Tutkimusraportti 102
Research Report 102

MINIMUM ERROR CONTRAST ENHANCEMENT

Jarkko Vartiainen, Pekka Paalanen, Joni-Kristian Kämäräinen,
Lasse Lensu and Heikki Kälviäinen

Lappeenranta University of Technology
Department of Information Technology
Box 20
FIN-53851 Lappeenranta

ISBN 952-214-215-8 (paperback)
ISBN 952-214-222-0 (PDF)
ISSN 0783-8069

Lappeenranta 2006

Minimum error contrast enhancement

Jarkko Vartiainen, Pekka Paalanen, Joni-Kristian Kamarainen, Lasse Lensu, and Heikki Kälviäinen

Department of Information Technology
Lappeenranta University of Technology
P.O.Box 20, FI-53851 Lappeenranta, Finland

Abstract. Contrast enhancement is an image processing technique where the objective is to preprocess the image so that relevant information can be either seen or further processed more reliably. These techniques are typically applied when the image itself or the device used for image reproduction provides poor visibility and distinguishability of different regions of interest in the image. In most studies, the emphasis is on the visualization of image data, but this human observer biased goal often results to images which are not optimal for automated processing. The main contribution of this study is to express the contrast enhancement as a mapping from N-channel image data to 1-channel gray-level image, and to devise a projection method which results to an image with minimal error to the correct contrast image. The projection, the minimum-error contrast image, possess the optimal contrast between the regions of interest in the image. The method is based on estimation of the probability density distributions of the region values, and it employs Bayesian inference to establish the minimum error projection.

1 Introduction

Contrast enhancement is an image processing technique where the objective is to preprocess the image so that relevant information can be either seen or further processed more reliably. These techniques are typically applied when the image itself or the device used for image representation provides poor visibility and distinguishability of different regions of interest in the image. The reason for the poor representation can be a high dynamic range of pixel intensities when displays cannot reproduce them, narrow dynamic range over the regions of interest, or even incompatibility with the human observer. The most important application areas of contrast enhancement are medical imaging [1] and visualization of images with high dynamic range [2]. The enhancement is usually performed to produce a better representation for a human observer, but properly enhanced images can also enable more accurate and more reliable results in general image processing tasks, such as segmentation, due to an enhanced signal-to-noise-ratio.

Often a reversible and fast contrast enhancement is preferred, e.g., in medical imaging. Therefore, the methods typically exploit histograms. The histogram can be multidimensional, such as a 3-D color histogram, but most methods assume a 1-D histogram resulting that the color components are separately processed in the enhancement process. The baseline method for the contrast enhancement is the histogram equalization [3,4], but the method does not utilize any information about which regions should be emphasized. This "blind" equalization often leads to a representation where image noise is amplified, and the properties aiding distinguishability of different image regions are declined [1]. The amplification of noise may be avoided to some extent by local processing [1], but an undesired result is the loss of rank order of image intensities. It seems that estimation of the regions of interest, either manual or automatic, is necessary for successful contrast enhancement. Manual estimation would make the solution trivial, but automatic estimation still remains a challenging problem. The automatic estimation may be based solely on the histogram information, or it can include also analysis of the spatial relationships.

The actual contrast enhancement is based on stretching the perceivable dynamic range between the estimated regions (visual enhancement). In the visual enhancement, pseudocoloring [5] or multiple images are used to artificially emphasize the contrast [2]. The use of multidimensional data is, however, difficult for many image processing techniques, such as frequency- or scale-space methods. Therefore, multi-representation is a good solution mainly for visualization purposes.

In this study, the authors present contrast enhancement as a method to project N-dimensional image data to a 1-dimensional (gray-level) image which possesses a maximal contrast between the regions of interest. The main reason for the gray-level image representation is to allow the use of standard image processing techniques, even though it can also be used for visualization purposes. In the simplest case, the image is bi-modal, i.e., there are two regions of interest. In this case, the maximal contrast is achieved by binarization. Binarization, however, does not provide the minimum-error contrast due to presence of noise. To achieve the minimum-error contrast, the authors utilize Bayesian inference and posteriori values in the enhancement. Similar approaches have been used, for example, in thresholding [6] and color segmentation [7]. The minimum-error enhancement is analytically studied with simulated data, and the efficiency is also demonstrated with real images.

