Methods for measuring and predicting the printability of paper

Marja Mettänen

KEYWORDS: printability, print quality, missing ink, surface topography, statistical analysis, image based measurement

SUMMARY: 2D measurements of local paper properties have been increasingly employed in paper research, as many small-scale print defects can directly be related to inhomogeneities in the paper structure. We have developed methods for analyzing local print anomalies and their probabilistic relations with paper properties and their abnormality. Given a large set of accurately aligned multivariate data, we estimate the joint probability density functions (pdfs) of the measured variables and analyze the tails of the pdfs, corresponding to the anomalies. The dependences between the small-scale print quality variations and paper properties are examined through visualizations of the distributions and anomalous regions, supported by quantitative measures of statistical dependence. The analysis procedure is applicable to various properties of paper and print. We demonstrate the procedure by two case examples where the data is from image based measurements of print density variations and the surface topography of unprinted paper. The results with offset printed newsprint and gravure printed supercalendered paper confirm that, while the correlation between the measured properties is weak, depressions on the unprinted paper surface cause significantly increased probability of abnormal print quality.

ADDRESS OF THE AUTHOR: Marja Mettänen (marja.mettanen@tut.fi): Tampere University of Technology, Department of Automation Science and Engineering, P.O.Box 692, FI-33101 Tampere, Finland.

Introduction

The printability of paper refers both to the runnability during the press operation and to the achievable print quality, controlled by the interaction between the critical properties of the paper and the printing parameters. Some relations have been established between physical measures of prints and the perceptually relevant subjective quality dimensions (Mangin, Dubé 2006; Eerola et al. 2008; Oittinen et al. 2008). At the same time these studies manifest the ambiguity of print quality as a target property to be optimized. When trying to ensure high print quality, it remains justified to minimize the number and severity of local abnormalities that are perceived as print imperfections, such as inadequate level of sharpness, contrast or gloss, or uneven appearance of the print. These issues can be partly controlled by competent printing press operation, but it is also important to be able to measure and control the paper properties that can cause print defects.

We have developed methods for quantifying how much information about printability is achievable through the small-scale measurement of paper properties before printing. We focus particularly on the surface topography of paper which we measure with a pixel size of 10 µm by 10 µm. Surface roughness is one of the most commonly reported characteristics of paper and it has straightforward implications for the quality category and price of the paper. The roughness of the substrate is known to explain print quality to some degree. Early studies on this relationship applied air-leak based roughness and porosity indices to describe the overall paper surface properties, and correlated them with print gloss and print uniformity (Lyne 1976b; Heintze, Gordon 1979). At the same time it was already recognized that air-leak methods are not optimal for the modeling; a more precise description of the surface properties is the distribution of surface heights, or the distribution of surface void sizes (Karttunen 1973; Parker 1973). Techniques for measuring the dynamic contact fraction were also studied (Lyne 1976a) to account for the compressed structure of paper in the printing nip.

The development in imaging and computer technology has opened the way for analyzing 2D and 3D images of paper and print properties. The common approach to utilize the surface topography images is to process them into scalar parameters that describe the roughness or porosity. Indices of print quality can be similarly derived from images of local print density or gloss. Relations between the parameters are then analyzed in experimental setups. Given the wide range of available techniques and resolutions for measuring surfaces, it is important to note that many of the parameters applied to characterize surface topography images are sensitive to the spatial and height resolution of the image (see, e.g., (Järnström et al. 2008; Poon and Bhushan 1995). The relevant size scales at which to compute the parameters depend on the application. Paper coatings present nanoporous features that affect their optical properties (Järnström et al. 2007). The size scales most

relevant for paper and print gloss relate to the micro or sub-micro scales of roughness, i.e., features of size from 2-4 µm to approximately 12 µm (Chinga et al. 2007; Järnström et al. 2008; Järnström et al. 2009). For coating and printing purposes a more relevant size scale originates from the individual fibers and fiber bundles (Kajanto et al. 1998). The focus in this work is on spots of missing ink which can be visually disturbing and cause unevenness of the print (Saarelma, Oittinen 1993). Studies on the causes of missing ink in solid print areas (Barros 2006; Barros, Johansson 2006; Mesic et al. 2006) and on missing dots in halftone production, particularly within gravure printing (e.g., by Metsäjoki (2005) and Preston et al. (2008)) have reported that the most critical size scale of defects in paper is from a few micrometers to one millimeter.

An alternative to the above mentioned parametric characterization is to search for pointwise relations directly from the spatially aligned images of paper and print properties. This approach has been successfully applied by Barros, who has explained uncovered areas of flexographic prints by the pits and valleys of the printed paper surface using aligned images with pixel size of 15.6 µm by 15.6 µm (Barros 2006; Barros, Johansson 2006). MacGregor et al. (1994) have aligned the images of gloss variation and surface topography of a printed paper, with pixel size of 4 µm by 4 µm, to show that the local surface inclinations can explain up to 78 % of the gloss. Mangin et al. (1993) have acquired aligned 3D topography presentations of paper surface under increasing load using confocal microscopy. They have then adopted the parametric approach and correlated the number of missing gravure dots with the compressibility and printing roughness. Kajanto (1989; 1991) and Dickson (2006) have found statistically significant - although not very strong - pointwise linear correlations between local print density variation, formation and surface topography, using 40 - 300 µm pixel size in the aligned images.

The computation of correlation coefficients and the fitting of (linear) regression parameters to the data represent methods of identifying deterministic relationships between the measured properties. With the large number of independently measured data points in the aligned images, it is possible to estimate the joint probability distributions. Our approach to the analysis of the dependence between print quality and the surface structure of paper is to use the pointwise aligned image data and to analyze the full joint probability densities of the measured properties. This is because the dependences are statistical and non-Gaussian, which is demonstrated by the asymmetric joint distribution in *Fig 1*. Most notably, the values of print reflectance spread towards the high end of their range. The joint distribution cannot be summarized with one expectation vector and covariance matrix, and regression between the surface topography and print quality does not provide full information about the dependences. A reliable estimate of the joint probability density function (pdf), on the other hand, is a complete description of the statistical behavior, and the interdependence of the variables can be inferred with conditional pdf derived from the joint pdf.

We examine the print result through images of print density variations, and concentrate on the imperfections in the print that are detected as small local areas of inadequate ink coverage or missing ink. We analyze the statistical dependence between the surface topography and print result, and determine the probability of encountering missing ink in locations where the surface topography shows exceptional values. We limit the wavelength range of our topography maps to shorter than 0.5 mm, which means that the largest topography features (e.g., pits) are of diameter 250 µm. The smallest features in the images are of diameter 20 µm, which is defined by the imaging resolution. The wavelengths of print density images are not limited by filtering.

The rest of this paper is organized as follows. A general level algorithmic description is first given of the probabilistic approach to characterizing the relationship between surface topography and print quality. It will be followed by case examples with newsprint and SC paper, together with the results of the analysis. Discussion about the methods and their scope of application will be presented after the case examples. The final section will conclude the work.

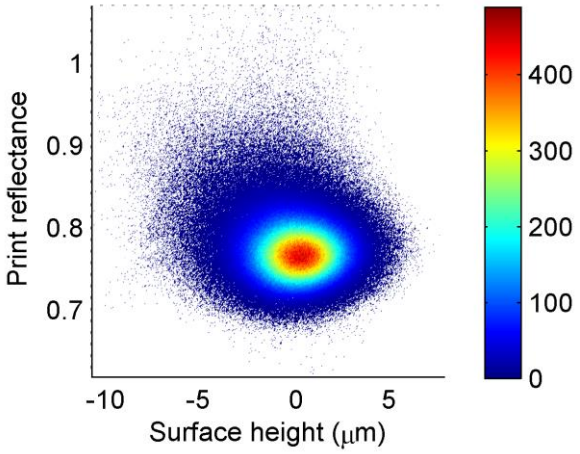

Fig 1: Joint histogram of surface topography before printing and reflectance after printing, measured from 22.5 mm x 15 mm area of a newsprint sheet printed with compact cyan in offset press. The pixel size in the measured maps is 10 µm x 10 µm.

