



Technique for Matching Similarity Using the Chromaticity Space

¹ Petya Em. Pavlova

² Ekaterina G. Borisova

¹ Technical University-Sofia, branch Plovdiv, 25 “Tzanko Djustabanov” str., Plovdiv, 4000.

² Institute of Electronics, Bulgarian Academy of Sciences, 72” Tzarigradsko chaussee” blvd, Sofia, 1784 Bulgaria



Corresponding author:

Petya Em. Pavlova

p_pavlova@tu-plovdiv.bg

Received: Aug 31, 2018

Revised: Sep 13, 2018

Published: Oct 30, 2018

ABSTRACT

The paper describes application of technique for estimation spectral distributions similarity on a set of diffuse-reflectance spectra obtained from different cutaneous benign and malignant lesions in vivo. A specific peculiarity of the colorimetric transformation to set same chromaticity coordinates for even spectral distributions is used for assessment of this similarity. Initial consequence is processed to obtain normalized diffuse-reflectance spectrum in the range from 380 to 780 nm region. The technique uses ratio between spectra that gives an opportunity to extract and estimate only variations between them. Final decision is based on two results: chromaticity coordinates for the normalized spectrum and chromaticity coordinates of the impulse sequences composed from the amplitude distribution for each wavelength. The practical experiment and check of the patterns is carried out on a new set of measured diffuse-reflectance cutaneous spectra from patients with different skin disorders. Each spectrum is compared with preliminary defined pattern distributions. A “complementary” diagnosis, based on the similarities with these patterns, was given to address the exact spectral data set, and the diagnostic accuracy was evaluated based on the histological evaluation of these lesions.

Keywords- Spectra comparison, Metamerism and chromaticity, Uniformity of spectral distributions, Similarity of spectra.

INTRODUCTION

The use of chromaticity plane space offers opportunities for different analyses that allow adaptation to working conditions [1]. Its application is connected directly with the colorimetric transformations and obtaining the chromaticity coordinates. In the colorimetry currently are used two types of transformations: on the basis of three stimuli and on the basis of spectral distribution [2]. The presented study is an extension of a method developed to automate classification in non-invasive diagnostics of skin diseases by diffuse-reflectance spectra in the visible range of optical radiation [3]. Initial analysis has shown that it is possible to separate spectra that characterize a particular disease and can be used as etalons in further comparison and diagnostics of newly measured diffuse-reflectance cutaneous spectra of lesions with preliminary unknown clinical diagnosis. The approach uses normalized distributions from normal and disease-altered skin areas recorded at the same time and light conditions. The ratio of these distributions reproduces one and the same value for the individual peculiarities (momentary state and pigment type) presented in the two areas and highlights the deviations resulting from the disease. If the forms of the initial spectra are similar, the resulting ratio-spectra would be a straight line parallel to the abscissa (x-axis) and some kind of broken line would be observed if differences are available. A scheme of automated disease classification was proposed with a criterion of similarity between the tested and the etalon distributions based on the peculiarities of the spectral colorimetry and the reproduction of the chromaticity coordinates. In this case, the measure of similarity is assessed by distance between chromaticity coordinates for the analyzed spectrum and the coordinates of pure white. The effectiveness of the approach was tested by a comparative assessment of the described method and a cross-

correlation method using currently defined etalons for some diseases. The results given in [4] show that for the specific purposes, the use of chromaticity plane gives better opportunities for determining similarity. In practice, however, it is possible for spectral distributions with different variations to reproduce the same co-ordinates of chromaticity. In colorimetry, this is known as metamerism [5]. In addition, the transformation from spectrum to the chromaticity adds an asymmetry in the analyzed space. The latter has been estimated by a single-pulse response [6] and clearly shows different sensitivity at different spectral intervals. Figure 1 shows a diagram of this asymmetry. For its calculation each distribution is treated as a set of single impulses. At the process of calculating chromaticity coordinated all values except one are equal to 1. The single one has amplitude between 0.1 and 0.9 and has localized at different wavelengths.

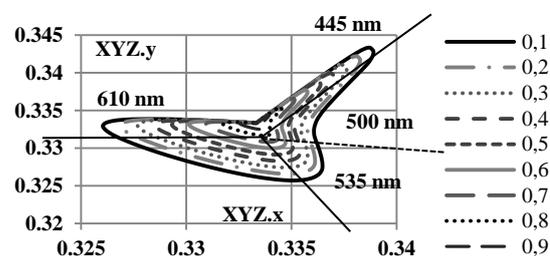


Fig. 1 Distribution of the chromaticity coordinates with different amplitude of the single impulse (The legend shows the value of the impulse)

According to the results obtained, see fig. 1, the greatest probability of false establishment of uniformity, when using the chromaticity plane, is in the area around 500 nm, below 420 nm and above 680 nm. Additional possibility for information loss in the analyzed distributions is the normalization to a peak value, which can lead to suppression of amplitudes and decreasing the accuracy of

data reproduction during a computational procedure.