2 Contrast enhancement

Successful contrast enhancement produces an output image where the dynamical range of output space values is organized to emphasize the separation between regions of semantically different classes. It should be noted that there exists no definition for optimal contrast enhancement, but the desired result depends on the application. A similar but more unambiguous problem is the image segmentation where regions of similar characteristics (texture, color, etc.) should be automatically labeled with the same label. For example, in color segmentation it is assumed that different colors belong to different semantic regions. The input space, typically the RGB or HSI color space, is mapped to discrete color labels, and if the assumption holds the method may reveal desired regions of interest, such as the human skin [7].

An optimal output space for the contrast enhancement cannot be defined either: some methods, such as histogram equalization, work on the gray level histogram domain and produce a new representation within the same domain, while other methods produce multichannel representation (pseudocoloring) or even multiple images.

In the following, the authors will address a specific mapping problem where an N-channel input image (e.g., $N = 1$ for gray-level images, $N = 3$ for RGB images) is mapped to a 1-D representation where the maximal contrast is optimized between M different regions of interest. The mapping of M different regions into a single variable is a distinct problem, and it will be evident that the minimum-error contrast, as it will be defined, provides the minimum error only as biased by the selected mapping method. How the biased error relates to the true error is an information theoretic problem how the region data should be optimally coded into a single variable. The true error can be achieved only in the bi-modal case ($M = 2$), but the approach is generalizable to any number of regions M, however.

2.1 Bi-modal image model

In the bi-modal image model, there are two different regions of interest, ω_0 and ω_1, in the image $f(x, y)$ where $f : (x, y) \to \boldsymbol{x}$. The output space values $\boldsymbol{x}$ of the two regions vary with respect to probability distributions $p(\boldsymbol{x}|\omega_0)$ and $p(\boldsymbol{x}|\omega_1)$, and the covered area of the two regions is defined by a priori probabilities $P(\omega_0)$ and $P(\omega_1)$. Since the maximal contrast can be achieved only by a representation where no ambiguity exists between the two regions, the maximal contrast for the bi-modal image model corresponds to a binary representation, e.g.,

$$\boldsymbol{x} \rightarrow \begin{cases} 0, \text{if } \boldsymbol{x} \in \omega_0 \\ 1, \text{if } \boldsymbol{x} \in \omega_1 \end{cases} . \tag{1}$$

Symbols 0 and 1 are selected here just for convenience – any other two different symbols agree with the definition. For a bi-modal image, the optimal contrast enhancement would produce a representation where the pixels belonging to the region ω_0 are denoted by one symbol, and the pixels belonging to the region ω_1 by another symbol.

It is obvious that the maximal contrast can be obtained by using binary thresholding methods. However, if the distributions of the two classes overlap, thresholding provides also the maximal error for a single pixel if a wrong decision has been made. Thresholding does not generally provide the minimum-error maximal contrast. Furthermore, the minimum-error maximal contrast depends on a penalty function for the wrong decision

$$penalty(\omega_i, \omega_j) = \begin{cases} 0, \ \omega_i = \omega_j, \\ \epsilon, \ \omega_i \neq \omega_j \end{cases} . \tag{2}$$

2.2 Multi-modal image model

In the multi-modal image model, there are M regions of interest, $\omega_0, \omega_1, \ldots, \omega_{M-1}$. The definition in Eq. 1 can be generalized for the multi-modal model, and also the penalty function in Eq. 2, now requiring $\binom{M}{2}$ elements if the penalties are not identical between all regions, can be defined. These definitions are applicable for visualization purposes, but it will be shown that M modalities would need $M - 1$ dimensions to represent the regions with their ambiguities and to provide the minimum-error contrast. In this study, contrast enhancement is defined as a mapping $\mathbb{R}^N \rightarrow \mathbb{R}$ ($\mathbb{Z}^N \rightarrow \mathbb{Z}$ in practice), and thus, a distinct problem is how to optimally represent the ambiguity between $M > 2$ regions in a single variable. Here, the emphasis will be on the bi-modal case since it is the most frequent case in practical applications, but the generalization will also be discussed.

