



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Fantastic plastic? An Image-based Test Method to Detect Aesthetic Defects in Batches Based on Reference Samples

Hansen, Anne Juhler; Knoche, Hendrik; Moeslund, Thomas B.

Published in:
Polymer Testing

DOI (link to publication from Publisher):
[10.1016/j.polymertesting.2020.106585](https://doi.org/10.1016/j.polymertesting.2020.106585)

Creative Commons License
CC BY-NC-ND 4.0

Publication date:
2020

Document Version
Early version, also known as pre-print

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Hansen, A. J., Knoche, H., & Moeslund, T. B. (2020). Fantastic plastic? An Image-based Test Method to Detect Aesthetic Defects in Batches Based on Reference Samples. *Polymer Testing*, 89, Article 106585. <https://doi.org/10.1016/j.polymertesting.2020.106585>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

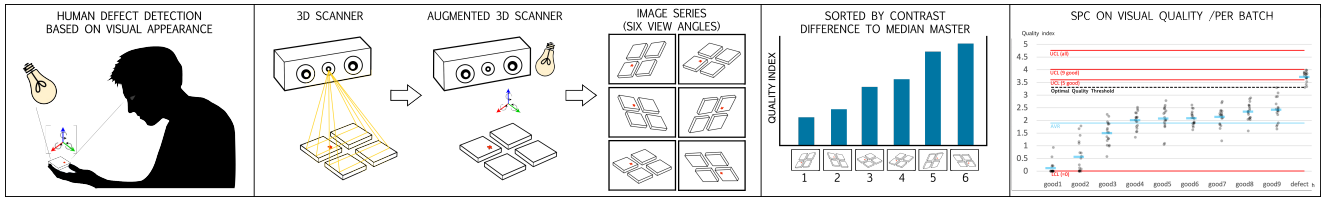
Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Graphical Abstract

Fantastic plastic? An Image-based Test Method to Detect Aesthetic Defects in Batches Based on Reference Samples

Anne Juhler Hansen, Hendrik Knoche, Thomas B. Moeslund



Highlights

Fantastic plastic? An Image-based Test Method to Detect Aesthetic Defects in Batches Based on Reference Samples

Anne Juhler Hansen, Hendrik Knoche, Thomas B. Moeslund

- Visual defect detection utilizing existing 3D scans to classify Aesthetic Quality.
- Image-based test method using median master comparison and contrast enhancement.
- Batch images are given a quality index and sorted based on difference to the master.
- The number of images in need of review by assessors is reduced by a factor of 13.
- High classification of defects in synergy with the principles of Six Sigma.

Fantastic plastic? An Image-based Test Method to Detect Aesthetic Defects in Batches Based on Reference Samples^{*}

Anne Juhler Hansen^{*}, Hendrik Knoche and Thomas B. Moeslund

Department of Architecture, Design, and Media Technology, Aalborg University, Rendsburggade 14, 9000 Aalborg, Denmark

ARTICLE INFO

Keywords:

Aesthetic quality
Visual appearance
Defect inspection
Machine vision
Image enhancement
Contrast sorting
Six sigma
Statistical processing control

ABSTRACT

The production of high-end manufactured products requires *Aesthetic Quality Control* (AQC) in the form of human visual inspection. Manufactures can reduce AQC costs by incorporating semi-automated visual defect detection in units with the existing 3D metrology scans. This paper demonstrates how an image-based test method for defect detection can reduce the workload related to human visual inspection by proposing a median master comparison of batch image series. Our contribution consist of a) contrast enhancing and sorting batch image series for human visual inspection and b) providing a quality index (nQI) incorporated into statistical process control (SPC) for monitoring and controlling the AQC process. Our data shows that the median master differencing together with the nQI is great for classification of defects in batch images series. We introduce a SPC design proposal where individual batches as well as aggregated data can be inspected in synergy with the principles of Six Sigma. Based on Six Sigma control limits we have reduced the number of images in need of review by AQC assessors by a factor of 13.

1. Introduction

High-end manufacturers of high volume premium products need *Aesthetic Quality Control* (AQC) [9] including inspecting for geometric and visual defects. Early defect detection reduces the risk of lost production. In high-volume manufacturing such as injection moulding assessors visually inspect several individual *units* (see Figure 1, B) as a *batch* (see Figure 1, A). Since in batch production items are produced together it is also logical to inspect units in batches to increase the efficiency of finding defective units. Visual defect detection requires inspecting units from various view angles given the non-constant relationship between illumination angle and the unit's surface for this task. Human defect detection entails: a) high labour cost and knowledge sharing between assessors and b) poor inter-assessor reliability [21]. A solution to these problems entails automation using machine vision piggybacked onto existing 3D metrology scans used for identifying geometrical deviations. So far, machine vision research usually focuses on single units from a single view angle and most importantly can only detect a subset of all visual defect types [19, 12]. Assessors, however, can detect all relevant defects and can be part of a human-in-the-loop solution. But they do require support in reviewing the multitude of view angle images obtained for a given batch.

Our semi-automatic aesthetic quality work-flow reduces the workload of AQC assessors (hereinafter assessors) by sorting a batch image series based on a master reference and a derived quality index. Our contribution consist of contrast-based image enhancement based on a median master and image sorting for human-in-the-loop batch defect detection. Our semi-automated visual inspections reduced the number of view angles (images) in need of review by an assessor by

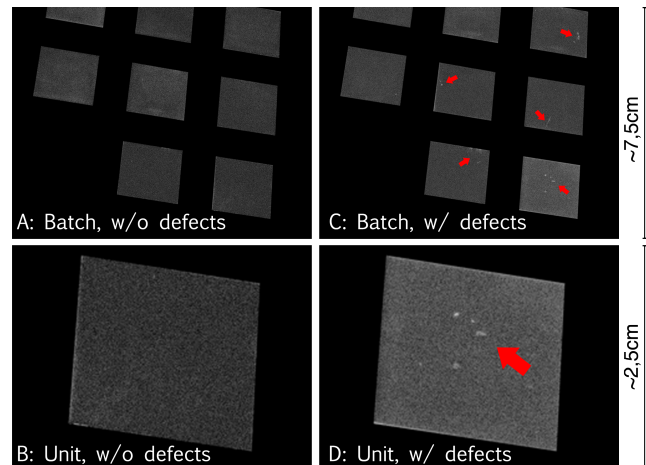


Figure 1: Batch level defect enhancement: A) batch image with 8 partially visible units, B) zoom-in on a single unit without defects, C) batch image where several units have surface imperfections, and D) zoom-in on a single defect unit. The assessor can identify defects (red arrows) based on the log-enhanced images. The red arrows indicate blisters. Notice the color change; the unit's surface (D) is brighter than the master unit (B).

a factor of 13 for visual defects (discolouration and blisters) when employing a Six Sigma statistical processing control (SPC) control limit. Our small scale feasibility study relied on realistic image footage counting 10 different batches. Each batch consisted of 12 units and 18 different view angle images were captured per batch. As a side effect our approach can presumably also improve inter-rater reliability.

^{*}Corresponding author.