Broad range of diffuse-scattering data is obtained in measurements *in vivo* of patients with different cutaneous diseases. The specific variations in the spectra received are correlated to unique combination of features for each pathology investigated – its type, stage of lesion growth, thickness, pigmentation, morphology, as well as from the general conditions such as skin phototype, anatomic place of the lesion' position for each particular patient [7,8]. Moreover, clinical picture (appearance, pigmentation, macroscopic structure of the surface, network) for different skin pathologies could be very similar and many false diagnoses could have place [9]. Large range of benign and dysplastic lesions, similar to the corresponding malignancies, also add their contribution to the picture of chaos, when one try to use spectral techniques for addressing and differentiation of cutaneous neoplasia. The variations of all these indicators do not allow developing of a simple discrimination technique for addressing of the lesion types. Standard discrimination techniques, such as principal component analysis, support vector machines, etc. lead to moderate improvement of the discrimination algorithms based on spectral data comparison and do not allow full automation of the decision process search of a proper diagnosis as well [10,11].

The aim of the present work is to complement the developed classification scheme for evaluation of spectral distributions in a plane of chromaticity to increase the reliability of the obtained results by eliminating effects of metamerism and normalizing to a peak value, which would allow to improve the diagnostic accuracy and proper addressing in a given lesion group of the spectral data

received using diffuse-reflectance cutaneous lesions' detection.

METHODS AND MATERIALS

Bases of colorimetric transformation

Transformations in colorimetry convert spectrum into the visible range in three-parametrical magnitude - color. According to CIE standard of colorimetry [2], colors visible to us are set with brightness values and a pair of color parameters. The CIE.XYZ system is a base of all currently used color models. The three basic parameters XYZ.X, XYZ.Y and XYZ.Z determine the brightness by value of XYZ.Y and chromaticity through coordinates XYZ.x and XYZ.y. Graphically, the color is represented by an asymmetric plane figure (Figure 2) defined in a range of coordinates between 0 and 0.8, and known as Locus. At its borders are localized the spectral colors (color of a certain wavelength) and the colors without spectral radiation- a mixture of red and blue, which we see as purple.

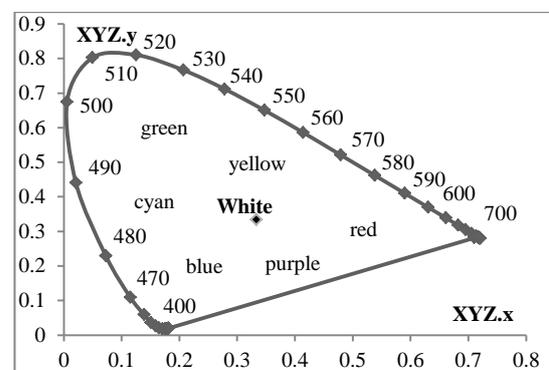


Fig. 2 Chromaticity diagram of XYZ system

Standardization of color measurement is achieved on the basis of the uniform perception for all spectral colors. The uniform power distribution over the whole visible interval of wavelengths always creates perception of pure white (E-source). Chromaticity coordinates of the E-source are fixed at the centrum of Locus and every displacement from this point

indicates the presence of coloration. Colors with coordinates counter in relation to the white point mutually neutralize and standard defines them as complementary. Referred to the form of the ratio-spectra, these peculiarities mean that a straight line parallel to the abscissa is equivalent to uniform spectral distribution and would reproduce the coordinates of the pure white, and any deviation reproduce coordinates of coloring. Besides, the distance from the white point would determine the degree of forms similarity. This allows the plane of chromaticity to be treated as a space of features [12] in which areas of uniformity can be limited around the point of the pure white. Theoretically, there is a possibility of using the chromaticity space of another color system. However, for the most systems with clearly defined chromaticity (Lab, Luv or HSV, HSI, HLS), the color parameters are relative to the values of the Reference White Parameters, and the Locus points are distributed between multiple planar planes of varying sizes whose areas depend of the value of the achromatic parameter. The latter complicates the setting of a similarity measure and makes the presence of one chromaticity plane the main advantage of the basic XYZ system.

Description of the method and the proposed additions

Briefly, the scheme for cutaneous disorders classification developed in [3,4] includes the following stages:

1. Preparation of spectral etalons. Each etalon defines a class. In the specific task for which the method is developed, these are averaged normalized spectral distributions between spectrum - ν of the harmless-state of the object and the spectrum after any changes - p , for all changes showing a different state. As a result, the etalons - $Et_i(\lambda)$ consist only the differences from the changes - p and can

be considered as multidimensional vectors of features in the terms of classification.