3 The method

As it was stated in the previous section, the main objective of the enhancement is to provide maximal contrast between different regions of interest in an image. It was shown how binary thresholding is the method for maximal contrast enhancement in the case of bi-modal images. However, it is obvious that if there is ambiguity in overlapping portions of regions, thresholding may produce maximal errors for values appearing in the areas where the distributions overlap, that is, the noise contrast is maximized as well. Next, we will introduce a notation for the minimum-error maximal contrast, and propose a method which minimizes the contrast error.

3.1 Minimum error contrast

For an image $f : (x, y) \rightarrow \boldsymbol{x}$ where for all spatial points (x, y) there exists a unique region label ω, we can define a maximal contrast description $c : (x, y) \rightarrow \Lambda$ where Λ is a set of discrete symbols λ_0, λ_1, …, λ_{M-1} corresponding to the regions in the image respectively. We can assign $\Lambda \in \mathbb{Z}$ or $\Lambda \in \{\mathbb{R}\}_M$ (a countable sub-set of real numbers) without any loss of generality. If the penalty function for symbol confusions is defined (e.g., the way in Eq. 2), there exists an infinite number of non-maximal contrast descriptions $\hat{c} : (x, y) \rightarrow \hat{\lambda}$ where $\hat{\lambda} \in \hat{\Lambda} = \mathbb{R}$ for which the penalties over the whole image, i.e., the error function $error(c, \hat{c})$ is ≥ 0. By these definitions it is clear that the non-maximal contrast description which minimizes the error function is the minimum-error (maximal) contrast estimation of c.

By utilizing domains where the maximal contrast description is given by a countable set of symbols belonging to the set of real numbers, and the non-maximal contrast description by an uncountable set of symbols belonging to the real numbers, we may utilize the standard arithmetics to devise the minimum-error contrast enhancement by seeking for $\hat{c}$ such that

$$\hat{c} \leftarrow \underset{\hat{c}}{\operatorname{argmin}}\, error(c, \hat{c}) \quad . \tag{3}$$

The standard way to solve the minimization problem in Eq. 3 is to utilize the L_2 norm (Euclidean distance) as the penalty term

$$penalty(\lambda, \hat{\lambda}) = \| \lambda - \hat{\lambda} \|_2 \tag{4}$$

and the mean of squared error (MSE)

$$error(c, \hat{c}) = \frac{\sum_x \sum_y \left(penalty(c(x, y), \hat{c}(x, y))\right)^2}{\sum_x \sum_y 1} \tag{5}$$

as the error function. $\sum_x \sum_y 1$ is the number of pixels in an image.

Contrast enhancement is a function which operates on the image output domain $\boldsymbol{x}$, and based on the output values produces a symbol belonging to the set $\hat{\Lambda}$ as $e : \boldsymbol{x} \rightarrow \hat{\Lambda}$. The minimum-error contrast enhancement method provides a description which satisfies Eq. 3 for the given error function, penalty function, and mapping procedure from the N-channel input space to the 1-channel contrast space.

3.2 Bi-modal enhancement

For the bi-modal enhancement, we may choose $\lambda_0 = 0$ and $\lambda_1 = 1$ ($\Lambda = \{0, 1\}$) without any loss of generality since any other choice of real values would just correspond to scaling of the values.

It has already been stated that for the bi-modal case, the maximal contrast can be achieved by binarization of the input image. Thresholding methods can be considered as methods for obtaining maximal contrast. If the classes' conditional probability density functions, $p(\boldsymbol{x}|0)$ and $p(\boldsymbol{x}|1)$, and a priories of the both regions, $P(0)$ and $P(1)$, are known, every image pixel can be assigned to one of the binary regions by utilizing Bayesian inference

$$\boldsymbol{x} \Rightarrow \begin{cases} 0, \text{ if } p(x|0)P(0) \geq p(x|1)P(1), \\ 1, \text{ if } p(x|0)P(0) < p(x|1)P(1) \end{cases} . \tag{6}$$

The Bayesian rule guarantees minimum error in binarization, and by following this principle, Kittler and Illingworth defined a method to select the optimal threshold value assuming normal distributions for $p(\boldsymbol{x}|0)$ and $p(\boldsymbol{x}|1)$ [6].