Print assessment algorithm

The following description of the analysis algorithm assumes image based measurements of paper and print properties. An image of print density variations can be recorded by scanner or camera. We use the term print reflectance to refer to a camera based measurement of the light reflected from the printed surface in uniform diffuse illumination with white light. Maps of paper properties can present, for instance, basis weight (with beta formation measurement), light transmittance, thickness, compressibility or surface topography. We propose the following algorithm for analyzing the print quality and its dependence on the measured paper properties. More detailed descriptions of the steps are given after the numbered list.

1. Register and align the images
2. Define the pixels that belong to the valid analysis area, *B*
3. Estimate the probability density functions (pdfs), using the observations in *B*
    a. one-dimensional pdf of each variable
    b. multidimensional joint distribution of print reflectance and paper properties
4. Make anomaly maps by defining the pixels in *B* that belong to the tails of the pdfs
5. Assess the dependences within anomaly maps
    a. define the proportion of overlap of anomalies detected in separate measurements
    b. characterize the conditional distribution of print reflectance values, given that the points are anomalous in the paper property measurements
    c. characterize the conditional distribution of paper properties, given that the points are anomalous in the print reflectance map
    d. estimate the conditional joint distribution of the variables, given that the points are anomalous according to a selected anomaly map; characterize the dependences by mutual information (MI)

Image registration and alignment

Accurate image alignment is a prerequisite for the meaningful analysis of the relationship between the measured variables. Alignment may require special registration marks, for instance laser marking, or it can be based on the similarity of the texture in the images, which is the case in our topography maps. Image alignment software is readily available today, supporting various rigid and non-rigid transformation types. For our purposes affine transformation is adequate. A description of our image alignment procedure can be found in (Mettänen et al. 2008). The aligned 2D maps constitute a multivariate image data set.

Defining the valid analysis area

There are two typical cases where parts of the imaged area need to be discarded from the analysis. The first is a result of imaging optics that can cause geometric distortions at the edges of the measured maps. Even small distortions, combined with the small pixel size and the warping of the images at the alignment phase, can destroy the accurate alignment of data points at the edges and corners of the images. Therefore it may be advisable to restrict the analysis are to the centre parts of the multivariate image. The second case occurs in the analysis of halftone print where not every point of the imaged area is expected to be covered by ink. In these cases the halftone pattern has to be recognized to separate the raster dots from the white paper. Our approach to defining the pixels that belong to the raster dots has been described by Mettänen et al. (2009b) and Lauri and Ihalainen (2010).

Estimating the probability density functions (pdfs)

It can be expected that, however accurate the measurement values at each pixel, there are no deterministic point-to-point relationships between local print density and the local structural properties of paper, such as surface height. Instead, the relationships are probabilistic and thus they are appropriately described with the full joint probability distributions of the measured properties (Papoulis 1990). A joint probability distribution of *N* variables can be described with an *N*-dimensional histogram or with a parametric distribution model such as the Gaussian mixture model (GMM) (Nabney 2002). The GMM approach has two attractive properties. First, it is a very compact representation of the large amount of data; being a weighted sum of Gaussian components, GMM is fully defined by its mean vectors, covariance matrices and the component weights. Second, the parametric representation enables the analytical calculation of conditional probability density functions (pdfs) of individual variables. The parametric nature of GMM can also be considered its downside. In principle, one would have to use as many GMM components as there are data points, in order to faithfully describe the observed data. This is obviously not meaningful, and thus the GMM may lose some details, such as an elaborate shape of the tail of a distribution. The histogram can better preserve the fine details in the tails but analysis results computed through histograms depend

heavily on the number of histogram bins used. A large number of histogram bins leads into seemingly high resolution in the presentation of the distribution, but the uncertainty about the height of each histogram bin may become large. Both the histogram and the GMM approach are approximations, and the true distribution that generated the observed data will always remain unknown. In this paper we demonstrate the analysis methods using mainly histograms.

Making anomaly maps

We use the term anomaly map, or mask, for binary images of the same size as the measured area, presenting by 1's (true values) the points that are deemed anomalous and assuming the value 0 (false) elsewhere. Anomaly maps indicate the points and areas that deviate most strongly from the typical statistical behavior of the measured properties. The simplest way of constructing an anomaly map is to base it upon a single measurement, such as the surface topography image. The pixels at which the surface height value is below[1] a certain threshold are marked as anomalies. We typically define the threshold to correspond to a certain percentile of the distribution whose extreme values are sought. For instance, a 2.5 % percentile of a topography image is the surface height below which only 2.5 % of all the surface height values in that image fall. In print reflectance images, high values denote missing ink, and thus a 2.5 % reflectance anomaly map indicates the pixels at which the print reflectance value is higher than the corresponding 97.5 % percentile value. An anomaly map can also be based on the joint distribution of several properties. In that case the map is implemented by evaluating the likelihood of each multivariate data point with respect to the joint distribution, and thresholding the resulting likelihood map by the percentile principle. The points that have very low likelihood value are deemed anomalous. We often refer to the anomaly maps shortly as masks since they determine the data points that are selected for further analysis in the next step of the algorithm.

Dependences within anomaly maps

The analysis of anomaly maps answers the question whether the abnormality of one property can be predicted based on the observation of the other property. It can thus provide information of print defects and their origins. At first the interrelation of anomalies in the measured properties is quantified by computing the degree of overlap between the points and areas of 1's in two masks that have identical mask percentage. The second approach is to examine the conditional distributions as described in the algorithm steps 5b and 5c. Selecting the data points according to a particular mask corresponds to applying a condition. For example, our interest is in the print defects caused by the exceptional values of surface topography. Thus we examine the conditional distribution of print reflectance values, given that surface height values are below the threshold that defines the topography mask. Finally, the joint distribution of several variables can be estimated using the data points indicated by a particular mask. This is the conditional joint distribution mentioned in the algorithm step 5d. It typically has a very distinctive shape due to the application of the anomaly condition, and cannot be summarized by a single Gaussian model or correlation coefficient. Mutual information (MI) (Bishop 2006) is a suitable measure of dependence in this case, as it describes the overall statistical dependence between an arbitrary number of variables and it is not restricted to linear dependences.

Case examples and results

This section demonstrates the print assessment algorithm with two case examples. The purpose is to show how the steps of the algorithm complement each other; the images of 2D and 1D pdfs provide realistic visualizations, and mask overlap analysis, tail analysis and mutual information serve as numerical quantification of the dependences. The first example compares two solid cyan areas printed on a newsprint sheet in different offset printing conditions. Every point is supposed to be covered by ink and thus the number of data points available for the analysis is very large. This example is optimal for demonstrating the joint distributions of print reflectance and surface topography. In addition, we determine the overlap of missing ink spots and topography anomalies with respect to the size of the anomalous areas found on the topography. The second case example complements the first one in two ways. Firstly, it deals with supercalendered (SC) paper and gravure printing, which is more relevant for commercial printing than sheet-fed offset with newsprint. The gravure technique also introduces the need to detect the screen pattern that is inherent in practically all printed products. Secondly, the latter case example demonstrates the dependence of print quality on paper roughness.

[1] Considering ink transfer, the deep pits on the surface of paper are more relevant anomalies than high elevations because they are more susceptible to being unattainable by the ink (Walker, Fetsko 1955).

In the printing experiments, the areas selected for examination have been imaged with a photometric stereo device that applies the principles described by Hansson and Johansson (2000). It is based on photographic imaging with slanting illumination and it provides the reflectance and surface topography maps of the paper sample from exactly the same area. It should be noted that our reflectance maps do not exactly conform to the definition of reflectance (Van den Akker 1982). The maximum value in our maps is larger than 1, but this is a matter of normalization and it is unimportant for the print analysis that concentrates on the relative values of reflectance. The size of the imaged area is 22.5 x 15 mm and each image contains 2268 x 1512 pixels, thus the pixel size is approximately 10 x 10 µm. The same area of each paper sample has been imaged before and after printing, and the images have been aligned at subpixel accuracy, using the unprinted and printed topography maps that contain similar texture (Mettänen et al. 2008). After alignment, the topography maps have been high-pass filtered to focus the analysis on the small-scale surface height variations.

