer. The rule viewer lets us simulate the FIS algorithm by entering an input vector. It shows the various calculations being carried out and gives us an output vector. The output vector is an array of the degrees of confidence with which the algorithm states that a sample is of a particular type. The element in the output vector with the highest value indicates the type of paper. Consider the following output vector:

- Newsprint confidence: 0.45
- White free sheet confidence: 0.54
- Coated confidence: 0.89
- Board confidence: 0.14
- Colored free sheet confidence: 0.27

From the output vector, we can interpret that the algorithm classifies the particular sample as a coated sheet because the confidence value is 0.89, the highest value in the output vector. After all samples were classified by the system, the prediction accuracy was calculated for each sample type, as we know what the samples are. **Table II** contains details about the number of samples in each paper type tested, number of

Sample Type	Number of Test Samples	Number Detected Correctly	Prediction Accuracy	Neural Network Accuracy
Newsprint	70	63	90.0%	35.4%
White Free Sheet	56	52	92.8%	37.8%
Coated Sheet	59	54	91.5%	39.6%
Board	21	21	100%	31.4%
Colored Free Sheet	18	14	77.8%	38.6%
Overall	224	204	90.4%	36.6%

II. Prediction accuracy of the fuzzy inference system model and comparison with modified neural network.

samples identified correctly, and the prediction accuracies for each type of paper and the neural network algorithm. To obtain the neural network prediction accuracies, we carried out simulations for one sample at a time, as described earlier. All other samples, except the one being tested, were used as the training data. Then the output obtained was compared to that required for that particular sample. This process was repeated for all other samples and the predictions of sample type were recorded. Because of the increase in the amount of the training data and the larger number of validation data, an increase in the prediction accuracy was observed with the modified neural network. Out of 224 test samples, 204 samples were identified correctly by the FIS algorithm, for a sorting accuracy of 90% for the FIS algorithm.

Some reasons why some of the samples were wrongly detected include the following:

- If for some reason a newsprint sample made from recycled paper (low lignin) somehow registers a moderate deflection value at the stiffness sensor, both the rules that are meant to identify the newsprint would fail and at the same time satisfy the requirements for a white lignin free sheet.
- When a white free sheet fails to deflect as expected at the stiffness sensor and registers a low value of surface gloss, it ends up satisfying all the conditions necessary for the identification of a board sample.
- When a coated sheet registers low gloss values because of the effect of color on the surface of the sample and produces a moderately high value of lignin, it is identified as a newsprint sample or white free sheet, depending on the extent of the lignin content measured.
- Colored free sheets are sometimes identified as white free sheet when the coloring is very close to white (i.e., when the a* and b* coordinates of the base color are very close to zero).
- Some colors register a very high value of gloss, even when printed on a low gloss paper. These papers can be identified as a coated sheet if the stiffness sensor reading also falls within range of those expected for coated sheets.

In all these cases, two parameters must simultaneously be wrong for a particular sample to be misidentified, the likelihood of which is not very high, thus providing high predictability, considering the variations typically found in a waste stream.

Commercial application

A sorting system that incorporates color, gloss, and lignin sensors recently has been developed for commercial use. A stiffness sensor is not integrated into the system currently; consequently, paperboard is sorted manually in a pre-sorting station. The manufacturer reports operating speeds of 6 m/s and a throughput of 15 tons/h. Although no data on sorting efficiency or the sensor fusion algorithm is available, color paper sorting is estimated at about 50% efficiency (compared with manual identification), while white paper appears to be near 90% efficient. A system that uses gloss, color, and lignin sensors to sort newspaper from other types of paper also is available [14]. Based on observation, the efficiency of that system is estimated to be near 60%-70%. With the addition of a stiffness sensor and the implementation of a fuzzy algorithm, efficiency could increase to near 90% with a mixed waste stream, without the need to pre-sort board grades.