If binary representation is assumed, the Kittler and Illingworth method can in this context be referenced as the minimum-error maximal contrast. It is clear that there exists a confusion factor in the binary Bayesian decision. For example, when the posteriories of both regions are 0.5, the decision favoring 0 would not be the optimal decision. In coin-toss gambling with a double reward, the optimal decision would be to bet on both regions, or neither one of them [8]. Binarization does not allow to utilize the confidence information, but if the non-maximal contrast by the real number space is sufficient, the confidence can be embedded into the contrast description. For values between $[0, 1]$, the minimum-error decision corresponds directly to the posteriori values [8]. The minimum-error contrast for values $[0, 1]$ can be provided by directly utilizing the posteriori values. Since 0 now represents strong certainty of the region ω_0 and 1 of ω_1, the posteriori of either region can be selected, and the new representation is now

$$\hat{c} \leftarrow \frac{p(x|0)P(0)}{p(x|0)p(0) + p(x|1)p(1)} \tag{7}$$

or

$$\hat{c} \leftarrow \frac{p(x|1)P(1)}{p(x|0)p(0) + p(x|1)p(1)} \tag{8}$$

where the difference between the two equations is only the polarization of the contrast enhancement. For the bi-modal image model, the posteriori values provide the true minimum-error contrast.

3.3 Multi-modal enhancement

It is obvious that the minimum-error contrast enhancement by posteriori values can be straightforwardly generalized from the bi-modal case ($M = 2$) to the multi-modal case. The generalization would, however, require $M - 1$ dimensional contrast description $\hat{c}$. The main objective of this study was to consider the optimal enhancement to a single variable output space, but for a multi-modal image, the description succeeds only if $\boldsymbol{x}$ values of the different regions do not overlap. If there is no overlap, there is no confusion, and consequently, the thresholding techniques can be straightforwardly extended for the multi-modal case as shown in, e.g., [6]. However, if more than two distributions overlap, the confusion information cannot be accurately embedded into a single variable. For visualization and other processing purposes, it becomes an important issue how the different regions should be optimally represented. It should be noted that the proposed theory in this study can be applied despite the used representation due to the generality of Bayesian optimal inference. However, the results would always be biased by the selected representation scheme. The bi-modal case is our main objective in this study since it is the most frequently encountered problem in practical applications and it can be extended to the cases where there is only one "interesting" region in the image. Therefore we will not yet proceed to the more general definitions covering also the multi-modal case. However, we claim that the optimal projection should obey the information theoretic optimal decision rules and furthermore, the optimal multi-modal contrast representation would correspond to the optimal data compression theorem [8].

3.4 Estimating probability distributions

Before the presented results can be applied, the distributions and their parameters for different regions of interest, $p(\boldsymbol{x}|\omega_i)$ and $P(\omega_i)$, must be estimated. In our only assumption, we generalize the assumption by Kittler and Illingworth and assume that the probability densities $p(\boldsymbol{x}|\omega_i)$ are multivariate normal distributions. If no prior knowledge of the distribution type is available, the multivariate normal distribution provides a good general solution playing a predominant role in many areas of mathematics [9].

For the estimation of M normal distributions in N-dimensional input space, the expectation maximization (EM) algorithm can be applied [10]. Furthermore, if the number of regions of interest is not know, the EM algorithm can be applied in an unsupervised manner, e.g., by using the Figueiredo-Jain algorithm [11].

4 Results

In this section, the optimality of the proposed approach is demonstrated using generated data and real image samples.

4.1 Color image enhancement

To demonstrate the method both quantitatively and visually, we selected the RGB color space as the image output value $\boldsymbol{x}$. An image is generated by placing two random points into the 3-D RGB space, assigning every image pixel to one of the points with respect to a fixed a priori values $P(0)$ and $P(1)$, and applying color space noise based on fixed normal distributions $p(\boldsymbol{x}|0)$ and $p(\boldsymbol{x}|1)$ to the pixels. The effect of distance of points in the color space and the variance, the points can be repeteadly generated from the uniform distribution, and by varying the distance and variance parameters.

