

YUV vs RGB – Choosing a Color Space for Human-Machine Interaction

Michał Podpora

Opole University of Technology
 Faculty of Electrical Engineering,
 Automatic Control and Informatics
 Institute of Control and Computer Engineering
 ul. Sosnkowskiego 31, 45-272 Opole, Poland
 m.podpora@po.opole.pl

Grzegorz Paweł Korbaś, Aleksandra Kawala-Janik

Opole University of Technology
 Faculty of Electrical Engineering,
 Automatic Control and Informatics
 Institute of Electromechanical Systems
 and Industrial Electronics
 ul. Prószkowska 76, 45-758 Opole, Poland
 {g.korbas, a.kawala-janik}@po.opole.pl

Abstract—This paper describes and compares two color spaces – YUV and RGB, taking into account possible human-computer interaction applications. Human perception-oriented properties are compared, including not only file size or bandwidth, but also subjective visibility of artifacts. 1700 tests on a group of 170 people were performed to describe the subjective quality of compressed YUV and RGB images. The paper shows that the use of the YUV color space for a machine vision implementation can give better subjective image quality than the RGB color space. The authors conclude that YUV is better for machine vision implementations than RGB due to the perceptual similarities to the human vision.

Keywords – YUV, RGB, color space, image processing, HCI, HMI, image understanding, computer graphic, machine vision, image compression

I. INTRODUCTION

MACHINE Vision and Computer Vision are dominated by the RGB color space, which seems to be the most intuitive programmer's choice, while it is being used by digital image acquisition hardware and in the majority of processing methods and algorithms. Red, green and blue optical filters, combined with Active-Pixel Sensors comprise the simplest and most popular color vision acquisition systems.

This article confronts the RGB color space with YUV. Although YUV was primarily introduced to add the color information to existing monochromatic channel, it turned out that YUV is also in a way similar to human vision – the "black and white" information has more impact on the image for human eye than the color information.

Therefore, it is worth considering which of these color spaces might be better for Human-Machine Interaction systems.

The authors have prepared a simple application to test perceptual capabilities of volunteers. The following chapters describe the research and present the conclusions.

II. RGB AND YUV COLOR SPACES

RGB and YUV color spaces are both based upon the perceptual capabilities of human eye. The RGB color space

is plainly based on the acquisition capabilities of cone cells in retina, which are able to react to different wavelengths. Electronic devices usually display three base colors (red, green and blue). Other colors and shades are achieved by combining these three colors using the additive color mixing. The cones' response to a specific wavelength is presented in fig.1. Fig.2 shows the spectrum of two exemplary computer monitors available on the market.

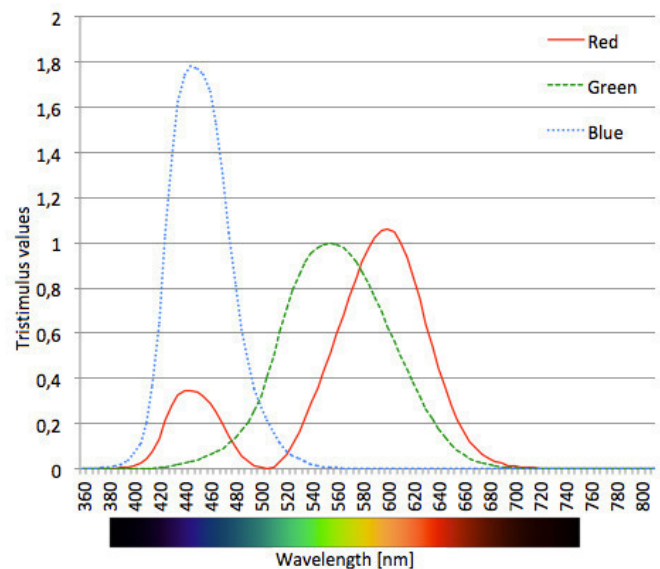


Fig. 1. CIE 1931 Color Matching Functions [4] [8]

Fig.1 (CIE 1931 CMF, corresponding to the acquisition capabilities of cone cells) and fig.2 (spectrograms of popular monitor technologies) are obviously different, but this difference, for a human eye, is perceptually negligible. Nevertheless, these figures show that there is a change in information when using RGB displays. Designers of Machine Vision systems should keep in mind that flattening the spectrum to three values is a huge simplification.