2. Investigation of the classes for distinctness.

In the selected classification scheme based on chromaticity, the test includes an establishment of normalized distribution of ratio between any two etalons (1).

$$g_{ij}(\lambda) = Et_i(\lambda) / Et_j(\lambda) \quad (1)$$

The criterion of distinctness is the distance of the chromaticity coordinates of $g_{ij}(\lambda)$ from the point of pure white.

3. Compilation of the test - spectrum. The tested distribution is calculated as a normalized ratio - $t_s(\lambda)$ between spectrum - ν of the harmless-state and the spectrum after any changes - p for one object.

4. Compilation of normalized spectra-ratio - $q_i(\lambda)$ between each etalon $Et_i(\lambda)$ and $t_s(\lambda)$.

5. Verification for normalization to peak value. Amplitude suppression is accepted when the maximum number of amplitudes of the same value in the test distribution is below 0.2

6. Determination of affiliation of the tested state to one of the defined classes (disorders) according to criterion of the maximal proximity of the chromaticity coordinates of the $q_i(\lambda)$ to the point of pure white.

Necessary calculations include:

- Determination of the average spectral distribution (2)

$$f_{av}(\lambda) = \sum_i f_i(\lambda) \quad (2)$$

Where f_i denotes all averaged distributions, f_{av} is the average distribution, and λ is the wavelength. f_{av} is normalized by dividing to maximum value.

- Control for available suppression of amplitudes due to normalization to peak value. Calculation the histogram h of the available amplitudes - a by formula (3) for the entire distribution.

$$h(a) = h(a) + 1 \quad a=[0, \dots, \text{max amplitude}] \quad (3)$$

For colorimetric transformation purposes, the data in the visible wavelength range is averaged at 5 or 10 nm, which reduces their number to 81 or 41. Since the distribution is normalized a limitation the levels of amplitude to a finite number of values is also needed. For example, for 11, amplitudes a are sufficient for each wavelength λ to be defined as:

$$a(\lambda) = \text{int}(f_{av}(\lambda) * 10 + 0.5) \quad (4)$$

Where int specifies conversion to integer value.

- Each distribution is converted to chromaticity through a standard transformation for the CIE-XYZ system [2]. Parameters are obtained by generic formula (5).

$$B = K \cdot \sum_{\lambda=380}^{780} f(\lambda) \cdot b(\lambda) \cdot \Delta\lambda \quad (5)$$

Where B denotes any of the X , Y or Z parameters ($XYZ.X$, $XYZ.Y$, $XYZ.Z$); $f(\lambda)$ is a spectral distribution; $b(\lambda)$ - is a spectral distribution of normalization curves for standard observer \bar{x}_{10} , \bar{y}_{10} , \bar{z}_{10} corresponding to each of the above three parameters; $\Delta\lambda=5$ - wavelength averaging interval; K - scaling factor for binding to the particular illuminant.

The chromaticity is determined by the coordinates $XYZ.x$ and $XYZ.y$ by formulas (6).

$$XYZ.x = \frac{XYZ.X}{XYZ.X + XYZ.Y + XYZ.Z} \quad (6)$$

$$XYZ.y = \frac{XYZ.Y}{XYZ.X + XYZ.Y + XYZ.Z}$$

The calculations give values independent of the scaling of the basic parameters, therefore, the need for determining the coefficient K in formula (5) is eliminated.

- The Euclidean distance, calculated between the pure white coordinates and the calculated chromaticity coordinates (7), is used towards to criterion of similarity.

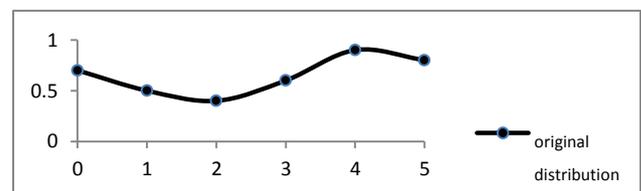
$$r_{(i=0..80)} = \sqrt{(XYZ.x_i - XYZ.x_w)^2 + (XYZ.y_i - XYZ.y_w)^2} \quad (7)$$

Where r is the distance; $XYZ.x_i$, $XYZ.y_i$ are chromaticity coordinates of the distribution; $XYZ.x_w=0.333333$; $XYZ.y_w=0.333333$ - coordinates of the pure white.

In order to achieve the goal of increasing the accuracy and eliminating the metamerism, an additional evaluation is proposed. That is performed through decomposition of the normalized ratio-spectra $q_i(\lambda)$ at series of impulses.

7. Compilation of distributions from the impulse decomposition. Determination of chromaticity coordinates for each component and its distance from the point of pure white.

Figure 3 illustrates schematically the procedure for decomposition of a normalized signal with 5 values at a set of distributions with maximal level 1 for all points except of one where is localized a single impulse that complements the signal amplitude to 1. Impulse decomposition gives an opportunity to obtain similarities for every wavelength.