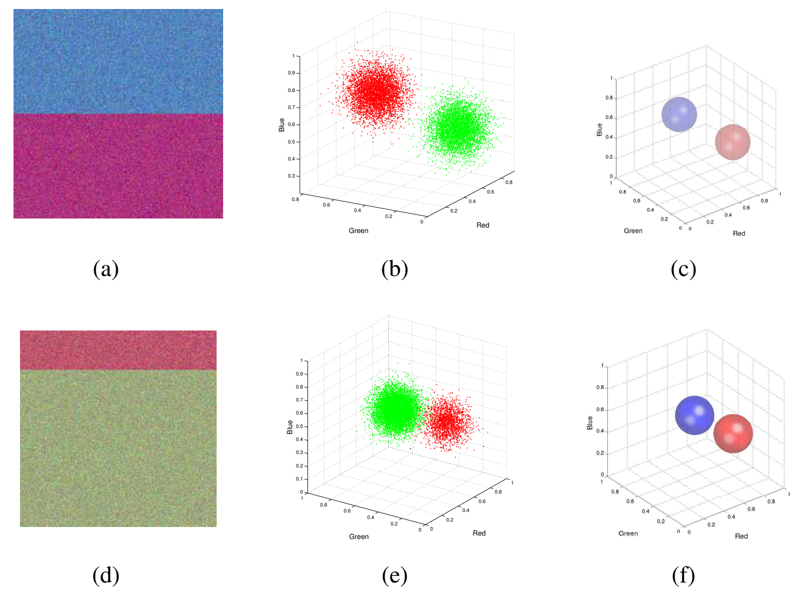


Fig. 1. Generated bi-modal image data in the spatial domain (left), generated points in RGB space (middle), and EM-estimated distributions (right); (a),(b),(c) $P(0) = 0.5$, $P(1) = 0.5$, $\mu_0 = [85, 134, 191]/256$, $\mu_1 = [174, 52, 126]/256$, $\sigma_0 = 20/256$, $\sigma_1 = 20/256$; (d),(e),(f) $P(0) = 0.2$, $P(1) = 0.8$, $\mu_0 = [192, 87, 111]/256$,$\mu_1 = [159, 172, 126]/256$, $\sigma_0 = 20/256$, $\sigma_1 = 20/256$.

In Fig. 1, two images with different parameter values are shown. It should be noted that the covariance is a diagonal with fixed variance in all dimensions.

The most popular general method to convert RGB data to a single variable image (gray-level image) is the standard RGB-to-gray transformation (intensity image), and one of the most successful maximum contrast (thresholding) methods is the minimum-error thresholding [6]. To compare the minimum-error contrast enhancement, we studied the MSE behavior of these methods for the generated data. The results are demonstrated in Fig. 2. The MSE is calculated with respect to the perfect maximal contrast image (binary).

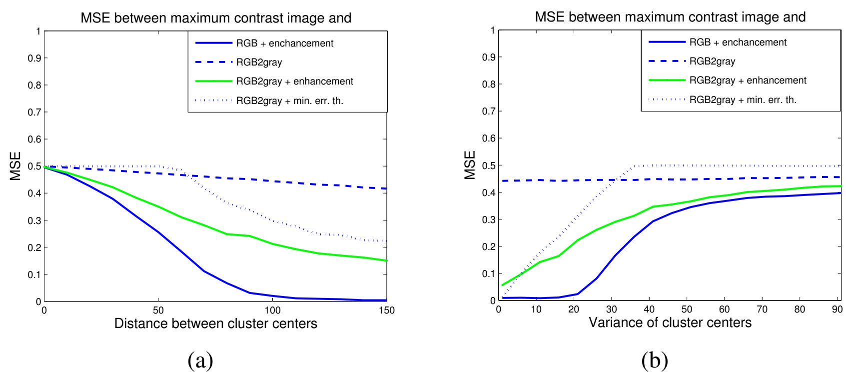
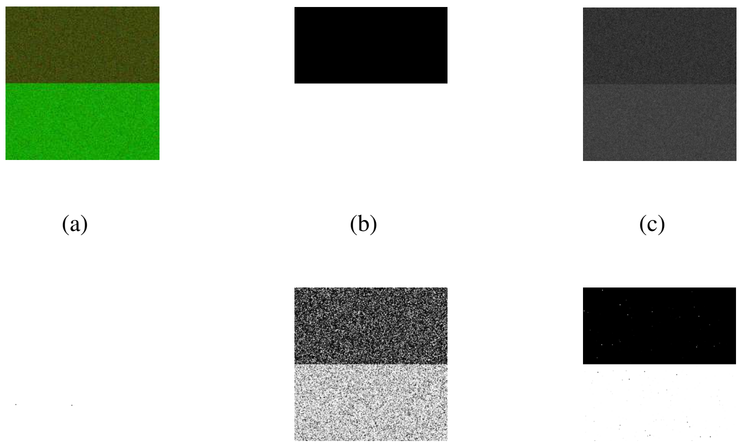