CONCLUSIONS

We developed a mechatronic system for high-speed identification of paper type. The system uses sensors for gloss, lignin content, and stiffness, plus a new sensing technique to sense and interpret color information to increase the efficiency and robustness with which paper samples can be identified. A large set of paper samples were collected to be tested with the sorting system. The samples were chosen such that they represented most of the variations possible in a typical waste stream. All the tests were performed under static conditions (i.e., the samples were all at rest with respect to the sensors for the purpose of algorithm development). The lignin and stiffness sensors were tested online to establish their dynamic behaviors. Specialized algorithms were developed to process the color sensor data to ensure compatibility with the sorting algorithms. Two algorithms based on artificial neural networks and fuzzy inference systems were evaluated using MATLAB. After extensive simulations, the algorithm based on fuzzy inference systems was found to be the most useful, with a sorting efficiency of nearly 90% when compared with manual sorting. This is far higher than any existing automated sorting system's overall efficiency. Commercial systems have been developed by a manufacturer using this sensor set and have proven successful. **TJ**

ACKNOWLEDGEMENTS

The authors thank the U.S. Department of Energy for supporting this research under the Industries for the Future Program, Forest Products Industry Agenda 2020; project number DE-FC07-00ID13880. The National Science Foundation IR/D Program provided travel support for Ramasubramanian to direct the research at North Carolina State University. Any opinion, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

LITERATURE CITED

1. U.S. Environmental Protection Agency, *Paper recycling: basic information details*, U.S. Environmental Protection Agency, Washington, DC, 2011. Available at http://www.epa.gov/osw/conserve/materials/paper/basics/index.htm.
2. U.S. Environmental Protection Agency, *Municipal solid waste generation, recycling, and disposal in the United States: facts and figures for 2010*, U.S. Environmental Protection Agency, Washington, DC, 2011. Available at http://www.epa.gov/osw/nonhaz/municipal/pubs/msw_2010_factsheet.pdf.
3. Merrild, H., Damgaard, A., and Christensen, T.H., *Waste Manag. Res.* 27(8): 746(2009).
4. Sigsgaard, T., Malmros, P., Nersting, L., et al., *Am. J. Respir. Crit. Care Med.* 149(6): 1407(1994).
5. Ramasubramanian, M.K., Venditti, R.A., Ammineni, C.M., et al., *IEEE Sensors J.* 5(5): 1132 (2005).
6. Ammineni, C.M., "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, 2001.
7. Mallapragada, V., "Online lignin sensor for high speed sorting of newsprint from mixed waste," Master's thesis, North Carolina State University, Raleigh, 2004.
8. Ramasubrmanian, M.K., Venditti, R.A., and Katuri, K.C., *J. Mech. Mat. Struct.* 2(2): 201(2007).
9. Katuri, K.C., "Development of online stiffness sensor for high speed sorting of recovered paper," Master's thesis, North Carolina State University, Raleigh, 2006.
10. Ramasubramanian, M.K., Venditti, R.A., and Katuri, K.C., *Appita J.* 60(5): 366(2007).
11. Jain, A.K., *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1989.
12. Haykin, S., *Neural Networks: A Comprehensive Foundation*, 2nd edn., Prentice Hall, Upper Saddle River, NJ, 1998.
13. Zilouchian, A. and Jamshidi, M., Eds., *Intelligent Control Systems Using Soft Computing Methodologies*, CRC Press, Boca Raton, FL, USA, 2001.
14. MSS Inc., "MultiWave, the world's fastest optical paper sorting system," MSS Inc., Nashville, TN, USA, 2012. Available at http://www.magsep.com/imgsmss/b/MULTIWAVE_English_2009_LR.pdf.

ABOUT THE AUTHORS

We chose this topic of research because, at the moment, downgrading paper in mixed wastes is an inefficient system to use for recovered paper. Great economic and environmental advantage arises if paper can be sorted and relatively homogenous streams of paper grades are used for high benefit applications.

This research group has as a goal to develop and characterize systems for the effective use of renewable resources to produce sustainable products, including polymeric materials and biofuels. The research here continues our project objectives of providing sensors and methods to improve paper sorting. The greatest challenge was to develop a system to collect multiple measurements of paper characteristics rapidly and to unambiguously identify the paper grade, despite complicating factors such as printing, coating, or presence of adhesives.

The most interesting aspect of the research has been the ability to create a human-like decision-making process with artificial intelligence and fuzzy logic algorithms. Already, aspects of these systems are used in operating sorting lines.

Future advancements in this area include improving the capacity of sorting lines. The efficient recovery of paper products from municipal wastes bound for landfills is an exciting challenge in this field, but currently not practically possible.

Ramasubramanian

Venditti

Gillella

Ramasubramanian is professor, Mechanical and Aerospace Engineering; Venditti is professor, Department of Forest Biomaterials, and has been selected for the 2012 class of TAPPI Fellows ; and Gillella is graduate student, Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC, USA. Email Ramasubramanian at rammk@ncsu.edu.