Julian BALCEREK, Adam DĄBROWSKI, Adam KONIECZKA

Poznan University of Technology, Chair of Control and System Engineering, Division of Signal Processing and Electronic Systems

Fast 2D to 3D image conversion based on reduced number of controlling parameters

Abstract. *In this paper a new approach to fast 2D to 3D image conversion based on the reduced number of individually adjusted controlling parameters and simplified depth maps is presented. With the reported experiments conducted among viewers, five simple 2D to 3D conversion schemes were examined and compared. The results of the first experiment indicate a linear dependence between the controlling parameters among the examined viewers. In the second experiment, for the assumed linear dependence between the parameters, obtained in the first experiment but for another test image, a clear 3D effect perception was confirmed.*

Streszczenie. *W artykule zaprezentowano nowe ujęcie problemu szybkiej konwersji obrazów dwuwymiarowych (2D) do trójwymiarowych (3D) oparte na zredukowanej liczbie indywidualnie dobieranych parametrów kontrolnych i uproszczonych mapach głębokości. W eksperymentach przebadano i porównano 5 prostych schematów konwersji, których parametry były ustawiane przez widzów. Wynik pierwszego eksperymentu wskazuje na liniową zależność pomiędzy tymi parametrami w populacji badanych osób. W drugim eksperymencie, dla założonej, uzyskanej w pierwszym eksperymencie liniowej zależności pomiędzy parametrami, ale dla innego obrazu testowego, potwierdzono postrzeganie wyraźnego efektu 3D. (**Szybka konwersja obrazów 2D do 3D oparta na zredukowanej liczbie parametrów kontrolnych**).*

Keywords: stereovision impressions, 2D to 3D conversion, anaglyph, depth map, image processing.
Słowa kluczowe: efekty stereowizyjne, konwersja 2D do 3D, anaglif, mapa głębokości, przetwarzanie obrazów.

doi:10.12915/pe.2014.12.03

Introduction

Presently, most of images and sequences of images are still recorded using two-dimensional (2D) techniques, i.e., with monocular devices. Professional stereovision or three-dimensional (3D) cameras are rare and expensive thus are rather seldom used. In some situations it is even impossible to use stereo recording. An example is a special movie effect called the "forced perspective", which is an optical illusion with an object appearing closer or farther away from the camera and perceived as larger or smaller than it really is. In other situations the typically recorded stereo effect is weak. This occurs, e.g., for far scene objects, due to too small stereo separation between the left and the right stereo images, as the camera lenses are typically placed to near to each other for the far recorded objects. Thus efficient 2D to 3D conversion is an important issue and there is an urgent demand for preparation of simple effective (i.e. real-time) tools for the 2D to 3D image or sequence of images conversion with some control of the perceived depth [1–4].

Image processing methods used for the 2D to 3D conversion allow for an approximation of the real 3D effect only but they may be sufficient having human ability to visual perception in the consideration. Taking simplicity of visualization into account, which is adequate for the easy and approximate approach, the authors have decided to consider the anaglyph technique [5–6]. Presented ideas may also be implemented with other visualization techniques like, e.g., the autostereoscopic screens. Using autostereoscopy, there is no need to use any special headgears or glasses by the viewers but the parallax barrier or lenticular based displays are required [7].

In this paper the authors analyze and verify five efficient real-time 2D to 3D conversion schemes, proposed in their previous publications [8–11], i.e.: direct shift, direct shift with interpolation, mirroring with differential filtering, segment shift, and finally the segment scaling. All these conversion schemes are based on horizontal shifts of the red color component of the chosen objects and the background.

The first and the simplest scheme is the direct shift, which consists merely in the red color component shifts, without any care taken to the information holes (i.e. to those regions, occurring after shifting the red color component, in which no information about this component is available).

In the other four proposed schemes different methods for filling the information holes are used. In the second scheme, i.e. the direct shift with interpolation, the information holes are filled in by means of the linear interpolation. In the third, i.e., mirroring with differential filtering scheme, the information holes are replaced with the respective regions of the mirror image followed by the differential filtering. In case of the fourth, i.e., segment shift scheme, the red color component information holes are filled with the rescaled neighboring segments of the image. In the last considered scheme, the segment scaling, the appropriate image segments are scaled up to cover the information holes.

