7. Головин О. В., Простов С. П. Системы и устройства коротковолновой радиосвязи / под ред. О. В. Головина. М. : Горячая линия – Телеком, 2006. С. 598.

8. Назаров С. Н. Общий подход к построению современных гибридных сетей беспроводной связи // Тр. Рос. науч.-техн. общества радиотехники, электроники и связи имени А. С. Попо-ва. М., 2009. Вып. LXIV, С. 22–24. (Сер.: Научная сессия, посвященная Дню радио).

9. Назаров С. Н. Использование стохастических моделей для оценки характеристик современной беспроводной сети передачи информации // Современ-ные проблемы создания и эксплуатации радиотехни-ческих систем : Тр. VI Всерос. науч.-практ. конф. (с участием стран СНГ). Ульяновск : УлГТУ, 2009. С. 170–174.

10. Назаров С. Н., Назаров А. С. Анализ методов моделирования беспроводной сети передачи информации // Современные проблемы создания и эксплуатации радиотехнических систем : тр. VI Всерос. науч.-практ. конф. (с участием стран СНГ). Ульяновск : УлГТУ, 2009. С. 174–177.

11. Назаров С. Н. Применение элементов декаметровой радиосвязи в современных беспроводных сетях // Тр. Рос. науч.-техн. общества радиотехники, электроники и связи имени А. С. Попо-ва. М., 2009. Вып. XI-1. С. 228–230. (Сер: Цифровая обработка сигналов и ее применение).

12. Назаров С. Н., Назаров А. С. Обобщенная модель беспроводной сети передачи информации авиационного предприятия // Современные научно-технические проблемы транспорта : сб. науч. тр. V Междунар. науч.-техн. конф. Ульяновск : УлГТУ, 2009. С. 108–111.

13. Назаров С. Н. Применение гибридной беспроводной сети передачи информации в автоматизированной системе управления воздушным движением // Современные научно-технические проблемы транспорта : сб. науч. тр. V Междунар. науч.-техн. конф. Ульяновск : УлГТУ, 2009. С. 112–116.

14. Назаров С. Н. Основные положения методики определения места расположения сети удаленных взаимосвязанных радиоцентров-ретрансляторов // ИКТ. 2009. Т. 7, № 2. С. 79–82.

15. Прохоров В. К., Шаров А. Н. Методы расчета показателей эффективности радиосвязи Л. : ВАС, 1990. С. 132.

A. I. Pakhirka
Siberian State Aerospace University named after academician M. F. Reshetnev, Russia, Krasnoyarsk

THE APPLICATION OF IMAGE ENHANCEMENT METHOD FOR FACE RECOGNITION SYSTEMS

Three-step face recognition algorithm which includes non-linear enhancement (dynamic range compression) and faces localization on the basis of skin color segmentation with subsequent extraction of anthropometric face points is proposed. The process of face recognition on the basis of principal component analysis is also considered.

Keywords: dynamic range compression, face localization, face recognition.

Face recognition has always caused great interest in computer vision, especially in connection with increasing practical needs such as biometrics, search engines, video compression, video conferencing systems, computer vision in robotics, intelligent security and access control systems.

Face recognition algorithms can be divided into two categories: methods based on extracting features of images and methods based on representation of a facial image. The first group of methods uses properties and geometric relationships such as areas, distances and angles between feature points of a facial image. The second group of methods considers global features of a facial image.

Usually these methods try to represent facial data more efficiently, for example, as a set of main vectors. Typically, a face recognition algorithm includes three steps: image preprocessing, face localization, face recognition. In this paper we present an algorithm which includes nonlinear image enhancement (dynamic range compression), face localization on the basis of skin color segmentation and face recognition on the basis of principal components analysis [1].

In practice images captured by digital devices often differ from what an observer remembers. It happens due to the fact that a camera captures the physical values of light data, while an observer's nervous system processes these data. For example, an observer can easily see details both in deep shadows and in illuminated areas while a capture device will get the given scene with too dark areas or light-struck areas. A human observer easily perceives scenes with a high range of light intensities while the ratio between the highest and the lowest luminance exceeds the capabilities of a capture or output device.

