Development of watermark quality measurement by image analysis

R.J. TREPANIER AND D. PERRON

ABSTRACT: A new technology for measuring the quality of watermarks is based on image analysis techniques used to analyze paper formation. The technique involves aligning a sample watermark and a reference image, normalizing the paper formation quality, and analyzing the quality of the match. The attributes quantified are the watermark intensity, the intensity ratio, and the sharpness difference. The intensity ratio is the intensity of the sample image divided by the intensity of the reference image. The intensity and sharpness of a watermark image have to be analyzed against the background of sheet formation. Measuring the watermark intensity and sharpness relative to a reference gave significant correlations with visual testing.

Application: An emerging technological method of analyzing watermarks could replace visual methods as the commonly accepted way to assess and compare the quality of watermarks.

Watermarks are security features that can also help identify the manufacturer of a paper product. Traditionally, watermarks are backlighted and inspected visually. However, visual techniques lack the precision and sensitivity to optimize quality and maintain the papermaking performance. Today, papermakers want technology that measures watermark quality precisely.

A watermark is imposed when the sheet is being formed, which disrupts the forming process locally. Quantifying the visual attributes of textured images, such as sheet formation, can be difficult because of the subjective nature of visual techniques [1]. Moreover, a number of attributes [2] other than formation quality must be considered in measuring watermark quality.

A survey of pertinent literature did not yield useful information about the visual attributes that are important to watermark quality. Consequently, we contacted papermakers and manufacturers of watermark wires in a survey of industry experts. We concluded that at least three attributes are commonly being considered: background formation, watermark intensity, and watermark sharpness.

In this study, our goal was to establish the optics and the analysis techniques that would precisely quantify all three attributes.

THE MEASUREMENT OF WATERMARK QUALITY

What defines a good watermark depends on the application. A "sharp" watermark, for example, may be appropriate for security paper but not for some writing papers. By definition, a watermark that is sharper than another has a steeper gradient of opacity change between adjacent light and dark areas.

To quantify watermark quality, we have to compare an example of the watermark with either a reference watermark image or the artwork that corresponds to that image. It is not a trivial task to compare two watermarks or a watermark and reference artwork. It is physically difficult to position two watermarks at the exact same point and orientation for imaging.

We developed a technique to orient the two images about a common reference point and normalize the background formation quality. Once the sample and reference watermark images are aligned and normalized, we can analyze the quality of the match. The attributes that affect watermark quality include:

1. Background formation
2. Watermark intensity
3. Watermark sharpness difference
4. Watermark intensity ratio.

Background formation. Sheet formation can interfere with watermark analysis. As the formation gets worse, the sharpness of the watermark will diminish, and the watermark will become less distinguishable from the surrounding paper.

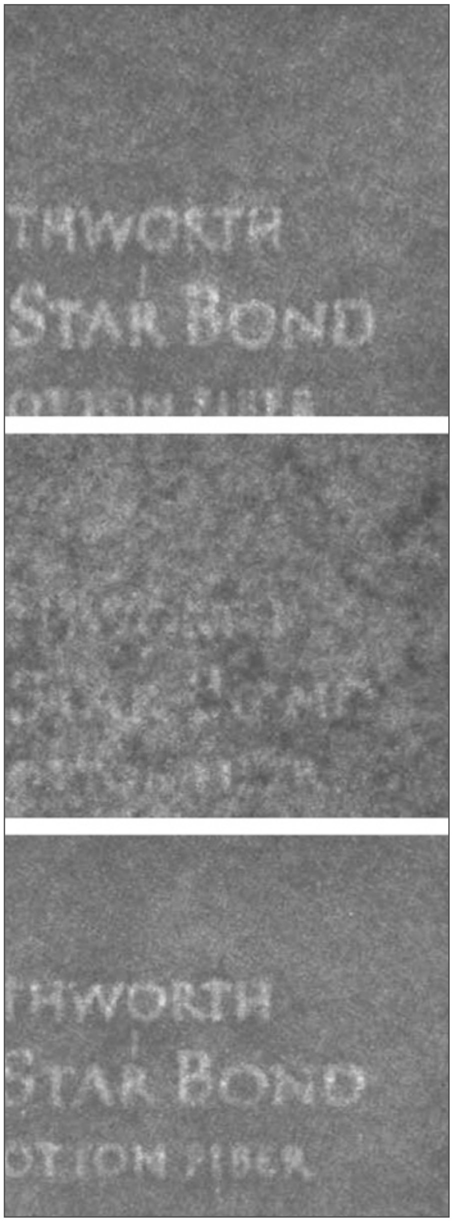
In formation analysis, the section of the sheet with the watermark must be analyzed separately from neighboring areas that do not have watermarks. A selected watermark area is inscribed within a box. The box size can be adjusted to accommodate watermarks of various sizes and design. Analysis of watermarks within the box are denoted as "in", and analysis outside the box are denoted as "out".

