Review paper

http://doi.org/10.24867/JGED-2020-2-047

Tone Reproduction Curve: rendering intents and their realization in halftone printing

ABSTRACT

Approaches to determining the Tone Reproduction Curve (TRC) which provides the reliable transfer of visual information in typical conditions of the halftone gray scale compression in relation to dynamic range of a graphic original or input image file are overviewed. The issues of such curve realization are also analyzed with taking into account the specifics of multiple stages of illustrative printing technology.

KEY WORDS

Tone Reproduction Curve, gray scale, rendering intent, transfer function

Yuri V. Kuznetsov [1]
Andrey A. Schadenko [2]
Vycheslav V. Vaganov [3]

[1] St. Petersburg State Institute of Cinema and Television, Saint Petersburg, Russia
[2] St. Petersburg State University of Industrial Technology and Design, Saint Petersburg, Russia
[3] St. Petersburg State Polytechnical University, Saint Petersburg, Russia

Corresponding author:
Yuri V. Kuznetsov
e-mail:
yurivk@mail.ru

First recieved: 17.04.2020.
Accepted: 16.06.2020.

Introduction

The brightness range of a print image is many orders of magnitude lower than for the outdoor objects. It is also much smaller in relation of slides or digital camera files at their use as originals for print reproduction.

Problem of the shortened halftone scale is somewhat solved by digital High Dynamic Range (HDR) photography, where sections of the scene of very different average brightness after being captured at different exposures are seamlessly joined in the resulting image (Reinhard et al., 2010). However, these copies are not completely reliable and natural, because in the conditions of outdoor lighting, the observer isn't able to simultaneously perceive all the gradations of such areas. Unlike to such a copy viewing, he can do it just separately after temporal local adaptation to the brightness level of an each part of a scene.

Therefore, as with the origin of photography itself (Jones, 1920), the development of approaches to determining the form of TRC, which would provide the most reliable transfer of visual information in the conditions of inevitable compression of a dynamic range, remains the key problem in theory of print reproduction. At the same time, even after finding such a curve in the light of certain tone rendering intentions, the task of its practical implementation is no less urgent, taking into account the non-linearity of numerous stages of the printing process.

Tone reproduction curve

It’s difficult to get physically identical reproduction all of whose pixels have the same reflection spectra as the original or imaged object. Due to restrictions, inherent in any means of display, is not always achieved the colorimetric, metameric identity based on the physiology of vision. Therefore, most of the practice is necessary deal with the so-called psycho-visual matching of the copy and original.

In these circumstances, special importance for the regulation of customer and printer relations has itself statement the goal of reproduction. Color Management Standard (CMS) partly accounts this (Has, 1995), suggesting four variants of the so-called rendering intents. However, it is for clarity useful to consider such intentions in the simplest case of a monochrome, b/w printing (Hunt, 1998), where they are completely determined by the curve which links the values of such copy optical parameters as the reflection coefficient, optical density, brightness... with the same of the original.

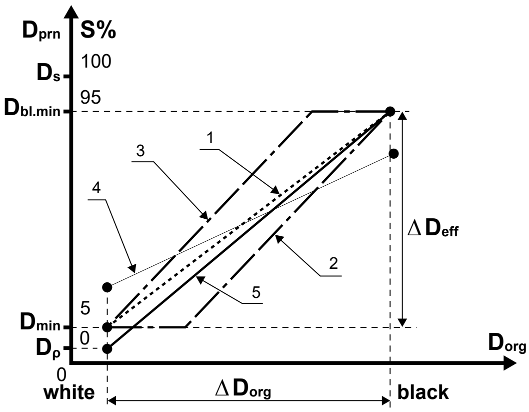
The first step in this characteristic generating is the finding of coordinates for its extreme points to match each other white and black levels of a copy and original. The next and no less important step is in finding the proper shape of the curve that will connect all the other values in effective, i.e. manageable interval of halftone printing. At last, there stay important the issues of a target TRC realization in multistage printing technology.

Range of the transferred gradations

Tonal content of CT originals is rather different. Reproduced objects may vary in contrast (ratio of ultimate reflectance values) or in density range (difference between the ultimate optical densities). The entire range may be also shifted left or right along the abscissa in figure1.

Paramount for the proper tones reproduction is possible precise alignment the interval $\mathbf{\Delta D_{org}}$ of an original with the effective print range $\mathbf{\Delta D_{eff}}$ (graph 1 in Figure 1). In the absence of such alignment, as can be seen from graphs 2 and 3, the tonal steps (details of brightness) are irreversibly lost in light (2) or dark (3) areas at the very beginning stage.

Assume that the original is a multi-step tone wedge and the reproduction task is the distinguished transfer of all its steps on the print. Then the line 3, passing at an angle of 45° from the origin will correspond the loss of some darker wedge patches, despite the absolutely accurate reproduction of the rest ones[1]. Much lighter, though with the same differences in density will be transferred just grey and dark patches of this wedge while following the graph 2. It is also obvious that the introduction of any nonlinearity in these graphs will not increase the number of distinguished patches and impact only on the lightness relationship between the ones, caught in the reproducible range.



» **Figure 1:** *Examples of combining optical density intervals (1, 5) and errors in setting the "white" and "black" levels: loss of brightness details in highlights (2) and shadows (3); decrease in overall contrast (4)*

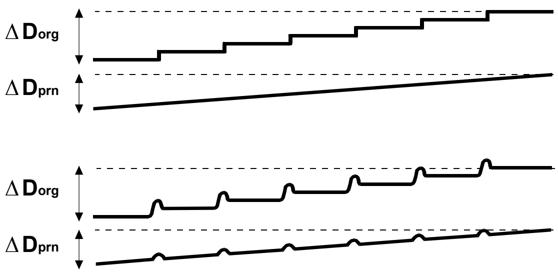
It is not always possible at least slightly extend this range, commonly reduced relative to that of the original, by assigning, as shown by graph 5, to minimal, “white” level of original the density $\mathbf{D_p}$ of substrate. In lithography, such assignment leads to loss of the brightest gradations. Halftone dots for the second, third and, may be for the fourth patch can occur smaller of the minimal printable one and the lightest reproduced patch will be the one whose density $\mathbf{D_{min}}$ corresponds to the size of “just printable”- reliably and consistently reproducible dot of, for example, 5%. The exception is, however, for originals containing the areas with specular reflection or self-illuminating objects for which, unlike given example, it is advisable assignment the Dp level of "whiter than white"- 0% value. Sun glare on the glasses can be transferred by clean paper because the rest of the image has no details of brightness or its differences with the values between the level of such glare and "white" of 5%. Otherwise, i.e. when assigning the latter to specular highlights, the contrast of the copy unjustifiably decreases.

The reproduction task is also not solved according to graph 4 with decreasing the image contrast when the differences between adjacent patches may become invisible. This case is additionally explained in Figure 2 (a, b), which shows such step wedge transform into degrade- continuous tone ramp on a print, which in essence distorts the meaning of the original data.