The human observer deals with high dynamic range scenes by adapting locally to each part of the scene and thus is able to retrieve details in low luminance as well as high luminance areas. Using a digital device is more problematic. The dynamic range of the scene has to be compressed, which often causes the captured image to lack details in areas of low and high illumination. Some recent developments made it possible the capture high dynamic range scenes.

The principle is to capture multiple pictures of the same scene with different exposure times. A so-called radiance map is built from the acquired pictures. This technique allows obtaining an accurate estimation of the scene despite a capture device limitation. Nevertheless, the problem of mapping the high dynamic range values expressed in floating point to the low dynamic range of the output device remains [2].

These problems can be solved by an algorithm which simulates a human visual system. For this purpose we can use Multi-Scale Retinex (from retina and cortex) algorithms. These algorithms compress dynamic range of images with saving (increasing) local contrast in areas of low and high illumination [3].

A classical multidimensional MSR-algorithm is a weighed sum of one-dimensional SSR (Single-Scale Retinex) algorithms for different scales. A univariate output function of i-th color channel $R_i(x, y, \sigma)$ is calculated like this:

$$R_i(x, y, \sigma) = \log\{I_i(x, y)\} - \log\{F(x, y, \sigma)*I_i(x, y)\},$$

where $I_i(x, y)$ is an output function for color channel i, σ is a scale factor, "*" denotes convolution, $F(x, y, \sigma)$ is Gaussian function given by:

$$F(x, y, \sigma) = Ke^{-(x2 + y2)/\sigma2}.$$

where parameter K is chosen so as to meet the requirement:

$$\iint_{\Omega_{x,y}} F(x, y, \sigma)\, dx\, dy = 1,$$

where $\Omega_{x,y}$ is a number of pixels of the whole input image.

Then a multidimensional output function of i-th color channel $R_{mi}(x; y; \mathbf{w}; \sigma)$ is calculated like this::

$$R_{M_i}(x, y, \mathbf{w}, \sigma) = \sum_{n=1}^{N} w_n R_i(x, y, \sigma_n),$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_m)$, $m = 1, 2, \ldots, M$ is a weight vector of univariate output functions of i-th color channel $R_i(x, y, \sigma)$; $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, $n = 1, 2, \ldots, N$ is a vector of scales of univariate output functions, and $\sum_{n=1}^{N} w_n = 1$. The length of scales vector is usually not less than 3. In different sources we can find different recommended scale values, in our experiments they were 15, 90, 180. A weight vector ***w*** has, as a rule, elements with equal values.

A block diagram of an image enhancement module is shown in fig. 1. Conversion from RGB to YCbCr is conditioned by the fact that in it color space luminance is presented separately. Therefore, the algorithm is applied only to *Y* component, without affecting Cb and Cr, which improves the performance of the algorithm. For Gaussian blur recursive implementation of Gaussian filter is used which approximates gaussian, with calculation of filter coefficients for a desired value of sigma (σ). Such representation of the filter is faster than standard filtering using convolution kernel [4].

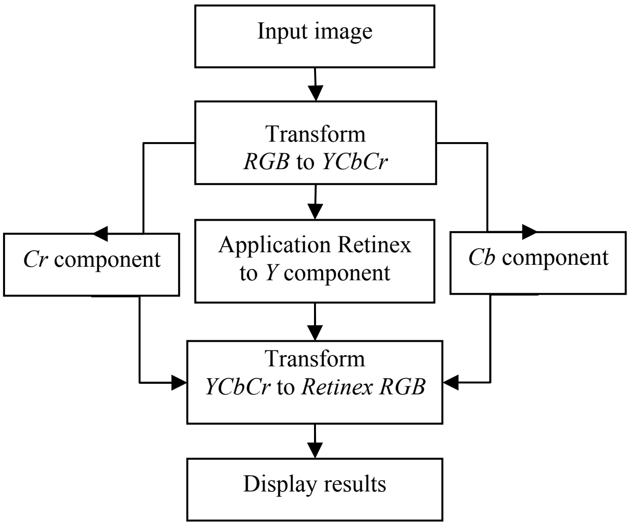

Fig 1. A block diagram of an image enhancement module

Fig. 2 shows the results of application of Retinex algorithm to low illumination image.