The YUV color space, on the other hand, can also be considered to be similar to human eye's retina, while the main

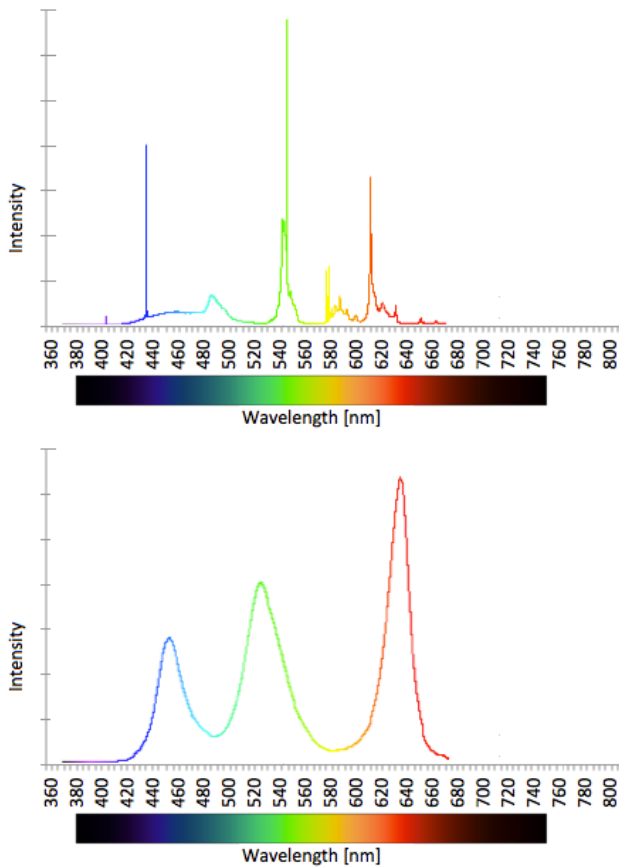


Fig. 2. Spectrograms of a CCFL-backlit LCD monitor (Samsung SyncMaster 913N) and a LED monitor (Samsung SyncMaster XL20), respectively [1]

channel – luminance (denoted as Y channel) or "luma" (denoted as Y') describes the intensity of light, just like rod cells in the retina. Rod cells are the primary source of information in the dark, when the cone cells do not have sufficient intensity of light for activation to distinguish colors. However, when the intensity slightly increases, additional information from cone cells become available. In the YUV color space, two additional channels – chrominance components called "U" and "V" – carry the color information (e.g. as blue-luminance and red-luminance, respectively, for a digital signal – in case of YCbCr).

In the YUV color space, the black and white information is separated from the color information. Primarily, YUV was used in analog television standards, when color information was added to the existing luminance channel. To enable backward compatibility for black-and-white transceivers, the chrominance channels were added in a separate subcarrier.

Using a YUV color space, also usually involves loss of information, but for a different reason than in RGB color space. In analog YUV it is popular to use interlacing in chrominance channels (the contrast in luminance channel is more significant information for a human eye than color in chrominance channels). In digital YUV, the signal is usually

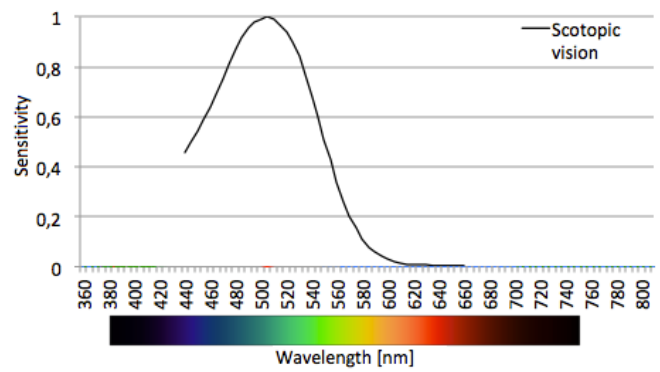


Fig. 3. Scotopic (rod cells) sensitivity [9] [10]

converted from RGB acquisition hardware, which involves a lossy conversion from RGB to YUV.