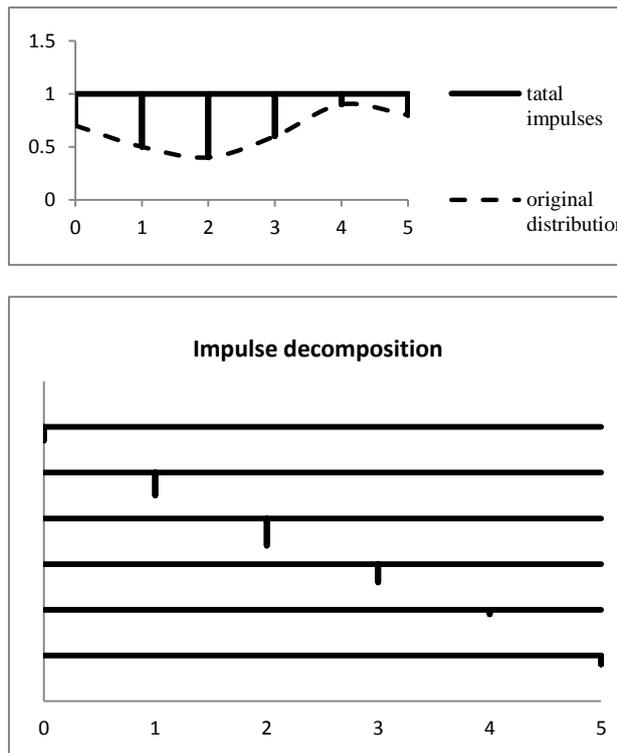


Fig. 3 Demonstration of impulse decomposition

8. Affiliation: for each wavelength is determined affiliation to disorder according to the etalon, for which there are chromaticity coordinates with minimal distance to the pure white. The final choice is made by the maximum choices of a particular disorder in the analyzed wavelength range.

A special computer program is developed for execution the steps of the sequence above.

Description of the test data

For the purpose of this study, reflection spectra from 50 patients with a total of 11 histological diagnoses were used. Every initial spectrum is result of three times measurement and subsequent averaging of the obtained values. Distributions were captured using "Ocean Optics" Inc. PC2000 fiber-optic micro spectrometer. The device is equipped with a diffractive grating with resolution of 600 1/mm, a linear CCD photodetector, and reproduces a total resolution of approximately 8.5 nm.

From the available 3648 values obtained in the wavelength range 346.32-1043.76 nm the data between 380 and 780 nm were separated and reduced to 81 values by averaging.

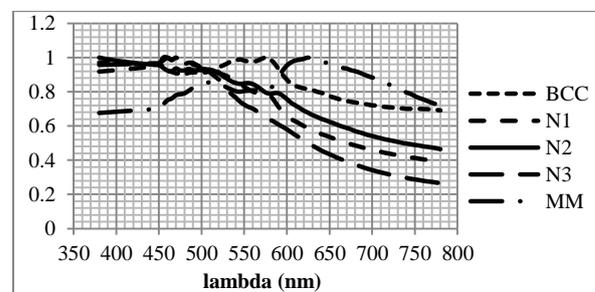
The etalon spectrum-ratios are defined for part of the disorders as follows: Basel cell carcinoma - BCC; Compound Nevus - N1; Dermal Nevus - N2; Dysplastic Nevus - N3; Malignant melanoma type 1 -MM. The test spectra are for a total of 12 diagnoses 7 of which do not have etalons and are marked with D1-D7.

It should be mentioned that because of the limitation of the sample, the specific compiled etalon distributions use data from a statistically insignificant number of patients, which makes them incomplete. The test-experiment provides a possible diagnosis for each patient falling within the reference group using the closest etalon.

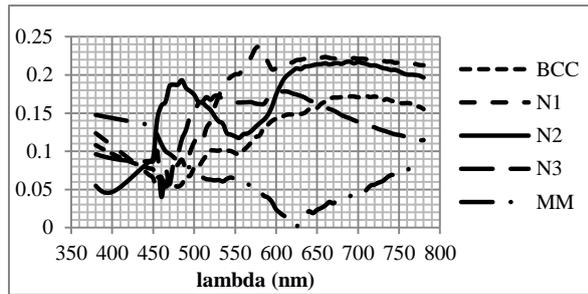
As long as the computational calculations are made with double precision data, for restriction the accuracy of the comparison, the distances are limited to the seven digits after decimal point by multiplying by 10⁷ and converting it to an integer value.

RESULTS AND DISCUSSION

Figure 4 (a) and (b) show the graphical appearance of the defined etalons and their standard deviations.



a)



b)

Fig. 4 Distributions- a) and standard deviation –b) of the defined etalons

It must be mentioned that there are areas where the standard deviations are evenly distributed.

An initial test procedure for distinguishing the etalons was carried out. Each pair etalons were compared according to the methodology for analysis of test-spectra described above. A right choice is obtained for all defined etalons.

Table 1 presented the statistical results of the check-up procedure for availability of suppression of amplitudes, which is due to a peak value for the all 50 spectra.