All considered schemes are based on shifts of the red color component separately for the objects and for the background according to the depth (disparity) map [4, 12–14] with the use of two parameters controlling the shifts measured with the numbers of pixels: the first parameter for the background shift and the second one for the object or group of objects shift.

After a short introduction in this Section an approach to the depth map generation and the 2D to 3D conversion is described in Section 2. In Section 3 the results of the first experiment are reported with a special attention paid to the linear dependence between the 3D effect controlling parameters among the examined viewers. In Section 4 the results of the second experiment are reported. The authors assumed and then experimentally confirmed the linear dependence between the indicated conversion parameters. A relatively strong human tolerance to various values of the 3D effect parameters was also shown. In the last Section 5, final remarks and directions for the future research on the presented subject are formulated.

Depth maps based 2D to 3D conversion

In the presented approach to the 2D to 3D conversion, simplified, i.e., binary depth maps were used. For the depth map generation, it was assumed that the metric depth value $Z_{\max}$ is a distance from the camera lens to the point of the farthest background plan, $Z_{\min}$ is a distance from the camera lens to the point of the nearest foreground, and Z is a distance from the camera lens to the point of the interest. The distance Z lies certainly between $Z_{\min}$ and $Z_{\max}$. Other designations are: $v_{\min}$ – the pixel depth value of the point of the farthest background, $v_{\max}$ – the pixel depth value of the point of the nearest foreground, and v – the pixel depth value

of the point of the interest. The depth value v is between $v_{\min}$ and $v_{\max}$ [8, 15]. For an N-bit depth map an inverse proportionality can be written as

$$\begin{aligned} & Z_{\min} \leftrightarrow v_{\max} = 2^N - 1 \\ & Z \leftrightarrow v \\ & Z_{\max} \leftrightarrow v_{\min} = 0 . \end{aligned} \tag{1}$$

The inverse proportionality from equation (1) can be rewritten to a direct proportionality:

$$\begin{aligned} & Z_{\max} - Z_{\min} \leftrightarrow v_{\max} - v_{\min} = 2^N - 1 \\ & Z - Z_{\min} \leftrightarrow v_{\max} - v = 2^N - 1 - v . \end{aligned} \tag{2}$$

Thus using the direct proportionality it can be written that

$$(Z_{\max} - Z_{\min})(2^N - 1 - v) = (Z - Z_{\min})(2^N - 1) . \tag{3}$$

For the depth map preparation using the known metric values of distances from the camera lens to the 3D scene objects, equation (3) can be transformed into a form

$$v = \text{round}\left((2^N - 1) \frac{Z - Z_{\max}}{Z_{\min} - Z_{\max}} \right), v \in \langle 0, 2^N - 1 \rangle . \tag{4}$$

Equation (4) describes translation of metric values to depth values for linear depth quantization. An equivalent equation for 8-bit gray scale levels is presented in [15]. In the simple schemes presented in this paper 1-bit (bivalent, binary) depth maps are used.

Using binary depth maps there are only two red color component shift parameters. As already mentioned, the first parameter is referred to the background and the second to the object or group of objects. Both parameters should be adjusted experimentally, separately for the object(s) and separately for the background for all considered schemes in order to obtain the best 3D effect. The results are described in the next paragraphs.

Adjustment of controlling parameters

In the first experiment the authors wanted to verify whether there occurs any dependence between the controlling parameters [16].

Three test images were used: *lena* (*lena_color*, the standard test image) with resolution 512×512 pixels, *bugatti* with resolution 518×389 pixels, and *slup* with resolution 509×382 pixels (Fig. 1).

a) b) c)

Fig. 1. Test images used for investigations during the first experiment: a) *lena*, b) *bugatti*, c) *slup*

56 students of the Poznan University of Technology participated in the first experiment. The distance from the viewer to the 24 inch diagonal LCD screen was about 0.8 m. Viewers wore anaglyph glasses with the red and cyan (green and blue) filters.

In the first experiment the test application implemented in the MATLAB programming environment, version 7.11.0 (R2010b) was used. The graphical user interface (GUI) is screened in Fig. 2.