Therefore, the YUV color space is also a compromise of perceptually-reasoned loss of information.

A. RGB formats

The RGB color space has many various representations, but they all have one in common: three separate color values are stored for three predefined colors: red, green and blue. The colors can be ordered starting from red (RGB) or starting from blue (BGR). If the fourth letter ("A") is present, the fourth channel contains the "alpha" (transparency) value for the pixel. If the name of the format contains any digits, the usually mean the amount of bits for every pixel – e.g. RGB24p (or RGB24bpp) means 24 bits for pixel's color information (i.e. 8 bits for red, green and for blue), and BGRA8888 means 8 bits for subsequent channels (i.e. blue, green, red and alpha, respectively). The formats are well explained [6] in OpenCV-related webpages, regarding converting image between two specific formats (functions like `cv_bgr2gray()`, `cv_rgb2ycrcb()` and other).

Some alternative RGB formats are also available in OpenCV (e.g. Bayer pattern) but these are not discussed in this paper.

B. YUV formats and conversion from RGB

Historically, the term YUV was used for analog encoding. Nowadays this term is frequently used for analog and digital encoding as well. There are many formulas to convert from RGB to YUV [5] [8]. In this article digital YCbCr defined by ITU-R BT.601 has been used. In this color space Y (denoted as Y') represents "luma" (the weighted sum of gamma-compressed RGB components) while Cb and Cr are blue-difference and red-difference chrominance components. Referring to the mentioned recommendation YCbCr is derived as follows:

$$\begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (1)$$

where Rd, Gd, Bd represent 8-bit values for red, green and blue color channels.

C. Popularity of formats

A trial research has been carried out to check the popularity of the most popular image representation formats. Authors studied the hit count of webpages containing 41 most popular OpenCV format-conversion functions (inter alia: `cv_rgb2yuv()`, `cv_bgr2hsv()`). Fig.4 shows the total hit count (number of webpages) for specific image format conversion methods, divided into 4 groups: grayscale, RGB, YUV, HSV. Fig.4 shows also average hit count (number of webpages divided by the number of queried conversion methods names).

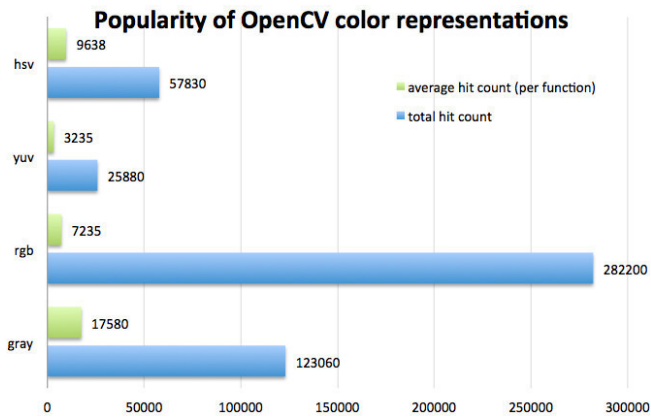


Fig. 4. Popularity of the most popular color representations based on the Google hit count for 41 most popular OpenCV [6] format conversion functions [own work]

Fig.4 clearly indicates that the OpenCV conversion functions for RGB color space have the greatest number of webpages/resources/queries, while the YUV color space is currently much less popular amongst OpenCV programmers.

III. RESEARCH INCENTIVES AND EXPECTATIONS

One of the authors in [7] has suggested that YUV color space might be more similar to human vision than RGB. However, the research [7] involved only a brief comparison of the color spaces and discussion on possible compression differences. The issue was analyzed in a Machine Vision aspect, no human factor/opinion was taken into account.

In this paper, the authors have decided to investigate if the subjective quality difference is real (i.e. if it is also visible/noticeable to other people), and carried out extensive tests to find out if the YUV representation of an image has perceptually better quality than RGB. It turned out that, indeed, people have noticed the difference in quality.