The photometric stereo images are in RGB (red, green, blue) colors but the topography maps have been computed from the mean of the color channels to minimize the noise in the topography map of unprinted paper. Print quality is assessed from a single color channel of the reflectance map of the printed paper. The selection of the color channel depends on the color of the printing ink. With cyan ink we use the red channel of the reflectance images, and with red ink we use the green channel, because these give the best contrast between white paper and inked areas.

Newsprint in sheet-fed offset printing

The newsprint case deals with newsprint sheets printed in a pilot scale sheet-fed offset press. Two rectangular areas of 100 % cyan are examined. One has been printed without water application and without back-trap conditions, i.e. only the cyan nip printing. The other area presents normal 4-color offset conditions. There are four trial points that differ in PPS roughness, and four sheets in each trial point. Both of the described cyan areas are present on each sheet. After aligning the images taken before and after printing, we have defined the valid portion of each area, free from geometric distortions. This limits the number of pixels from the original 3.4 million to a minimum of 2 million, depending on the sample. All image analyses are restricted to the valid areas. *Table 1* summarizes the roughness and density data of this printing experiment and reports the correlation coefficient between print reflectance and unprinted topography at each trial point. The roughness parameters computed from the topography map are the root mean square deviation, S_q, and skewness, S_{sk} (Stout, Blunt 2000).

The correlation coefficients indicate that the linear dependence between surface topography and print reflectance is weak. Only 6 % (R^2 at $|R| = 0.24$) or less of print reflectance variation can be explained by the topography of the unprinted paper in a linear model. In addition, the roughness parameters and the average reflectance and density readings, fail to describe the details of the print result. Instead of compressing the dependence information into single numbers, we prefer more expressive ways of characterizing the interrelations of print reflectance and surface topography.

The surface topography maps are high-pass filtered with wavelength limit 500 µm, which emphasizes the sharp pits and elevations on the surface. A local variance map and a map of the local gradients in the printing direction are computed from the high-pass filtered topography and used in the analyses as well. This allows also other surface properties than height to explain the print quality. Our objective is to determine how the observation of an exceptional topography point on the unprinted paper changes our information about whether the print quality attainable at that point will be exceptional or not. The exceptional points are indicated by binary masks. To test whether the gradient and variance maps of topography provide useful information for predicting or explaining missing ink, the joint distribution of topography (T), gradient (G) and

Table 1: PPS and density measurements together with parameters computed from unprinted topography maps (high-pass filtered with cut-off wavelength 500 µm) and printed reflectance maps. The parameters are reported as the averages of four parallel samples (sheets). The density readings present averages of 20 samples. BT stands for back-trap.

PPS10 (µm)	S_q (µm)	S_{sk}	refl. normal	refl. non-BT	density normal	density non-BT	corr.coef(topo, print refl.) normal	non-BT
2.61	1.65	-0.34	0.77	0.77	0.94	0.99	-0.09	-0.17
2.77	1.71	-0.36	0.76	0.75	0.95	1.00	-0.13	-0.15
3.16	1.66	-0.36	0.78	0.78	0.93	0.99	-0.16	-0.24
3.63	1.91	-0.37	0.79	0.78	0.90	0.93	-0.17	-0.24

Table 2: Summary of masks.

Mask name	Points selected according to
Topo	lowest values of unprinted topography
TGV	most unlikely combinations of topography, its gradient and its variance
Refl.	highest values of print reflectance
Random	uniformly and randomly distributed points

variance (V) is used as a basis for constructing a mask. In addition, two masks are based simply on the marginal distributions of the observed topography and print reflectance values, respectively. We also use a random mask to evaluate the overall dependences in the data. *Table 2* presents a summary of the masks.

We have previously analyzed the whole set of 16 newsprint sheets and reported that the typical overlap of topography based (i.e., Topo or TGV) masks with the points of unusually high print reflectance (i.e., Refl. mask) varies from 7 % to 12 % in the normally printed cyan samples and from 10 % to 18 % in the non-back-trap cyan samples (Mettänen et al. 2009a). In that study, only mask percentages smaller than or equal to 1.5 % were used, and the wavelength limit in the high-pass filtering of the topography map was 250 µm. We have examined the dependence of mask overlap on the high-pass filtering limit of the topography image, and applied mask percentages from 0.5 % to 10 %. The overlap of topography based masks and reflectance masks diminishes when the wavelength limit is raised to 1000 µm or longer. With wavelength limits of 250 µm and 500 µm the results vary depending on the strategy of constructing the topography based mask. The results also suggest that mask percentages larger than 4 % do not significantly improve the overlap of reflectance and topography anomalies. This is an expected result in prints where missing ink is not very common.

In the following, we demonstrate the analysis procedure by figures and tabulated results, concentrating on the normally printed and the non-back-trap cyan areas of one paper sheet. One of the smoothest sheets, with PPS roughness of 2.61 µm, has been selected for this purpose. *Table 3* lists the overlap percentages computed from the two cyan areas of this sheet, using 2.5 % masks and three wavelength limits in topography filtering. As mentioned above, the rest of the analyses use the 500 µm wavelength limit. *Fig 2* and *Fig 3* present the overlap results in varying size scales of uniform spots (objects) formed by adjacent mask pixels in Topo and TGV masks, using 2.5 % mask percentage. The range of spot sizes is from a few pixels to the maximum size of a circular spot allowed by the filtering, which equals $\pi*12.5^2 \approx 490$ pixels.

Table 3: Overlap of reflectance based and topography based masks in the studied areas with three different wavelength limits applied to the topography map. Mask percentage, 2.5 %, has been subtracted from the overlap percentages to eliminate the effect of random coincidence of masks.

Area	Masks	Wavelength limit		
		250 µm	500 µm	1000 µm
non-BT	Topo & Refl.	22.0 %	22.6 %	21.9 %
	TGV & Refl.	22.1 %	20.7 %	19.8 %
normal	Topo & Refl.	11.3 %	11.5 %	11.2 %
	TGV & Refl.	12.6 %	11.6 %	10.9 %

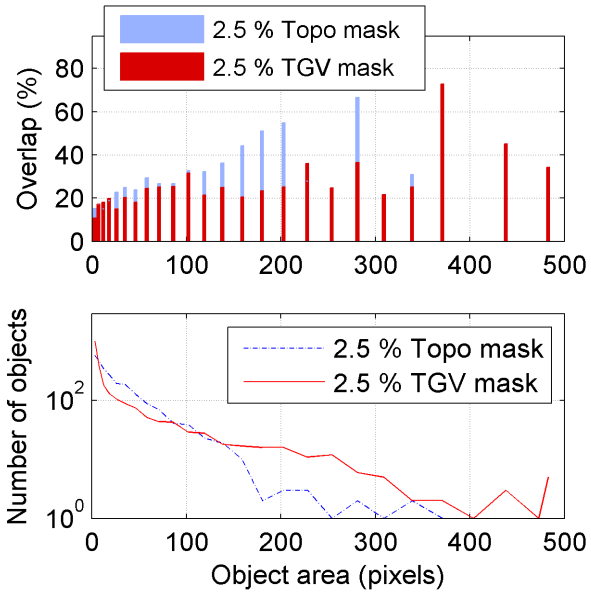

Fig 2: Top: Overlap of Topo mask objects and TGV mask objects with the reflectance mask in the non-back-trap sample. The results are reported with respect to the size of the objects found in the topography based masks. Bottom: Size distribution of the objects found in the topography based masks.