The background formation is defined as the percent coefficient of variation (% COV) of the pixel grey levels (GL) outside the box. The formation nonuniformity is calculated as the GL standard deviation of the *out* pixels divided by the average GL and is reported as a percentage. This percentage is also called the "percent contrast of formation" [3] and is a measure of the intensity of flocs relative to voids in the sheet. A higher value corresponds to poorer formation. The average pixel gray levels both inside and outside of the inscribing box are also measured and used in subsequent calculations.

Watermark intensity. The watermark intensity is determined for the sample and reference images separately. To obtain sensible results, it is necessary to factor out the background formation. The % COV calculated out of the box is subtracted from the % COV of the watermark image in the box. The resulting difference is defined to be the percent intensity of the watermark.

Watermark sharpness difference. This measure is defined as the difference between the percent intensity of the reference and that of the sample watermark images. If the value is positive, then the sharpness of the reference image is greater than the sample watermark, and vice versa for a negative value.

Watermark intensity ratio. This ratio is calculated as the percent intensity of the sample image divided by the percent intensity of the reference image.



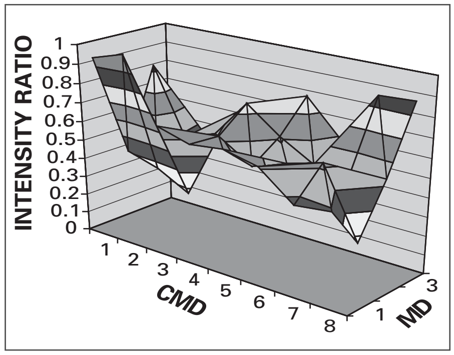
1. Typical watermark used in this study.

If the intensity ratio is greater than 1, then intensity of the sample watermark image is greater than that of the reference image, and vice versa.

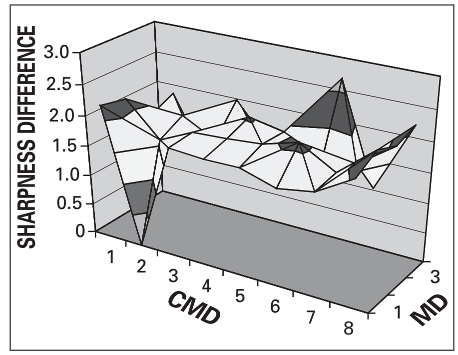
EXPERIMENTAL

Image alignment technique

We developed and tested several different techniques to align the watermark images. In the most successful technique, the two watermarks are assumed to be the same size but are considered to be located and oriented differently.



2. Variation in the commercial watermark quality in the cross-machine direction.



3. Variation in the commercial watermark quality in the machine direction.

The alignment technique entails the following steps:

Step 1. Inscribe the area of the watermark test image within a box.
Step 2. Select a point within the box as a target point.
Step 3. Select the "same" point on the reference watermark image.
Step 4. Align the target point of the test image with the target point of the reference image. Rotate and translate the image about the target point until a minimum difference in gray level is calculated for all the pixels within the inscribing box.

During the course of this work, we observed that the average GL in the inscribing box was usually different from that outside the box. The difference depended on the quality of the formation and the intensity of the watermark. Furthermore, the differences varied with different images of the same watermark.