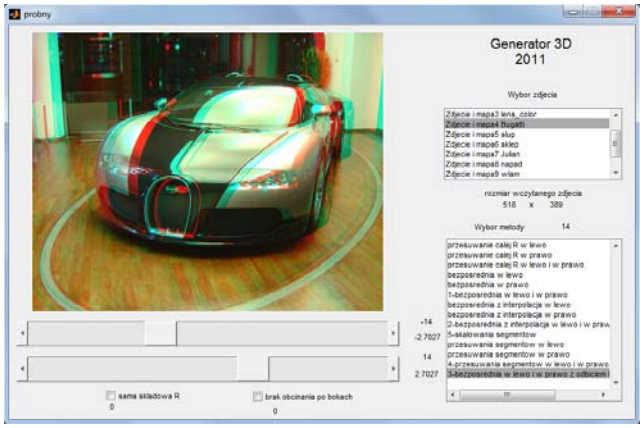


Fig. 2. Application window used in the first experiment

Each viewer saw three test images for five conversion schemes. They were numbered form 1 to 5 in the following order: 1) the direct shift, 2) the direct shift with interpolation, 3) the mirroring with differential filtering, 4) the segment shift, and 5) the segment scaling.

For each image the examined person had to adjust two best parameter values (in the sense of the best subjective 3D effect) referred to each scheme or to affirm that the adjustment of the optimal parameters is impossible. For adjustment of the background shift and the object shift the top slider and the bottom slider were used, respectively. These parameters were referred to the red color component shift or the level of rescaling [8–11]. Viewers using two sliders arbitrary operated in the range from –50 to 50 pixels for parameters of the schemes 1–4 and in the range from 1 to 6 for parameters of the scheme no. 5. In the segment scaling scheme, each image line is a sum of background and object segments a according to the background width and objects areas. Image horizontal resolution is denoted by x according to

$$a_1 + a_2 + \ldots = x. \tag{5}$$

Scaling ratio w is specified separately for background and objects:

$$a_1 \cdot w_1 + a_2 \cdot w_2 + \ldots = x_n . \tag{6}$$

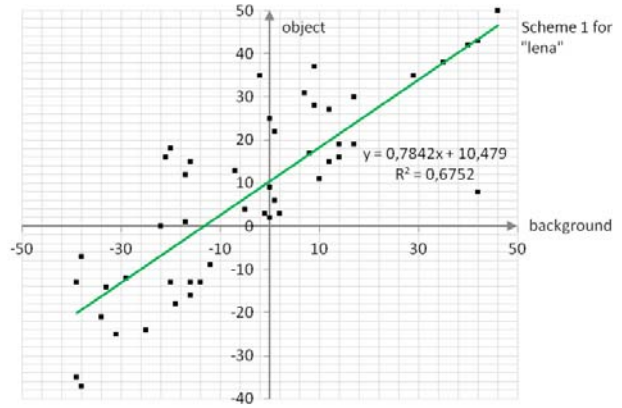
In order to avoid the new image resolution x_n, original image resolution is preserved if the segment scaling scheme is used according to equation:

$$a_1 \cdot w_1 \cdot \frac{x}{x_n} + a_2 \cdot w_2 \cdot \frac{x}{x_n} + \ldots = \frac{x_n}{x_n} x. \tag{7}$$

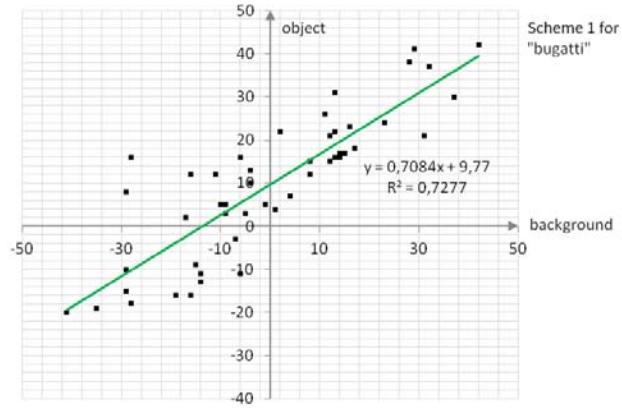
Scaling ratio w is adjusted by users separately for background and objects. The results of each parameter change were visualized on the application window in real time or in almost real time.