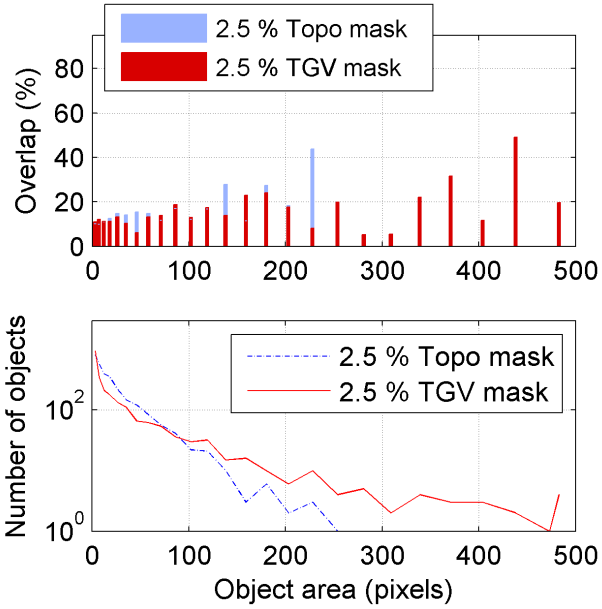

Fig 3: Overlap of Topo mask objects and TGV mask objects with the reflectance mask in the normally printed sample.

The coincidence of topography anomalies and missing ink is clearly more common in the non-back-trap area than in the normal cyan area, which is seen both in *Table 3* and in *Fig 2* and *Fig 3*. The effect of surface roughness on the ink coverage is attenuated by the successive impressions in the 4-color offset press. However, the normal printing conditions also produce significantly higher than random coincidence of anomalies. The TGV mask contains larger anomaly areas than the Topo mask because the gradient and variance typically reach extreme values at the edges of pits, and this enlarges the area marked as anomalous at the pit locations. *Fig 4* presents a close zoom of the surface topography and print reflectance maps of the non-back-trap cyan area together with 2.5 % masks. The anomalies detected from the topography map using the Topo and TGV techniques (*Table 2*) are shown in subfigures 4(c) and 4(e) by white, and the print defects are shown in 4(d) and 4(f) using both black and white. In 4(d) and 4(f) the white color is used to indicate the portion of print defects that overlap with the topography anomalies, and the black areas cover the rest of the print defects. As illustrated by the figures, both Topo and TGV masks predict the locations of missing ink with moderate success. This is encouraging particularly because the surface topography maps have been acquired before the sheets were printed. Thus the detection of anomalies from the topography of the unprinted paper provides information about print anomalies.

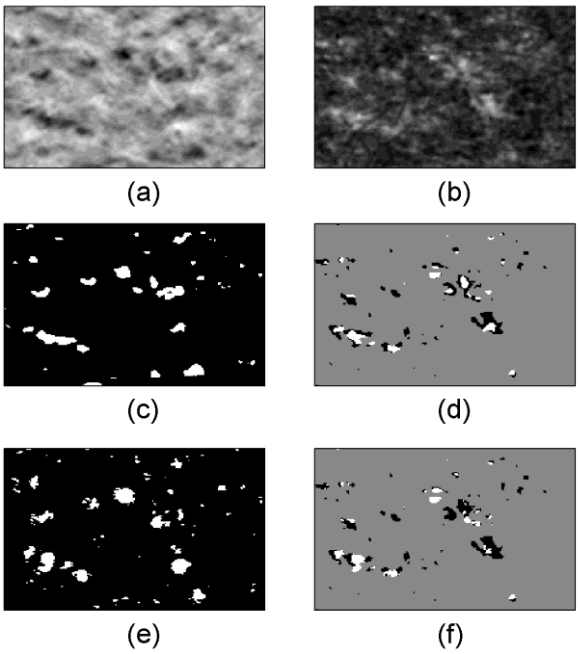

Fig 4: Series of 2.1 mm by 1.3 mm images from the non-back-trap cyan area. (a) Unprinted surface topography, high-pass filtered with wavelength limit 500 µm, (b) print reflectance, (c) topography anomaly map, (d) print anomaly map and its coincidence with the topography anomalies, (e) TGV anomaly map, (f) print anomaly map and its coincidence with the TGV anomalies.

The joint probability distributions of surface topography and print reflectance on the areas under examination are shown in *Fig 5*, corresponding to the algorithm step 3b. We illustrate the distributions by 2D histograms to give a detailed picture of the tails. The measurements made on the non-back-trap area show a clearly longer tail towards the high print reflectance values than the corresponding data from the normal cyan area. The tail is to some degree related with low surface topography values.

Fig 6 shows how many percent of the points that have surface height lower than the value given on the horizontal axis have print reflectance value higher than the 97.5 % limit computed from all the print reflectance values in the image. For example, selecting from the non-back-trap cyan area all the data points at which surface height is lower than -5 µm results in a data set in which 40 % of the points have print reflectance value higher than the

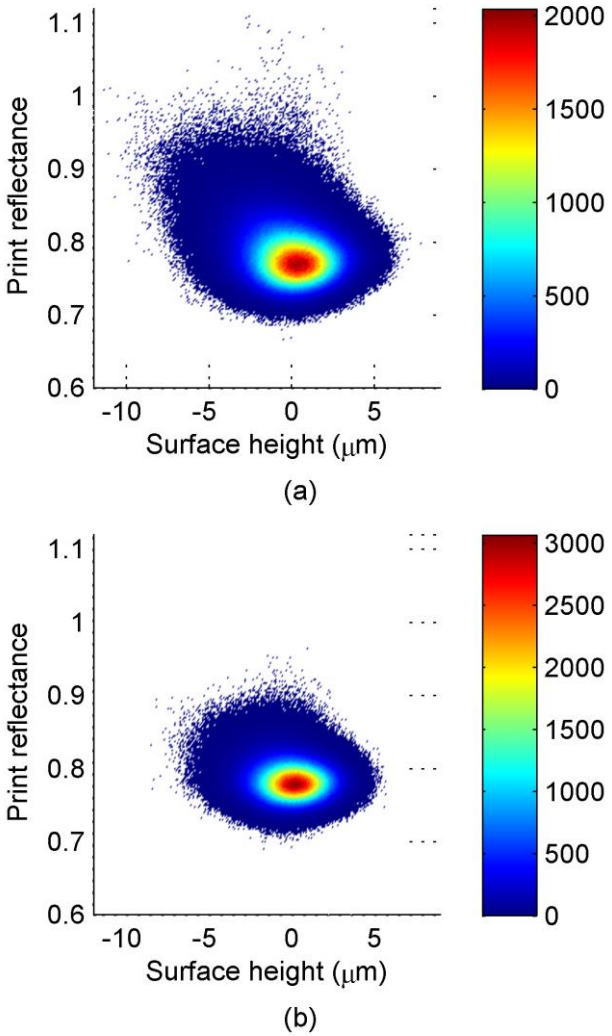

Fig 5: 2D joint histograms of unprinted surface topography and print reflectance in the newsprint case. (a) Non-back-trap cyan area, (b) normal cyan area.