The average GL in the inscribing box has to be normalized to a nominal value for proper alignment if we are to obtain good results. Of the different procedures developed, the best involved a modification of Step 3: First, rescale the average GL inside the box to 128 for both the test and reference images, and then apply the rescaling function to corresponding images outside the box.

Image analysis

To obtain watermark images, we used an OpTest Paper PerFect Formation Analyzer. The optical resolution of the charge-coupled device (CCD) camera was 150 μm/pixel. The field of view was 70 mm × 70 mm, with a pixel array of 256 × 256 corresponding to an effective aperture of 0.27 mm. An adjustable transmitted light source was used with suitable infrared filters to ensure that only visible light was being measured. The light reaching the CCD camera was nominally adjusted to 50% of full scale (128 GL). This set-up was consistent with published guidelines [4].

It took less than 5 min to collect and analyze a set of 32 watermarks images. The all pairs ranking, with four judges examining only 12 watermarks, took 2 h to complete.

We established the technique's ability to reproduce results by repeating the test five times with a reference watermark and two different sample watermarks. The coefficients of variation of the percent difference in sharpness and the intensity ratio were within 2% and 5%, respectively.

Validation of watermark algorithms

In the initial validation study, we used 75 g/m² paper containing watermarks of the letter "S" inscribed within a disk. Here, we refer to these images as PS for "positive sample" and PR for "positive reference." (See the appendix for the meaning of "positive.") In one case, the sample and reference images were the same; in the other case, the sample reference was different from the reference image.

Same image for sample and reference. The PS watermark was used for the sample and the reference image. The results are reported in the first row of **Table I.** The sharpness difference was zero, and the intensity ratio was 1 because the watermarks were the same. In other words, the alignment technique was operating correctly for identical images.

SAMPLE ID	REFERENCE ID	REFERENCE BACKGROUND FORMATION, % COV	% INTENSITY		REFERENCE–SAMPLE	
			SAMPLE	REFERENCE	SHARPNESS DIFFERENCE	INTENSITY RATIO
PS	PS	7.30	1.03	1.03	0.00	1.00
PS	PR	7.50	1.03	0.78	-0.38	1.32
PR	SIR	0.45	0.64	4.87	3.10	0.12
PR	DIR	0.45	0.64	4.25	2.73	0.16
PS	SIR	0.45	1.03	4.87	3.27	0.27
PS	DIR	0.45	1.02	4.25	2.79	0.24

I. Summary of the results from the watermark image analysis for two actual sample watermarks, PS and PR, compared to actual and synthetic reference watermark images.

Different sample and reference images. Two different watermarks, PS with PR, were compared, and the results are reported in the second row of the table.The sample watermark, PS, had a higher percent intensity than the reference watermark, PR. Consequently, the sharpness difference values were correspondingly negative, and the intensity ratio values were greater than 1.

The % COV values of the in box, containing the sample (PS) and reference (PR) watermarks, were both 8.3%. The percent COV values for background formation of both the sample and reference were about 7.5%. The sample and reference images were aligned and normalized according to our procedure. After alignment, the reference and sample images were subtracted from each other, yielding a third image called the "reference–sample" residual image. The percent COV values of the in box, containing the reference–sample residual image, was 7.13%. This value was comparable to the percent COV value outside the box and was significantly lower than the corresponding values for the PS and PR watermark images. These results indicated that the watermark alignment technique was operating properly.

Comparing watermarks with a synthetic reference image

The preferred way to assess the quality of a watermark is to evaluate it relative to the corresponding artwork or to an idealized image "synthesized" from the watermark. A method for synthesizing watermark images is described in the appendix.

We generated two synthetic watermark images from the PR watermark. Synthetic watermark SIR had sharp edges. The DIR synthetic watermark was identical to SIR, except with more diffuse edges. ("SIR" denotes "sharp image reference," and "DIR" stands for "diffuse image reference.")

Comparing watermark PR with synthetic reference images

Rows 3 and 4 of Table I show that the sharpness difference was 3.10 between PR and SIR and was 2.73 between PR and DIR. The positive values indicate that both of the synthetic reference images were sharper than the PR watermark. Furthermore, the higher sharpness value with SIR was consistent with its having a shaper edge than DIR had.