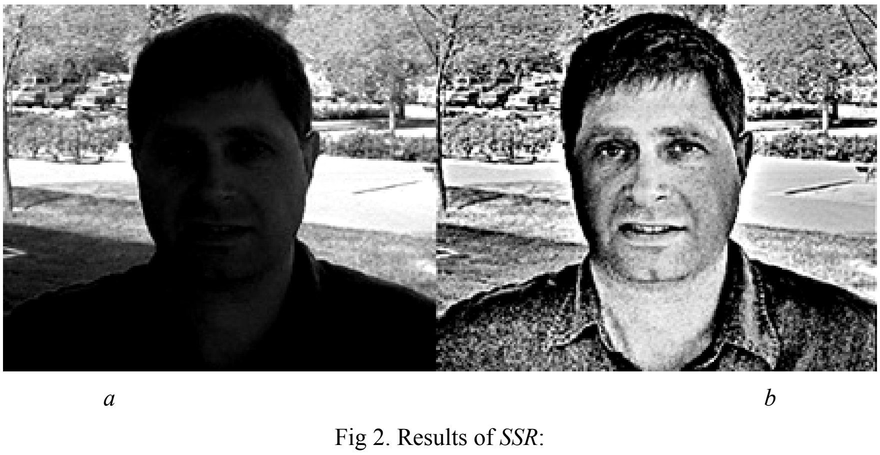

Fig 2. Results of *SSR*:
a – input image; *b* – output image

The next step is faces localization on the basis of skin color segmentation. The process of face localization can be divided into two stages:

– extraction of image areas with the color similar to human skin color (skin color segmentation);

– analysis of extracted regions after segmentation.

Determination of skin color can significantly reduce a search area and is the first step in many methods of face localization.

Human skin has a characteristic color which allows to successfully segment skin in color images. The independence of hue color component on the face orientation as well as its small dependence on brightness make color a stable characteristic of skin. The advantages of skin color segmentation are:

– low computational complexity;

– stability to change in scale and face rotation;

– stability to change in lighting;

– stability to change in facial expressions and face overlapping.

Skin color segmentation requires building of some rules which distinguish between facial color pixels and pixels not related to skin color. For this purpose a metrics is introduced which allows to measure the distance between pixel color and skin hue. This metrics is a model of skin color distribution in a selected color space.

We use the metrics for NCC RGB color space. Skin color distribution for NCC RGB is shown in fig. 3:

$$Skin(r,g)=\begin{cases}1 & if\ \ (g<g_u)\cdot(g>g_d)\cdot(W>0.0004)\\ & 0\ \ otherwise\end{cases},$$

where u is an upper boundary, d is a bottom boundary. Values g_u, g_d, W are defined as:

$$g_u = J_u r^2 + K_u r + L_u\ ,\ g_d = J_d r^2 + K_d r + L_d\ ,$$
$$W = (r-0.33)^2 + (g-0.33)^2\ ,$$

coefficients take the following values:

$$J_u = -1.377,\ K_u = 1.074,\ L_u = 0.145,$$
$$J_d = -0{,}776\ \ K_d = 0.560,\ L_d = 0.177.$$

Results of skin segmentation are shown in fig. 4. The segmented image is processed morphologically (compression with subsequent expansion) which allows to separate poorly connected regions and delete regions of small sizes (noise). Later on the connected areas are marked (fig. 5) and anthropometric points are detected in each area (eyes, lips, nose).