As it can be seen from the table, the maximum numbers of levels in the test ratio-spectra are located above the value 2. It means that there is no suppression of the amplitudes which is due to a normalizing to a peak value.

Table-1 Number of test spectra with maximum at mentioned value

| value | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------|---|---|---|----|----|---|----|----|----|----|----|
| etalon | | | | | | | | | | | |
| BCC | 0 | 0 | 0 | 3 | 2 | 1 | 9 | 11 | 10 | 12 | 2 |
| N1 | 0 | 0 | 0 | 1 | 7 | 9 | 12 | 4 | 7 | 8 | 2 |
| N2 | 0 | 0 | 0 | 3 | 2 | 7 | 5 | 10 | 10 | 7 | 6 |
| N3 | 0 | 0 | 1 | 13 | 10 | 5 | 5 | 5 | 5 | 6 | 0 |
| MM | 0 | 0 | 3 | 1 | 3 | 3 | 4 | 7 | 12 | 13 | 4 |

Table 2 shows an example of the output list of the comparison produced from the computer program.

If there is possible to choose more than one diagnose/etalon, the program gives a message *impossible choice*.

Table-2 Output data from the computer program. Content: distances x10⁷

| nm | 0 | 1 | 2 | 3 | 4 | * | ** | | 0 | 1 | 2 | 3 | 4 | * | ** |
|-----|-------|------|------|------|-------|---|----|-----|-------|-------|-------|-------|------|---|----|
| 380 | 33 | 26 | 37 | 26 | 122 | 1 | 1 | 585 | 3881 | 13874 | 10661 | 23061 | 6885 | 0 | 0 |
| 385 | 60 | 40 | 35 | 25 | 209 | 3 | 0 | 590 | 8662 | 19241 | 13152 | 27064 | 6582 | 4 | 0 |
| 390 | 135 | 86 | 29 | 41 | 424 | 2 | 0 | 595 | 14030 | 24527 | 17330 | 31366 | 6631 | 4 | 0 |
| 395 | 266 | 168 | 25 | 82 | 781 | 2 | 0 | 600 | 17657 | 28234 | 20514 | 34410 | 5770 | 4 | 0 |
| 400 | 527 | 326 | 58 | 170 | 1481 | 2 | 0 | 605 | 20371 | 30843 | 23312 | 36774 | 5468 | 4 | 0 |
| 405 | 879 | 529 | 132 | 286 | 2409 | 2 | 0 | 610 | 20870 | 31231 | 24011 | 37402 | 4431 | 4 | 0 |
| 410 | 1691 | 985 | 316 | 551 | 4533 | 2 | 0 | 615 | 20632 | 30916 | 24161 | 36901 | 4095 | 4 | 0 |
| 415 | 3072 | 1724 | 664 | 985 | 8099 | 2 | 0 | 620 | 19406 | 29532 | 23355 | 35259 | 3301 | 4 | 0 |
| 420 | 5401 | 2907 | 1298 | 1694 | 14040 | 2 | 0 | 625 | 17383 | 26792 | 21427 | 32237 | 2387 | 4 | 0 |
| 425 | 8768 | 4511 | 2290 | 2676 | 22518 | 2 | 0 | 630 | 14775 | 23397 | 18733 | 28221 | 1849 | 4 | 0 |
| 430 | 11767 | 5767 | 3287 | 3484 | 29882 | 2 | 0 | 635 | 12791 | 20256 | 16225 | 24534 | 1626 | 4 | 0 |
| 435 | 13852 | 6448 | 4097 | 3971 | 34781 | 3 | 0 | 640 | 10562 | 17094 | 13681 | 20752 | 1332 | 4 | 0 |
| 440 | 14963 | 6594 | 4648 | 4147 | 37146 | 3 | 0 | 645 | 8564 | 13937 | 11132 | 17017 | 879 | 4 | 0 |
| 445 | 15303 | 6362 | 4962 | 4095 | 37558 | 3 | 0 | 650 | 6864 | 11091 | 8886 | 13581 | 626 | 4 | 0 |
| 450 | 15229 | 5949 | 5128 | 3931 | 36955 | 3 | 0 | 655 | 5288 | 8645 | 6897 | 10622 | 402 | 4 | 0 |
| 455 | 15774 | 37 | 6828 | 1962 | 33119 | 1 | 0 | 660 | 4003 | 6577 | 5277 | 8128 | 260 | 4 | 0 |
| 460 | 17390 | 54 | 8405 | 135 | 28591 | 1 | 0 | 665 | 2896 | 4846 | 3881 | 6015 | 180 | 4 | 0 |
| 465 | 16228 | 707 | 7525 | 989 | 24475 | 1 | 0 | 670 | 2079 | 3513 | 2816 | 4394 | 93 | 4 | 0 |
| 470 | 14451 | 1678 | 5224 | 37 | 19005 | 3 | 0 | 675 | 1500 | 2578 | 2062 | 3227 | 72 | 4 | 0 |
| 475 | 11835 | 1617 | 5184 | 1541 | 14942 | 3 | 0 | 680 | 1089 | 1907 | 1537 | 2396 | 44 | 4 | 0 |