The results in the table also show that the intensity ratios were less than 1. This outcome suggests that the intensity of the sample PR was less than that of the references DIR and SIR.

Comparing watermark PS with synthetic reference images

The same trends in the values of sharpness difference and intensity ratio of PR were also observed with PS. Rows 5 and 6 in Table I show that the sharpness difference between PS and the synthetic reference image was 3.27 for SIR and was 2.79 for DIR. The positive values indicate that both of the synthetic reference images were sharper than the PR watermark. Furthermore, the higher sharpness value with SIR is consistent with its having a shaper edge on the "S" than DIR had.

In both cases, the intensity ratio was less than 1. This outcome suggests that the intensity of the sample PS was less than that of both DIR and SIR. The intensity ratio was slightly higher for DIR than for SIR.

TRIAL RUN	MILL TESTER	IMAGE ANALYSIS	
		INTENSITY RATIO	SHARPNESS DIFFERENCE
1.00	2.5	0.6	1.7
2.00	1.5	0.5	2.0
3.07	0.5	0.5	2.2
3.13	0.0	0.4	2.4
3.15	0.7	0.4	2.0
5.03	2.5	1.0	1.8
5.04	1.7	0.9	2.0
6.05	1.0	0.8	2.0
7.01	3.0	0.7	1.9

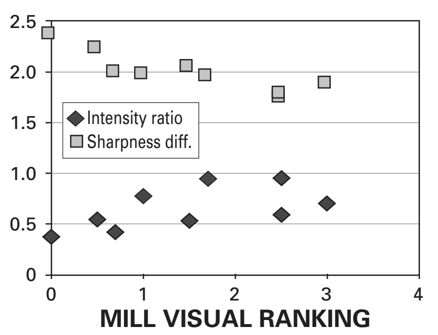
II. Mill test scores and watermark quality analysis results.

RESULTS AND DISCUSSION

We referenced the watermarks analyzed against a watermark judged to be superior. All watermarks were made on paper produced during forming trials on a commercial paper machine [5]. Watermark B from Strip 1 of Run 1 (labelled 1.00/01/b) was used as the reference and is shown at the top of **Fig. 1**.

Figures 2 and 3 illustrate that the variation in the commercial watermark quality was significant in both the cross-machine direction and the machine direction (at 7 m intervals). The mill's method was to check a few watermarks at a particular CD position after each reel turn-up. Image analysis has the advantage of being able to derive analysis statistics with greater confidence over more specimens.

Thirty-two watermarks (8 per CD × 4 MD strips) were averaged per run. The watermark analysis results are presented in **Table II**. A better watermark quality corresponds to a higher intensity ratio and a lower value for sharpness difference. If the test and reference watermarks are completely equal in quality, we get an intensity ratio of 1 and a sharpness difference of 0.



4. Watermark quality using image analysis as a function of mill testing.

RUN NUMBER	FORMATION % COV	ALL PAIRS RANKING[1]	INTENSITY	INTENSITY RATIO	SHARPNESS DIFFERENCE
3.06/01/g	10.05	7.0± 0.8	1.48	0.28	3.02
3.06/07/d	9.81	10.0± 1.4	3.30	0.72	2.10
3.06/10/d	11.22	4.5± 0.6	0.92	0.37	1.37
6.00/01/c	9.86	10.5± 1.3	2.84	0.60	2.45
6.00/10/b	10.20	9.8± 1.5	2.88	0.55	2.38
6.05/01/b	9.46	13.0± 0.8	3.42	0.69	2.00
7.01/10/c	7.60	18.8± 1.0	4.38	0.85	1.44
7.00/04/f	8.52	15.8± 2.3	2.51	0.53	1.11
7.01/04/c	7.79	20.3± 1.8	4.22	0.87	1.28
3.15/01/a	11.02	0.8± 1.0	0.93	0.31	2.35
5.03/10/b	10.85	19.3± 2.2	8.56	1.17	1.06
3.15/07/a	12.06	1.8± 0.5	0.51	0.05	1.73

[1]Average score, where higher = better.

III. Results of the watermark quality analysis of specific watermarks selected from the trial samples, along with the all pairs ranking scores

We compared the image analysis technique with visual testing in terms of sensitivity and precision. Two methods of visual ranking were used—straight ranking and all pairs ranking.