E-mail: ajha@create.aau.dk (A.J. Hansen)

<https://vbn.aau.dk/en/persons/128031> (A.J. Hansen)

ORCID(s): 0000-0001-8112-8599 (A.J. Hansen)

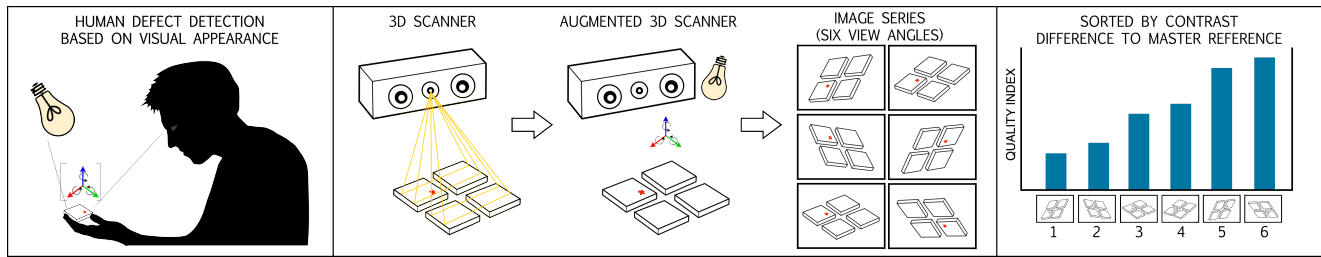


Figure 2: Manual aesthetic quality control model with illumination light source, unit surface (including potential defects) and human visual inspection. The visual appearance of the unit changes with both the angle of light hitting the surface and the assessors viewing angle to the surface. The 3D scanner employs two cameras for metrology measurements and as a by-product we get a batch image series which can be used for semi-automated visual inspection. The batch image series is sorted based on the absolute difference to the median master (see e.g. D in Fig. 4).

2. Aesthetic Quality Control

Mass-manufacturing produces large amounts of standardised products at both high and low volume using assembly line techniques. The difference between high and low volume is expressed in number of units subjected to quality assurance. In low volume productions all individual units often undergo AQC whereas high volume production quality control is performed by sampling batches. In quality management focus is on preventing non-conformity (i.e. deviation from a standard or specification) and we examine this problem by creating a test method for batch defect detection of injection moulded plastic surfaces.

We draw on Levitt’s definition of manufacturing quality as *conformance to specification* [17] i.e. absence of defects, deficiencies, and significant variations to prevent non-conformity. Several standards exist for testing mechanical properties of injection molded products but product appearance is difficult for assessors to quantify [31]. Imperfections in materials and failure during manufacturing produce observable aesthetic defects that customers find aesthetically displeasing [8]. Manual inspection of visual appearance is expensive and both accuracy and speed are difficult to improve for human quality assessors. An alternative is automatic and potentially non-destructive testing. Nondestructive inspection technology do not permanently alter the inspected object but the non-destructive methods should accurately and rapidly detect defects for quality assurance [31]. Non-destructive inspection technology includes industrial radiography, ultrasonic testing, magnetic particle inspection, visual testing and more [15].

3. Automatic Visual Inspection (AVI)

To enable faster and more accurate inspection industrial production nowadays routinely employs machine vision to conduct automatic visual inspection (AVI) e.g. of circuit boards [16] in which functional quality thresholds are easy to define. Aesthetic quality evaluation is more complex due to the number of dimensions involved, i.e. perception and cog-

nition. Aesthetic quality is part of a continuum where both the holistic understanding as well as discrete measurements comprise the product specification. This is reduced to a binary decision of OK/NOT OK represented by defect characteristics. Profiling different defect types within manufacturing can help in automating AQC for machine vision using quantifiable anomaly descriptors. Many different materials (e.g. leather [28], ceramic [3], stone [18], metals [19], [23], and plastic etc. [12]) make AVI subject to real-time and in situ testing and can contribute detection, localization and classification of defects in polymer products. Highly reflective materials (e.g. plastic) gives varying reflectance and is part of the problem for AVI since specular highlights on objects create noise in the images and eradicate homogeneity. Homogeneity is particularly relevant when evaluating visual appearance of injection moulded objects with uniform surfaces [27]. One way of introducing image enhancement is to use structured backlighting with binary stripe patterns to enhance contrast and make visual defects stand out from the noise [8]. Contrast-based methods have also been successfully employed for quantitatively assessing scratch visibility in transparent polymer films [5, 2]

Commonly AVI approaches take CCD-images as input for fast evaluation of visual appearance. For AVI the focus is often on noise reduction, contrast enhancement, and image segmentation [3] and often the detected number of defect types is limited to only a selected few [19, 12] unless the approach is to detect surface homogeneity [27]. Recent advances within AVI use varying approaches including texture analysis methods [30], statistical analysis methods (i.e. spatial distribution of texture values) [24], filter-based approaches such as Fourier, Gabor, and Wavelet transform etc. [29], and model-based methods including fractal, autoregressive, random field, and texem models [4].

4. Visual inspection within companies

Six Sigma is a methodology used for increasing product quality and eliminating defects in the manufacturing process. By deploying a structured, scientific, and data-driven pro-

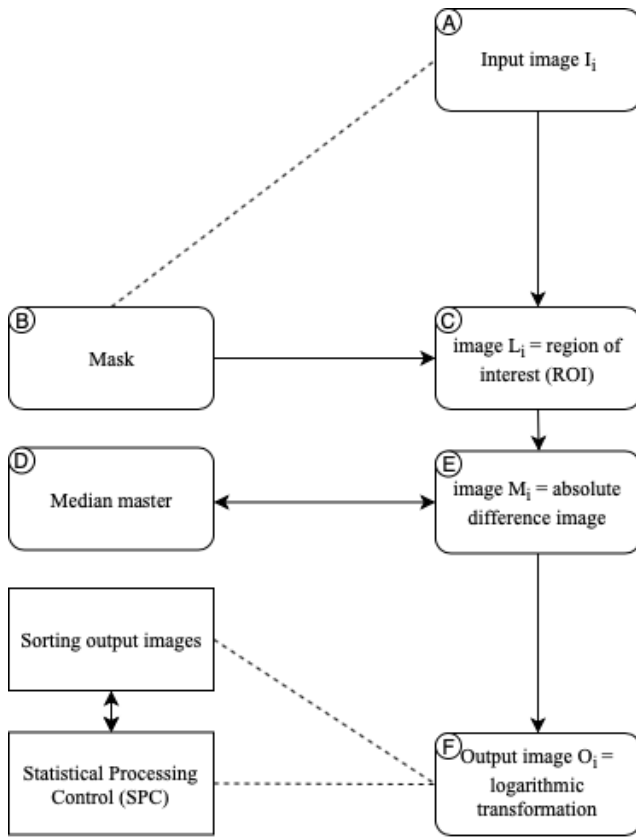


Figure 3: Flowchart; starting with input image I_i we first create the ROI image L_i of a batch consisting of several units. We then compare L_i to the median master and calculate the absolute difference image M_i for each viewing angle. M_i is then contrast enhanced with a logarithmic transformation giving the output image O_i which can be sorted for visual inspection (see Figure 6) and used for SPC (see Figure 9).