These general ideas underlie the assignment of print tone values to ultimate gradations of the reproduced original. Loading illustrations in DTP usually involves pre-scan of the original with rather low resolution just sufficient for its display on a monitor. If in areas with

[1] In terms of Color Management this option corresponds to the *Absolute Colorimetric* rendering intent and to Gamut Mapping strategy (Morovic, 2003) with *clipping* the out-of-gamut colors.

values close to extreme, the scanner does not distinguish gradations (graphs 2 and 3 in Figure 1), the operator repeats the scan with changing such parameters "brightness" and "contrast". Management "brightness" corresponds to the vertical parallel shift of these graphs while the "contrast" control changes their slope. With sufficient experience the appropriate levels setting, i.e. the positioning of original range between the extreme points of graphs 1 or 5 is achieved in a few iterations.



» **Figure 2:** *The steps of the original wedge (a), which are close to the brightness sensitivity threshold or print noise level, are lost on a print (b) as a result of the density range compression; illusion of such steps (d) can be created just by retouching the original (c)*

There is also possible the automatic finding of original extreme densities for assigning them the prescribed tone values by analysis for histograms of the pixel meanings occurrence in the pre-scan image file. This analysis allows, for example, for finding out how representative is the particular value for the reproduced object, i.e. can it be taken as a basis for the range setting or is a random outlier caused by a particle of dust or damage of original.

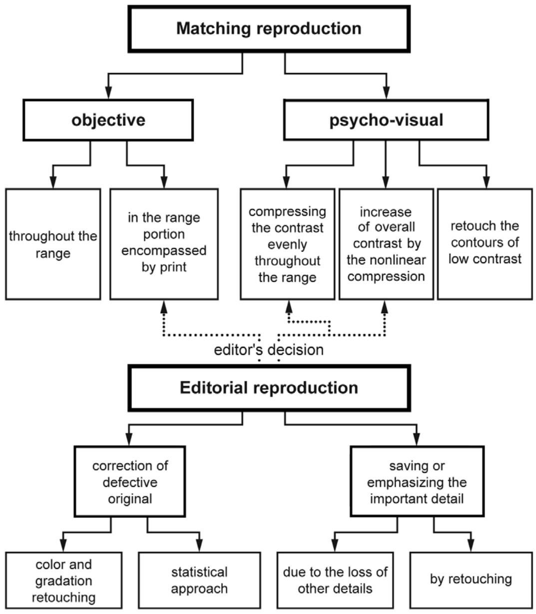
Form of tone curve

Next stage of creating a curve is the assigning to it a certain form within the above set range. Replacing camera by electronic scanner allowed for manipulating the TRC form highly flexibly and several scientific schools actively developed in the 70-ies the approaches to finding some optimal, universal curve suitable for "all cases" as the solution of cornerstone theoretical problem of print reproduction. For example, it was proposed (Namakura & Namakura, 1994) its analytical calculation accounting the substrate and ink film reflections, optical density of the reproduced detail, ratio of the density ranges on print and original.

However, such empirical equations can have only a limited application, since for the proper connection of the original and copy parameters is necessary, first of all, to determine the nature of reproduction goal. It may be rather different in such, for example, contradictory approaches as the matching original and editorial (with certain changes) one. However, the line between these, it would seem, fundamentally different approaches is blurred even when the original comprises not a natural scene or object but their intermediate representation as a hard copy or digital file.

For the most part, the observer judges the image reliability or quality in the absence of the original, i.e. from memory and according to personal experience, expectations or visual preferences. In addition, under the conditions of printing technology restrictions, the identical reproduction of certain property of an original makes impossible to match some its other parameter and vice versa. Therefore the mere decision about which of these properties should be saved on a print once again requires a certain editorial intervention.

The scheme in Figure 3 (Kuznetsov, 1998) differentiate approaches in the formulation of tone rendition intentions for only one of the attributes of color- its brightness.[2] Along with the figures 1 and 4 it points to interrelation of these options by the given below examples.



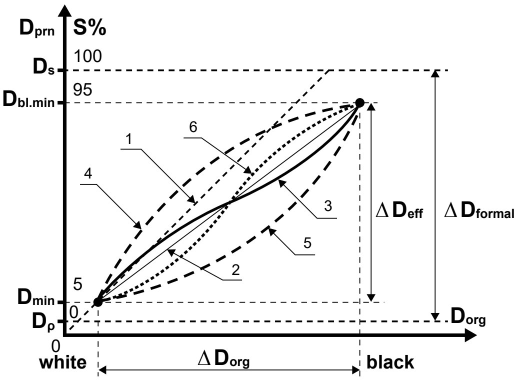
» **Figure 3:** *Variants of rendering intents within the original matching and editorial approaches to the tone reproduction goal*

Reproduction matching the original

Objective (physical, facsimile) matching takes place, when the brightness coming from original and print are the same for the viewer. Identity, in which the measured, for example, densitometer values of their corresponding areas are equal, is rarely met in practice and is possible just for the low-contrast, opaque originals, the range of which does not exceed the effective one of a printer. The graph under discussion here represents a straight line 1, coming at 45° in Figure 4, and correspond to reproduction variant 1 on scheme of Figure 3. In most cases, the interval of the original $\mathbf{\Delta_{Dorg}}$ significantly exceeds the interval $\mathbf{\Delta D_{eff}}$ of print and just the tone values that do not

[2] Applying to three attributes of color (brightness, hue, chroma), the ICC color management standard limits to cases of only four variants of such intentions (Morovic, 2003).

fall out of it can be transferred unchanged. At the same time all the brightness differences are clipped, according to curve 1, in deep shadows and matching the original as a whole is not achieved (case 2 in the Figure 3).



» **Figure 4:** *Tone transfer curves: 1 – objectively identical (facsimile) reproduction; 2 – linear, 3 - subjective-identical, (4, 5, 6) - editorial compressions of the density range of original*

In the conditions of range compression there is possible only the subjective identity, i.e. creation to certain extent possible the apparent similarity of two images, when all gradations acquire other meanings. Such cases are defined in options 3, 4 and 5 of figure 3. The aim is to ensure on a print the same relationship of lightness meanings which has place when looking at an object. If, for example, on an original step wedge the fourth and fifth patches differ from each other much more than seventh and eighth ones, in result of printing these differences should not be equal or in the opposite sense. Under this condition, the meaning of information is preserved, despite the change of patches values themselves and their between differences.

The most important is to preserve the image contours as reference data of image recognition. Priority of the correct contour strength and its geometry transmission over the brightness variation in stationary image area is confirmed through the whole experience of the visual arts, ranging from about the works of rock art and ending with the practice of retouching in photography and graphics. Therefore, following above example with a tonal step wedge the aim of reproduction can be interpreted as transmission of contours — brightness differences at the borders edges of adjacent each image areas.

In terms of optical density, finding the TRC form, which will better retain such differences, can be discussed with taking into account coarsening of the visual perception threshold at coming from lighter to grey and darker print areas from 0.02 up to 0,01 and 0,04 (Rabinovich, 1976).

It can be taken for example the aim of identical reproduction of a multiple, for example, 32 step greyscale with retention the visual differences between the all adjacent patches, i.e. when their tone values look equally meaningful. At assuming the 0.04 density differences between the neighboring patches, the entire original scale range comprises 1.28. Its linear compression at ¼ corresponds in figure 4 to curve 2 connecting the points ($D_{org.min}$, D_{min}) and ($D_{org.max}$, $D_{bl.min}$). Effective range on a print is of 0.96 units and all adjacent fields are different only 0.03 (instead of 0.04 on the original). As result, with the grey and lighter patches staying quite distinguishable the several darker ones will merge on a print copy.