Technically (and mathematically) compressing RGB (RGB24p) and (YUV888) color spaces images should give comparable quality, while in both of them there are three bytes describing every single pixel. However, the "black and white" detail has more impact on the image for a human eye because of its rather low color sensitivity. Manipulating with red, green or blue value always gives a perceptibly different image, while converting an image to the YUV color space gives possibility to process the luminance and

chrominance signals independently. The luminance channel is surely considered the most useful one for image processing in YUV, therefore reducing the chrominance signal quality may pass unnoticed to a human.

The idea of chroma subsampling has been formerly used for image coding in YUV formats, e.g. YUV422, but in every case the output image quality was aimed at a human recipient. If the image is to be analyzed by a Machine Vision system, the RGB color space is nowadays considered to be the most useful form of the visual information. This is not necessarily true. If a robot is supposed to work and to co-exist along with human, it should "see" the world the way we do – with similar inaccuracies. The threshold and the ability of "not recognizing" an object is the key issue of learning and robot's better understanding of it's environment. [7]

Converting an RGB image to the YUV color space is a lossy operation – this might be the reason (according to [7]) compressed YUV image files are often smaller than compressed RGB image files. However, the difference in the file sizes seems not to be proportional to the difference in quality of the images. It is difficult to discuss the change in quality as the quality loss is usually defined as the difference between original image and the compressed image. In that case, the YUV representation of the image should be considered as the worse one (due to additional lossy operation - RGB-YUV conversion). If (after the same lossy compression/conversion) the RGB representation would be of better quality than YUV, it would be natural to try to lower the threshold for compression/conversion of YUV to improve its quality (so that the quality of RGB and YUV would become similar). Surprisingly, the quality of the RGB representation did not seem to be of better quality than YUV. Contrary, it seemed to be of worse quality than YUV. Authors have named it as the subjective quality difference.

To confirm the existence of the subjective quality difference, and to assess it's extent, a simple test application has been developed. A brief description of the application, test images, testing procedures and the discussion of the results are included in the next section (Research Methodology).

A. The preparation of test images

The authors expected to find and estimate the subjective quality difference between RGB and YUV images. To ensure the quality of the research, all images have been processed using the same algorithms and settings. The basic image preparation procedure, presented in fig.5, consisted of following steps:

- choosing an interesting image with representative thus unique image attributes (contrast, quality, saturation, edges, gradients, etc.) and satisfying ppi (pixels per inch) resolution,
- cropping the image to 256x256 pixels,
- saving the image as an RGB 24bpp format bmp file.

Both images, visible on the form of the application (fig.6) are created in runtime, basing on the same RGB 24bpp test image.

One of ten predefined test images is loaded and processed in following steps:

- the image is cloned in the application's memory,
- one of the images is converted to YUV888 using algorithm from equation (1),
- both matrices are converted using DWT (discrete wavelet transform) [2] with a specific threshold,
- both images are recovered using IDWT (inverse DWT),
- the YUV image is converted back to RGB,
- both images are displayed on the application's form.

The threshold of the DWT algorithm of one of the images is random (in a predefined range), whereas the threshold of the second image's DWT algorithm is available on the front-end of the application to user as a trackbar - the user is able to modify it's value (and re-run the DWT-IDWT algorithm with the new setting).