cess the Six Sigma methodology limits product defects to six standard deviations i.e. six sigmas. DMAIC is a Six Sigma improvement method which advance an existing process by reducing the variation through five phases: Define, Measure, Analyze, Improve, and Control [6]. As part of the control phase, Statistical Processing Control (SPC) is used to analyse, monitor and control a process performance. SPC use statistical tools to observe the performance of the production process by involving the use of control charts in order to verify if certain process variables remain close to their desired values [20]. If the process has a normal distribution, 99.7% of the population is within three standard deviations from the mean. A measurement value beyond 3 standard deviations indicates that the process has shifted since we see more variability. Hence, the control limit is typically three standard deviations above the process mean. The Six Sigma improvement methodology is a statistical-based and data-driven approach to quality management, but in situ the assessors are prone to influencing the evaluation (e.g. attention, training etc.), which is why they need help with the procedure in order to reduce variability. Different methods to reduce vari-

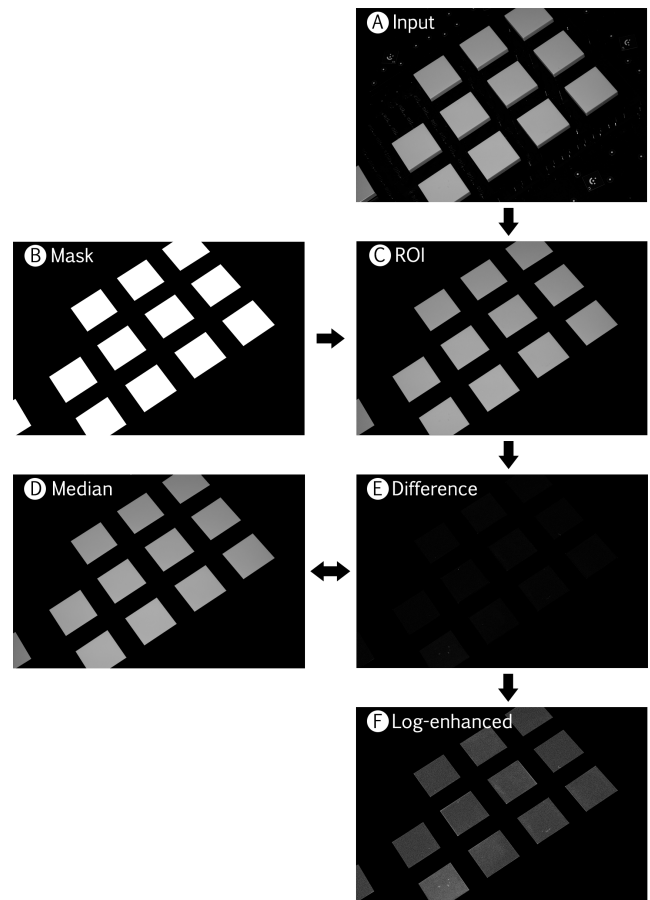


Figure 4: Pipeline; starting with input image (A) and a mask (B) we create the ROI image (C). We can then compare it to the median master (D) and calculate the absolute difference image (E). A contrast enhancement on the difference image employed log-transformation (F) to intensify the defects on the individual units in the batch.

ability have been developed [1] and international standards on surface imperfections (e.g. ISO8587 [10]) describe many specific types of defects and their characteristics. Existing measures used within manually performed quality assessment include mean opinion scores (MOS) [25], binary acceptability ratings [22] etc. According to the CIE TC1-65 standard on visual appearance [26] the AQC process can be divided into two stages: a) detection/exploration i.e. searching for anomalies using either a random or systematic exploration strategy and b) decision/evaluation i.e. evaluating anomaly type and deciding the intensity of the anomaly. Expert quality assessors represent a measuring instrument in AQC [1]. In regular visual inspection the assessors judge the anomaly based on their *perception* and *cognition*. Deviations within AQC are represented as surface inhomogeneity in the form of various defect types (e.g. scratches, dents, holes, color changes, etc.) [26] which can consist of both geometrical changes (e.g. hole) and visually perceived differences (e.g. color changes) [9]. A basic classification of aesthetic defect includes various defect types (e.g. impuri-

ties, stains, scratches, dents etc.) whereas a more complex classification would consider the location, size and shape of each defect [8]. One perceptual categorization include differentiation between open and closed shaped anomalies; closed shapes contain clear contours with unbroken edges (often constituting geometrical defects as dents, sink marks, etc.) whereas open shapes do not have to be bounded across a clearly defined region of space (e.g. discolouration, flow marks etc.) [13].

One of the main obstacles to detecting defects in polymer plastic is that the illumination, material reflectance and view angle influence the perceived appearance of a surface. Hence, assessors account for the incoming light angle by rotating the inspected objects in order to better detect defects. The human visual systems excels at visually distinguishing between materials by estimating the material properties of surfaces [11]. Our perception usually makes us distinguish between a variety of materials despite significant changes in light, viewpoint, shape etc. In other words, glossier surfaces manifest more salient specular reflections than less glossy surfaces, and these salient features (size, contrast, and distinctness of highlights) help us characterize and distinguish between materials and surfaces. Appearance of injection moulded plastic parts is important, e.g. colour evenness, surface texture, smoothness, gloss etc. Thus standard measurement methods need to be precise, repeatable, reproducible, and produce results which correlate with human inhomogeneity perception (see Section 4). In other words, a defect detection test method should encompass all types of visual deviations. A geometrical difference is one way of measuring conformance to specification, but evaluation of visual appearance includes all visually perceived deviations.

Many defects are only visible under a specific angle of light and orientation of the object (see Figure 2). Thus defect detection needs several viewpoints/rotation of the object: a) for all unit surfaces subject to AQC, b) for defects only visible from certain viewing angles, and c) for accounting for reflections/highlights. An image-based test method for defect detection should incorporate these principles by varying the angle of viewpoints over a series of images. This aligns well with the process required for metrology measurements where we can obtain a series of images from different viewpoints without additional cost. We obtained several batch image series from the 3D scanner setup where the image series demonstrate different view angles of the same batch of units. The images from different view angles represents the human rotation of the inspected object in relation to the surface characteristics and the light source. Performing defect enhancement on images for human-in-the-loop visual inspection (i.e. model requires human interaction) not only help the assessors to detect defects, but the image series create new libraries of references. Therefore, an image-based test method can foster an improved judgment process for the assessors. By comparing reference images (defect free) with test images (with potential defects) we can compute an objective quality index (see Section 8). The quality index can help assessors by ordering the images according to an ob-

jective calculated quality index based on the contrast difference between reference and test image and therefore assessors only need to evaluate a subset of the available images.

5. An Image-based Test Method within AQC

The developed imaging enhancement mechanism strengthens the visibility of defects for human visual inspection. The implementation include several stages; noise reduction using region of interest (ROI), creating a median master as defect-free reference, enhancing each individual view angle image using log transformation, sorting the batch image series, and evaluating the results through SPC using an objective quality index (nQI).