Distributions of the best parameter values among viewers are illustrated in Fig. 3 as functions of the object versus background shifts for all conversion schemes and all tested images. In each case five outliers (most distinct points) were removed for the calculation of the linear regression and the square of the Pearson product-moment correlation coefficient.

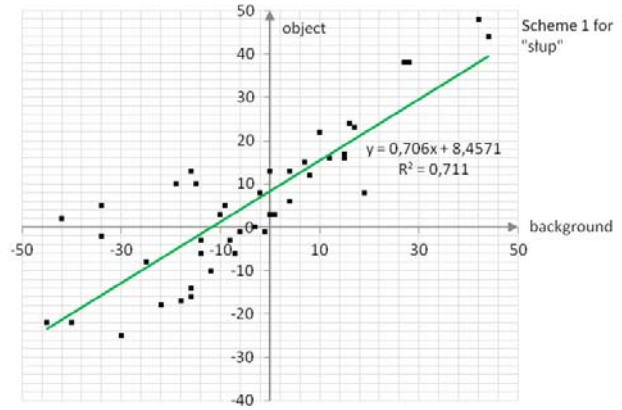
a)


object
Scheme 1 for "lena"
y = 0,7842x + 10,479
R² = 0,6752
background

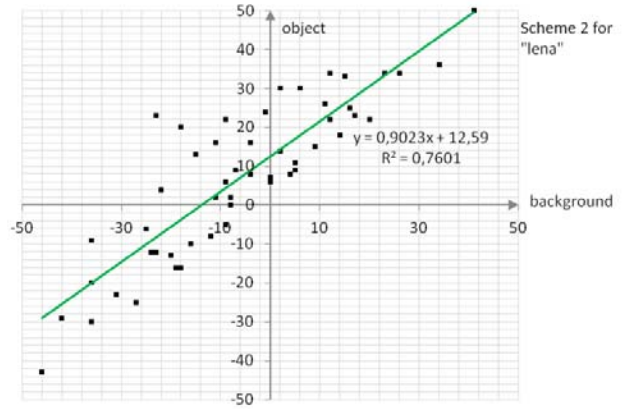
b)


object
Scheme 1 for "bugatti"
y = 0,7084x + 9,77
R² = 0,7277
background

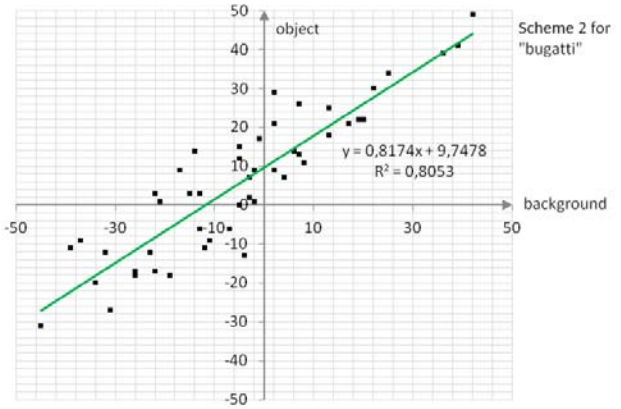
c)


object
Scheme 1 for "słup"
y = 0,706x + 8,4571
R² = 0,711
background

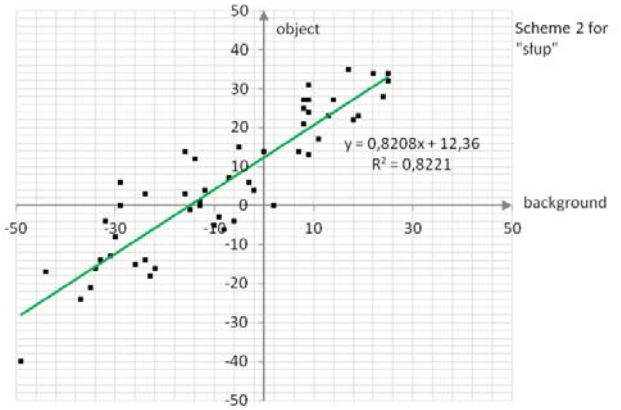
d)