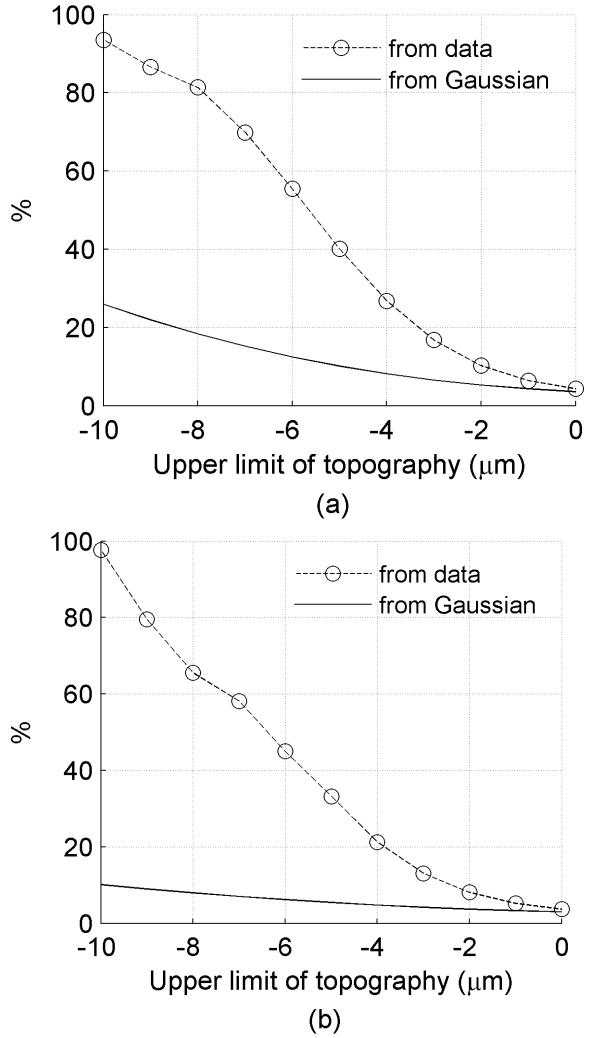

Fig 6: Tail analysis of the newsprint case with 97.5 % reflectance limit. (a) Non-back-trap cyan area, (b) normal cyan area.

97.5 % limit of all reflectance values (0.86 in this case). If the data were normally distributed, the corresponding portion of high reflectance values would be only 10 %, as indicated by the solid curve in *Fig 6*(a). The parameters of the reference Gaussian distribution have been determined from the data. The correlation coefficient between unprinted surface topography and print reflectance among the valid analysis pixels is -0.23 in the non-back-trap area and -0.10 in the normally printed area. The weaker correlation in the normal cyan explains why the tails of the Gaussian models in the two samples are different. If the correlation between the variables is strongly negative, the slope of the curve representing the Gaussian case is steep. Similar effect is caused by small variance of surface height values. In the samples presented here the variances are almost equal.

Fig 5 and *Fig 6* demonstrate that the tail of the joint distribution of surface topography and print reflectance is strongly skewed and deviates from Gaussian in both the non-back-trap sample and the normally printed sample. This implies that the mechanism by which the extreme topography values cause missing printing ink should not be modeled by a Gaussian stochastic process.

Finally, we evaluate the mutual information between print reflectance and the variables related to surface topography: the topography itself, or gradient or variance derived from the topography. The joint pdf of the multivariate observations within a selected mask is estimated using the Gaussian mixture model (GMM), and mutual information (MI) is estimated from this continuous pdf estimate. MI describes the statistical dependence between two or more variables, and it is always non-negative. The minimum value, zero, indicates that the variables are statistically independent of each other. MI results for the whole data set acquired from this offset printing trial have been previously reported in (Mettänen et al. 2009a). Using the two cyan areas presented in the current case example, we have increased the maximum mask percentage from the previously used 1.5 % to 5 %. The minimum mask percentage is 0.5 %. We have also studied how the high-pass (HP) filtering of the topography map affects the MI. Our tests with the varying wavelength limits of the HP filter have indicated that using larger than 250 μm limit is generally advantageous for the dependence measures between print reflectance and the filtered surface topography. Thus the size scale of the surface topography variations that explain local print density is not limited to 125 μm in diameter, but higher. This is in line with the scale considerations presented in the introduction. *Table 4* presents the MI results with three different mask percentages. The data points have been selected to the GMM fitting using either topography ('Topo') or TGV masks. The cut-off wavelength of topography is 500 μm.

The new results confirm the finding reported in (Mettänen et al. 2009a) that the strongest mutual information between print reflectance and surface topography is obtained by selecting the data points by the TGV mask. MI is also reasonably strong in Topo mask, and relatively small mask percentages lead to the highest MI in both topography based masks. This implies that a point in the measured surface topography map should be deemed anomalous only if the data value is really extreme. In addition, the supremacy of the TGV mask over

Table 4: Mutual information computed from GMMs.

Area	Mask	Mask percentage		
		0.5 %	2.5 %	5.0 %
non-BT	Topo	0.114 bit	0.126 bit	0.121 bit
	TGV	0.380 bit	0.243 bit	0.188 bit
normal	Topo	0.081 bit	0.065 bit	0.060 bit
	TGV	0.173 bit	0.094 bit	0.065 bit

the Topo mask means that, to select areas where the local print density is statistically dependent on the topography of unprinted paper, the local variations in topography should be observed and used alongside with the surface height measurements.

SC paper in laboratory gravure printing

The second case example presents two SC paper strips that were a part of a laboratory printing trial with the IGT gravure press (IGT 2003). The papers are from a calendering series and therefore the main difference between the studied strips is the roughness. We refer to the sample with PPS10 roughness of 1.24 µm as sample A, and the one with PPS10 roughness of 1.15 µm is sample B. Both have been printed with red Heliotest ink using 600 N printing force. Electrostatic assist (ESA) is not available in this laboratory test. We have imaged a 22.5 x 15 mm area of the conventional screening area of the Heliotest layout with 10 µm pixel resolution before and after printing to analyze the effect of surface topography on the print quality. Since the surface of SC paper is smooth, the topography maps are high-pass filtered with a 250 µm wavelength limit instead of the 500 µm that was used with the newsprint sheets.

Due to the screening, the regular pattern of the printed dots (which we call raster dots) has to be recognized from the image of print reflectance before the analysis. This enables the concentration of the analysis to the points that were intended to be covered by ink. We have implemented a set of robust image analysis methods for automatically detecting the regularity from the print reflectance image and locating with subpixel accuracy all the coordinates where raster dots are supposed to be situated (Mettänen et al. 2009b; Lauri, Ihalainen 2010). Some dots are partly or totally missing but still their coordinates are found. *Fig 7*(a) illustrates the print quality in the smoother sample B and shows the detected raster dot locations. *Fig 7*(b) presents the unprinted surface topography from the same zoomed area. Some fibers that have turned dark in the calendering are clearly visible in both the reflectance and surface topography maps. As dark shades in the topography map are meant to indicate depressions, the blackened fibers give erroneous topography observations. They are detected from the topography, reflectance and light transmittance images taken before printing. This information is combined with the detected raster pattern to define the valid analysis area. Consequently, a pixel is valid for analysis only if it lies within a radius of 5 pixels from a raster dot centre and if it does not present a part of a blackened fiber. The raster dots are approximately 13 pixels apart and thus the white

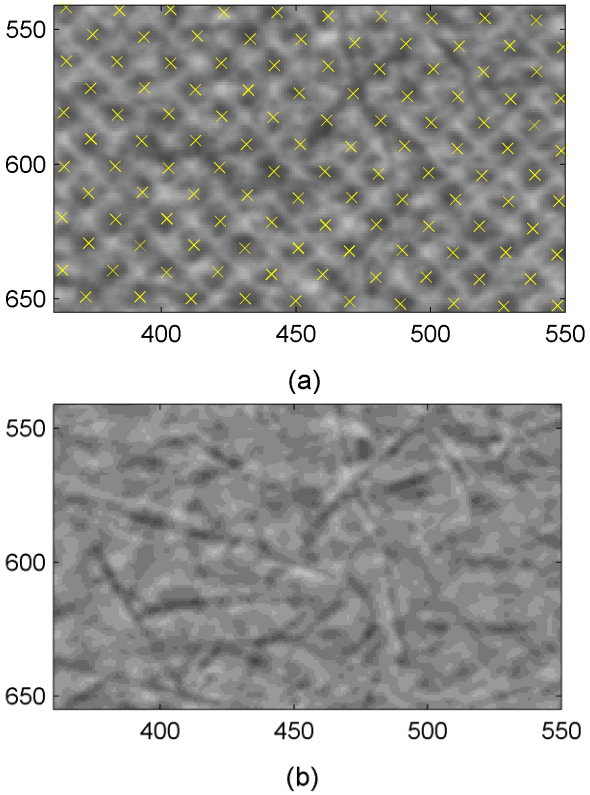

Fig 7: (a) Print reflectance of SC sample B with raster dot locations marked by yellow crosses. (b) Unprinted surface topography measurement from the same area. The units on the axes are pixels.