The experimental system consisted of a GOM ATOS Capsule. The ATOS Capsule is an optical precision measuring system with dimensions of 200×140 mm. The structural light 3D scanner was used as an integrated part of the quality control process for metrology measurements outputting a 3D reconstruction based on an obtained image sequence. Since handling is expensive it is impractical and infeasible to inspect single units and, thus, we collect batch images since batch inspection simulate the real-life setup. Capturing the batch image from various views are constrained by the amount of units visible in each image, but this variation affords a strength that causes our method to be robust and effective in a production setup. In our setup we focused on quality inspection of visual appearance using machine vision on the image sequence. The only light used came off the shelf from the central projector (see Figure 2, augmented 3D scanner) and was fixed to the moving camera. The image resolution was 4248×2832 pixels and in the grey scale color space.

6. Image Processing Pipeline

We developed our prototype in Python using several libraries (e.g. the OpenCV library for image processing, Pandas for data manipulation and structure, NumPy for multi-dimensional arrays etc.). The pipeline starts with an input image I_i (image I of image series i) from where we find the region of interest (ROI) L_i (see Figure 3 and 4, A-C). We then compare L_i to a master reference (the median master was created as an average of pixel values of several good batch images) and calculate the absolute difference image M_i (see Figure 4, D-E). M_i is then contrast enhanced with a logarithmic intensity transformation giving the output image O_i which makes it easier to visualise differences with both high and low intensity (see Figure 4, F).

We start by explaining the pipeline so the sections are represented as the numbers from figure 4, A) input image, B) mask, C) ROI, D) median master, E) absolute difference image, and F) output image.

6.1. Input Image (A)

A GOM ATOS Capsule 3D scanner took images of a batch from different view angles creating a *batch image series*. This process emulated the human AQC process of ro-

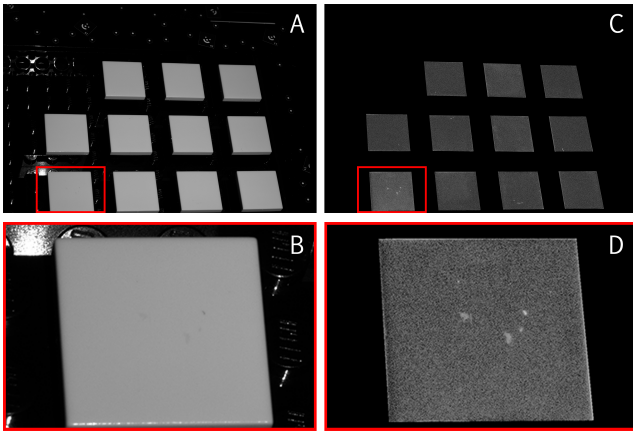


Figure 5: Defect enhancement; A) unprocessed grey scale batch image, B) zoom-in on a single unprocessed defect unit, C) processed ROI and log-enhanced batch image, D) zoom-in on ROI enhanced defect unit where the defects can be seen as white stains.

tating an object to identify defects using the illumination to calculate visual appearance. Not all units were fully visible in all views (minimum 9 out of 12 units).

6.2. Masking (B) and Region of Interest (C)

We need a mask to differentiate between unit surfaces and unwanted noise from the background and hence we create a mask that covers only the surfaces of the units (see Figure 4, B). The region of interest (ROI) can be calculated either from automatically generated masks or with a pre-defined mask. In our case using a simple handcrafted mask was efficient since 1) we always compare images from the same view angle, and 2) the masks are created once and reused for all batches hereafter. The individual units are not completely aligned (e.g. due to human error when placing the units in the fixture). However, in this paper image alignment was not necessary since the difference between defected and OK units were strong enough even without image alignment (see Section 10). To remove noise from the background (e.g. highlights, fixture markers used in 3D scanning, etc.) we mask out the region of interest (ROI) to remove the non-important parts (see Figure 5 where B has a strong highlight in the upper left corner of the background which has been masked out in the ROI image D).

6.3. Median master (D)

To visually strengthen the signal of the defects we compared a current batch to a master reference to focus on the differences in pixel values. The master symbolizes a perfect batch and is created from the median of several good batch images. Using the median as a master reference removes noise and increases the signal-to-noise ratio. The median master image is created by taking several good batch images and then running through all images pixel by pixel and calculating each entry based on the median.

6.4. Absolute difference image (E)

The defects were then visually enhanced by creating the absolute difference image between a current batch image and the median master. The calculation of the absolute difference image is calculated as the absolute value of the pixel difference between the current unit and the median master for each pixel value. The absolute difference image was a grey scale image only showing any differences in pixel values. The difference image pixels are black where no pixel difference existed between the batch image and the master and any change in pixel value will be indicated by the pixel intensity in the grey scale spectrum. Therefore, any non-black pixel indicate a difference to the master and a potential defect (see Figure 5, D).

6.5. Log Intensity Transformation for image contrast enhancement (F)

To further enhance the signal for human visual inspection on the absolute difference images we perform contrast enhancement on the images using the logarithmic intensity transformation. Contrast enhancement works well for low contrast images where the contrast of the image is stretched. We expect a low intensity difference image and, hence, the logarithmic transformation enhance the brightness values of low intensity pixels more than high intensity pixels. The low intensity pixels are enhanced using log transform (see Figure 4, E-F) so enhancement occurs at each pixel location $f(x, y)$. We used a log10 transform in which the base 10 logarithm of the grey value input pixels is returned pixel-wise for the 2D image;

$$g(x, y) = \log_{10}[f(x, y)] \quad (1)$$

This process makes the image features stand out more clearly by changing the range of values in an image in order to increase contrast. Making optimal use of the available pixel values makes the defects not only more noticeable for human assessors but also increases the contrast of the defects for image processing. For a faster human inspection, we assume imaging the defect as a set of light pixels will improve the human detection by reflecting more light compared to the dark background where more of the light will be absorbed (see Figure 5).

7. Sorting batch image series

Manual inspection is expensive and both accuracy and speed are difficult to improve for human quality assessors. Consequently, our focus is to reduce the workload related to quality assessments and support the attention of assessors by sorting images based on the pixel brightness sum to reduce the number of images in need of review. Although the contrast-enhanced images are improved for human visual inspection, we still have many different images (i.e. different view angles) that needs to undergo AQC. Hence, by sorting the images for human visual inspection we create a helping tool, which assessors can use to evaluate AQC. Jointly, the image sorting works well with SPC charts (see Section 9).

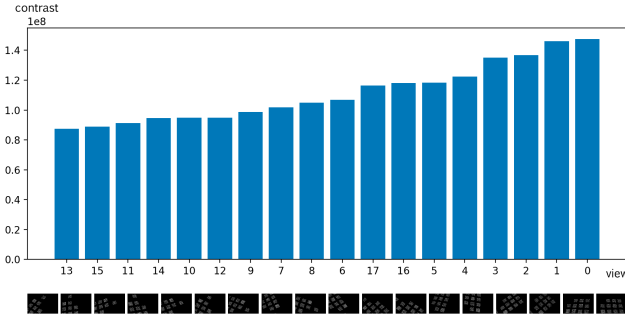


Figure 6: The x-axis shows image/image number and the y-axis the pixel contrast sum per image. The batch image series is sorted based on the sum of all non-black pixel. Images with the highest sum differ most from the master and hence visual inspection is performed on images with highest contrast.