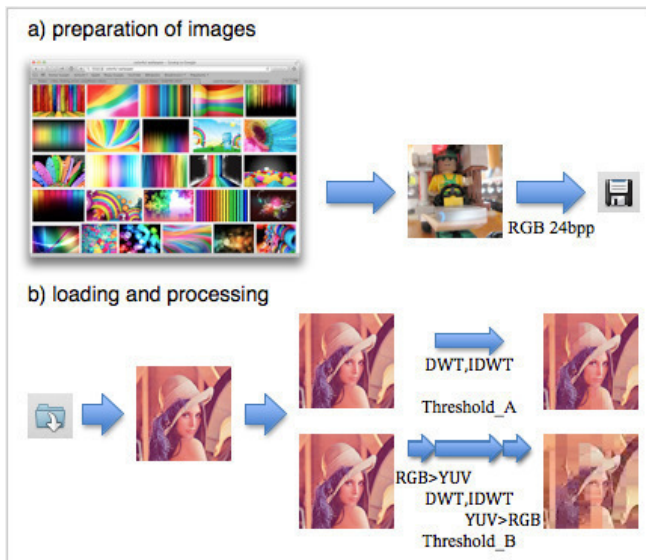


Fig. 5. Image preparation procedure for the application's test images

Since the quality/distortion threshold parameter, as the authors suggest, should be modeled on human perception rather than simply as a variance of difference between input and output image, some perceptual distortion measures should be developed. Audio compression perceptual models are relatively advanced (mp3, ogg), the perception aspect is also present in some of the compression algorithms of image data (usually available to users as a quality threshold value), but in the aspect of color space, currently there seem to be no research.

IV. RESEARCH METHODOLOGY

In order to investigate the subjective difference in images, an application has been implemented. The graphical user interface of the application is presented in fig.6. The application form includes two panels and a slider. Each of these two panels show an image, based on the same source but converted in a different way. The source images have good quality, but the

images in GUI have significantly lower quality (so that the quality drop would be clearly visible to a human). The source images were transformed using different DWT threshold value for each panel. One of the GUI images was transformed directly from the RGB color space, and the second one was transformed to the YUV color space first. During the tests, the threshold value of the DWT transformation for the RGB-based image was set to a (random) fixed value, while for the YUV-based image the user was able to modify the DWT threshold using a slider. The tests included also an inverted scenario: a fixed threshold for YUV and a slider for RGB.

Ten tests were conducted in every test scheme (five various images were loaded and presented in two following procedures: at first the threshold for RGB was fixed and then the threshold for YUV was fixed). Pictures that were chosen for the research, comprised a set of interesting features, inter alia: variable complexity, clear edges as well as some gradients, good color saturation, etc. The source images were 256x256 pixels, 24 bits per pixel, RGB, uncompressed, BMP images.

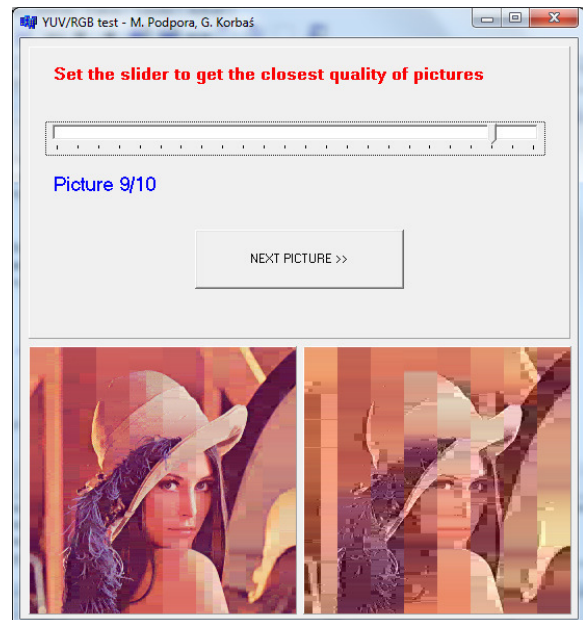


Fig. 6. Graphical user interface of the application

During the trials, users were asked to set the slider in such a way that the quality of the two images would seem similar. The slider offered 24 positions, translated into 24 threshold values of the wavelet transform, affecting the quality of one of the images. The default slider position was either 1 or 24, while the user was supposed to set the slider to 7–17 (the DWT threshold of the second image was randomized from the range: 7–17). If the user did not modify the default sidebar position (and left it on 1 or 24) it clearly indicated that the test results should be rejected. By moving the slider to an intermediate position, the user could subjectively ascertain if the quality of the two images was at a comparable level.

The study involved 170 people, aged from 10 to 40 years

(most aged 17-24 years). This has provided 1700 test results. 7% of the tests were rejected because of the extreme positions of sliders (indubitable quality difference). However, further analysis was carried out for both situations: not only for tests marked as correct, but for all tests (without rejecting any) and the results were very similar - the users that did not bother to use the slider, did not use it in both test configurations: fixed-RGB and fixed-YUV, influencing both result data sets in the same way.