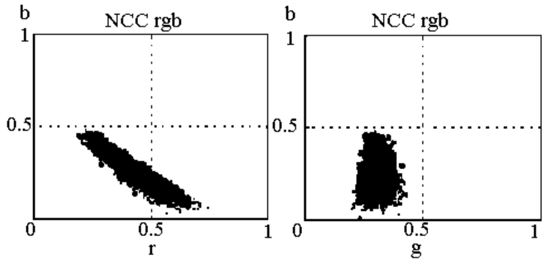

Fig. 3. Skin color distribution for NCC rgb

We use face recognition based on principal component analysis (PCA). Principal component analysis is a standard technique used to approximate the original data with lower dimensional feature vector. PCA is probably the most widely used subspace projection technique for face recognition. This method projects image space into the space of smaller signs. The main idea of PCA is to represent human faces images as a set of image principal components called eigenfaces. The calculation of principal components is reduced to the calculation of eigenvectors and eigenvalues of a covariance matrix which is calculated form image [5].

Any image may be considered as a vector of pixels each value of which is presented by a value in gray scale gradation. For example, an 8×8 image may be unwrapped and treated as a vector of length of 64 pixels. Such vector representation describes the image input space. To present and recognize faces we use a subspace created by eigenvectors of a covariance matrix of investigated images. Eigenvectors corresponding to nonzero eigenvalues of a covariance matrix form an orthogonal basis which rotates and/or reflects the images in the N–dimensional space. Specifically, each image is stored in a vector of size N.

$$X^i = \left[x_1^i \cdots x_N^i\right]^{\mathrm{T}}, \qquad (1)$$

where x^i are master images, X is a matrix of master images. The images are centered by subtracting the mean image from each image vector

$$\overline{x}^i = x^i - m, \qquad (2)$$

where $m = \frac{1}{P}\sum_{i=1}^{P} x^i$

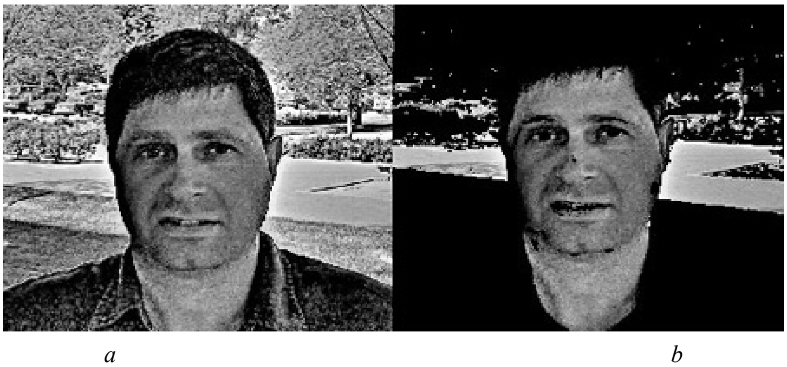

Fig. 4. Skin segmentation:
a – input image; *b* – output image after the use of metrics

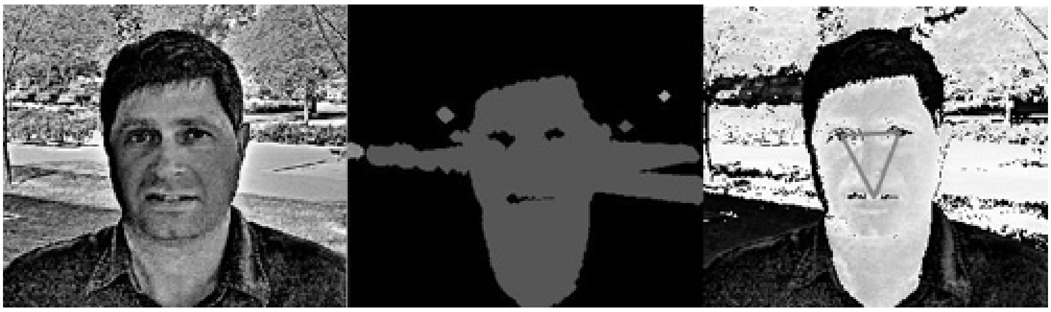

Fig. 5. Face localization:
a – input image; *b* – morphological processing with marking of connected components; *c* – anthropometric facial points detection

