A LOW-COMPLEXITY REDUCED-REFERENCE PRINT IDENTIFICATION ALGORITHM

Matthew D. Gaubatz, Robert A. Ulichney

Hewlett-Packard, Co.
matthew.gaubatz@hp.com, u@hp.com

David M. Rouse

Cornell University
dmr58@cornell.edu

ABSTRACT

In a print production system, the ability to match a printed document with its original electronic form enables services that improve robustness of the production process, such as content tracking mechanisms, or highly targeted quality assurance checks. One approach to this problem is to use overt markings that are later removed. This work, however, adopts a method that uses no markings relying instead on fast image matching criteria. Though general image matching can be a difficult problem, the nature of this environment allows constraints to be imposed that yield simple, efficient solutions to the matching problem. A reduced-reference quality assessment algorithm is presented that is used to match every document in a collection of printed and scanned documents with its original electronic form, based on down-sampling, equalization, quantization and per-pixel comparison. Using a 1000 image test set, a 100 % matching rate is achievable by representing each image with only 5 bits-per-pixel at 8 dots-per-inch, which compares favorably with the performance achieved by other current reduced-reference quality assessment algorithms in terms of accuracy, speed and storage.

***Index Terms*—** Print production, quality assurance, page identification, reduced-reference assessment.

1. INTRODUCTION

This paper examines a constrained whole-image matching problem, one of associating each image from a *set* of query images to a *unique* image in from a set of reference images. One motivation is that in a print production environment, any time between when a page is printed and when it is part of a finished product, it is useful to be able to associate a printed page with its original electronic form. This functionality facilitates job tracking, page-specific quality assurance measurements, detailed error reporting, and other quality-of-service tools on any device with a scanning mechanism. It also has applications areas such as security, recognition or electronic indexing. One approach to this problem is to use overt markings that are later removed, but this work adopts a method without markings that relies on fast matching criteria.

Research in image matching generally focuses on the problem of associating a query image with one image from a set of reference images by comparing descriptors computed from the images. Examples that have proven effective include the scale-invariant feature transform (SIFT) [1], shape context [2], spin images [3], gradient location and orientation histogram [4], complex filters [5], Hausdorff distance [6], color scheme and layout [7], and even cross-correlation of sampled pixel values [8]. Many of these approaches are for more difficult problems (such as matching images captured from different perspectives), and thus have not been optimized for efficiency. *Reduced-reference* quality assessment focuses on how to efficiently store descriptors of reference images for comparison purposes [9–12]. Many such methods are computationally intensive and/or require training.

This paper proposes using a reduced-reference approach consisting of down-sampling, equalizing and quantizing images prior to comparison, where the reduced-reference descriptors are essentially thumbnails. Mean-squared-error (MSE) is used to compare reference (original) and query (printed+scanned) descriptors, and match images between these two sets. Analysis suggests that this approach may be tuned, that is, down-sampling, equalization and quantization parameters may be selected, to yield an effective trade-off between computation time, storage required and accuracy achieved. Simulation with 1000 test images consisting of text, photographic and composite data corroborates this result. An optimized version of the matching metric that achieves 98 % is implementable with 80 bits per image and a simple XOR comparison operation.

This paper is organized as follows. Section 2 summarizes the page-matching problem, and proposes an algorithm for a solution. The algorithm is analyzed in section 3, and comparisons are given in section 4. Section 5 concludes the paper.

2. PRINT IDENTIFICATION PROBLEM DESCRIPTION

This section introduces notation used to describe the print identification problem, and proposes a solution as well.

2.1. Notation

Let $k \in [1, K]$ index a sequence of $N \times M$ digital images, and let $\mathbf{I}^k(n, m)$ represent the k^{th} image in the sequence. Assume

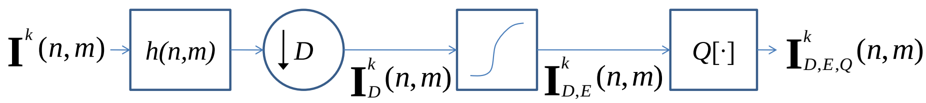

Fig. 1. System diagram, illustrating the relationship between $\mathbf{I}^k(n,m)$, $\mathbf{I}^k_D(n,m)$, $\mathbf{I}^k_{D,E}(n,m)$ and $\mathbf{I}^k_{D,E,Q}(n,m)$.