object
Scheme 2 for "lena"
y = 0,9023x + 12,59
R² = 0,7601
background

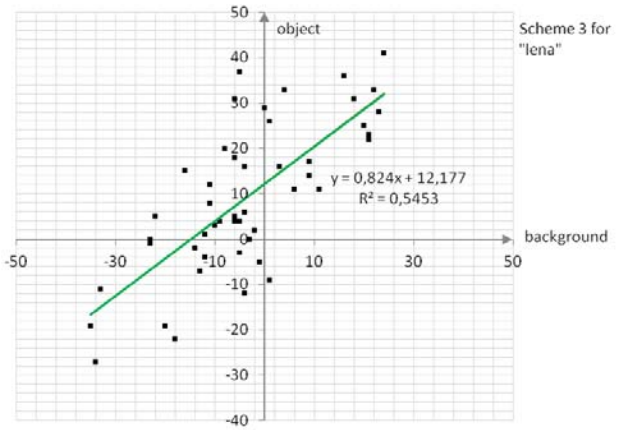
e)


object
Scheme 2 for "bugatti"
y = 0,8174x + 9,7478
R² = 0,8053
background

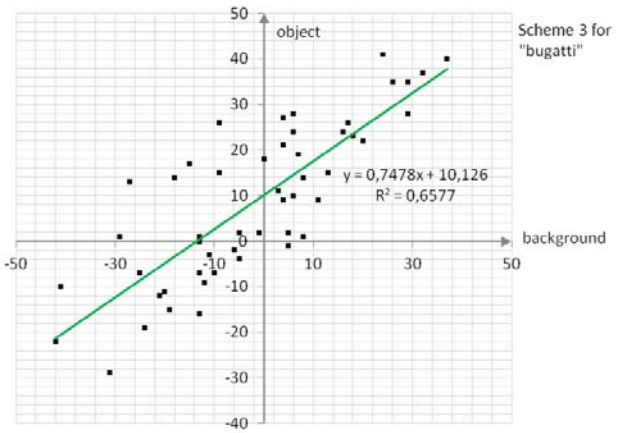
f)


object
Scheme 2 for "słup"
y = 0,8208x + 12,36
R² = 0,8221
background

g)


object
Scheme 3 for "lena"
y = 0,824x + 12,177
R² = 0,5453
background

h)