Straight ranking . Mill testing was based on visual ranking over a 17 point scale with values between 0 and 4. A higher score corresponded to a better quality of watermark. The mill tester examined one watermark per run. The results are reported in Table II. Also shown are the averages for the watermark image analysis, where a better watermark corresponds to a higher intensity ratio and a lower difference in sharpness.

Figure 4 presents a plot of the watermark analysis results as a function of the mill tester rankings. A linear regression yields R^2 values of 0.34 for the intensity ratio and 0.75 for the sharpness difference. These R^2 values are statistically significant at a 95% confidence level.

A stronger correlation was expected between the intensity ratio and mill visual ranking. The poor agreement was probably caused by the high variability of watermark quality and the limited amount of visual testing performed. As a result, we decided to perform an all pairs ranking study [1].

All pairs ranking

This study was carried out on 12 specific watermarks selected from the nearly 1000 collected during the paper machine trials. These 12 watermarks were selected to represent a range of qualities. For each watermark, the intensity ratio and sharpness difference results were correlated with the average all pairs ranking score. The average scores were calculated from the rankings of four judges and are reported in **Table III**. Linear regression analyses yielded R^2 values of 0.77 for the intensity ratio and 0.30 for the sharpness difference. Both of these correlations were significant at 95% confidence. The correlations were weakest for the sheets with the poorest formation.

The three sheets with the worst formation—3.15/07/a, 3.15/01/a (middle image in Fig. 1), and 3.06/10/d—had watermark intensities of less than 1, as shown in Table III. An intensity of less than 1 indicates that the intensity of the background formation was equivalent to the intensity of the watermark. Under these conditions, the judges would be unable to visually discriminate watermark quality, as explained by the Webber-Fechner Law [1]. These poorly formed watermark samples were included to determine the lower limit of watermark intensity required for visual judging. These samples were subsequently removed from the analysis.

With these spurious samples removed, the analysis yielded R^2 values of 0.63 for the intensity ratio and 0.86 for the sharpness difference. Both of these correlations were significant at 99.5% confidence.

CONCLUSIONS

The measurement of watermark sharpness difference and intensity ratio by image analysis, when combined with the appropriate optics platform, correlates significantly with visual assessment. Image analysis can analyze and average a large number of watermarks quickly, providing greater accuracy and precision. Furthermore, poor sheet formation can obliterate watermark quality to the point at which visual assessment is dubious. **TJ**

LITERATURE CITED

1. Trepanier, R. J., *J. Pulp Paper Sci.* 13(4): J111(1987).
2. Frith, C. D., and Nias, D. K., *Gen. Psych.* 91: 163(1975).
3. Trepanier, R. J., *TAPPI J.* 72(12): 153(1989).
4. "Basic guidelines for image analysis measurement," TIP 0804-09, TAPPI PRESS, Atlanta, 2001.
5. Trepanier, R. J., and Turner, N., reprints of the *PAPTAC 88th Annual Meeting*, Montreal, Canada, 2002, p. C87.

Received: October 24, 2001
Revised: September 24, 2002
Accepted: September 24, 2002

This paper is also published on TAPPI's web site <www.tappi.org> and summarized in the August *Solutions! for People, Processes and Paper* magazine (Vol. 86 No. 8)

INSIGHTS FROM THE AUTHORS

Papermakers who manufacture papers with watermarks have frequently asked us if there is any technology that can quantify watermark quality. Having done some research in characterizing sheet formation by image analysis, we knew that analyzing watermark quality would be similar in many ways. However, the first problem was to define what makes a good watermark.

We found that there are many definitions of what constitutes a good watermark. Our first challenge was to quantify these aspects of quality into the minimum number of image attributes to measure. On the other hand, we knew that image analysis would be much faster, more precise, and more reproducible than the traditional method of assessing watermark quality visually.

One thing we learned through this research is that the quality of a watermark is related to the end use of the paper. Consequently, there are more attributes to quantify than just sheet formation. Surprisingly, the conditions under which the best formation is obtained may not coincide with the best watermark quality.

