Hybrid Target for Camera-Based Document/Object Capture System *

Yang Lei [a], Peter Majewicz [b], Kurt R. Bengtson [c], Lisa Li [c], and Jan P. Allebach [a]; a: School of Electrical and Computer Engineering, Purdue University; West Lafayette, Indiana USA, b: Flextronics, Inc.; San Jose, California USA, c: Hewlett-Packard Company; Boise, Idaho USA

Abstract

The use of mobile phone camera technology in systems for capture of documents or 3D objects by a single exposure is becoming increasingly popular. Compared to conventional flat-bed scanner technology, such systems pose special image quality concerns due to the relatively uncontrolled environment within which they operate. In particular, it is not possible to control ambient lighting, and there is more variability with the internal illumination source and the placement of the document or object to be captured. This leads to higher levels of noise and greater non-uniformity across the image plane. In addition, since the underlying camera and internal illumination technologies are inherently low-cost, each unit is designed to operate closer to the threshold of acceptable quality and there is greater unit-to-unit variability. This makes it necessary to inspect each individual product for acceptable image quality as it comes off the manufacturing line. In order to maintain efficiency in the manufacturing process, this inspection must be completed as autonomously as possible.

This paper proposes a specially designed hybrid target and an automatic analysis tool for image quality inspection of camera-based document/object capture systems. The tool can be used in both the research laboratory and on the manufacturing line. The criteria and thresholds for final production are chosen based on visual inspection of image data from hundreds of pre-production units. Currently, our solution is used in mass manufacturing, and shows effectiveness in failing units with low image quality, as well as saving manufacturing time.

Introduction

The capture system of the product of interest adopts mobile phone camera technology to scan documents and objects by one shot. Since it is based on a low-cost camera technology, and works in a highly open environment compared to a flat-bed scanner, new image quality concerns arise: higher noise level due to the ambient lighting conditions, non-uniformity of the page, and increased unit-to-unit variability. Thus, image quality testing systems for traditional scanners are not applicable to such a product. The key image quality metrics of concern include image sharpness (Modulation Transfer Function), uniformity of the page, noise level, gray balance (Tone Reproduction Curve), color accuracy, and geometric distortion.

Unfortunately, available commercial digital image quality software packages are not well suited for this particular application. For example, Imatest Master Edition [1] is a very popular and powerful software package for testing image quality of digital cameras. However, human interactions are needed for analysis, which is unrealistic in a manufacturing setting.

Therefore, one single hybrid target and a highly automated analysis tool for camera-based capture systems must be developed, so that no target changing and human interactions will be needed during the analysis.

The rest of this paper is organized as follows. The next section describes the design of the hybrid target in detail. Then we illustrate the algorithm used for extracting all different regions of interest from the target. The analysis process and metrics used are described as well, followed by experimental results and conclusions.

Target Design

At first, several existing targets are used for image quality inspection, such as those shown in Figs. 1 and 2. Human judgments are necessary as well. On the manufacturing line, it is not cost-effective to change between targets, and it is impossible for researchers to make judgments on each single unit built. Therefore, one single hybrid target and a highly automated analysis tool must be developed.

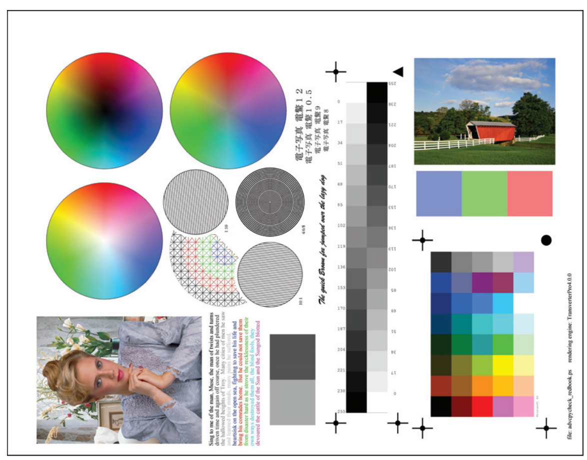

Figure 1. Target previously used for visually checking color accuracy, gray balance, and noise level of product units.