that each image is printed then scanned, but not necessarily in the same order. The function $\pi(k) : [1, K] \rightarrow [1, K]$ denotes a permutation of $[1, 2, ..., K]$ that represents the order in which the images are printed and scanned. Let $\mathbf{I}^k[n, m]$ represent the sequence of scanned images. Note that $\hat{\mathbf{I}}^k(n, m)$ is a printed and scanned version of $\mathbf{I}^{\pi(k)}(n, m)$. The goal of the print identification problem is to estimate the quantity $\pi(k)$. This estimate is given by $\hat{\pi}(k, \{\mathbf{I}^k[n, m]\}, \{\hat{\mathbf{I}}^k[n, m]\})$, since it is calculated as a function of image sequences $\{\mathbf{I}^k[n, m]\}$ and $\{\hat{\mathbf{I}}^k[n, m]\}$. (This term is abbreviated as $\hat{\pi}(k)$ for convenience.) The success of any estimation procedure is measured with error fraction, that is, the number of entries of $\hat{\pi}(k)$ that match entries in $\pi(k)$ exactly:

$$\text{error fraction} = \frac{|\{k|\pi(k) = \hat{\pi}(k)\}|}{K}. \tag{1}$$

2.2. Proposed Solution

A straightforward solution to the problem is to essentially encode the index k onto each page, and to decode this index when needed. This approach is fairly simple, but either requires modification of the image printed, or removal of the marking encoding the index at a later stage. Therefore, a method is adopted which estimates the index without requiring and additional infor5ion. The chosen approach is to develop an image assessment algorithm $f(\cdot, \cdot)$ such that

$$\hat{\pi}(k) = \underset{j}{argmin}\, f(\hat{\mathbf{I}}^k[n, m], \mathbf{I}^j[n, m]). \tag{2}$$

Many algorithms have been designed for a similar purpose, and some have even been designed with computational efficiency as a consideration [13, 14]. In order to reduce the amount of information required from $\{\mathbf{I}^k[n, m]\}$ for this process, a reduced-reference assessment algorithm is used. It has been previously shown than a number of general quality assessment algorithms are robust to changes in image size [15]. Furthermore, the constraints associated with the production environment, i.e., where pages are stored in stacks (resulting in stable lighting conditions and minimal perspective differences), increase the possibility of successful low-complexity algorithms. Reduced-reference descriptors used herein are formed by down-sampling, equalizing, and (uniformly) quantizing the original image. Subscripts D, E and Q represent, respectively, the application of these "reference-reducing" operators; combinations of subscripts separated by commas denote the application of multiple operators. Equations describing their behavior are given as follows:

$$\mathbf{I}^k_D(n, m) = \sum_{i,j} h(i, j) \cdot \mathbf{I}^k(\frac{n}{D} - i, \frac{m}{D} - j), \tag{3}$$

$$\mathbf{I}^k_E(n, m) = 255 \cdot \sum_{p=0}^{\mathbf{I}^k(n,m)} \Pr_{i,j}\{\mathbf{I}^k(i, j) == p\}, \tag{4}$$

$$\mathbf{I}^k_Q(n, m) = Q \cdot \lfloor \frac{\mathbf{I}^k(n, m) + 0.5}{Q} \rfloor. \tag{5}$$

Since a reduced-reference solution is desired, it is important to limit the amount of information involved in estimating $\hat{\pi}(k)$ as early as possible in the process. The proposed solution is to perform down-sampling, equalization, quantization and comparison (DEQC) of candidate images, that is, to solve (2) with a function of the form

$$\begin{aligned} & f_{D,E,Q}(\hat{\mathbf{I}}^k[n, m], \mathbf{I}^j[n, m]) \\ & \quad = \quad f(\hat{\mathbf{I}}^k_{D,E,Q}(n, m), \mathbf{I}^j_{D,E,Q}(n, m)). \end{aligned} \tag{6}$$

The relationship between original and compared images is illustrated in Figure 1. Success of this approach is based on how the re-sampling rate D (as well as the method), the type of equalization selected, and the quantization step-size Q, and the comparison function $f(\cdot, \cdot)$. Due to the ubiquity and the simplicity of the mean-squared-error (MSE) criterion, this metric is selected for $f(\cdot, \cdot)$.

3. EFFECTS OF REFERENCE-REDUCING OPERATORS

The proposed method was evaluated using a large test set consisting of 1085 images. Image content consisted of pages from documents such as journal papers, catalogs, event programs, and contained text, photographic data, and compositions of both. Each page was printed at full resolution, then scanned at 200 dots-per-inch (DPI), with 8 bits to represent each sample, that is at 8 bits-per-pixel (bpp). This section illustrates the trade-offs in error fraction achieved by the proposed method on this data set, due to different amounts of resolution in space versus resolution in sample value.