Fig. 2. MSE of different contrast enhancement methods: (a) as function of cluster distance $\sqrt{(\mu_0 - \mu_1)^2}$ ($P(0) = P(1) = 0.5$, $\sigma_0 = \sigma_1 = 20/256$); (b) as functions of cluster variance ($\sqrt{(\mu_0 - \mu_1)^2} = 100$, $P(0) = P(1) = 0.5$).

From the results in Fig. 2, it is evident that the behavior of different methods corresponds to what was expected: the standard RGB-to-gray conversion provides the weakest contrast, the minimum-error thresholding method provides sufficiently good contrast if the overlap is not significant, but the cases where the minimum-error contrast enhancement is applied are the most successful in the contrast representation. It should be noted that utilizing the color information provides the most accurate results. The results are visualized in Figs. 3 and 4. The only difference between the images is the location of cluster centers, the cluster distance $\sqrt{(\mu_0 - \mu_1)^2} = 100$ and noise variance $\sigma_0 = \sigma_1 = 15/256$ are equal.

4.2 Real image samples

There exists a practical machine vision problem where missing dots must be detected from images of paper samples containing a printed dot pattern [12]. To achieve a successful detection, the image contrast must be enhanced to magnify the distinguishability



(a) (b) (c)

(d) (e) (f)

Fig. 3. Examples of contrast enhancement images: (a) original RGB image; (b) maximal contrast image; (c) RGB to gray converted image (MSE=0.4726); (d) minimum-error thresholding applied to the intensity image (MSE=0.5); (e) minimum-error contrast enhancement applied to the intensity image (MSE=0.2628); (f) minimum-error contrast enhancement applied to the RGB image (MSE=0.0072).

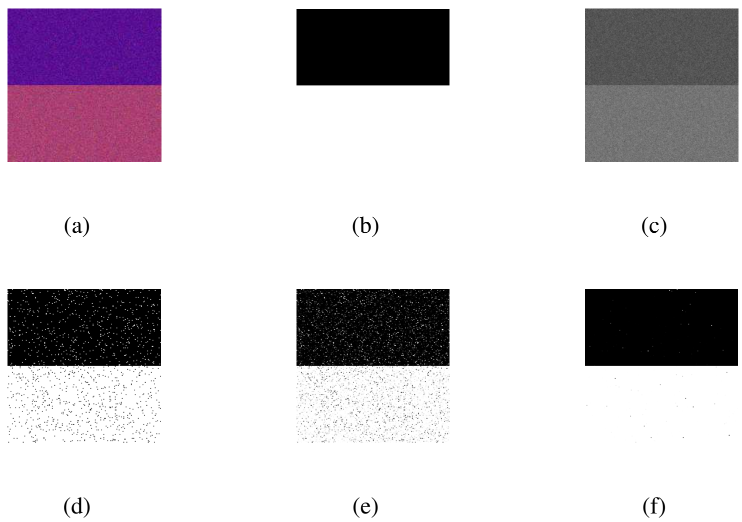


Fig. 4. Examples of contrast enhancement images: (a) original RGB image; (b) maximal contrast image; (c) RGB to gray converted image (MSE=0.4383); (d) minimum-error thresholding applied to the intensity image (MSE=0.0308); (e) minimum-error contrast enhancement applied to the intensity image (MSE=0.0466); (f) minimum-error contrast enhancement applied to the RGB image (MSE=0.0006).

of printed dots from the background. In Fig. 5, the performance of contrast enhancement has been demonstrated.