The authors decided that the basic parameter analyzed in this study is the number of zeros present in the matrix describing the image after wavelet transform. If the two images are subjectively the same quality, the greater the number of zeroes in the DWT matrices of one of the images means that less information is needed to describe the image. A greater number of zeroes also enables a possibility of higher level of image compression, which translates into the potential application of the results. In order to compare the usefulness of a color space, following factor has been defined:

$$P = \frac{Z_{YUV} - Z_{RGB}}{Z_{RGB}} \cdot 100 \quad (2)$$

where Z_{RGB} is the number of zeroes in the DWT matrices when using the RGB color space, and Z_{YUV} is number of zeroes when using the YUV space. The P coefficient indicates percentage – the number of zeroes that have been found using YUV compared to the number of zeroes that have been found using RGB. A positive coefficient value means that (after the user had set the slider position to set similar quality of both images) more zeroes occurred in the YUV, while a negative value of the coefficient indicates the superiority of RGB color space.

It is worth noting that some of the slider positions allowed to obtain positive and some – negative values of the P coefficient. An exemplary dependence of P upon the slider’s position (for one of the tests with fixed DWT threshold for the RGB and modifiable one for YUV) is shown in fig.7. The distribution of the slider’s positions set by users for this particular test is shown in fig.8. (Positions 20 and 22 were excluded from evaluation.)

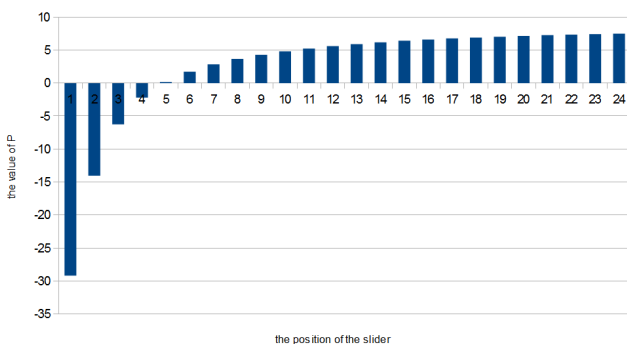


Fig. 7. Dependence of the P coefficient upon specific slider positions (i.e. specific threshold values) for one specific exemplary test

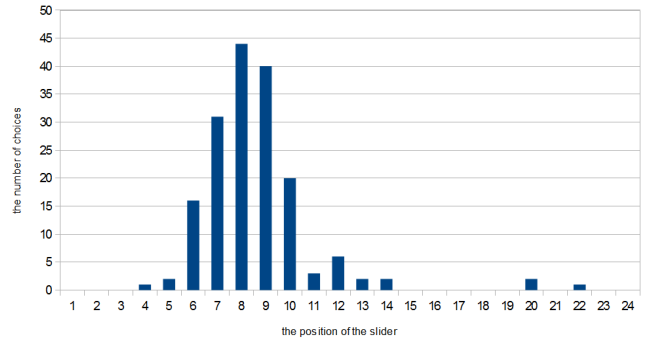


Fig. 8. Exemplary statistics of slider positions for one specific test

V. RESULTS

The results were analyzed separately for each test. In each case, the average value of the coefficient P was positive - ranging from 0.59 to 4.40. After merging the results of all tests, the average of the P coefficient has been obtained. Statistical results for all images and tests are shown in fig.9.

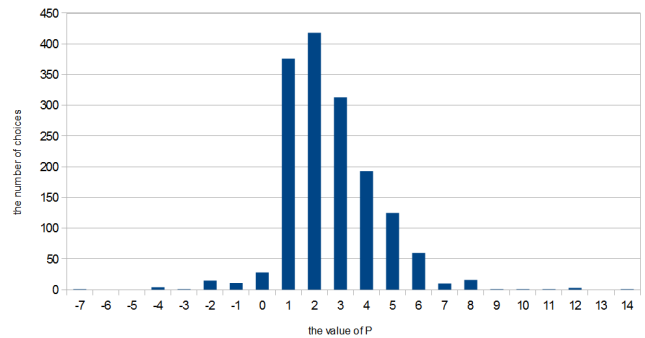


Fig. 9. General statistics of the P coefficient (all tests)

The results indicate that the use of the YUV space in some cases may be more effective than using the RGB color space. These conclusions are of a general nature, but they are also true for each image individually. The use of the YUV space results in a larger number of zeroes in the DWT matrices.