We propose to sort the image sequences based on image contrast (see Figure 6), since image contrast indicate a deviation from the master. The absolute difference image give us the difference between the current batch image and the master, hence non-black pixel exist only when there is a difference to the pixel value in the master reference (e.g. due to a defect) and these areas should be visually inspected by an assessor (see Figure 5). We determine the contrast in each image by summing the value of non-black pixels per image and then sort the images based on contrast intensity sum. Images with the highest sum have the highest contrast and differ most from the master. The contrast sorting is based on the following summation per image which is then ordered for all different image views:

$$C = \sum_x \sum_y I_{xy} \quad (2)$$

where C is the image contrast sum, I_{xy} is the input image, x is the pixel rows and y the pixel columns of the image providing an x,y -position of each pixel. We sort based on the grey scale pixel values as a proof of concept whereas other measures (e.g. quality index (nQI), see Section 8) could likewise create basis for the image sorting.

8. Calculating the Quality Index (nQI)

For each batch image view angle we propose a quality index to help assessors making fact-based decisions. We expect the defect to resemble a small part of the image, and hence, we want to normalize the image according to the defect and the typical noise-signal we get in the image. This means that if we have a high contrast count value we want to subtract the minimum noise value that occurs in every image and divide by the variation we see in the data set. Therefore, after creating the contrast difference images the quality in-

dex (nQI) is based on the following equation:

$$nQI = \frac{C - \min_c(X_c)}{\sqrt{\frac{\sum_c (C - \mu_c)^2}{n_c}}} \quad (3)$$

where C is the total image contrast (i.e. pixel brightness sum) per image view, X_c is the contrast intensity of the population of the same view angles across batches, μ_c is the contrast intensity sample mean of the same view across batches, and n_c is the sample size. To determine the nQI we calculated the standard deviation σ from the contrast values (i.e. pixels being different from the master reference) of all images from the same view angles. The standard deviation gives a measure of how widely values are dispersed from the average and is calculated using the "n" method [7] for each individual batch image.

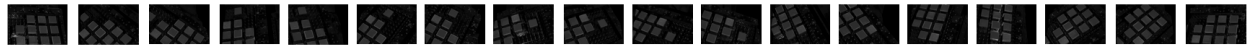
The nQI is calculated for all 10 different batches and their 18 respective view angles. A low nQI index signifies a good batch with small deviations to the master whereas a high nQI signifies that the batch have high contrast to the master (i.e. which might be caused by defects). We used a simple colour coding mapping a low nQI index (i.e. good batches) to white and a high nQI index (i.e. potential defects) to red (see Figure 7).

9. Measurement Results

Based on our data set nQIs were calculated for all 10 batches and 18 view angles for a total of 180 images. Approximately 100 individual units (some units were used in more than one image) were captured, processed and analyzed in ten different batches, where each batch consisted of 12 blue plastic units. While not validated by AQC assessors we identified 5 defected units with easily noticeable defects within the defected batch *defect*. At least one defected unit was visible in each view for the defected batch and an average of approximately 10 partially visible unit surfaces (with min=9 and max=12 visible units in each image for all batches).

Our data shows that the median master differencing together with the nQI is great for classification of defects in batch images series (see Figure 8). We visualize our results to the assessors using control limits to illustrate a quality cut-off threshold. The SPC chart presents our data in relation to the control limits (see Figure 9).

We have created three different upper control limits (UCL) based on *good1-good5* (UCL5), *good1-good9* (UCL9) and all data points *good1-good9* and *defect* (UCLall) in our limited sample set. The control limit is set three standard deviations above the process mean. The upper control limit (see Figure 7) is e.g. 3.61 for UCL5 whereas the lower control limit (LCL) is set to 0 since we operate with no negative nQI values. The purpose of the SPC chart is to check whether data points lie above or below the control limit where data points lying above three standard deviations away from the master indicate unacceptable deviations within the batch.



	View 1	View 2	View 3	View 4	View 5	View 6	View 7	View 8	View 9	View 10	View 11	View 12	View 13	View 14	View 15	View 16	View 17	View 18
defect	3.80	3.98	3.91	3.68	3.98	3.53	3.42	3.33	3.70	3.89	3.69	3.67	3.86	3.30	3.84	3.85	3.82	3.77
good1	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.94	0.00	0.57	0.00	0.20	0.00	0.00	0.00	0.00
good2	0.00	0.58	0.72	0.09	1.09	0.25	0.03	0.13	0.00	0.00	0.51	0.00	0.73	0.00	1.79	1.67	1.16	1.41
good3	1.85	1.53	2.16	1.87	2.24	1.90	1.29	1.53	1.02	1.69	1.36	0.99	1.02	0.57	1.28	1.95	1.64	1.20
good4	2.30	1.34	2.13	1.80	2.29	2.52	2.13	2.40	1.68	2.45	2.12	1.76	1.97	1.96	1.56	2.02	2.09	1.57
good5	1.78	1.05	2.38	1.92	2.60	2.31	2.03	2.27	1.10	1.85	1.75	2.00	2.11	2.03	2.53	2.78	2.48	2.45
good6	1.74	1.63	2.27	2.10	2.49	2.19	1.90	2.22	1.89	2.61	2.34	1.95	1.80	2.07	2.11	2.48	1.88	1.92
good7	1.84	1.19	2.10	1.69	2.75	1.87	2.23	2.27	2.18	1.79	2.36	2.28	2.09	2.20	2.40	2.69	2.37	2.20
good8	2.03	2.08	2.59	2.24	2.55	2.41	2.12	2.30	1.58	2.34	2.36	2.47	2.10	2.04	2.61	2.89	2.84	2.75
good9	1.66	1.77	2.37	1.90	2.53	2.23	2.37	2.45	1.82	2.66	2.44	2.40	2.47	2.31	3.08	3.29	2.94	2.94

Figure 7: Color coded table of the nQI for all batch images and all view angles. The x-axis shows the different view angles and the y-axis shows the different batches. White cell color coding signifies a low nQI and red cells a high nQI indicating that these batch images or view angles should potentially undergo visual inspection.

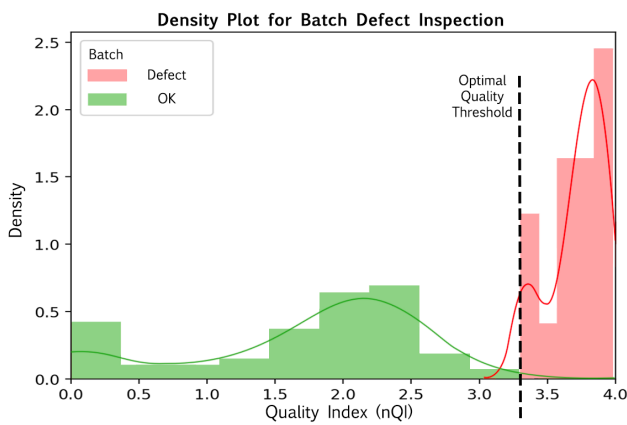


Figure 8: Density plot of all batch images; *good1-9* $n=162$ (green) with area under the curve (AUC)=1 and *defect* $n=18$ (red) with AUC=1. The calculated nQI divides the good batches and the defected batches into two data groups that can be separated by a set quality threshold. For our data set a favorable quality threshold can be set at $nQI=3.3$.