Table 5: Overlap of the map of missing ink (Refl. mask) and the map of deep surface depressions (Topo mask) in the studied SC samples A and B. The mask percentages have been subtracted from the overlap percentages of each column to eliminate the effect of random coincidence of masks.

Sample	Mask percentage				
	1.0 %	2.5 %	4.5 %	8.0 %	10.0 %
A	8.9 %	12.1 %	12.9 %	13.2 %	13.0 %
B	6.9 %	8.9 %	9.4 %	9.6 %	9.5 %

grid between the raster dots, visible in *Fig 7*(a), is not included in the valid analysis area. The final number of pixels on the valid area of both samples A and B is approximately 0.85 million.

The overlap of print anomalies and surface depressions is not as strong in the SC paper strips as in the previously analyzed newsprint case example. The rougher sample (A) gives higher overlap results than the smoother one (B), as expected. *Table 5* reports the overlap results. The correlation coefficient between surface heights and print reflectance in the valid analysis area is -0.23 in sample A and -0.18 in sample B.

Tail analysis results (cf. *Fig 6*) for the SC samples are given in *Table 6*. It shows that, for instance, of all the points of sample A where surface height is lower than -4 µm, 89 % have a high print reflectance value. The corresponding portion of high

Table 6: Tail analysis results for the SC samples A and B with 97.5 % reflectance limit. The 'Tail' column indicates whether the percentage is based on the true histogram of the data (Hist.) or on the Gaussian approximation (Gauss.).

	Tail	Topography limit					
		-6µm	-5µm	-4µm	-3µm	-2µm	-1µm
A	Hist.	100 %	100 %	89 %	51 %	20 %	8 %
	Gauss.	40 %	30 %	21 %	14 %	9 %	6 %
B	Hist.	100%	94 %	69 %	36 %	15 %	7 %
	Gauss.	28 %	21 %	15 %	11 %	7 %	5 %

reflectance values would be 21 % if the data were normally distributed. Even though the SC paper samples are smooth and present moderate print quality, the results clearly show that the joint distribution of surface topography and print reflectance is non-Gaussian, and depressions on the paper surface cause significantly increased probability of exceptionally high print reflectance.

Finally, we use 1D histograms to visualize the distribution of print reflectance values in points where topography is anomalous, and similarly we visualize topography values in points where ink is missing. These distributions are shown in *Fig 8* and *Fig 9*, using 1.5 % masks. The comparison with the overall distribution (shown by the black curves) has been implemented by selecting the same number of data points, 0.015*850000 = 12750, with a mask of randomly located points. If all pixels from the valid analysis area were selected, the curve would be smoother, but now it is better comparable with the distributions in the topography and reflectance based masks.

Whereas the tail analysis presented for the SC samples in *Table 6* served as a numerical quantification of the joint distribution tails at low topography values, *Fig 8* provides the whole marginal and conditional print reflectance distributions. They help to visualize the dependence of print reflectance on the selected degree of abnormality of surface height values. This example applies 1.5 % Topo masks, resulting in topography threshold values of -1.91 µm and -1.87 µm in samples A and B, respectively. As indicated already by *Table 6*, the probability of finding missing ink within the topography based mask points is higher in the rougher sample A than in sample B. However, *Fig 8* also shows that the variance of reflectance values depends on the mask. The largest variance occurs among the data points selected by the Topo mask.

The marginal distributions of surface topography values, shown by the black curves in *Fig 9*, are rather similar in samples A and B. Both have a slight tail towards the negative values due to the expected log-normal distribution of pore heights (Niskanen et al. 1998). The conditional distributions

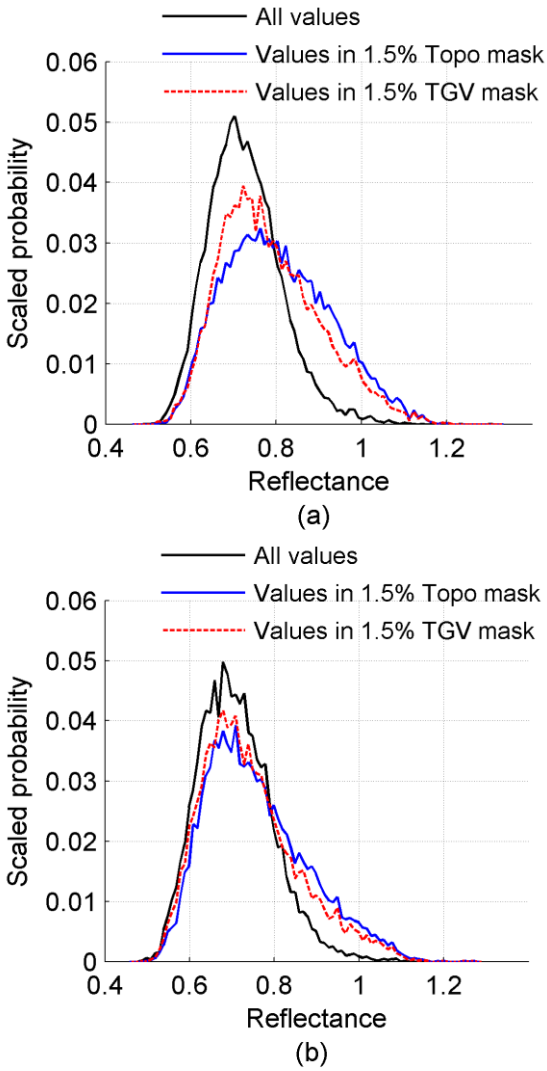

Fig 8: Distribution of print reflectance values in (a) sample A and (b) sample B.

of topography values, given that ink is missing, are also very similar in samples A and B. Surface height is, on average, lower in the print defect points than in randomly selected points, but the roughness difference between samples A and B cannot notably affect the conditional distributions in *Fig 9*, because the surface topography at the missing ink spots does not depend on the overall distribution of surface topography.

Discussion of the properties of the analysis methods

As our focus has been on the small-scale properties of paper and print, the areas imaged and analyzed in this work are rather small (22.5 x 15 mm). In both the newsprint and SC cases they have been captured from a larger area that consists of the same printed pattern. This has enabled the visual assessment of the print quality of the studied samples. Although informal, the visual assessment has confirmed what was expected: the normal offset conditions have

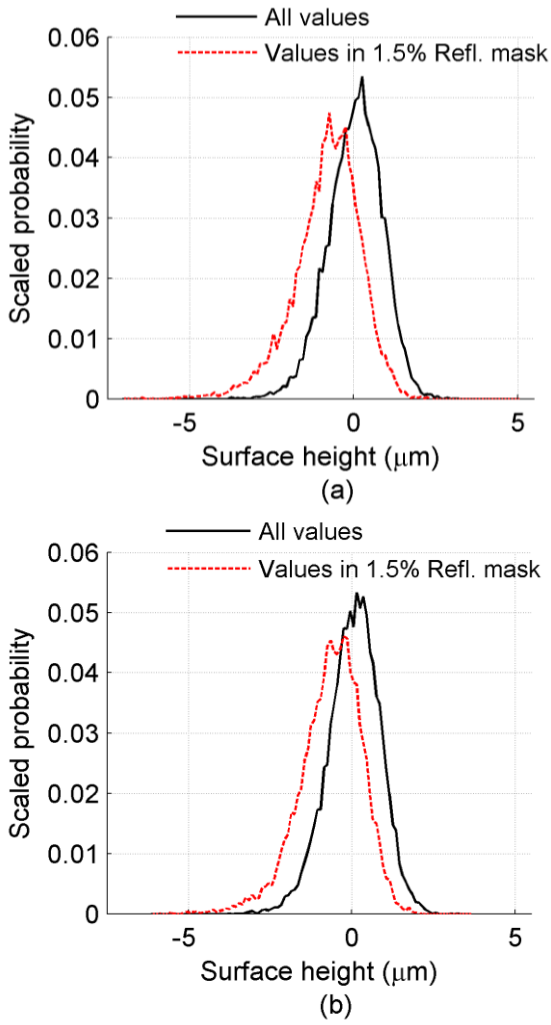

Fig 9: Distribution of surface height values in (a) sample A and (b) sample B.

produced more satisfying print quality on the newsprint than the non-back-trap conditions, and the smoother SC sample is of higher print quality than the rougher one. The small size of the examined area is not problematic in cases where the printing process can be considered sufficiently stable and thus the small printed region represents the average behavior of the whole print.