VI. ADVANTAGES FOR HMI SYSTEMS, FUTURE WORK

The study showed that people qualify the images as qualitatively similar even though the images are described using a different form of information. Comparison of the RGB and YUV color spaces in conjunction with the wavelet transform shows that the use of the YUV space enables efficient reduction of the amount of information necessary to represent the image of subjectively similar quality. Therefore, systems designed to map the human factor in the field of image processing should use YUV color space. A smaller number of data needed to make a decision may result, among others, in faster performance and/or reduction of the size of transmitted data in the system. This perceptual difference in quality can be used in another way - by modifying the subjective image quality (if image size reduction is not required [7]) by adjusting the

luminance channel compression threshold value to improve perceptual quality while preserving comparable file size.

It is noteworthy that the conversion between RGB and YUV was performed using the specific conversion coefficient values defined by ITU-R BT.601.

The analysis of other coefficients, their values, thresholds and their impact on the effectiveness of the use of YUV space should be subjected to further research.

The positive results related to the analysis of the human factor, of course do not preclude the benefits of the YUV space in traditional decision-making systems. There are many algorithms related to the recognition of shapes of objects, searching for specific parameters, motion detection, etc. It is possible that the use of the YUV color space can bring many benefits also in the classical cases. The authors consider further research in this field is to allow a broader view and more accurate analysis of application areas of YUV instead of RGB color space.

REFERENCES

- [1] Artamonov O., Xbit Laboratories, 'Contemporary LCD Monitor Parameters - Objective and Subjective Analysis', 2007, available on-line (accessed 2014-03-01) <http://www.xbitlabs.com/articles/monitors/display/lcd-parameters.html>
- [2] Berowski P., 'Transformata Falkowa', Instytut Elektrotechniki, Warszawa, available on-line (accessed 2014-04-01) http://bambus.iel.waw.pl/pliki/ogolne/studia_doktoranckie/wyklady/wyklad_falki.pdf
- [3] Color and Vision Research Labs, 'CIE Standards', 2006, available on-line (accessed 2014-03-01) <http://cvrli.ioo.ucl.ac.uk/cmfs.htm>
- [4] Glynn E. F., 'CIE Chromaticity Diagrams - CIE Color Matching Functions', 2009, available on-line (accessed 2014-03-01) <http://www.efg2.com/Lab/Graphics/Colors/Chromaticity.htm>
- [5] Hamilton E., 'JPEG File Interchange Format, Version 1.02', available on-line (accessed 2014-03-08) <http://www.jpeg.org/public/jfif.pdf>
- [6] Open Source Computer Vision Library, 'OpenCV online documentation', available on-line (accessed 2014-03-01) <http://opencv.org>
- [7] Podpora M. A., 'YUV vs. RGB – A comparison of lossy compressions for human-oriented man-machine interfaces', *III SWD conference proceedings*, Glucholazy 2009, Oficyna Wydawnicza Politechniki Opolskiej, ISSN 1429-1533, vol. 62, no. 329/2009, Opole, 2009
- [8] Szeliski R., 'Computer Vision: Algorithms and Applications', Springer, 2011
- [9] Wald G., Brown P. K., 'Human rhodopsin', *Science*, 1958
- [10] Wandell B. A., 'Foundations of Vision', Sinauer Associates Inc., ISBN 978-0878938537, Sunderland, 1995, available on-line (accessed 2014-03-01) <https://www.stanford.edu/group/vista/cgi-bin/FOV/>

[1] Artamonov O., Xbit Laboratories, 'Contemporary LCD Monitor Parameters - Objective and Subjective Anal-