Transfer of all 32 steps and, therefore, the informative matching of the print to the original can be achieved by increasing the curve gamma in shadows at the expense of the respective reduce it in the midtones, i.e. using the N-shape curve 3 related to case 3 in the diagram of Figure 3. This example shows that the linear (over optical densities) transfer is inaccurate and only adds to the loss of information, collateral to the tone range compression in printing.

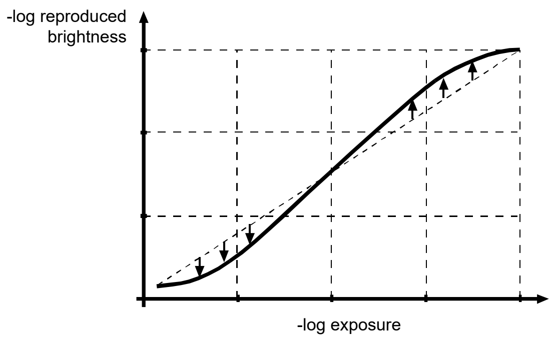
If the number of fields or contours on the scale is large, more than for example, sixty-four, then the brightness differences that form these contours are extremely small already on the original and are close to the threshold contrast sensitivity of vision or level of the process own noise. So, even the use of mentioned N-shaped curve will result in this case, to the disappearance of all steps of the scale (Figure 2, a, b) and it will look on the print as a continuous tone ramp ("vignette"). The complete absence of steps within it says the linearity of the entire reproduction system. However, thus reached the linear lightness transfer retains only one part of the semantic image content but the information about the presence of steps or contours there between are lost.

Save the situation in this, exaggerated example is possible, as already noted, just by retouching — introducing into the original or representing it signal of pre-emphasis to compensate for the reduced print range (option 5 in Figure 3). It can be, for example, the slight dividing lines drawing between the patches on an original (Figure 2, b, d). A similar result may give the unsharp masking, organized in such a manner as in the greater extent influence the weak tone steps than anyway saved the larger ones (Gast, 1981).

Tone rendition dilemma

Due to mechanism of brightness adaptation the flower stays looking white when taken indoors where its brightness decreased in the hundred or thousand times but there are no other, lighter reference objects. At the same time, the certain incompleteness of such adaptation creates the dilemma in subjectively identical tone rendering (Giorgianni & Madden, 1998). In spite of preserving lightness relationships along the whole input grey scale on a print, the above concerned subjective matching is accompanied by reducing the overall, visually perceived contrast of an image. Despite the relative nature of brightness perception, the copy, preserving the greatest amount of a small local contrast details, looks faded and its colors unsaturated. Therefore in photographic, television and other systems the brightness, especially of natural scenes, is reproduced very non-linear.

As the exemplary curve in Figure 5 shows, under conditions of a range compression, the compensation of adaptation incompleteness is achieved by the redistribution brightness logarithms relationship with reducing the gradient in the lower and upper parts of a curve. Image contrast increases because its light areas become lighter and dark ones darker. However, the brightness differences are reduced or completely disappear there.



» **Figure 5:** *Because of the incomplete brightness adaptation, when looking at a natural scene reproduction, the appropriate overall contrast is achieved by the use of S-wise curve, thus losing tone steps in the light and dark areas*

Such a transformation and associated losses may have place in the manufacture of illustrative original in its material or electronic form due to the S-shaped sensitivity curves of the photo film or paper, scanner, TV and digital cameras. However, in the graphic arts reproduction may be necessary the additional compensatory control. This happens when the original is a slide with a density range two or more times greater than that in print and initially oriented, for example, on viewing its projection in a dark room. Insufficiently nonlinear, with respect to the overall contrast preservation, can be the lightness of a natural scene or in a file of computer graphics intended to be displayed not on a print but with the use of some other means, the brightness of which is considerably higher than in standard print viewing conditions.

In units of optical density, an example of such a complement compensation the lack of general brightness adaptation (case 4 in the diagram of Figure 3) can serve S-shaped curve 6 in figure 4 inverse to N- shaped curve 3. Wedding photo, taken in bright sunlight, might look quite contrast and saturated if made using an S-shaped characteristic. But it is unlikely to reproduce all the folds of the bride white dress and groom's black suit. Conversely, the use of N- shaped curve, which safes all the lightness differences within the tonal range, may make an impression that the event took place on a cloudy day. This is the dilemma of the very formulation of the tone rendition goal.

In the light of above, the choice of matching reproduction variant within the printing process limitations is ambiguous, stays beyond the technology bounds and requires creative intervention of an artist or editor.

Editorial tone rendering

Editorial approach to image tone rendering puts aside the demand of its matching an original. The TRC form is set in attempt to achieve the identity of a copy not to original itself but, for example, to the imagined by operator, technologist or editor visual object which the original copy is only trying to convey (variants 6, 7 in figure 3). This may be the case when the original is clearly inferior due to improper lighting of scene, errors in film processing, long term or poor storage conditions and other reasons. This may also happen to original in form of the digital photo file as a result of inadequate interpretation of its meanings values.

The exemplary characteristics 4 and 5 of such transmission are shown in Figure 4. The first of them allows for correcting the faded "underexposed" original and the second brightens "overexposed", too dark one.

Such an approach to the optimal reproduction is sometimes construed more widely, on the assumption that the vast majority of originals are in essence a distorted, subjective replica of the visually perceived world, at least for the reason that to their production the man had a hand (Ovchinnikov, Fainberg & Litvan, 1971). Based on this premise, they consider the best gradation curve, which would lead to lightness distribution, in the so-called information component of an image, to the normal, Gaussian law inherent in this world. The distribution deviation from such law on the original is taken into account only as confirmation of the validity of this approach.

To another case of correcting the image tonal content can be related the deliberate distortion of lightness distribution of the original, even when its quality does not cause claims. The higher gradient is attached to the curve portion relating to brightness detail, which, according to the mind of customer, operator or art editor, is the most important for this scene. Local contrast of some detail can be strengthened disproportionately to contrasts of others, if, for example, in the illustrated text the author pays special attention to it. The tradeoff of such an accent again brought the contrasts of the details of other, reducing their slope curve areas as far as to strengthen or at least save contrasts of important (informative) details is possible only by reducing the gradients for other ones.

As a case of such kind editorial rendition is the curve 2 of linear optical densities compression in Figure 4. The steps in midtones are disproportionately transferred due to the mentioned in above example their loss in shadows.

It should be also noted that the interrelation of such non similar quality parameters of an image as sharpness and overall contrast still stays out of the set of tone rendition variants in Figure 3 scheme. It's recommended to preserve the image naturalness by reducing contrast with lowering the TRC gamma up to 40% at the growth of the sharpness (Field, 2001).