| | | | | | | | | | | | | | | | |
|-----|-------|-------|------|-------|-------|---|---|-----|-----|------|------|------|----|---|---|
| 480 | 8957 | 1864 | 3720 | 791 | 10902 | 3 | 0 | 685 | 747 | 1339 | 1077 | 1692 | 34 | 4 | 0 |
| 485 | 6818 | 1501 | 2784 | 1076 | 7949 | 3 | 0 | 690 | 503 | 924 | 741 | 1173 | 34 | 4 | 0 |
| 490 | 5683 | 1459 | 2747 | 1769 | 6154 | 1 | 0 | 695 | 341 | 640 | 515 | 819 | 34 | 4 | 0 |
| 495 | 5786 | 2143 | 2889 | 2171 | 6106 | 1 | 0 | 700 | 240 | 460 | 368 | 590 | 35 | 4 | 0 |
| 500 | 6720 | 2851 | 3455 | 3158 | 6657 | 1 | 0 | 705 | 162 | 323 | 258 | 418 | 37 | 4 | 0 |
| 505 | 8511 | 4061 | 4290 | 4367 | 7950 | 1 | 0 | 710 | 112 | 228 | 182 | 296 | 36 | 4 | 0 |
| 510 | 9913 | 5004 | 5126 | 6106 | 9334 | 1 | 0 | 715 | 76 | 157 | 125 | 206 | 36 | 4 | 0 |
| 515 | 11446 | 6890 | 6546 | 8473 | 10819 | 2 | 0 | 720 | 54 | 110 | 88 | 145 | 37 | 4 | 0 |
| 520 | 11235 | 7817 | 7173 | 10383 | 12254 | 2 | 0 | 725 | 41 | 75 | 61 | 98 | 36 | 4 | 0 |
| 525 | 10122 | 7689 | 8362 | 11977 | 13063 | 1 | 0 | 730 | 35 | 53 | 45 | 68 | 37 | 0 | 0 |
| 530 | 8301 | 7538 | 8624 | 14107 | 13580 | 1 | 0 | 735 | 34 | 43 | 38 | 51 | 37 | 0 | 0 |
| 535 | 5962 | 7105 | 8846 | 15691 | 13498 | 0 | 0 | 740 | 33 | 34 | 33 | 39 | 37 | 0 | 1 |
| 540 | 5082 | 7900 | 9133 | 16945 | 13897 | 0 | 0 | 745 | 35 | 35 | 34 | 36 | 37 | 2 | 0 |
| 545 | 4244 | 9363 | 9183 | 18545 | 13651 | 0 | 0 | 750 | 35 | 33 | 34 | 33 | 37 | 1 | 1 |
| 550 | 4536 | 10365 | 8589 | 19339 | 12757 | 0 | 0 | 755 | 37 | 36 | 36 | 36 | 37 | 1 | 1 |
| 555 | 4588 | 11071 | 7881 | 19492 | 11933 | 0 | 0 | 760 | 37 | 36 | 36 | 36 | 37 | 1 | 1 |
| 560 | 3940 | 11106 | 7554 | 18868 | 10740 | 0 | 0 | 765 | 36 | 34 | 35 | 33 | 37 | 3 | 0 |
| 565 | 3192 | 10871 | 7903 | 18857 | 9808 | 0 | 0 | 770 | 36 | 34 | 35 | 33 | 37 | 3 | 0 |
| 570 | 1197 | 9587 | 7794 | 18616 | 8552 | 0 | 0 | 775 | 36 | 34 | 35 | 33 | 37 | 3 | 0 |
| 575 | 37 | 8742 | 8317 | 18878 | 8053 | 0 | 0 | 780 | 37 | 37 | 37 | 37 | 37 | 0 | 1 |
| 580 | 430 | 9641 | 8780 | 19990 | 7417 | 0 | 0 | | | | | | | | |

Legend: Left column – wavelength;

First raw: column 0-4 –index for etalon 0-BCC, 1-N1, 2-N2, 3-N3, 4-MM; *-choice;

**-flag: 0=used 1=unused

| | | | | | | | |
|-------------------------------------|---------------|--------|--------|--------|-----------|--------|------|
| Etalon index | 0 | 1 | 2 | 3 | 4 | choice | flag |
| Total choices by decomposition Imp. | 13 | 10 | 12 | 12 | 28 | 4 | 0 |
| XYZ distances -XYZ | 109785 | 383222 | 257610 | 577081 | 275797 | 0 | 0 |

Decisions: Imp. choice: **MM** XYZ choice: **BCC**

Table 3 gives the data for obtained correspondences for the tested patients.