In the hybrid target proposed in this paper, shown in Fig. 3, one basic pattern repeats on an 8×11 grid throughout the page, which covers the largest common area of A4 and Letter size pages. Based on the slanted edge target, each basic pattern includes four gray level regions, four colors regions and four slanted

*Research conducted while Yang Lei was an intern with HP and Peter Majewicz was an employee of HP.

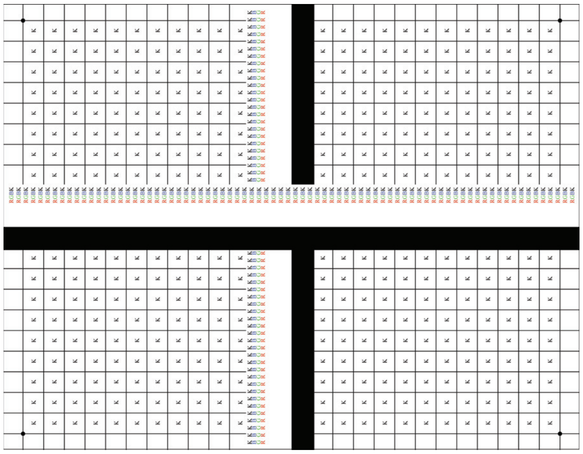

Figure 2. Target previously used for visually checking geometric distortion, color accuracy, and image sharpness of product units.

edges.

We summarize the design as follows:

1. Four slanted edges between dark and light gray regions are used for MTF (image sharpness) inspection.
2. Four gray levels (0%, 33%, 66%, and 100% reflectance) are designed for gray balance measurement.
3. Four color dots (red, green, blue, yellow) are designed for color accuracy inspection.
4. All smooth regions are tested for noise level.
5. The same region in different locations are examined for uniformity performance across the page.
6. The centroids of basic patterns on the 8-by-11 grid are used for testing geometric distortion.

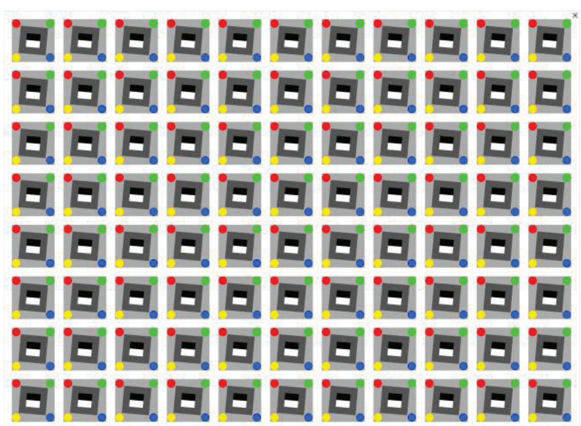

Figure 3. Digital frame of the hybrid target.

In practice, the procedure of printing the target is handled very carefully. All hard copies in use are printed from the same HP inkjet printer and all settings remain the same. Every hard copy is attached to a thin plastic backing board to make sure it remains flat during the inspection process.

Region Extraction

Above all, the location of each region of interest, i.e. four gray levels and four colors, must be precisely extracted for accurate key metrics evaluation. An image captured by the product of interest is shown in Fig. 4. Under the observation that the white rectangle within each pattern stands out from its surroundings, we extracted all the white rectangles first. The preset global threshold of 50% reflectance and areas of regions are used in this step, which are determined based on the data of pre-production units and can be adjusted if needed. The result is shown in Fig. 5.

Figure 4. Image captured by the product of interest. The white rectangles surrounded by black and dark gray area clearly stand out and are easy to extract.