object
Scheme 3 for "bugatti"
y = 0,7478x + 10,126
R² = 0,6577
background

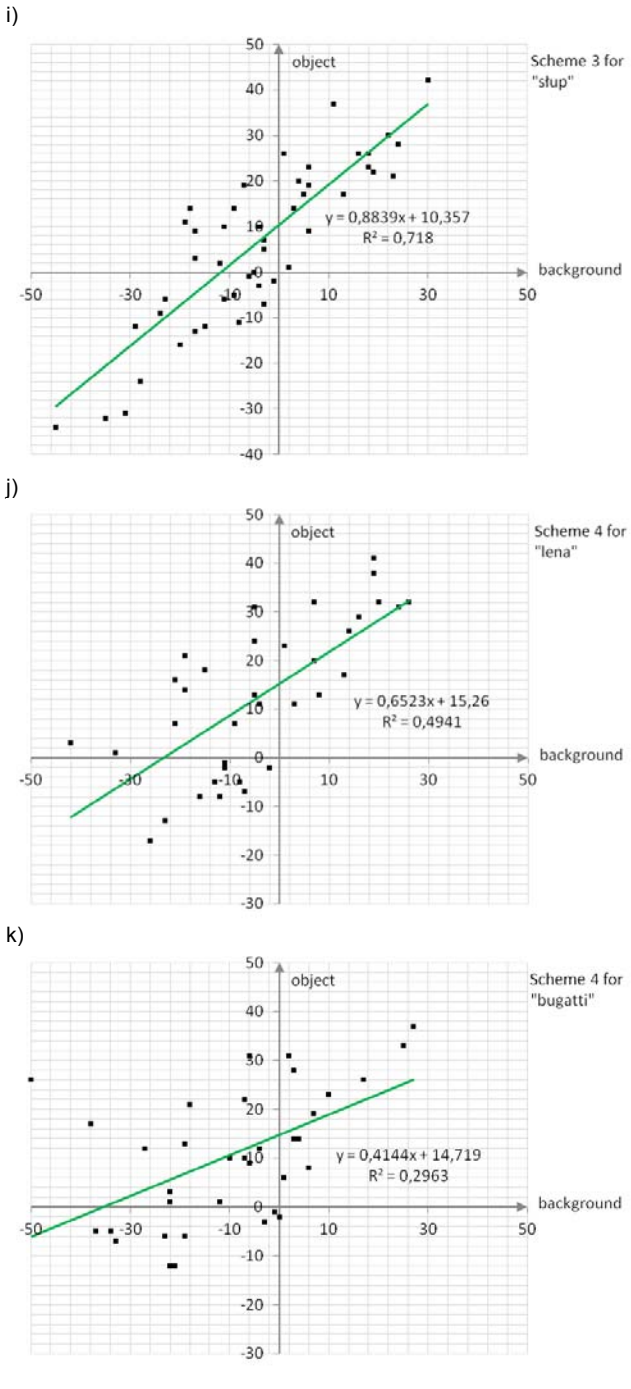


Fig. 3. Distributions of the best parameter values as functions of the object versus background shifts for five conversion schemes: a)–c), d)–f), g)–i), j)–l), m)–o) and three test images: a), d), g), j), m) *lena*, b), e), h), k), n) *bugatti*, c), f), i), l), o) *slup*

In Table 1 a comparison of the best parameters for the proposed conversion schemes is presented. In cases of the direct shift, direct shift with interpolation as well as for the mirroring with differential filtering schemes, the viewers perceived clear 3D effects and could adjust the optimal control parameter values with the mean of at least 95 %. On the contrary a smaller number of people (with the mean less than 71 %) perceived the 3D effect and could adjust the optimal parameter values for two last, i.e., the segment shifting and the segment scaling schemes. Thus these two schemes occurred to be less effective than the others.

A clearly linear relationship between parameters for the control of the object and the background can be observed. Equations resulting from the linear regression and squares

of the Pearson product-moment correlation coefficients are also presented in Table 1.

Table 1. Comparison of 2D to 3D conversion schemes as results of the first experiment

Schemes	Image	Percentage of people who saw 3D effect and could adjust optimal parameters [in %]		Linear regression	Square of Pearson product-moment correlation coefficient (R^2)
1) direct shift	*lena*	100.00	mean: 95.24	$y = 0.78x+10.48$	0.68
	bugatti	94.64		$y = 0.71x+9.77$	0.73
	slup	91.07		$y = 0.71x+8.46$	0.71
2) direct shift with interpolation	*lena*	100.00	mean: 98.81	$y = 0.90x+12.59$	0.76
	bugatti	96.42		$y = 0.82x+9.75$	0.81
	slup	100.00		$y = 0.82x+12.36$	0.82
3) mirroring with differential filtering	*lena*	96.43	mean: 97.02	$y = 0.82x+12.18$	0.55
	bugatti	98.21		$y = 0.75x+10.13$	0.66
	slup	96.43		$y = 0.88x+10.36$	0.72
4) segment shift	*lena*	73.21	mean: 70.83	$y = 0.65x+15.26$	0.49
	bugatti	73.21		$y = 0.41x+14.72$	0.30
	slup	66.07		$y = 0.77x+13.49$	0.62
5) segment scaling	*lena*	58.93	mean: 57.74	$y = 0.83x+1.26$	0.73
	bugatti	53.57		$y = 0.90x+0.70$	0.80
	slup	60.71		$y = 0.76x+1.33$	0.65

Dependence of linear parameters

In the second experiment the authors wanted to check a tolerance of the human 3D perception (perception of depth) on the 2D to 3D conversion control parameters. Therefore the linear relations of the conversion control parameters obtained in the first experiment were averaged for particular conversion schemes (Tab. 2.). Then the test images were again processed but using these averaged linear relations.

25 students from the Poznan University of Technology participated in the second experiment. This was another group of students than that participated in the first experiment. In comparison to the first experiment a new image namely the *wintercar* image with resolution 768x576 pixels was also tested. This is an example of a typical frame from the video monitoring (Fig. 4).