From UCL5 (see Figure 9) one defected batch (*defect*) were identified based on 14 data points above the UCL5 resulting in 14 true positives but with 4 false negative view images. Based on this we calculate the precision and recall of defected views based on true positives (TP=14), true negative (TN=162) false positives (FP=0) and false negatives (FN=4). Precision and recall are calculated as follows; precision = $TP/(TP+FP)$ and recall = $TP/(TP+FN)$. This resulted in a precision of 100% (14/14) while the recall was 78% (14/18) for our limited batch sample size. The accuracy is 98% given accuracy = $(TP+TN)/(TP+FP+FN+TN)$ = $(14+162)/(14+0+162+4)$. The number of view angles in need of review by assessors is reduced from 180 images to the 14 images whose values exceeded the control limit in *defect*. This decreases the number of images in need of review by a factor of 13 (180/14). The control limit can also be calculated from all good batches *good1-9* (UCL9) giving a process mean of 1.90 and a control limit of 4.76 (see Fig-

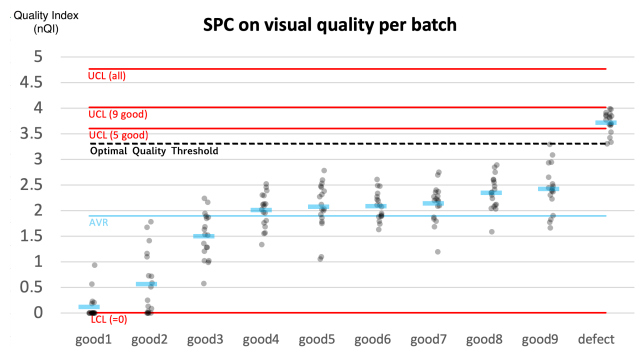


Figure 9: Statistical processing control chart; The red lines represents the upper control limit (UCL) three standard deviations from the process mean and the lower control limit set to 0. The process mean and control limit is based on *good1-good5* for UCL5, *good1-good9* for UCL9 and includes all good batches and the defected batch in UCLall. The blue data markers indicate the averaged quality index of image views and the blue line the average (AVR) over all nQIs. All good batches (*good1-good9*) are below the control limits whereas the image series with defects lies above UCL5 but below UCL9 and UCLall.

ure 9, UCL9). Here the SPC control limit lies above all data points and does not provide a clear threshold between good and defected batches. The same results are evident when calculating the control limit based on all data points *good1-good9* and *defect* where UCLall is set too high to function as a proper quality threshold. The best fit quality threshold for our data set lies at $nQI = 3.3$ (see Figure 8) where we have a clear cut-off between *good1-good9* and *defect*.

10. Discussion

Our calculated nQI indices (see Figure 7) are remarkably low for the batches *good1-2* (i.e. the scores are close to zero for both individual batch images and for the different image views) indicating very close resemblance to the master reference. Since the median master was created based on *good1-9*

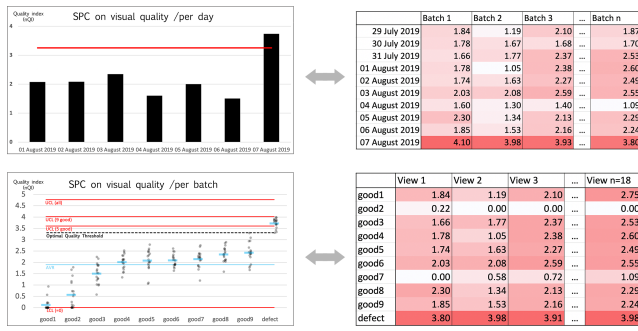


Figure 10: Alternate views: The SPC charts show a clear relation to the control limit (red line) whereas the color coded table provide an easy look-up table for nQIs over image series and individual image view angles. Aggregated data can be used to compare quality over different days and the nQI values can be used to explore the quality of specific batches.

(including *good1-2*), this is most likely an artifact from including *good1-2* in the median master. This means that the pixel values for these views lie close to the median values and that the baseline for some batches can be zero due to $\min(X_c)$ being based on the same values as those being assessed. The remaining data points (*good3-good9*) have a higher average (see Figure 9). Since our ground truth validation is not performed on individual units (but instead on batches containing multiple units), all good batches might include smaller surface deviations on a few units and/or noise in the form of dust, particles, hair etc. on the surface. This should be taken into account when creating a quality threshold based on a Six Sigma control limit since it affects the calculated process mean. The SPC control limit is calculated as 3 standard deviations from the process mean and may need to be continuously updated in order to maintain the most efficient quality threshold. This implies that the human assessors also need to focus on evaluating the process mean which the control limit is based on, but on the other hand this also provides extra control if the production quality intentionally is changed. When the control limit is based on *good1-5* (UCL5) we get a cut-off limit that excludes the *defect*, but when the control limit is based on *good1-9* (UCL9), all batches (including the *defect*) lies below the control limit. The UCLall is presented to show the relative difference since a control limit should not contain defected batches. A control limit based on Six Sigma SPC limits might be arbitrary and the cut-off threshold should presumably be tweaked by quality assessors to obtain the best cut-off limit per product (see Figure 8). This argues in favour of a human-in-the-loop approach where batches with nQI values both above and below a given control limit requires human judgment. Even though it might still be a challenge to decide on the best possible quality threshold our data shows that the nQI can be a valuable estimator for the quality spectrum of batch images series (see Figure 8). The nQI is not only beneficial during the human-in-the-loop defect detection but creates an auditable document trail that can help assessors standardize by using

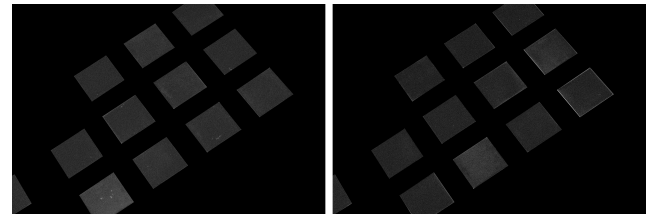


Figure 11: A defected batch image (left, *view 15 defect*, $nQI=3.84$) and a non-defected batch image (right, *view 15 good9*, $nQI=3.08$). The nQI values are in the same range and the high nQI is most likely due to misalignment of units creating noise around the edges.

the nQI archive as reference (see Figure 10).

Prior studies have compared image brightness values of scratch damaged areas on transparent polymer films with background areas and found contrast as a useful digital image analysis method [5, 2]. Other studies have found that an optimal combination of image pre-processing steps in defect detection is accomplished through noise reduction, contrast enhancement, and image segmentation [3]. We implemented this through the use of masking ROI to reduce noise and segment the image. Contrast enhancement was accomplished through log-enhancement of images computed from image differencing to a median master reference image. Not all steps in our process included complete automation (e.g. for proof of concept we have manually created masks, we have no image alignment, etc.) and the defect detection pipeline can be further improved by automating tedious manual steps.

The UCL5 calculated for our batch image series yielded an overall classification of visually defected batches with 100% precision, 78% recall and an accuracy of 98%. Even though these results are only preliminary they give higher accuracy than previous comparable results for defect detection within e.g. rolled steel (87% on the three defect types; welding, clamp and identification holes) [19] or ceramic ground workpieces (93% on grinding surface damage such as breaks and cracks) [3] potentially due to our limited sample size of 10 batches (compared to sample sizes of several hundreds).