These vectors are combined to create a data matrix of size *N*×*P* (where *P* is the number of images):

$$\overline{\mathbf{X}} = \left[\overline{x}^1 \ \ \overline{x}^2 \ \cdots \ \overline{x}^P\right]. \qquad (3)$$

The data matrix $\overline{\mathbf{X}}$ is multiplied by data transposed matrix to calculate the covariance matrix:

$$\Omega = \overline{X}\overline{X}^{\mathrm{T}}. \qquad (4)$$

This covariance matrix has up to *P* eigenvectors associated with non-zero eigenvalues, assuming *P* < *N*. The eigenvectors are sorted from higher to lower value according to their associated eigenvalues. The eigenvector with the largest eigenvalue represents the greatest dispersion in images.

Images recognition through eigenspace projection has three basic steps.

1. Eigenspace must be created using master images (training stage).

2. Master images are projected into the eigenspace (training stage).

3. The projected input image is compared with the projected test images (recognition stage).

Let's consider the first stage – creation of eigenspace consisting of the following steps:

– data centering: each image is centered by subtracting an averaged image from each master image. An averaged image is a column vector including mean pixel values of all pixels of master images (2);

– creation of a data matrix: once the input images are centered, they are combined into a data matrix of size *N*×*P* (equation (3));

– creation of a covariance matrix: the data matrix is multiplied by its transposed representation (4);

– calculation of eigenvalues and eigenvectors: eigenvalues and corresponding eigenvectors are calculated from the covariance matrix:

$$\Omega\mathbf{V} = \Lambda\mathbf{V}$$

where $\mathbf{V}$ is a set of eigenvectors associated with the eigenvalues Λ;

– ordering eigenvectors: the eigenvectors $v_i \in \mathbf{V}$ are ordered according to their corresponding eigenvalues $\lambda_i \in \Lambda$ from higher to lower value. The eigenvectors associated with non-zero eigenvalues are kept. This matrix of eigenvectors is the eigenspace V , where each column of V is an eigenvector:

$$\mathbf{V} = \left[v_1 \ \ v_2 \ \cdots \ v_P\right].$$

Projecting of master images into eigenspace takes place at the second stage. Each centered input image ($\overline{x}^i$) is projected into an eigenspace:

$$\tilde{x}^i = \mathbf{V}^T \overline{x}^i.$$

Identifying input images takes place at the third stage. Each input image is first centered by subtracting the averaged image, and is then projected into the same eigenspace defined by $\mathbf{V}$:

$$\overline{y}^i = y^i - m,$$

where

$$m = \frac{1}{P}\sum_{i=1}^{P} x^i$$

and

$$\tilde{y}^i = \mathbf{V}^T \overline{y}.$$

The projected input image is compared with every projected master image. Images can be compared using any simple metrics, for example, Euclidean.

Currently a nonlinear image enhancement system in different color spaces is being developed. We plan to use MultiScale Retinex algorithm with color restoration for capture and processing video streams having a big range of brightness. A system of face capturing from video sequence with subsequent processing and "averaging" of images, reducing the influence of illumination, correction of face position, choosing the best face image from video data is being developed.

In this paper we propose an improved approach to face recognition in images, using a nonlinear image enhancement algorithm that allows compensating shadows and highlights. Also, the analysis of color spaces can improve the quality of skin color segmentation recognition and anthropometric face points.

References

1. Jain K., Flynn P., Ross A. Handbook of Biometrics. Springer, 2008.

2. Meylan L., Susstrunk S. Bio-inspired color image enhancement // SPIE Electronic Imaging. San Jose, 2004, P. 46–56

3. Tao L., Asari K. V. Nonlinear enhancement of color images // SPIE Journal of Electronic Imaging. 2005. Vol. 14.

4. Young T., Van Vliet L. J. Recursive Implementation of the Gaussian filter : Signal Processing 44. Elsevier, 1995.

5. Yambor W. Analysis of PCA-based and Fisher discriminant-based image recognition algorithms : Technical Report CS-00-103. 2000.