Table-3 Results

| patient | classification by: | | | Number of used impulses | Limited interval 450-645nm |
|---------|--------------------|-----|------|-------------------------|----------------------------|
| | histology | XYZ | Imp. | | |
| 1. | D1 | BCC | N1 | 77 | N1 |
| 2. | BCC | BCC | BCC | 77 | BCC |
| 3. | BCC | MM | MM | 77 | MM |
| 4. | BCC | BCC | BCC | 77 | N1 |
| 5. | BCC | N2 | N2 | 76 | imposs. choice |
| 6. | BCC | MM | MM | 72 | MM |

| | | | | | |
|-----|-----|-----|-----|----|----------------|
| 7. | BCC | N2 | MM | 74 | BCC |
| 8. | BCC | MM | MM | 74 | imposs. choice |
| 9. | N1 | N3 | N3 | 71 | N3 |
| 10. | N1 | BCC | BCC | 73 | BCC |
| 11. | N1 | N1 | N3 | 74 | N3 |
| 12. | N1 | BCC | MM | 75 | N1 |
| 13. | N1 | N1 | N3 | 74 | N3 |
| 14. | N1 | N3 | N3 | 76 | N3 |
| 15. | N1 | BCC | MM | 75 | BCC |
| 16. | N1 | BCC | MM | 76 | N1 |
| 17. | N2 | N2 | N2 | 76 | BCC |
| 18. | N2 | N2 | N2 | 79 | N2 |
| 19. | N2 | MM | MM | 74 | MM |

| | | | | | |
|-----|----|-----|-----|----|----------------|
| 20. | N2 | N1 | N3 | 74 | N3 |
| 21. | N2 | N3 | N3 | 70 | N3 |
| 22. | N2 | N3 | N3 | 70 | N3 |
| 23. | N2 | BCC | MM | 75 | MM |
| 24. | N3 | N1 | N3 | 76 | N3 |
| 25. | N3 | N1 | N3 | 74 | N3 |
| 26. | N3 | N2 | N1 | 76 | N1 |
| 27. | N3 | N3 | N3 | 70 | N3 |
| 28. | N3 | MM | MM | 73 | N1 |
| 29. | N3 | N3 | N3 | 72 | N3 |
| 30. | N3 | N3 | N3 | 69 | N3 |
| 31. | N3 | N1 | N3 | 72 | N3 |
| 32. | D2 | N3 | N3 | 73 | N3 |
| 33. | D3 | N1 | N3 | 76 | N3 |
| 34. | D3 | BCC | BCC | 73 | BCC |
| 35. | D3 | MM | MM | 77 | MM |
| 36. | D3 | BCC | BCC | 73 | BCC |
| 37. | D4 | MM | MM | 75 | MM |
| 38. | D4 | MM | MM | 74 | MM |
| 39. | D5 | MM | MM | 72 | MM |
| 40. | D6 | MM | MM | 74 | MM |
| 41. | D6 | N2 | N3 | 78 | N3 |
| 42. | D6 | BCC | MM | 78 | MM |
| 43. | MM | MM | MM | 74 | MM |
| 44. | D7 | N1 | N3 | 79 | N1 |
| 45. | MM | MM | MM | 76 | MM |
| 46. | MM | MM | MM | 73 | MM |
| 47. | D7 | BCC | N2 | 76 | N2 |
| 48. | MM | MM | MM | 75 | MM |
| 49. | D7 | N1 | N3 | 76 | N1 |
| 50. | MM | BCC | MM | 75 | imposs. choice |

Note: diagnoses D1-D7 – without etalons

The last column gives the number of the wavelengths on which the selection was made. As you can see, there is no choice for a full

number of impulses. The area in which there is no ability to make selection (flag = 1) is found to be for wavelengths above 650 nm. Compared to the areas in the asymmetry chart (Figure 1), it shows the coincidence of the areas with conditions of suspicious similarity. In the chart of standard deviation for the etalons, these are the areas for which the deviations are evenly distributed. This suggests the use of a limited area of wavelengths for searching similarity. The last column of Table 4 shows the results of limiting the interval of wavelengths into impulse decomposition between 450 and 645 nm. In practice, limiting the range approximates the diagnosis to the histological one. In the most cases, however, there is no difference from choice made by full interval.

Unfortunately it is impossible to fix etalons and similarity that give 100% accuracy of the results. One of the main reasons is the inability to be procured spectra of all the possible variations of skin disorders.

CONCLUSION

The demonstrated method has specific features but also provides scope for other tasks beyond the specific one for which it was originally developed. The main advantage is searching in a two-dimensional space. If it is used a method in which each wavelength is treated as a dimension of the features space, averaging of the visible range from 380 to 780nm to 5nm forms 81 dimensions. In the accepted practice for spectrum [13,14], in such cases, diminution is sought primarily by eliminating the features with the same values. In the method described here, the equal values mean available similarity.