The number of view angles in need of review by assessors is reduced from 180 images to the 14 images whose values exceeded the control limit, but potentially those 14 data points could be considered as one since assessors might only need to look at one or maybe a few images to classify a batch as OK or defected. The different view angles do not include all units i.e. in some images some units might be outside the image view. This simply means that not all units are evaluated in all images which evidently affect the results. When capturing batch images (i.e. images covering several units) the number of units affect the contrast of the image, since an image with more individual units is assumed to have higher contrast from the master due to noise caused by the total surface area of all visible units in an image. Also, this approach is prone to error since units with defects might not be present in all images. This is problematic since it will affect the im-

age sorting and Six Sigma control limit. Since our testing method is based on a calculated nQI for all image views we argue that as long as all units are captured in at least one image the provided nQI will deviate for that respective view which should then be evaluated by a human assessor. In our case sorting the batch images based on contrast and using control limits on the nQI scores worked even though all defected units where not visible in all images (see Section 5). This indicate a strong starting point for our approach and for the SPC design proposal.

We demonstrated that image series can be employed to detect defect in batches but the pipeline could potentially be extended to also be used within individual unit defect detection. The individual units were not completely aligned in our setup which means that when the absolute difference images are created from a current batch image and the median master reference we can expect noise, especially around edges, due to images not being completely aligned (see Figure 11). Also, if the used pre-defined mask is not making a perfect overlap we might miss some defects close to the unit edges due to the masking. The signal-to-noise-ratio in the difference image could be strengthened by making an exact overlapping match between the images that are compared. This signifies that our preliminary results could be even further improved by image alignment since we get a lot of noise from the misalignment of units. This goes hand in hand with the fact that we get no false-positives from any of our set control limits (UCL5, UCL9 or UCLall), but if we look at image views with values close to the defected batch, we see that the high nQIs is most likely due to misalignment and noise around the edges (see Figure 11). It could be beneficial to perform feature extraction and detection for creating an automatic ROI and conducting individual unit alignment. A homography transformation (i.e. a perspective transformation of a plane reprojecting points from one view point to a different view point) [14] could align individual units and create better comparison between batch images. This approach would allow for processing individual units since we already have the units divided into ROIs that can be operated on individually.

Future work needs to focus on image alignment based on individual units to improve the signal-to-noise ratio and remove noise causes by physical misalignment of units. Automated classification within AVI is mostly limited to a finite number of defect [19, 12], but for AQC it is essential to detect all types of defects which means that we need the human-in-the-loop approach. Potentially the joint 3D metrology and visual appearance inspection could be further improved by visualizing the aesthetic defects directly on a 3D model. This could be done by highlighting, for example through projecting, the white pixel areas in the difference images to the planes of the 3D model and hence creating a fully compatible geometric and aesthetic quality control process.

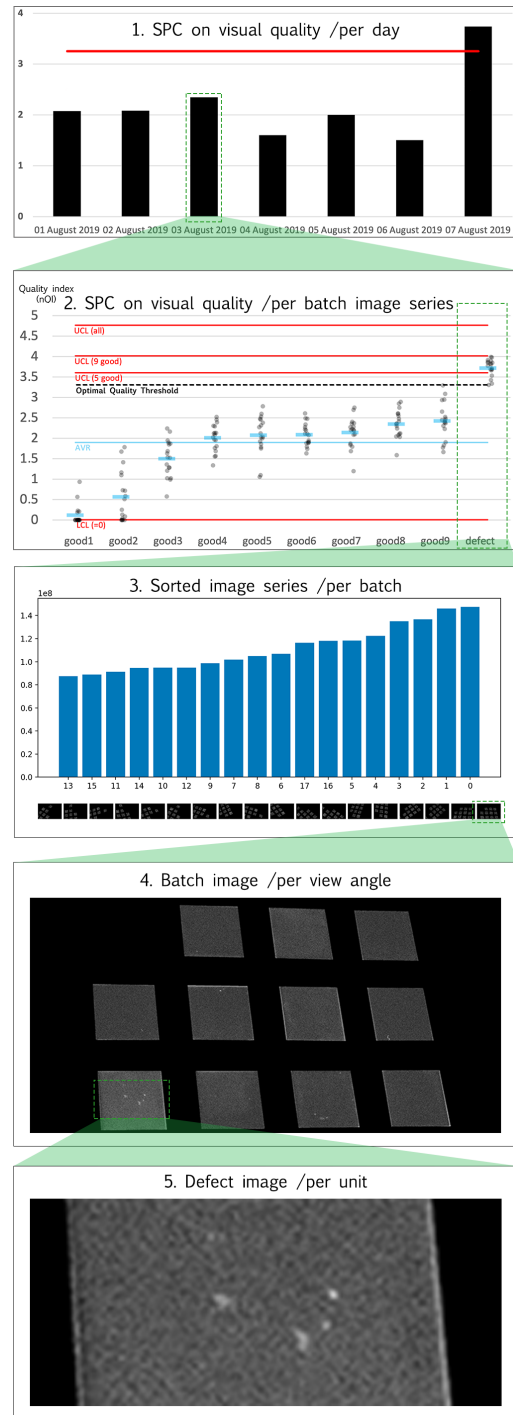


Figure 12: The SPC design proposal include overviews; 1) per day, 2) per batch image series, 3) sorting, 4) image per view angle, 5) defects per unit.

11. SPC Design Proposal

We propose an SPC design which can use enhanced image data for monitoring and controlling processes in AQC. The goal of our SPC design proposal is to assist in standardising the defect detection process by automating parts of the inspection process with machine vision and statistical pro-

cessing control. The nQI can be used to create an overview of aggregated data (e.g. examining and comparing quality over different days, see Figure 12, 1), reduce the image views in need of review (see Figure 12, 2-3), and evaluate a single image via image differencing and contrast image enhancement (see Figure 12, 4-5). Using image contrast enhancement, the nQI calculated for each image, and quality control limits assessors are able to navigate and zoom in/out on potential defects e.g. mouse clicking an image to enlarge it for full inspection.

All data (base data plus aggregated data) is logged and used for statistical processing (see Figure 10 and 12): 1) The assessors get an overview of all collected data per day. They can zoom in on specific dates to access data on all individual batch image series. 2) data per batch image series gives a clear overview of certain batches that need to be evaluated by human AQC (i.e. nQI data above the control limit (see Figure 12, 2, *defect*)). They can zoom in on individual batches to get access to the image data. 3) the sorted image series per batch. All different view angles for one batch are sorted and plotted based on contrast. Highest contrast sum shows most difference to the master and hence we are likely to unfold image views with higher contrast first. 4) we investigate the contrast enhanced difference image and search for potential defects. 5) we zoom in on individual defects and evaluate AQC. We consider a solution were it is possible to unfold data on the go (e.g. the data could be linked to a digital twin) for both process control and as assistance for human assessors. We propose to alternate between different data representations to achieve overview of the entire AQC process as well as individual data points (see Figure 10).