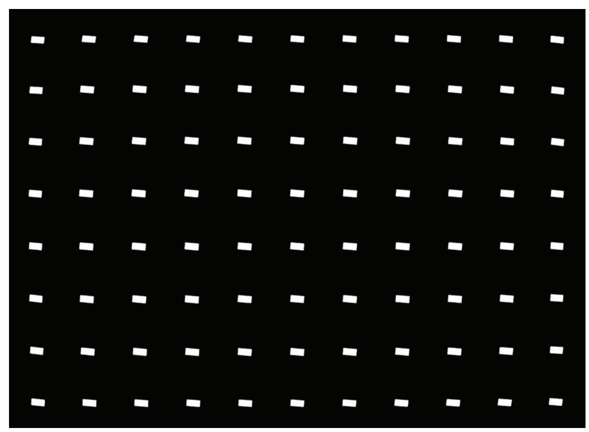

Figure 5. White rectangles extracted from Fig. 4. In this binary image, the bright areas correspond to the locations of interest, the white rectangles in Fig. 4. We call it the white rectangle region map.

Extract gray regions

Once we have the positions of the white rectangles, it is easy to crop the sub-patterns from the target. Then, for the other three gray level regions, two methods are used to extract them: location approximation based on offsets and thresholds. The results from these two methods are combined. The location approximation method is quite straightforward, since each black rectangle is located right above a white rectangle. As we can observe from Fig. 4, the uniformity of the captured image is not perfect. Also, the image quality varies from unit to unit. Experiments show that preset global thresholds do not work well for all units, not even within the image of a single unit. The purpose of the region extraction is to get as much correct information as possible, so that we can have input information that is good enough for the next step, which is key metrics evaluation. Therefore, adaptive thresholds must be determined. The key idea of our solution is to first perform local histogram equalization within each subpattern, then apply preset global thresholds. This works equivalently to applying local thresholds, and makes the procedure much simpler.

To guarantee accurate calculation of key metrics, safe region erosion is applied to the region maps we have obtained. We perform image erosion to the boundary of regions, in order to eliminate potential interference from neighboring regions. We summarize the process in Fig. 6.

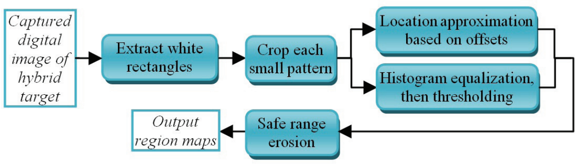

Figure 6. *Block diagram of the region extraction algorithm for gray regions.*

Extract color regions

Similar to extracting gray regions, location approximation is also used to prevent getting irrelevant regions, while the thresholding method is a little different. Since all color regions are isolated from each other and are surrounded by gray regions, the saturation of each pixel is calculated. Then color regions can be obtained by applying global thresholds to the saturation channel of the captured image. Safe region erosion is performed on the region maps, as well. We summarize the process in Fig. 7.

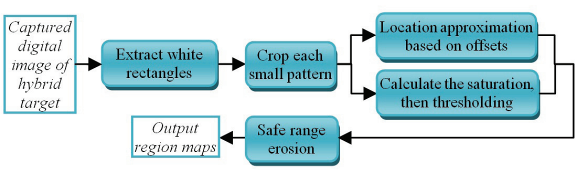

Figure 7. *Block diagram of the region extraction algorithm for color regions.*

At this point, very neat region maps have been obtained for each region of interest, along with the centroid of each basic pattern. All the information is used for calculating and evaluating the key metrics.

Key Metrics Evaluation

After getting the maps of all regions of interest, all the key metrics, i.e. MTF [2], uniformity, noise level [3, 4], gray balance, color accuracy, and geometric distortion can be measured and evaluated.

MTF

ISO 12233 Standard recommends using a slanted edge for MTF analysis. The four slanted edges in each grid are extracted to do MTF calculation at each spatial location. At each grid location, the slanted edge analysis can give MTF values across a range of spatial frequencies. Figure 8 gives an example from the manufacturing data. However, to get a full view of image sharpness across the page, the specific spatial frequency 2 lp/mm (line pairs per millimeter) is chosen based on the lens design. Then the corresponding MTF values are coded with color. This gives us the 2D view of the MTF distribution across the page, as shown in Fig. 9.