Tone value transfer

Interrelation of transfer curves

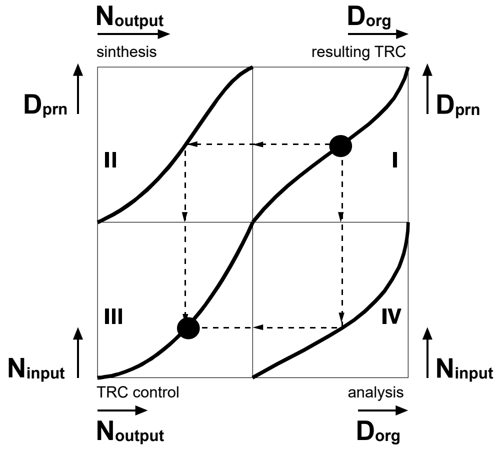
It has been shown above that the tone rendition problem is in its very statement ambiguous and the compromise editorial decision is necessary. However, for whatever reasons it was set the shape of resulting TRC, the following inevitable and purely technological stage is its practical implementation.

Illustrative original-to-print transformation is multistage. Each the step, whether it be electro-optical analysis, transparencies recording, plate making and printing itself, has its own transfer characteristic, linking the input and output values. Their product results in the whole process TRC and the issues of its realization should be discussed with accounting the characteristics of these multiple stages.

The generalized tone reproduction scheme is presented in Figure 6 by four interrelated curves with the target TRC- $\mathbf{D_{prn} = f_1(D_{org})}$ in its first quadrant.

In the third quadrant is located curve $\mathbf{N_{out} = f_3(N_{in})}$ of flexible control of gradation converter. Its dialog box gives it by default as a straight line running at 45 degrees and corresponding to the unchanged transfer of some of the 256 input and output tone scale quantization levels. However, the resulting graph of the first quadrant may be quite different without full compensation for the nonlinearity of the characteristics of other stages.[3]

Their performance is conventionally summarized in this diagram in two groups. The first of them includes a sequence of nonlinear transforms the optical density of original to the quantization level of the input signal entering the gradation converter: $\mathbf{N_{in} = f_4(D_{org})}$. It accounts the scanner characteristics and a number of other stages associated with the signal digital coding and perceptually uniform presentation.



» **Figure 6:** *Generalized scheme of tone transmission which shows that the managing conversion of an image signal (III) must consider the input (IV) and output (II) characteristics of a system.*

The other group is represented in the second quadrant as the generalized characteristics of the synthesis-link of the converter output signal with the print density: $\mathbf{D_{prn} = f_2(N_{out})}$. It combines the results of nonlinear transformations accompanying the screening, transparency recording, platemaking, printing and continuous perception (descreening) of a halftone when viewing.

The diagram clearly shows that the control function of the third quadrant is a derivative of the other three ones which is indicated by the lines of one of its dots finding. So, accounting the characteristics of all other stages it subject to mandatory correction when changing any of them.

In more detail the process of tone transmission reflects the multi-quadrant scheme in Figure.7 the quadrants 1 and 7 of which are identical to 1 and 3 ones of Figure 6 but the rest ones reveal the generalized characteristics 2 and 4 of the latter. By analogy with the previous scheme, subscripts of the transformation functions **f** coincide here with the corresponding quadrant number. An image capturing step, where the optical densities of original are converted into voltages or currents proportional to the luminous flux, and, consequently, the reflection coefficient **ρ**, is given in the fourth quadrant by the link of the level number of a linearly quantized signal with density of an original $\mathbf{N_\rho = f_4(D_{org})}$. To go from the $\mathbf{N_\rho}$ signal proportional to reflections, to the signal $\mathbf{N_d}$ of more visually uniform optical densities, the logarithm transform is used schematically indicated by curve $\mathbf{N_d = f_9(N_\rho)}$. The sixth quadrant presents further the characteristic $\mathbf{N_{in} = f_6(N_d)}$ of signal pre-distortion to compensate for the dot gain.

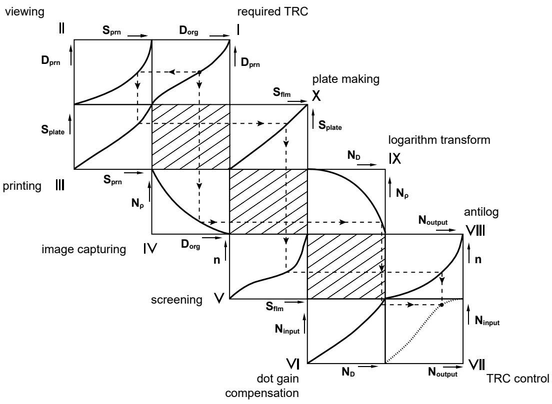
At following from the ordinate axis of the first quadrant to the left and further along the scheme, it should be reminded that the print density is related to the

[3] Some of the prepress programs include, for example, the Dot Gain Compensation for such compensation. In the context of the Color Management System discussed below, a similar task is assigned to the so-called Color Profiles.

area **S** of its halftone dots (tone value) in quadrant 2 ($\mathbf{D_{prn} = f_2(S_{prn})}$) logarithmically, namely, according to the Yule-Nielsen equation or other model of halftoning.

Curve $\mathbf{S_{prn} = f_3(S_{pl})}$ of the third quadrant is the print characteristics connecting the dot areas on a print and plate. Next, also nonlinear one is the function $\mathbf{S_{pl} = f_{10}(S_{flm})}$ of the dot area transfer from transparency onto a plate.

Neglecting the non-absolute nature of the ink film absorption and paper reflection, it can be assumed that the relative areas on a print are essentially its absorption coefficients connected, by definition, with uniform (proportional to density) output signal $\mathbf{N_{out}}$ by antilogarithm law. Because in most modern software and hardware screening the halftone dots are formed by even smaller elements (microdots), this is accounted at the halftoning stage. Therefore in this scheme the antilogarithm procedure is presented in quadrant 8 by the nonlinear connection of amount **n** microdots forming the dot with the level number of signal $\mathbf{N_{out}}$ controlling the halftone dot generator: $\mathbf{n = f_8(N_{out})}$.



» **Figure 7:** *Interrelation of tone transfer curves in printing.*

Essentially nonlinear in practice, the relationship $\mathbf{S_{flm} = f_5(n)}$ between the microdots amount n and therefrom formed halftone dot area Sflm in mechanicals recording displays the fifth quadrant.[4] Curve $\mathbf{N_{out} = f_7(N_{in})}$ of the converter, solving the gradation crucial task in quadrant 7 accounts the characteristics of all the other stages which is once again demonstrated by one of its points finding.

Step-by-step examined above the relationship of nonlinear image signal transformations allows to clarify their in-principle distinction. So, if the curve of quadrant 1 defines the goal of reproduction, then, being a derivative of all the others, the characteristic of a flexible, control function of gradation converter in quadrant 7 acts as a means of its destination.