12. Conclusion

The existing visual quality control process of polymer surfaces in high-volume manufacturing can benefit from examining visual appearance joined with 3D metrology scans. Our SPC design proposal suggested a human-in-the-loop approach since we believe that assessors need a tool that can help in the existing quality control process. We did this by sorting contrast-enhanced batch images for human visual inspection and utilizing statistical methods to monitor and control the defect detection process of injection molded products. We have proposed a quality index nQI based on median master image differencing which works great for classification of defects in batch images series.

This paper presented an image-based test method for batch defect detection for semi-automated human visual inspection. Specifically, our SPC design proposal established a day-to-day overview of a nQIs over a batch image series in interdisciplinary cooperation with the principles of Six Sigma. Based on the inexpensive image series created from 3D metrology, we proposed image defect enhancement, batch image series sorting and a pipeline assisting statistical processing control limits to reduce labour costs within visual inspection. Our contribution contains contrast-based image enhancement in combination with image differencing with

a median master. Human-in-the-loop assessors can use this approach to inspect and compare several images; both individual series from different viewpoints and over time for several different batches. Our analysis showed that batch defect detection can be automated using machine vision to accommodate less biased quality judgments through Six Sigma control limits compared to a human AQC inspection process without our SPC design proposal. The precision and recall of defected views are high and was computed 100% (14/14) for precision while the recall was 78% (14/18). The assessors can employ our semi-automated SPC design approach as a tool in the existing visual inspection process, where the semi-automated approach can guide assessors and improve the use of objective quality measures in the human visual inspection process through a quality control feedback loop. The number of view angles in need of review by an assessor has been reduced by a factor of 13 and, consequently, created a tool that assessors can use to advance the existing visual inspection process.

Acknowledgment

This work was partially funded by the Manufacturing Academy of Denmark (MADE) through the research program MADE Digital supported by the Innovation Fund Denmark.

References

- [1] Baudet, N., Maire, J.L., Pillet, M., 2013. The visual inspection of product surfaces. *Food quality and preference* 27, 153–160.
- [2] Browning, R., Hossain, M.M., Li, J., Jones, S., Sue, H.J., 2011. Contrast-based evaluation of mar resistance of thermoplastic olefins. *Tribology international* 44, 1024–1031.
- [3] Chen, S., Lin, B., Han, X., Liang, X., 2013. Automated inspection of engineering ceramic grinding surface damage based on image recognition. *The International Journal of Advanced Manufacturing Technology* 66, 431–443.
- [4] Choi, J., Kim, C., 2012. Unsupervised detection of surface defects: A two-step approach, in: 2012 19th IEEE International Conference on Image Processing, IEEE. pp. 1037–1040.
- [5] Chrisman, J., Xiao, S., Hamdi, M., Pham, H., Mullins, M.J., Sue, H.J., 2018. Testing and evaluation of mar visibility resistance for polymer films. *Polymer Testing* 69, 238–244.
- [6] De Mast, J., Lokkerbol, J., 2012. An analysis of the six sigma DMAIC method from the perspective of problem solving. *International Journal of Production Economics* 139, 604–614.
- [7] Dekking, F.M., Kraaikamp, C., Lopuhaä, H.P., Meester, L.E., 2005. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media.
- [8] Deng, Y.L., Xu, S.p., Lai, W.w., 2017. A novel imaging-enhancement-based inspection method for transparent aesthetic defects in a polymeric polarizer. *Polymer Testing* 61, 333–340.
- [9] Eugène, C., 2008. Measurement of “total visual appearance”: a CIE challenge of soft metrology, in: 12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement, pp. 61–65.
- [10] European Committee for Standardization, 1999. *Geometrical Product Specification (GPS) – Surface imperfections – Terms, definitions and parameters, EN ISO 8785*. Technical Report. International Standardization Organization, ISO Standard.
- [11] Fleming, R.W., 2014. Visual perception of materials and their properties. *Vision research* 94, 62–75.
- [12] Gruber, D.P., Macher, J., Haba, D., Berger, G.R., Pacher, G., Friesenbichler, W., 2014. Measurement of the visual perceptibility of sink

- marks on injection molding parts by a new fast processing model. *Polymer Testing* 33, 7–12.
- [13] Hansen, A.J., Knoche, H., Moeslund, T.B., 2018. Getting crevices, cracks, and grooves in line: Anomaly categorization for AQC judgment models, in: 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), IEEE. pp. 1–3.
- [14] Hartley, R., Zisserman, A., 2003. *Multiple view geometry in computer vision*. Cambridge university press.
- [15] Hellier, C., 2001. *Handbook of nondestructive evaluation*. McGraw-Hill Professional Publishing.
- [16] Huang, S.H., Pan, Y.C., 2015. Automated visual inspection in the semiconductor industry: A survey. *Computers in industry* 66, 1–10.
- [17] Levitt, T., 1972. Production-line approach to service. *Harvard business review* 50, 41–52.
- [18] Liu, J.J., MacGregor, J.F., 2006. Estimation and monitoring of product aesthetics: application to manufacturing of “engineered stone” countertops. *Machine Vision and Applications* 16, 374.
- [19] Luiz, A.M., Flávio, L.P., Paulo, E.A., 2010. Automatic detection of surface defects on rolled steel using computer vision and artificial neural networks, in: IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society, IEEE. pp. 1081–1086.
- [20] MacGregor, J.F., Kourti, T., 1995. Statistical process control of multivariate processes. *Control Engineering Practice* 3, 403–414.
- [21] Maire, J.L., Pillet, M., Baudet, N., 2013. Measurement of the perceived quality of a product: Characterization of aesthetic anomalies. *International Journal of Metrology and Quality Engineering* 4, 63–69.
- [22] McCarthy, J.D., Sasse, M.A., Miras, D., 2004. Sharp or smooth?: comparing the effects of quantization vs. frame rate for streamed video, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 535–542.
- [23] Mera, C., Orozco-Alzate, M., Branch, J., Mery, D., 2016. Automatic visual inspection: An approach with multi-instance learning. *Computers in Industry* 83, 46–54.
- [24] Ng, H.F., 2006. Automatic thresholding for defect detection. *Pattern recognition letters* 27, 1644–1649.
- [25] Pinson, M.H., Wolf, S., 2004. A new standardized method for objectively measuring video quality. *IEEE Transactions on broadcasting* 50, 312–322.
- [26] Pointer, M., 2006. CIE TC1-65 - A framework for the measurement of visual appearance. Technical Report. CIE Publication.
- [27] Tolba, A.S., 2011. Fast defect detection in homogeneous flat surface products. *Expert Systems with Applications* 38, 12339–12347.
- [28] Tsai, D.M., Chen, M.C., Li, W.C., Chiu, W.Y., 2012. A fast regularity measure for surface defect detection. *Machine Vision and applications* 23, 869–886.
- [29] Tsai, D.M., Huang, T.Y., 2003. Automated surface inspection for statistical textures. *Image and Vision computing* 21, 307–323.
- [30] Xie, X., 2008. A review of recent advances in surface defect detection using texture analysis techniques. *ELCVIA: Electronic letters on computer vision and image Analysis* 7, 1–22.
- [31] Zsíros, L., Suplicz, A., Romhány, G., Tábi, T., Kovács, J., 2014. Development of a novel color inhomogeneity test method for injection molded parts. *Polymer Testing* 37, 112–116.