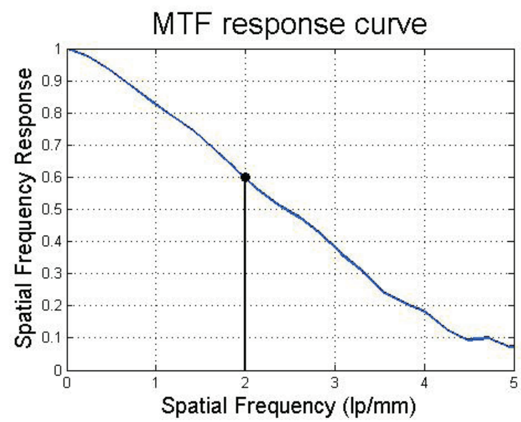

Figure 8. *MTF response curve for the sub-target located on row 3, column 5 in Fig. 4. X-axis is the spatial frequency from 0 to 5 lp/mm. Y-axis represents the spatial frequency response ranges from 0 to 1. The response value at 2 lp/mm is selected for evaluation.*

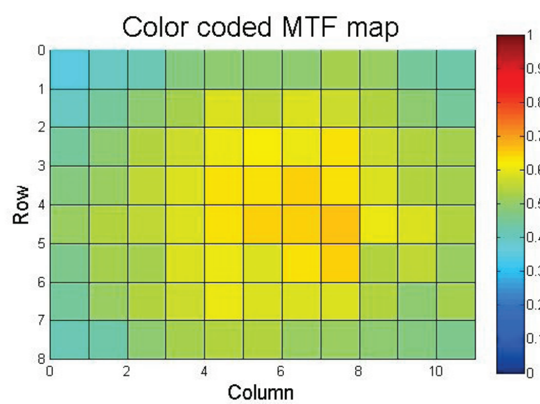

Figure 9. *Color coded MTF map corresponding to the sub-targets shown in Fig. 4. Blue represents a low MTF value and low image sharpness. Red represents a high MTF value and high image sharpness.*

Based on data from pre-builds, we take the 2D matrix and get the average response value at 2 lp/mm. The number of values that are less than 0.4 but larger than 0 is also counted.

$$\text{Average MTF} = \frac{\text{Summation of each non-zero value}}{\text{Number of items we are summing}}. \quad (1)$$

$$\text{N} = \text{the number of MTF values} < 0.4\text{, but} > 0. \quad (2)$$

The two values calculated from Eqs. 1 and 2 are compared with preset thresholds to determine whether a unit is good or bad.

Uniformity

After getting all region maps, the mean intensity value of each region at each location is calculated using the following weights (R, G, B) = (0.299, 0.587, 0.114). To check variations across the page, the range of mean intensities for each region is calculated as in Eq. 3 and similarly for the other seven colors. All the eight ranges are averaged, and compared to a preset threshold.

$$\text{Range}_{white} = \text{max}_{white} - \text{min}_{white}. \quad (3)$$

$$\text{Range}_{avg} = \frac{\text{Range}_{white} + \text{Range}_{black} + ... + \text{Range}_{yellow}}{8}. \quad (4)$$

Noise level

An algorithm based on [3] and [4] is developed to measure the granularity of a gray-scale or RGB patch taking into account the characteristics of the HVS (Human Visual System). We cropped image patches from the color circles and black, white rectangles in each sub-target to measure the noise throughout the whole page. The maximum values of the noise across all the regions of a given color are averaged across the eight different colors and compared to a preset threshold.

$$\text{Max}_{avg} = \frac{\text{max}_{white} + \text{max}_{black} + ... + \text{max}_{yellow}}{8}. \quad (5)$$

Tone Reproduction Curve

The RGB values for each gray level are transferred from sRGB space to CIE XYZ space first, and then to the CIE L*a*b* space. The L* values of pixels from the same region are averaged across the whole page as in Eq. 6. We call this the measured value. Then four average values are obtained for all four gray levels. On the other hand, the actual L* values (L^*_{white}, $L^*_{lightgray}$, $L^*_{darkgray}$, L^*_{black}) are measured from the hard copies of the hybrid target. We call these the the standard values. Then we have four corresponding points for the Tone Reproduction Curve. A straight line going through the origin with slope 1 is expected for perfect performance. In order to quantify the results, we fit the four coordinate pairs with a second order polynomial, and obtain three parameters [p_2, p_1, p_0] corresponding to the coefficient for the second order, first order and the constant terms, respectively. Then the differences between parameters [p_2, p_1, p_0] and the expected value [0, 1, 0] are chosen as the criteria. Thresholds are set for these differences. One example of the expected performance, measured results, and fitted curve is shown in Fig. 10.