A completely different role is played by all the other nonlinear transforms that make up the third, their most extensive category. For the most part, they reflect the relationship of output and input values inherent in specific physical nature of the various technological stages. Their modes are set and optimized according to their own criteria, and not in order to obtain some predetermined shape of the whole process TRC. For example, the printing of a halftone is adjusted by the criterion of ensuring an effective tone range. When recording a halftone dots on film the most important is achievement of high copying properties of the resulting transparency for platemaking, etc. Being associated with the properties of materials and equipment and are not changed at the will of the operator, i.e. not performing TRC control function, these characteristics nevertheless, have a direct impact on the final result and are therefore subject to strict checking. In a continuous or lengthy cycle, they are monitored periodically, but each time when changing the consumables or equipment. Such technology stages should be stable. Only if their parameters are maintained within the specified tolerances, the prepress is effective in controlling the final result. For example, at the last stage, the required stability and repeatability is provided on a press by the visual, instrumental or automatic evaluating the prints and their control scales.

It is also seen from this scheme that each step, having an individual nature, contributes to the resulting transfer. The nonlinearity of one stage, having the same or opposite shape (convex, concave, N-shaped, S-shaped, etc.) may enhance or compensate for the nonlinearity of the other.

In CMYK printing the achromatic scale is also greatly influenced by the applied strategy of the fourth, black ink use within such prepress options as Under Color Removal- UCR, Gray Component Replacement- GCR, Under Color Addition – UCA (Kuznetsov & Ermoshina, 2016).

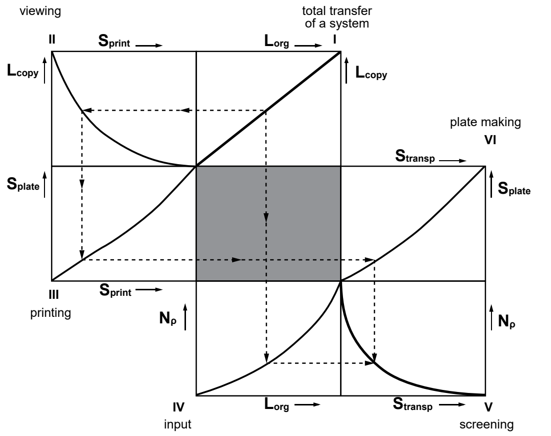
Transfer characteristic of the system

Of particular interest is the end-to-end characterization of the system itself as a cumulative result of all stages. With taking it linear in units, for example, of the Lab lightness L, the other stages effect on this linearity can be estimated according to the scheme in Figure 8. It differs from the previous one in that it includes only the characteristics of the materials using stages while those associated with the signal conversions are omitted.

According to this scheme, it can be seen that with fixed parameters of other operations, the overall linearity of the system can be provided only by an appropriate selection of the screening characteristics placed in the fifth quadrant- the connection of halftone dot area with the image signal value. Modern tools allow for flexible specifying it analytically, by the alphabet of halftone dots and other methods in the RIP, digital halftone generator or in the corresponding module of a computer program.

[4] Such function stays as well nonlinear in the CTP and various kinds of direct digital printing.

If the initial, at least approximate, overall linearity is achieved by the process stages agreement, then such corrections as, for example, shown in quadrant 6 of Figure 7 for Tone Value Increase (TVI) compensation, will be trimming, accounting only the transfer nuances of other steps. Otherwise, deeper nonlinear transformations of the discrete (quantized) image signal may be necessary, which are fraught with interference and loss of information.



» **Figure 8:** *Screening stage characteristic (quadrant* ***V****), linking reflection* ***Nρ*** *of the original with tone value* **S***, is responsible for the TRC linearity (quadrant* ***I****) of a system. (In the CtP technology the quadrant* ***VI*** *is excluded of this scheme, while in digital printing -* ***VI*** *and* ***III****)*

Contouring capacity

In addition to the overall linearity is also important the number of responses- distinguishable steps of the Human visual System (HVS) contrast sensitivity thresholds or amount of just noticeable differences (JND) in the reflection copy viewing conditions. It, according to various estimates, is from 80 to 100. However, printing, as a rule, does not provide such a number of steps because of its own noise caused by the substrate roughness and fluctuations in the ink transfer. In the letterpress newsprint it is, for example, difficult to get a step wedge of more than 16 patches. However, it is obvious that the configuration of any printing system should provide the maximum possible their number.

Another obstacle to the tone steps transfer may comprise the second, after analog-to-digital conversion, tone scale quantization at the screening stage caused by the halftone dot discrete formation from microdots. The number of these quantization levels depends on the screen cell dimensions and is directly related to the output resolution of a film/plate recorder or digital printer.

It was reminded (Kuznetsov, 2019) that the basic criterion for substrate-ink-plate system optimizing is in providing of minimal possible, stable printing element. Taking into account its size and effective tone range, the screen ruling value is set (Kouznetsov, 1999). The tone rendering within this interval is affected by the TVI, which depends not only on the ruling, but also on the screen geometry and is determined by the degree of dot area distortion when transferring its image, idealized by a computer program or a bitmap, to a print.

The most universal optimization criterion of the entire system behavior in illustrative printing can be, in this respect, the so-called contouring capacity $\mathbf{K_c}$ - the number of combinations of two in **N** steps of the *ultimate step wedge*, possible for a certain printing: $\mathbf{K_c = N (N - 1)/2}$ (Rabinovich 1976, Kuznetsov 1998).

The number **N** is limited by the noise level of printing. Theoretically, the *ultimate* can be considered a wedge of **N** patches, if an attempt to supplement it with one else distinguishable patch by reducing the differences between each pair of the others by **1/N**, converts it into a *vignette* throughout whole tone range.

In contrast to previous years, the creation of such wedges with freely adjacent patches and with an arbitrary step in up to 256 values of the uniform signal has become possible today thanks to the prepress computerization. Printing of such scales is not only the effective means of controlling the optimal adjustment of the screen-plate-paper-ink-press system, but also an operational visual indicator of the degree of its linearity, normalization and stability. Any setting deviation leads to the merging of more or less patches on a particular scale area. So, if the system noise increases as, for example, a result of using less smooth paper, some patches are perceived conjoint, but the other separate, this indicates the need for use of corrective "profile" at renewed print settings. If all the patches merge in the new conditions in a vignette, then there is reason to assume that the system remains linear.

The evaluation of small tone steps in the brightness logarithms, optical densities or the CIE Lab lightness is rather approximate and may not match the visual perception in the specific viewing conditions, since these metrics, in themselves, are not known to be absolutely uniform. It may happen that the Lab lightness distribution measured for the "ultimate", and, consequently, linear and visually uniform scale is nonlinear. However, this will just indicate the inadequacy of the used metric to the task. In such a case the calculation formula for L or other attribute of color can be replaced with a new one which will account the printing specific and such particular viewing conditions as the type of light source, illumination intensity, image background, the overall level of brightness adaptation …

Uniform representation of the image signal

A signal, formally proportional on the sensor output to the light flux, and thus the reflectance or brightness, is used to be transformed at input into the perceptually uniform one either by logarithms, or by CIE non-linear empirical transforms. Naturally, this takes also into account a certain nonlinearity of the light characteristics of an input device. Because ultimately, at the system output, there is used the signal of ink amount (dot area), so again associated with reflections (this time of a print), the need for such a transform may seem, at first glance, not obvious.

The representation of an image signal with equal contrast, i.e. by the values most closely related to visual perception, makes it more resistant to interference that inevitably accompanies any functional transformations of such a signal. In analog part of imaging channel the noise is generated by the dark current of a sensor, fluctuations of the electronic circuitry p-n transitions, and in digital – due to the quantization errors in analog-to-digital and digital-to-analog transforms. Such errors also accompany the screening with halftone dots discrete formation by microdots.