A. N. Pakhomov, M. V. Krivenkov, V. I. Ivanchura
Siberian Federal University, Russia, Krasnoyarsk

MODAL REGULATORS OF A DIRECT CURRENT ELECTRIC DRIVE WITH A PULSE-WIDTH CONVERTER

The technique of modal regulators synthesis of instantaneous coordinates values of direct current digital electric drives with a pulse-width converter by a state space method taking into account the influence of variable pure delay in the control channel is presented.

Keywords: a modal regulator, a direct current electric drive, a pulse-width converter.

The theory of digital repeated systems of the subordinated regulation of the electric drive [1–3] is developed well enough. Less attention has been paid to electric drive systems constructed on the basis of summation of feedbacks on a state vector that allows to expect the decrease of sensitivity to variations of control object parameters. In domestic literature such regulators are usually called modal because factors of a feedback vector directly influence eigenvalues (modes) of the closed system matrix. In the present paper the task of designing of such regulators is set. The task includes:

– reception of discrete equations of a control object condition according to its differential equations;

– definition of feedback factors by state variables according to the set spectrum of a matrix of the closed digital control system dynamics;

– introducing the corrective amendments connected with some features of the real pulse-width converter (PWC). It is supposed that the influence of variable pure delay in the control channel of a double digital system which is rather typical for an electric drive system with microprocessor control is taken into account.

The reasons of occurrence of two periods of discreteness as well as the nature of pure delay in microprocessor systems are considered in works [4; 5]. It is supposed that the interruption period (IP) of work of a control microcomputer contains integer N of the switching periods (SP) PWC: $T = NT_k$, where T and T_k are values IP and SP accordingly. In view of two periods of discreteness it is expedient to introduce two types of relative time – global τ and local one inside IP θ, whereby

$$\tau = \frac{t}{T}; \quad \theta = \frac{t}{T_k} = \frac{N}{T}t = N\tau; \quad \theta \in [0, N]. \qquad (1)$$

Below only thus determined relative values of time are used. Let the calculation of a control signal $u[n]$ on IP with number n is completed after computing delay τ_d on k of the SP (fig. 1). Local computing delay θ_d is counted off from the beginning of SP with number k. If the local time delay θ_{PWC1} necessary for realization of $u[n]$ is more than θ_d (fig. 1) then this realization can be carried out already on k of the SP. The previous value of a control signal $u[n–1]$ is realized *o*n the first k of the SP.

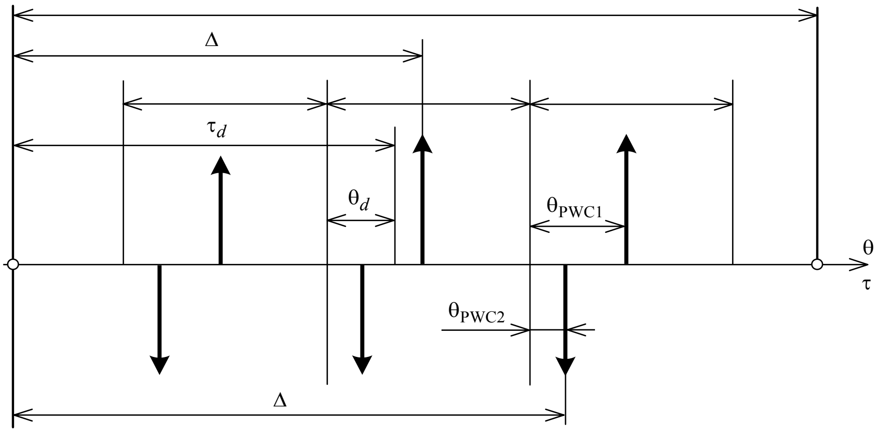

Fig. 1