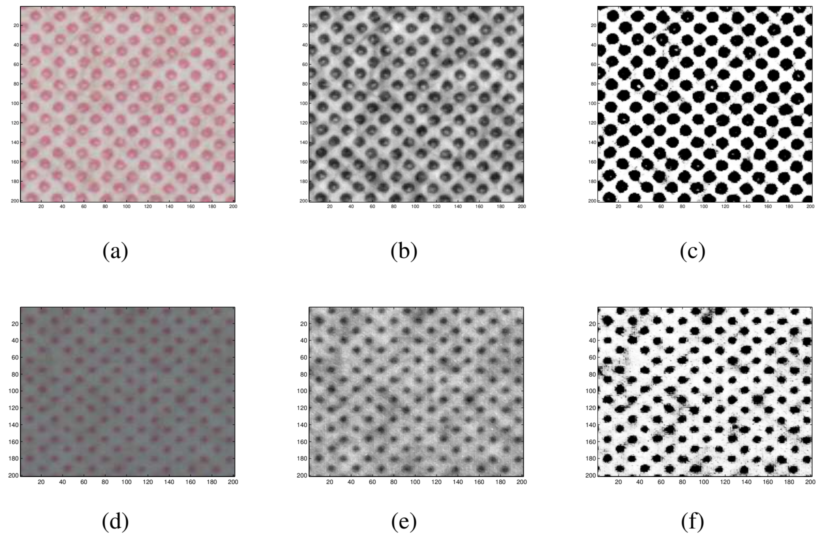


Fig. 5. Heliotest images: (a),(d) original color image; (b),(e) intensity image; (c),(f) minimum-error contrast enhanced image.

5 Conclusions

The problem of optimal contrast enhancement was addressed in this study. The desired enhancement result depends on the application, but as a general result, the optimality was defined as the minimum error to the correct maximum-contrast image. An efficient solution to the problem of minimum-error contrast enhancement was derived under the assumption that the region of interest values are multivariate normally distributed. The principal idea behind the method is to optimize the average pixel-contrast error rate. The provided solution provides the optimal contrast for N-channel bi-modal images. The general solution for multi-modal images was discussed, but the problem will be addressed in future research.

References

1. Zimmerman, J., Pizer, S., Staab, E., Perry, J., McCartney, W., Brenton, B.: An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. IEEE Trans. on Medical Imaging **7** (1988)

2. Pardo, A., Sapiro, G.: Visualization of high dynamic range images. IEEE Trans. on Image Processing **12** (2003)
3. Hall, E.: Almost uniform distributions for computer image enhancement. IEEE Trans. Computers **C-23** (1974)
4. Hummel, R.: Histogram modification techniques. Computer Graphics and Image Processing **4** (1975)
5. Smith, S.: Color coding and visual separability in information displays. Journal of Applied Psychology **47** (1963)
6. Kittler, J., Illingworth, J.: Minimum error thresholding. Pattern Recognition **19** (1986) 41–47
7. Phung, S., Bouzerdoum, A., Chai, D.: Skin segmentation using color pixel classification: Analysis and comparison. IEEE Trans. on Pattern Analysis and Machine Intelligence **27** (2005) 148–154
8. Cover, T., Thomas, J.: Elements of Information Theory. John Wiley & Sons (1991)
9. Tong, Y.: The Multivariate Normal Distribution. Springer Series in Statistics. Springer-Verlag (1990)
10. Bilmes, J.: A gentle tutorial on the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models (1997)
11. Figueiredo, M., Jain, A.: Unsupervised learning of finite mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 381–396
12. Sadovnikov, A., Vartiainen, J., Kamarainen, J.K., Lensu, L., Kälviäinen, H.: Detection of irregularities in regular dot patterns. In: Proc. of the IAPR Conf. on Machine Vision Applications, Tsukuba Science City, Japan (2005) 380–383