It is known from signals theory that the concentrating noise or errors of communication channel effects most damaging the data. For an image, this means that at the same resultant power, they may be most noticeable if they are concentrated in some its part, in one of the areas of tonal range or band of spatial frequency spectrum.

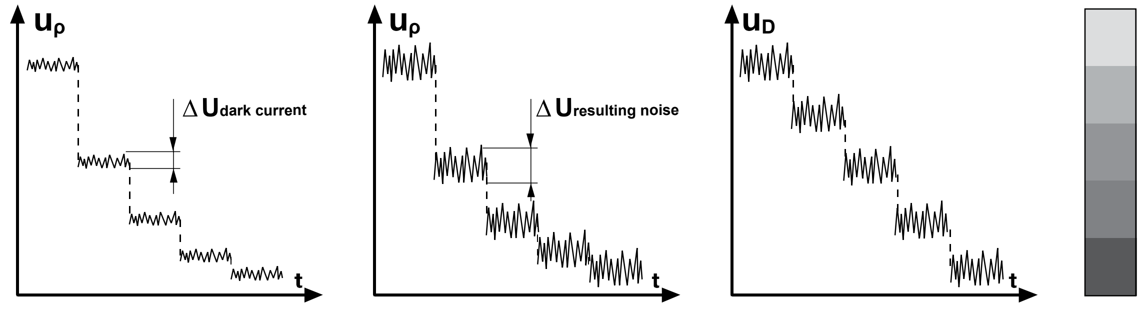
Noises affect the image additive, summing up those previously accumulated into each of the subsequent stages. Therefore, the signal is lead to uniform values at the very beginning of the channel, where the noise level is still relatively small being determined just by the dark current of a sensor. Effect of total process noise on the image, represented by such a signal, is uniform in the reproducible range and less affect the quality of resulting copy. Otherwise, it is accumulated in the shadows and leads to a noticeable loss of gradations as illustrated in Figure 9 by the logarithm processing the analog signal of a scanner. Figure 9 (a) shows a form of the voltage at the sensor output for a visually uniform step-wise tone scale (the adjacent patches of which the same differ in optical density in Figure 9 (d). Steps of this curve are in contrast very uneven, since this voltage is directly related to the patches reflection and not with their densities. Part of gradations, as conditionally shown in Figure 9 (b), will be "drown" in noise, the level of which can significantly increase as a result of the accumulation of noise accompanying subsequent functional transformations of a signal. And vice versa, from Figure 9 (c) it's seen that for the same resulting noise all the scale patches are equally discernible, if the signal is initially subject to logarithm conversion. Noise is uniformly distributed in this case over the tone range and, despite the same rate, its effect less negative.

Actuality of the brightness values conversion into optical density or lightness grows with the tone range increase of an original. Therefore, the input signal entering the analogue-to-digital converter (ADC) is usually of 10 and more digits. Reserve of bits (quantization scale steps) in the nonlinear conversion is needed to keep informative all the 256 levels of standard eight-bit output signal. In the scanner, equipped, for example, with 30 bit ADC, each of the captured RGB signals is quantized on a scale of the $2^{10} = 1024$ levels. On the part of tone curve transform with a gradient of 4/1 each of output eight-bit values is, thanks to this reserve, corresponded to at least one level of an input signal. Professional scanners can use the 16-bit ADC in each color channel to process the originals of up to 4 density units.

Characteristic of the screening stage

In Figure 7 it's presented by quadrants 5 and 8 comprising together the function $\mathbf{S_{flm} = f_{5,8}(N_{OUT})}$ which connects the dot area $\mathbf{S_{flm}}$ and signal $\mathbf{N_{OUT}}$ recording the halftone transparency.

The halftone dot in most cases is formed from discrete microdots exposed by the output device. If to assume that its area is directly determined by the number **n** of



» **Figure 9:** *Forms of* ***Uρ*** *and* U_D *signals for visually uniform tone scale (d): at the sensor output (a) and at the end of a channel without (b) and with the use (c) of logarithm transform* such sub-elements, it appears that it linearly expresses the print absorbance/reflection and not optical density or lightness values of a uniform output signal $\mathbf{N_{OUT}}$. Therefore, the screen mesh of, for example, of 4x4 microdots can formally transmit 16 plus one, corresponding clean paper, gradations. However, with its use of the sixteen visually uniform steps with considerable distortions of contrasts will be printed just 10-12. The rest, little differing from each other, will merge in shadows. So, the relationship between dot area and uniform signal needs to formally submit antilogarithm law what is indicated in quadrant 8 of Figure 7 as function $\mathbf{n = f_8 (N_{OUT})}$.

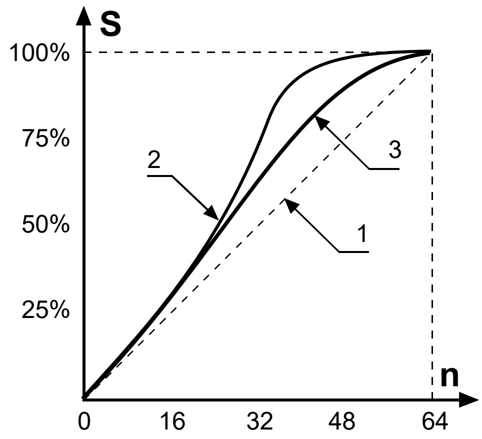
For this reason, it is not correct to calculate the number of reproduced tone levels according to the size of matrix or screen cell. Such widely recommended calculation (Kipphan, 2001) as the square of the output device resolution (**Dpi**) and screen ruling (**Lpi**) ratio: $\mathbf{(Dpi/Lpi)^2 + 1}$, lacking the logarithm or the cube root extraction, looks rather confusing.

And nevertheless, the question may arise: why is the dimension of matrices used in practice, although it exceeds the number of reproducible reflection coefficients, but still not as much as it suggests the reverse, for example, antilogarithmic, transition from the signal of optical density to halftone dot area on the print? After all, if direct conversion has assumed at the input, for example, 256 times the margin by the number of levels (to 2^{16} levels in ADC compared with 2^8 in the image file), then the output would be logical to use a screen cell size 256 x 256 = 65536 = 2^{16} microdots?

Several factors contribute to the practical limitation of the output record resolution and thereby the number of dot fonts of the halftoning alphabet.

First, the number of brightness levels, distinguished in print viewing conditions does not over of 100 and gray scale of a print rarely exceeds 1.8 density units (against to 3-4 units on slides). Second, even on very smooth papers because of the process noise are not achieved even 64 distinguishable steps. Therefore the uniform eight-bit signal, with its 256 meanings, enables in this relation the almost a fourfold margin.