Of the three operator mentioned in the previous section, only two (down-sampling and quantization) explicitly reduce the storage cost of the associated descriptors. The purpose of this section is to illustrate the cost of these reference-reducing operators. Accordingly, the following reported results were

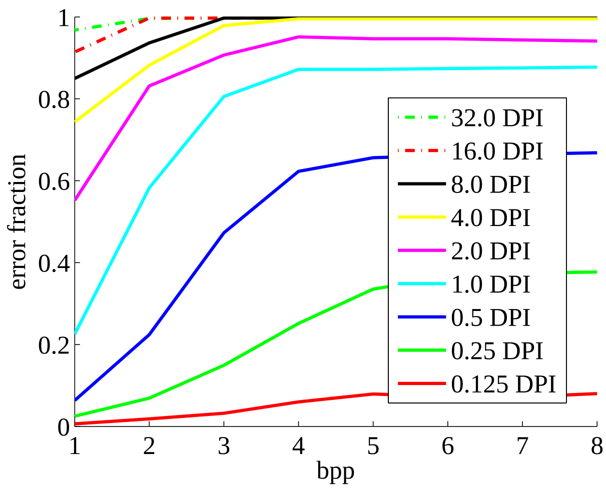

Fig. 2. Trade-off achieved between error fraction and bit depth in bits-per-pixel (bpp). Notice that the error fraction is the same for the last 3 numbers of significant bits, i.e., that the error fractions achieved for data quantized to 6, 7, or 8 bit of precision are virtually identical.

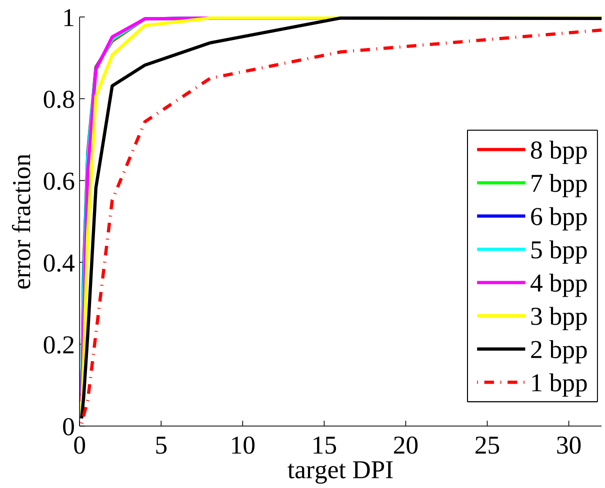

Fig. 3. Trade-off achieved between error fraction and DPI. Each line corresponds to a different bit depth. Note that the results for bit depths 5-8 are almost identical.

generated *without* applying any equalization. Testing reveals that the proposed method is capable of achieving perfect matching using images re-sampled to 8 DPI. Figure 2 illustrates performance as a function of bpp, and Figure 3 gives performance as a function of target resolution. Several trends are immediately obvious from these plots. First, performance is not influenced by the two least significant bits of precision. (The plots in Figure 2 are relatively flat for all bit depths greater than or equal to 6; similarly, the plots associated with the top three bit-depths are nearly indistinguishable in Figure 3.)

Another trend that becomes obvious is that resolution in space is more important than resolution in value. In other words, with respect to error fraction, there is (almost) always a benefit associated with increasing the spatial resolution (DPI) of a descriptor, but not necessarily with increasing the sample resolution (bpp). This result implies that the optimal error fraction may be determined by picking the appropriate spatial resolution at a low bit depth. Aspects of the performance improvements achieved due to equalization are discussed in the next section.

4. COMPARATIVE RESULTS AND DISCUSSION

This section compares results achieved using the proposed method with those due to state-of-the art reduced-reference quality assessment algorithms. One previously proposed approach is based on a steerable pyramid decomposition, and represents each reference image with only 162 bits [10]. The advantage of this approach is that is uses a very small number of bits to represent an image, but one disadvantage was that it was not designed (explicitly) for performance in the severe distortion regime. This particular algorithm was chosen because it has demonstrated strong performance as a general purpose reduced-reference assessment algorithm. Another previously proposed approach is computationally simpler, and essentially only computes the mean-square-error between LL bands of a wavelet decomposition [11]. This approach is more easily tunable (since any LL band from a wavelet decomposition can be used), that is, the amount of storage required to implement this approach can be manipulated in different ways. The same data selected to evaluate different properties of the proposed algorithm (in section 3) was used for the basis of this comparison.