We expect to work with papermakers to improve the technology as we apply it. For papermakers, the ability to monitor watermark quality accurately should prove advantageous. When a watermark is changed or when the same watermark is to be used on a different grade of paper, the image analysis method should prove beneficial in comparing the quality of watermarks made before and after the change.

Trepanier and Perron are with OpTest Equipment Inc., 900 Tupper St. Hawkesbury, ON, K6A 3S3, Canada. Email Trepanier at trepanier@optest.ca.

Trepanier

Perron

APPENDIX: PROCEDURE FOR SYNTHESIZING A WATERMARK IMAGE

We followed a five-step process to develop a synthetic watermark image. **Figure 5** shows the results of these steps for the PR watermark.

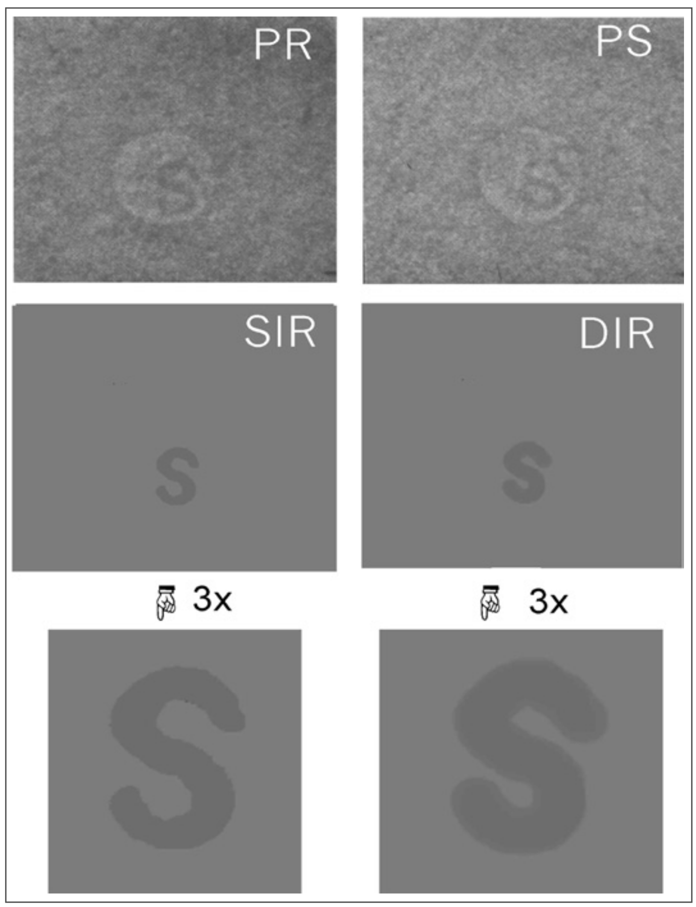
Step 1. An image of a selected watermark is collected, preferably from the same image analyzer on which all the other images will be collected. Otherwise, the number of x and y pixels must be the same for both the synthetic image and the watermark sample images. If printed artwork is provided, then the image should be collected with reflected light to minimize formation problems. The watermark image is imported as a bitmap into any commercial graphics software that can perform the subsequent steps.

Step 2.The pixels in the imported watermark image are converted to binary numbers such that the all pixels above the mean GL are assigned a maximum white value (255 GL) and those below are assigned a minimum black value (0 GL). This procedure gives us a positive image. If the watermark features of interest are lighter than the mean GL, then a negative image should be used.

Step 3. The watermark features (letter “S”) are enhanced, or corrected, using the graphics software drawing tools. Other features, such as background formation, are removed with the eraser tools. This image has the highest contrast, with white for the background and black for the watermark feature.

Step 4. For us to make sensible comparisons with actual watermarks, the background GL must be similar to that in the paper. The background of the synthetic watermark is “filled” to a mean level of 128 GL. In the image, the watermark contrast is reduced from 255 GL to 128 GL.

Step 5.This final step involves aesthetic issues. The level of watermark intensity and sharpness depends upon the application. In this case, the intensity of the synthetic water mark was set to about 4.5%. The intention was to mimic an actual watermark.



5. Watermark validation test images.