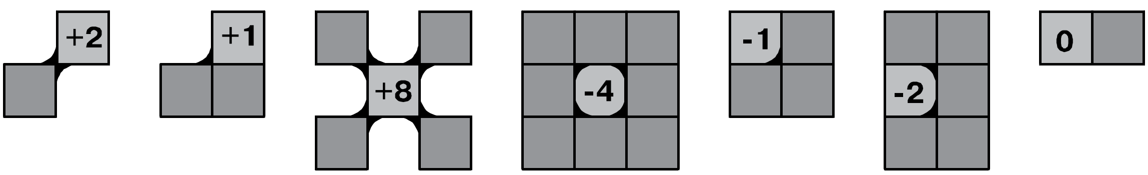
Another to be taken into account objective dependence is in the nonlinear connection of resulting tone value (dot area) with the number **n** of microdots forming the print elements in a regular screen or their conglomerates in its FM, stochastic version. This proportionality failure of $\mathbf{S_{film} = f_5(n)}$ function was first mentioned at the very initial steps of electronic halftoning (Hallows & Klensch, 1968). It takes place at the stages of a latent image formation in the electrophotography as well as in halftone transparency recording amplified thereafter at the plate making and especially in printing. It's explained in Figure 10 by the example of randomly filling the 8x8 matrix with 64 assumedly square microdots.



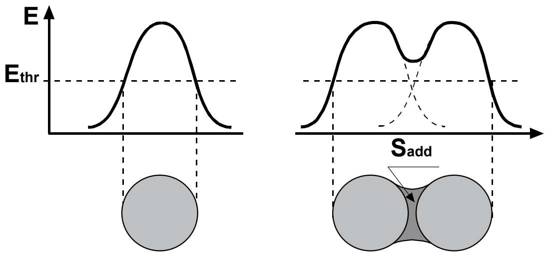
» **Figure 10:** *Proportionality failure between the resulting ink coverage **S** and the number **n** of its constituent microdots (1 – formal; with the random 2 and in cluster 3 microdots placement)*

While the amount of microdots is relatively small, they are located mostly in isolation of each other and at first the curve is linear, because each added sub-element brings an equal value to the resulting coverage. But even with an ordered fill, this happens up to a tone value not exceeding 25%. The rate of coverage sharply growths as soon as the microdots begin to touch each other. In their contact zones, as shown in figure 11 (a, b, c), there are formed the additional, not geometrically provided areas. Their number depends on the amount of simultaneous contacts of introduced microdot (shown lighter in Figure 11) with neighboring, previously set ones and is accompanied by the formation of one (Figure 11 b) and sometimes of eight (Figure 11 c) such areas.

Formation of such extra area has the different physical nature. When recording onto a film or on sensitive layer of a plate, it is due to the summation of light scattered



» **Figure 11:** *Ink fill-in with the formation of unsolicited positive (a-c) and negative (d-f) coverage (marked dark) in different variants of mutual positioning the ideal, assumedly square microdots* at the edges of neighboring exposed elements (Figure 12, b) while for isolated, taken separately elements this light stays insufficient for blackening (Figure 12, a).



» **Figure 12:** *Summation of exposure in the halos of neighboring elements leads to the formation of an additional area in the film or plate recording*

Additional areas occur further in the press due to the action of surface tension forces in the ink layer, pressure and other reasons.

In the electrophotography, for example, in a copier, laser or LED printer, the formation of these areas, in addition to the marked summation of exposures, is due to the spreading of the charge within the optical photoconductor (OPC). Sharp fractures of the electrostatic field equipotential lines in the zones of microdots contact are accompanied by an excessive increase in electric field strength, which is offset by the charge leakage through the OPC finite resistance.

For darker gradations, the rate of coverage growth is reduced because, as Figure 11 (d, f, g) shows and the shape of curves 2 and 3 in Figure 10 reflects, still blank areas already include the filled-in portions occurred when touching the elements earlier, for lighter tone values. So, the introduction of each new microdot gives here an increase smaller of its own area. As result the extra area with a single microdot adding depends on position of the latter in relation to others and ranges from +8 to-4 such unit areas. Each of them is absolute, i.e. not related to the size of a microdot, and therefore its share becomes greater at higher resolution thus strengthen the proportionality failure of curves 2 and 3 in Figure 10, i.e. their difference from the theoretical graph 1. The latter is true only when using such large microdots, compared to which the surplus areas are negligible.

The shape of curves is somewhat opposite to the logarithm characteristic. This partly explains the possibility of uniform transfer of the original values with a fewer number of microdots in the screen cell than it follows from purely geometric considerations.

Effect of the surplus inked areas is especially pronounced in stochastic screening (curve 2 in Figure 10) making the tone value close to that of ink solid (100%) at filling the matrix just by 70 – 80 % of its microdots. This explains the relatively low printability and tone responses ($\mathbf{K_c}$) of non-periodic halftoning as compared to its regular counterpart and, as result, rejection of the former by wide printing practice in spite of the great number of scientific publications and promotion efforts of last decades.

Forming additional areas was discussed above as applied to a simplified variant of filling the orthogonal matrices by conditional, square microdots with a uniform distribution of ink or radiation. In practice, their shape is closer to rounded and in addition to touching, they somehow overlap each other when positioned the orthogonal lattice idealized in an image output bitmap. Such resulting coverage dependence from of the neighboring microdots location is accounted in the analytical modeling of halftoning (Carrara, Analoui & Allebach, 1992; Allebach, 1994).

The non-linear transformation of uniform eight-bit signal to proportional the brightness or reflectance is not used in prepress software options. During the transition to tone values (a relative ink coverage areas — CMYK) the 256 levels of this signal are usually linearly converted to percent. Antilogarithm transformation is provide by the threshold function connecting the number of **n** microdots in bitmap dot font with the level of $\mathbf{N_{OUT}}$ signal in a RIP or printer driver (quadrant 8 in Figure 7). To achieve the maximum possible contour printing capacity, this function should also account other nonlinearities discussed above, for example, related to the shape, order of microdots positioning in a screen cell…

Optical and physical dot gain

Even more complexly the geometry of elements mutual positioning effects on the optical dot gain (Gustavson, 1997). In the Figure 7 it's formally taken into account in the second quadrant.

Unlike this apparent dot gain the physical Tone Value Increase (TVI) means the extension of size of a halftone dot at transfer it from transparency onto a plate and further to substrate.[5] For regular screens, the maximum of both optical and physical TVI occurs when adjacent dots are touched, i.e. in medium and darker tones. In non-periodic halftones such a touch is more or less common throughout the gray scale and the maximum is shifted to a lighter area. In the most general case, the TVI is considered as an ink coverage change in the entire technology chain: bitmap-film-plate-paper. It is obvious that in the CtP and digital printing from this chain fall respectively the second and third links. Curve $\mathbf{N_{IN} = f_6(N_D)}$ in the sixth quadrant of scheme 7 collectively compensates for these nonlinear distortions, denoted for platemaking and printing stages in quadrants 10 and 3.

Causes of physical dot gain are in the light scattering in transparency recording, its transfer onto a

[5] In the second edition (ISO, 2004) of ISO 12647 it's recommended to replace the term *dot gain* by *Tone Value Increase -TVI*.

plate, pressure in the printing contact, filling-in... These "natural" factors may be supplemented by the avoidable causes such as slur and doubling associated with improper press adjustments.