Mottling has been measured by a scanner system from the newsprint sheets printed in offset – not exactly from the areas imaged and analyzed here, but from parallel sheets of the same paper. According to the mottling measurement, print density variation is consistently higher in the non-back-trap area than in the normal printing area in the studied wavelength range, 0.35 mm to 20 mm. This must be connected with the more severe occurrence of missing ink in the former area. Our analysis showed that the dependence between surface topography and print reflectance is stronger in the non-back-trap area than in the normal area. Clearly, surface topography is not the key property determining offset print quality in either case. However, even the normal printing conditions left some spots of the compact print layout unprinted and these could be partly connected to deep pits on the paper surface. A prior contribution of Saarelma and Oittinen (1993) also supports the conclusion that spots of missing ink, even when they are small enough to correspond to micro scale noise (wavelengths from 10 µm to 100 µm), can cause print unevenness that is visually harmful.

The SC paper case with gravure printing is perhaps more relevant for commercial applications than the newsprint case. The determination of the raster dot locations has been a success despite the irregular shape of the dots and occasionally severe absence of ink. The analysis has shown that points of low surface height, measured from the unprinted paper, present high print reflectance values with clearly higher probability than random chance. It has been reported earlier (Mettänen et al. 2009b) that this is particularly the case with high pressing forces in the printing nip. Although the higher compression improves the visual print quality in the IGT test, some depressions on the surface of SC paper are so "hard" that they prevent perfect contact of the paper with the printing cylinder despite the pressure. It can be expected that a map of dynamic contact between the paper and the printing nip (e.g., (Preston et al. 2008)) would yield additional value in the analysis of missing ink in gravure. Further analysis of the currently available static topography maps may also shed light on the role of surface gradients and variance in the occurrence of missing ink in gravure.

The case examples indicated that the steps of the proposed analysis algorithm can be applied flexibly according to the goals of the study. It is not always meaningful to complete each step of the analysis algorithm. We emphasize the importance of visualizing the measurement data, anomaly maps and pdfs, in order to understand the measured phenomena. However, numerical quantification of the dependences is necessary to summarize the results from large data sets. The percentage overlap of print defects and surface topography anomalies is the most straightforward measure for this purpose. The tail analysis, as presented in *Fig 6* and *Table 6*, provides a more detailed expression of the skewness of the conditional print reflectance distributions. The interpretation of mutual information (MI) as a dependence measure may be less intuitive. MI is expressed in 'bits' or 'nats' and it describes statistical dependences without restricting the analysis to the Gaussian statistics as the standard correlation analyses does. In practice the MI between two (or more) variables often increases along with the correlation coefficient. The added value in MI is then in the comparison of the MI

result obtained from the non-Gaussian pdf model with that obtained from a Gaussian approximation. This gives a measure of the nonlinearity of the dependence. The tails of the distributions typically contain nonlinear statistical dependences between the variables. We have also presented a simulation procedure for confirming the statistical significance of the MI results computed from multivariate data (Mettänen et al. 2009a).

It is worth noting that the probabilistic analysis setup, particularly when concentrating on the tails of the distributions, is best applicable to large data sets. However, it is not limited to surface topography and local print density measurements. A map of the local variations of print gloss is also relevant for print quality analysis (MacGregor et al. 1994), and the measurements of paper properties could include, for instance, local compressibility (Mangin et al. 1993), dynamic contact pattern (Preston et al. 2008), apparent density (Sung et al. 2005) or beta formation measurement (Kajanto 1991). Although the description of the analysis methods assumed multivariate image data, this data format is not an absolute requirement. A sufficient number of data points might also be obtainable from line scans on a running paper machine or printing press. We expect our tools to find use in print quality analysis and troubleshooting tasks within modern printing technologies where delicate small-scale analysis with statistically meaningful results is required.

Conclusions

We have presented probabilistic methods for assessing the information that the small-scale measurement of unprinted paper properties can carry about the printability of the paper. Image based measurements are natural means to acquire data for such analysis. When failing to explain, for instance, local print density variations, through the averages or other compact parameters of paper structural properties, and explaining the dependences through pixel by pixel regression does not work either, the large number of independent observations in each measured 2D map is best utilized by characterizing the joint probability distributions of the measured properties. We believe that the analysis of the distributions and particularly the analysis of their tails can provide valuable information about the print and paper structure, and reveal causes of print defects. The essential requirement regarding the data is the accurate alignment of the multivariate measurements.

So far we have used aligned images of the print and surface topography to develop our analysis procedures. We have demonstrated the steps of our print assessment algorithm by two case examples, first with offset printed newsprint and then with gravure printed SC paper. The joint probability density function of unprinted surface topography and print reflectance has in both cases shown clear differences between the samples of moderate and imperfect print quality. Conditional pdfs derived from the joint pdf have successfully characterized the interdependence of the variables in points and areas where the surface topography of the unprinted paper shows anomalous behavior. Visualizations have been supported by quantitative results, confirming that depressions on the paper surface cause significantly increased probability of exceptionally high print reflectance. The overlap of these print defects with topography anomalies is as high as 22 % in the weaker quality offset sample, and clearly higher than the probability of random coincidence in all cases. With the offset printed samples we have also been able to show that the local gradient and variance maps of topography can provide useful information for explaining the print density variations.

In future work, our attention will be directed to halftone printing. We have two research ideas that are based on utilizing the coordinates of the printed dots. First, we are interested in quantifying the dependence of mottling on the variation of dot gain between small local areas of the paper sheet. The second idea concerns the analysis of missing dots in gravure printed products. Based on the coordinates of the raster dots, information about the depths of the engraved cells on the printing cylinder can be included in the analysis. This allows the separation of two important factors that affect the quality of printed dots: the amount of ink feed and the surface structure of the paper. A simple test case for this idea is a heliotest strip where the decrease of the ink cell depths towards the end of the test strip is standardized.

Acknowledgements

The author wishes to thank the Forest Pilot Center Oy (FPC) and UPM Oyj, for supporting the data acquisition. The co-workers of the "PapSurface" project are gratefully acknowledged. The author would also like to express her gratitude to Professor Risto Ritala for supervising this work.

Literature

Barros, G.G. (2006): Influence of substrate topography on ink distribution in flexography, Ph.D. Thesis, Karlstad University, Karlstad, Sweden.

Barros, G.G. and Johansson, P.-Å. (2006): Prediction of UnCovered Area occurrence in flexography based on

topography - A feasibility study, Nord. Pulp Paper Res. J. 21(2), 172.

Bishop, C.M. (2006): Pattern recognition and machine learning, Springer, Singapore, pp. 48-58.

Chinga, G., Johnsen, P.O., Dougherty, R., Berli, E.L. and Walters, J. (2007): Quantification of the 3D microstructure of SC surfaces, J. Microscopy 277(3), 254.

Dickson, A.R. (2006): Evaluating the relationship between grammage, topography and print properties in newsprint, Proc. 60th Appita Annual Conf., Melbourne, Australia, April 3-5, 2006, Appita Inc., pp. 19-23.