The error fractions achieved using the different test algorithms are listed in Table 1. Notice that when an equivalent number of bits are used to represent each image, the proposed method outperforms the other tested approaches. In terms of speed, the proposed method and the one in [11] are roughly the same; the difference in performance, due to the equalization step, however, is significant, i.e., it can result in a 10-30 percent improvement. The comparison with [10] is unfair in a sense, since it is applied to images at 32 DPI instead of at native resolution (200 DPI). On the other hand, it should be noted that applying a steerable pyramid decomposition at native resolution is completely unrealistic for some real-time identification systems (though the final version of this paper will incoporate this comparison anyway).

One benefit of the proposed approach, when utilized in conjunction with data at 1 bpp, is that it can be implemented with an XOR operation instead of a summation (consider cal-

assessment algorithm	bit depth (bpp)	resolution (DPI)	storage required (total bits)	accuracy (error fraction)	computational complexity
proposed	2	2.0	640	0.993	equalization
proposed	2	1.0	160	0.990	equalization
proposed	1	1.0	80	0.980	equalization
[11]	8	1.0	640	0.878	2-D DWT
[11]	8	0.5	160	0.668	2-D DWT
[11]	1	1.0	80	0.226	2-D DWT
[10]	8	32	162	0.185	steerable pyramid

Table 1. Comparison between performance achieved with the proposed method and other reduced-reference assessment algorithms. Notice that at equivalent bit rates, the proposed method outperforms all tested alternatives.

culations the MSE of a one bit array). In addition, for increased efficiency, the equalization procedure can be incorporated into the quantization stage, if the reduction in bit-depth is achieved by choosing whether or not each candidate pixel value is above or below the pixel mean. Similar savings can be achieved when representing data with other low bit-depths.

5. CONCLUSIONS AND FUTURE WORK

This work presents a fast, storage efficient method of performing print identification. The trade-off between resolution in space and in sample value is examined via extensive simulation. Most reduced-reference assessment algorithms are designed to detect large or small differences between a reference image and a distorted image. It has already been shown that, for a number of quality assessment algorithms, the resolution at which quality is assessed can be decreased with only a minimal effect on performance [15]. The results herein suggest that in the case of extreme distortion, that is, when an assessment algorithm is used to identify different images, the bit depth of the image can be reduced as well. The proposed method compares favorably with other previously proposed reduced-reference general quality assessment algorithms. Future work will involve development of more sophisticated performance prediction models, application of bit-depth reduction to general quality assessment, and comparisons with more reduced-reference assessment algorithms.

6. REFERENCES

[1] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[2] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 4, pp. 509–522, 2002.

[3] S. Lazebnik, C. Schmid, and J. Ponce, "Sparse texture representation using affine-invariant neighborhoods," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2003, pp. 319–324.

[4] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1/2, pp. 43–72, 2005.

[5] F. Schaffalitzky and A. Zisserman, "Multi-view matching for unordered image sets, or "How do i organize my holiday snaps?"," in *Proceedings of the Seventh European Conference on Computer Vision*, 2002, pp. 414–431.

[6] Chyuan-Huei Thomas Yang, Shang-Hong Lai, and Long-Wen Chang, "Reliable image matching via modified hausdorff distance with normalized gradient consistency measure," in *3rd International Conference on Information Technology, Research and Education*, June 2005, pp. 158–161.

[7] C. Chen, T. Kurozumi, and J. Yamato, "Poster image matching by color scheme and layout information," in *IEEE International Conference on Multimedia and Expo*, July 2006, pp. 345–348.

[8] Krystian Mikolajczyk and Cordelia Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[9] M. Carnec, P. Le Callet, and D. Barba, "Visual features for image quality assessment with reduced reference," in *Proceedings of the IEEE International Conference on Image Processing*, Septemeber 2005, vol. 1, pp. I – 421–424.

[10] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1680–1689, June 2006.

[11] M. Sendashonga and F. Labeau, "Low complexity image quality assessment using frequency domain transforms," in *Proceedings of the IEEE International Conference on Image Processing*, October 2006, pp. 385–388.

[12] Qiang Li and Zhou Wang, "General-purpose reduced-reference image quality assessment based on perceptually and statistically motivated image representation," in *IEEE International Conference on Image Processing*, October 2008, pp. 1192–1195.

[13] S. M. Weiss, B. F. White, C. V. Apte, and F. J. Damereau, "Lightweight document matching for help-desk applications," *IEEE Intelligent Systems and their Applications*, , no. 2, pp. 57–61, March/April 2000.

[14] J. You and P. Bhattacharya, "A wavelet-based coarse to fine image matching scheme in a parallel virtual machine environment," *IEEE Transactions on Image Processing*, , no. 9, pp. 1547–1559, September 2000.

[15] M. D. Gaubatz and S. S. Hemami, "On the nearly scale-independent rank behavior of image quality metrics," in *IEEE International Conference on Image Processing*, October 2008, pp. 701–704.