Specifying the TVI (ΔS) by a single number indicates the difference between the tone value on print from the 50% on a plate, film or in digital file. In sheet fed offset printing, the TVI reaches 15% (at 150 Lpi) and 20% (at 135 Lpi), respectively, for coated and uncoated papers. In web offset at 135 Lpi it is 22% while for newspapers at 100 Lpi achieves 30% (Southworth, McIlray & Southworth, 1992). If it is written in two numbers, they correspond to the similar differences for the margins of 40% and 80% of standard print control wedges. However, the most complete picture gives a TVI curve $\mathbf{S_{PRN} = f(S_{PL})}$ in the third quadrant of the scheme 7. The position of its maximum depends on both the dot shape and screen geometry. If the square elements are staggered, the greatest TVI occurs at about 50%. If they are round, the maximum is shifted to 79%. With irregular placement of fixed size dots, the same maximum is shifted to the light (25%- 35%) region.

One of the operational parameters of densitometry monitoring this phenomenon is the *print contrast* **K** measured on the control scale for solid ink layer ($\mathbf{D_S}$) and patch of 80% ($\mathbf{D_{80}}$): $\mathbf{K} = 100\% \cdot (\mathbf{D_S - D_{80}}) / \mathbf{D_S}$. Its zero value indicates the complete absence of a gap between the dots of 80%. If in such case the ink solid density greater of norm the cause may be in the excess of ink supply. Otherwise, the reason for this control parameter decrease should be found in the excessive pressure, appearance of slur or doubling, etc.

Above were considered the characteristics of the "material" stages associated with the production of mechanicals, plates, proofs and prints themselves, as well as the principles of accounting for their specifics in the prepress software when creating the output image file. However, the parameters of each of these stages are maintained only within certain specified technological tolerances. Their margins depend on the quality of materials used, equipment and degree of its wear, level of regimes normalization and stabilization, which, in turn, depends on both the economic opportunities and the overall technological culture of the production site. Deviation occurrences within these tolerances can be unpredictable and contribute to the corresponding uncertainty in the values obtained in printing.

Conclusions

The TRC defining is comprised of aligning the tone ranges of original and print copy and finding the shape of the curve connecting tone values of original and print within the range with the priority of a first stage as far as the ranges mismatch can result in irreparable loss of an image data.

There are two principally different approaches in setting the tone curve shape as a reproduction goal: matching an original and editorial. Physically (objectively) matching is possible for only those details of original whose values are within the printable range. In typical conditions of the latter compression relative to an original just the psycho-visual matching is possible.

Trade off in visually identical transfer of the overall contrast of an image in conditions of gray scale compression is the loss of brightness differences in the darkest and/or lightest image areas. Connection between the original and resulting print gray levels is greatly influenced by the transfer functions of all technology stages optimized according their own criteria.

In “closed” reproduction systems there is possible the end-to-end calibration compensating for the tone transfer nonlinearity of all stages while in the open reproduction environment the tone rendition programming involves individual records of separate stages characteristics. Image signal related to the visually uniform metrics is less susceptible to interference being evenly distributed over the reproduced scale. At the prepress stage, the maximal contour capacity of halftone printing is ensured by sufficiency of the dots alphabet and shape of the threshold function.

Balance provision for transfer characteristics of numerous reproduction stages in relation to the real conditions is a prerogative of technology. The accuracy of these stages parameters coordination directly depends on the degree of their normalization, general technological culture of the print site and, ultimately, determines the level and stability of print quality parameters.

References

Allebach, J. P. (1994) Model-Based Halftoning via Direct Binary Search. In: Eschbach, R. (ed.) *Recent Progress in Digital Halftoning.* Springfield, Virginia, IS&T, pp. 50-56.

Carrara, D. A., Analoui M. & Allebach J.P. (1992) Recent progress in digital halftoning. In: *Proceedings of 8th International Congress on Advances in Non-Impact Printing Technologies, IS&T, 25-30 October 1992, Williamsburg, Virginia.* Springfield, IS&T.

Field, G. (2001) High resolution color. *TAGA Newsletter.* 138, 8-9.

Gast, U. (1981) *Method and circuit arrangement for correcting contrast intensification*. United States Patent and Trademark Office US4484232 (Patent).

Giorgianni, E. J. & Madden, T. E. (1998) *Digital Color Management.* Boston, Addison-Wesley.

Gustavson, S. (1997) *Dot Gain in Color Halftones*. PhD Thesis. Linkoping University.

Hallows, R. & Klensch, J. (1968) Electronic halftones. *IEEE Spectrum*. 5 (10), 64-72. Available from: doi: 10.1109/MSPEC.1968.5215383

Has, M. (1995) Color Management – Current Approaches, Standards and Future perspectives. In: *Proceedings of 11th International Congress on Advances in Non-Impact Printing Technologies, IS&T, 29 October-03 November 1995, Hilton Head, South Carolina.* Springfield, IS&T. pp. 441-446.

Hunt, R. (1998) Why is black and white so important in color? In: Buckley, R. (ed.) *Recent Progress in Color Management and Communications.* Springfield, Virginia, IS&T, pp. 189-192.

Jones, L. A. (1920) On the theory of tone reproduction, with a graphic method for the solution of problems. *Journal of the Franklin Institute.* 190 (1), 39-90. Available from: doi: 10.5594/J08066

Kipphan, H. (2001) *Handbook of Print Media, Technologies and Production Methods*. Berlin, Springer.

Kouznetsov, J. V. (1999) Does some philosophy still exist for the halftone frequency selection? In: *Proceedings of 15th International Congress on Advances in Non-Impact Printing Technologies, IS&T, 17-22 October 1999, Orlando, Florida.* Springfield, IS&T. pp. 362-365.

Kuznetsov, Y.V. (1998) *Basic of the prepress image processing. Halftoning.* Moscow, Mir Knigi.

Kuznetsov, Y.V. & Ermoshina M. (2016) Does the use of black ink still comprise the "darkest" issue of CMYK printing?. *Journal of Print and Media Technology Research.* 3 (16), 215-230. Available from: doi: 10.14622/JPMTR-151

Kuznetsov, Y.V. (2019) To take for granted or question the technology fundamentals in research and learning?. *Journal of Print and Media Technology Research.* 4 (19), 227-238. Available from: doi: 10.14622/JPMTR-1908

Morovic, J. (2003) Gamut mapping. In: Sharma, G. (ed.) *Digital Color Imaging Handbook.* Boca Raton, CRC Press, pp. 639-686.

Namakura, T. & Numakura, I. (1994) *Tonal conversion method of a picture.* United States Patent and Trademark Office US5371616 (Patent).

Ovchinnikov, Y. M., Fineberg I. S. & Litvan R. I. (1971) Problems of optimal programming the graphic reproduction process. *Poligraphia.* 3.

Rabinovich, A. D. (1976) Optimization goal and image transform in graphic electronic machines. In: P*roceedings of VNIIOPIT Electronic systems of information transforms.* Moscow, 1976.

Reinhard, E., Ward, G., Pattanaik, S. N., Debevec, P. E., & Heidrich, W. (2010) *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. 2nd ed. Cambridge, Academic Press.

Southworth, M., McIlroy, T. & Southworth, D. (1992) *The Color Resource Complete Color Glossary: From Desktop to Color Electronic Prepress.* Graphic Arts Pub Co.