$$\text{Lavg}^*_{white} = \text{average all } L^*_{white} \text{ in the image.} \quad (6)$$

$$[p_2, p_1, p_0] = polyfit_{2nd}(\text{standard } L^*, \text{ averaged } L^*). \quad (7)$$

Color accuracy

As the TRC measurement, we perform the transformation for color regions as well. With the average CIE L*a*b* value of each region, the ΔE_{94} distance between the captured copy and the actual target are calculated to represent the differences between the two. For each color, the maximum difference in the measured value across all sub-targets is computed; and these maximum values are averaged across the four colors. This average should not exceed a preset threshold.

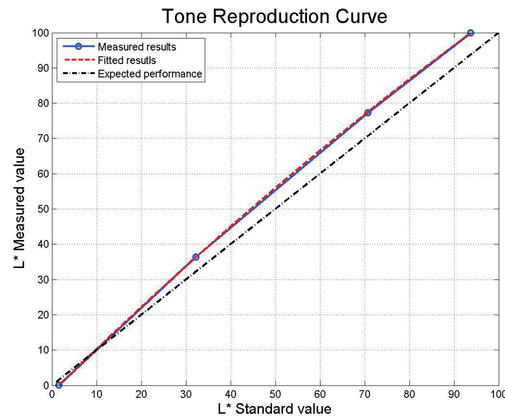

Figure 10. An example Tone Reproduction Curve. The measured result, fitted curve, and expected performance are shown respectively.

Geometric Distortion

Barrel distortion is observed in images from the camera-based capture system. In order to measure this effect, we took advantage of the centroids of basic patterns and the 8×11 grid. Specifically, for each row of the patterns, we take the X and Y coordinates of their centroids, and fit them with a second order polynomial. For each column, we flip the X and Y coordinates and treat them as rows. As we know, the expected values for the second and first order terms are both 0. Thus, maximum thresholds are set for these two parameters.

Experimental Results

The hybrid target and the automated analysis tool are used on an actual manufacturing line for picking out bad units. The tool was originally written in Matlab, then compiled to executables to use on the manufacturing line. As mentioned in the target design section, all targets in use are printed with the same inkjet printer, and mounted on plastic boards. In addition, all images are taken under a black box with no ambient lighting, to reduce interference caused by the environment.

After reviewing data from the manufacturing line, the current failure rate caused by bad image quality is less than 5%. Units that are failed in our tests do perform badly in their intended application.

Conclusions

A newly designed hybrid target and an automatic analysis tool for inspecting the image quality of a camera-based document/object capture system is proposed in this paper. Specific criteria and thresholds are illustrated for key aspects of image quality. This solution has been tested on hundreds of data sets from pre-production capture systems, and is being used on an actual manufacturing line. It has been shown to be very effective in picking out bad units. Future work includes maintenance and optimization of the analysis tool, and continuing review of the data from manufacturing.

References

[1] Imatest LLC official website: http://www.imatest.com/.
[2] Fundamentals of MTF: http://www.imatest.com/docs/sharpness.html/.

[3] M. McGuire and R. Shaw, ”Granularity in Digital Images”, HP Labs internal report, 1995.
[4] H. Mizes, ”Graininess of Color Halftones”, IS&T International Conference on Digital Printing Technologies, NIP16: 2000.

Author Biography

Yang Lei received her B.S. degree in electrical engineering from Sichuan University, Chengdu, China, in 2009. She is currently pursuing her Ph.D. degree in electrical and computer engineering at Purdue University. Her current research interests include digital image processing, image quality inspection, and 3D modeling with structured light systems.