Eerola, T., Kämäräinen, J.-K., Leisti, T., Halonen, R., Lensu, L., Kälviäinen, H., Oittinen, P. and Nyman, G. (2008): Finding best measurable quantities for predicting human visual quality experience, IEEE Int. Conf. Systems, Man, and Cybernetics, Singapore, October 12-15, 2008, IEEE, pp. 733-738.

Hansson, P. and Johansson, P.-Å. (2000): Topography and reflectance analysis of paper surfaces using a photometric stereo method, Optical Engineering 39 (9), 2555.

Heintze, H.U. and Gordon, R.W. (1979): Tuning of the GRI proof press as a predictor of rotonews print quality in the pressroom, Tappi 62(11), 97.

IGT Testing Systems (2003): IGT Information leaflet W41, Amsterdam (available at: www.igt.nl).

Järnström, J., Ihalainen, P., Backfolk, K. and Peltonen, J. (2008): Roughness of pigment coatings and its influence on gloss, Applied Surface Sci. 254(18), 5741.

Järnström, J., Ihalainen, P., Lemström, A., Toivakka, M. and Peltonen, J. (2009): The influence of different roughness scales of pigment coated papers on print gloss, Nord. Pulp Paper Res. J. 24(3), 327.

Järnström, J., Sinervo, L., Toivakka, M. and Peltonen, J. (2007): Topography and gloss of precipitated calcium carbonate coating layers on a model substrate, Tappi J. 6(5), 23.

Kajanto, I.M. (1989): The effect of formation on print quality with woodfree offset papers, Nord. Pulp Paper Res. J. 4(1), 8.

Kajanto, I.M. (1991): Pointwise dependence between local grammage and local print density, Paperi Puu 73(4), 338.

Kajanto, I., Laamanen, J. and Kainulainen, M. (1998): Paper bulk and surface, In: Niskanen, K. (ed.), Paper Physics, Papermaking Science and Technology, Book 16, Fapet Oy, Jyväskylä, Finland, pp. 89-115.

Karttunen, S.T.P. (1973): Structure and behaviour of a paper's surface in printing, In: Bolam, F. (ed.),Trans. 5th Fund. Res. Symp., The fundamental properties of paper related to its uses, Cambridge, UK, September 1973, BPBIF, London, pp. 544-560.

Lauri, M. and Ihalainen, H. (2010): Measuring periodic patterns in noisy spatial data, to be published in ENBIS - IMEKO TC21 Workshop on Measurement Systems and Process Improvement, Teddington, UK, April 19-21, 2010.

Lyne, B. (1976a): Measurement of the distribution of surface void sizes in paper, Tappi 59(7), 102.

Lyne, B. (1976b): Machine-calendering: implications for surface structure and printing properties, Tappi 59(8), 101.

MacGregor, M.A., Johansson, P.-Å. and Béland, M.-C. (1994): Measurement of small-scale gloss variation in printed paper: topography explains much of the variation for one paper, TAPPI/CPPA Int. Printing and Graphic Arts Conf., Halifax, Nova Scotia, Canada, October 18-20, 1994, Canadian Pulp & Paper Association, pp. 33-43.

Mangin, P.J., Béland, M.-C. and Cormier, L.M. (1993): A structural approach to paper surface compressibility - relationship with printing characteristics, In: Baker, C.F. (ed.), Trans. 10th Fund. Res. Symp., Products of papermaking, Oxford, UK, September 1993, FRC / Pira International, Leatherhead, pp. 1397-1427.

Mangin, P.J. and Dubé, M. (2006): Fundamental questions on print quality, In: Cui, L.C. and Miyake, Y. (ed.), Proc. of SPIE-IS&T Electronic Imaging Conf., Image Qual. and System Perform. III, San Jose, California, USA, January 17-19, 2006, SPIE/IS&T, pp. 605901-1-605901-12.

Mesic, B., Lestelius, M. and Engström, G. (2006): Occurrence and causes of uncovered areas in water-borne flexographic print on PE-extrusion-coated packaging papers, Tappi Int. Printing and Graphic Arts Conf., Cincinnati, Ohio, USA, September 20-22, 2006.

Metsäjoki, T. (2005): Investigation of correlation between surface structure and print defects in rotogravure printing of SC paper, M.Sc. Thesis, Tampere University of Technology, Tampere, Finland.

Mettänen, M., Hirn, U., Lauri, M. and Ritala, R. (2009a): Probabilistic analysis of small-scale print defects with aligned 2D measurements, In: l'Anson, S.J. (ed.), Trans. 14th Fund. Res. Symp., Advances in pulp and paper research, Oxford, UK, September 2009, FRC, pp. 1293-1323.

Mettänen, M., Ihalainen, H. and Ritala, R. (2008): Alignment and statistical analysis of 2D small-scale paper property maps, Appita J. 61(4), 323.

Mettänen, M., Lauri, M., Ihalainen, H., Kumpulainen, P. and Ritala, R. (2009b): Aligned analysis of surface topography and printed dot pattern maps, Papermaking Res. Symp., Kuopio, Finland, June 1-4, 2009.

Nabney, I.T. (2002): Netlab: Algorithms for pattern recognition, Springer, London, pp. 79-113.

Niskanen, K., Kajanto, I. and Pakarinen, P. (1998): Paper structure, In: Niskanen, K. (ed.), Paper Physics, Papermaking Science and Technology, Book 16, Fapet Oy, Jyväskylä, Finland, pp. 14-53.

Oittinen, P., Halonen, R., Kokkonen, A., Leisti, T., Nyman, G., Eerola, T., Lensu, L., Kälviäinen, H., Ritala, R., Pulla, J. and Mettänen, M. (2008): Framework for modeling visual printed image quality from the paper perspective, In: Farnand, S.P. and Gaykema, F. (ed.), Proc. SPIE-IS&T Electronic Imaging Conf., Image Qual. and System Perform. V, San Jose, California, USA, January 27-31, 2008, SPIE/IS&T, pp. 68080L-1-68080-12.

Papoulis, A. (1990): Probability & Statistics, Prentice-Hall, Englewood Cliffs, NJ, p. 136.

Parker, J.R. (1973): Fundamental paper properties in relation to printability, In: Bolam, F. (ed.),Trans. 5th Fund. Res. Symp., The fundamental properties of paper related to its uses, Cambridge, UK, September 1973, BPBIF, London, pp. 517-543.

Poon, C.Y. and Bhushan, B. (1995): Comparison of surface roughness measurements by stylus profiler, AFM and non-contact optical profiler, Wear 190(1), 76.

Preston, J., Bollström, R., Nylander, K., Toivakka, M. and Peltonen, J. (2008): A new test method to measure dynamic contact in a printing nip for rotogravure papers, 35th Int. Res. Conf. of IARIGAI, Advances in printing and media technology, Valencia, Spain, September 7-10, 2008, pp. 147-153.

Saarelma, H.J. and Oittinen, P.T. (1993): Paper and print noise as limiting factors of information capacity. In: Baker, C.F. (ed.), Trans. 10th Fund. Res. Symp., Products of papermaking, Oxford, UK, September 1993, FRC / Pira International, Leatherhead, pp. 351-365.

Stout, K.J. and Blunt, L. (2000): Three-dimensional surface topography (2nd ed.), Penton Press, London, pp. 157-159.

Sung, Y.J., Ham, C.H., Kwon, O., Lee, H.L. and Keller, D.S. (2005): Applications of thickness and apparent density mapping by laser profilometry, In: l'Anson, S.J. (ed.), Trans. 13th Fund. Res. Symp., Advances in paper science and technology, Cambridge, UK, September 2005, FRC, pp. 961-1007.

Van den Akker, J.A. (1982): Optical properties of paper, In: Rance, H.F. (ed.), Handbook of paper science, Vol. 2, The structural and physical properties of paper, Elsevier, Amsterdam, pp. 127-139.

Walker, W.C. and Fetsko, J.M. (1955): A concept of ink transfer in printing, American Ink Maker 33(12), 38.