The evaluation is not necessarily set in the XYZ system. Any other color system which has presentation of chromaticity plane could be used to complete an automated process of searching similarity.

It is assumed that consistency could be made without transformation to chromaticity, directly on the spectra-relation. However, in

this case, there should be a single criterion for maximum approximation to the etalon. An advantage of the choice here is the independence of precise values for the restrictive conditions. A prerequisite for applying these criteria to the classification scheme is the number of possible reference states (etalons) to be known in advance.

On the other hand, given that the necessary for comparison information is derived from the form of the discrete function distribution, the method provides good application capabilities for tracking objects with varying characteristics. The conditions that need to be available are:

- The objects under investigation are the same, but have variations in the reflectance spectra for their harmless-state ν . These variations are independent of the searching change and are superimposed on it.
- Identical variations p in the states gives identical changes at of reflective spectra of the main harmless-state.

This means that in the spectrum of the tested change are combined results of ν and p .

An additional condition is the gradual change in wavelengths to avoid the normalization effect to a peak value.

A similar task occurs with the use of fluorescence spectra for diagnosis [15]. The test spectrum is the result of stimulated light emission in the subject by irradiating with a specific color of light. The reaction of the sample depends on object's current state.

The application of the technique could be expanded outside the visible wavelength range, it is sufficient to have a pair of spectra from the same object: one used for etalon / or baseline state and a second one for creation a ratio-distribution between them. By counting the number of wavelengths, the range can always be originally translated to the range of 380-780 nm. The latter shows that the method offers

good opportunities for adapting to a particular task.

REFERENCES

1. ed. by G.R. Jones, A.G. Deakin, J. W. Spencer, Chromatic Monitoring of Complex Conditions, Taylor & Francis Group, 2008, ISBN: 978-1-58488-988-5
2. International commission on illumination, Colorimetry. Third Edition. Publication CIE N15.3:2004
3. Pavlova P., E. Borisova. L. Avramov. "Automation of the Skin Cancer Diagnostics on the Bases of Reflectance Spectroscopy". IX International Conference on "Biomedical physics and engineering", Sofia, 2004, Proc. Pages: 260-265.
4. Pavlova P., E. Borisova, L. Avramov, El. Petkova, P. Troyanova. Development of spectral etalons for skin cancer diagnosis. 20th International Laser Physics Workshop, (LPHYS'11) July, 2011, Sarajevo, Bosnia and Herzegovina.
5. CIE, Special metamerism index: change in observer, First edition, Publication CIE No.80, (Central Bureau of the CIE, Vienna, Austria, 1989).
6. Pavlova P. Application of chromaticity as a space for comparing of spectral distributions. International Conference AUTOMATICS AND INFORMATICS'17. Oct. 2017, Sofia, Bulgaria, pp: B-169-B-172.
7. Borisova E., Pavlova P., Pavlova E., Troyanova P., Avramov L., "Optical biopsy of human skin – tool for cutaneous tumours' diagnosis", Int. J. Bioautomation, 16(1): 53-72, (2012).
8. Tomatis S., Bartoli C., Bono A., Cascinelli N., Clemente C., Marchesini R. (1998), Spectrophotometric imaging of cutaneous pigmented lesions: discriminant

analysis, optical properties and histological characteristics, *J Photochem Photobiol B: Biology*, 42:32-39

9. Troyanova P., A beneficial effect of a short-term formal training course in epiluminescence microscopy on the diagnostic performance of dermatologists about cutaneous malignant melanoma, *Skin Res. Tech.* 9, 269-273 (2003).

10. Wallace V., Bamber J., Crawford D., Ott R., Mortimer P., Classification of reflectance spectra from pigmented skin lesions, a comparison of multivariate discriminant analysis and artificial neural networks, *Phys Med Biol* 45, 2859-2871 (2000).

11. Prigent S., Descombes X., Zugaj D., Zerubia J, Spectral analysis and unsupervised SVM classification for skin hyper-pigmentation classification, *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (Whispers)*, Jun 2010, Reykjavik, Iceland. 2010.

12. Daugherty G., *Pattern Recognition and Identification, an introduction.* Springer, 2013.

13. Tao Wang, Kang Shao, Qinying Chu, Yanfei Ren, Yiming Mu, Lijia Qu, Jie He, Changwen Jin and Bin Xia, Automics: an integrated platform for NMR-based metabonomics spectral processing and data analysis. *BMC Bioinformatics* 2009 10:83, <https://doi.org/10.1186/1471-2105-10-83>

14. Gautam R., S.Vanga, F. Ariese, S. Umapathy. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Techniques and Instrumentation* (2015) 2:8 DOI 10.1140/epjti/s40485-015-0018-6

15. Borisova E., P. Troyanova, P. Pavlova, L. Avramov, Diagnostics of pigmented skin tumors based on laser-induced autofluorescence and diffuse reflectance spectroscopy, *Quantum Electronics* 38(6), 597-605 (2008).