Fig. 4. *Wintercar* test image – an illustrative frame from the video monitoring

A modified version of the test application from the first experiment was used (Fig. 5). Each examined person watched four test images for five schemes numbered from 1 to 5 in the same way as in the first experiment. For each image the viewer had to adjust a single the best parameter value referred to the background for each conversion scheme (the object shift controlling parameter was adjusted automatically according to the respective linear relation).

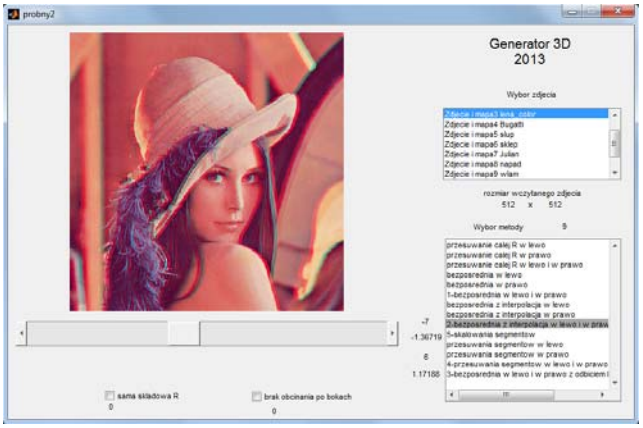


Fig. 5. Application window used in the second experiment

In Table 2 a comparison of percentages of people is presented, who saw the 3D effect and could adjust the control parameter in order to find its optimal value.

Table 2. Comparison of 2D to 3D conversion schemes from the second experiment

Schemes	Image	Linear dependence	Percentage of people who saw 3D effect and could adjust optimal parameter [in %]	
1) direct shift	*lena*	$y = 0.73x+9.57$	100.00	mean: 91.00
	bugatti		96.00	
	slup		100.00	
	wintercar		68.00	
2) direct shift with interpolation	*lena*	$y = 0.85x+11.57$	96.00	mean: 90.00
	bugatti		92.00	
	slup		88.00	
	wintercar		84.00	
3) mirroring with differential filtering	*lena*	$y = 0.82x+10.89$	92.00	mean: 93.00
	bugatti		100.00	
	slup		92.00	
	wintercar		88.00	
4) segment shift	*lena*	$y = 0.61x+14.49$	80.00	mean: 84.00
	bugatti		96.00	
	slup		76.00	
	wintercar		84.00	
5) segment scaling	*lena*	$y = 0.83x+1.10$	24.00	mean: 24.00
	bugatti		28.00	
	slup		24.00	
	wintercar		20.00	

In cases of the direct shift, direct shift with interpolation as well as for the mirroring with differential filtering the viewers perceived the clear 3D effect and could adjust the optimal value of the control parameter with the mean of at least 90 %. Even in case of a test image newly prepared for this experiment, percentage of cases with observable and optimally adjustable 3D effect did not severely decreased for direct shift with interpolation, mirroring with differential filtering, and segment shift schemes.

A smaller number of people (the mean of 84 % and 24%) perceived the 3D effect and could adjust the optimal parameter values for the segment shifting and the segment scaling schemes, respectively. Thus these two schemes occurred to be rather ineffective.

This experiment indicates on the best observable effects of conversion and visualization for mirroring with differential filtering scheme with mean of 93 % of cases.

Concluding remarks

The results of two conducted experiments indicate that the best 3D impression visibility is obtained for the direct shift, direct shift with interpolation, and mirroring with

differential filtering schemes. These schemes are simple, with a low computational cost, and offer a relatively good 3D quality anaglyph images. There is a possibility of implementation the 2D to 3D conversion using authors schemes in the real time.

Experimentally determined linear correlations based on the obtained values of the square of the Pearson product moment correlation coefficient, show strong linear dependences between the adjustable parameters for all examined 2D to 3D conversion schemes.

Occurrence of a linear dependence between parameters was experimentally confirmed. Even for assumed approximated linear dependences and even for other test image there were not severely decrease with 3D effect perception. People can choose optimal parameters of 3D perception but for slightly different values of parameters, 3D effect is still perceived. This indicates on a relatively high tolerance of human sense of sight on various parameters of 3D effect.

As a result, in presented approach, it is possible to reduce the number of the adjustable parameters to only one with the second being controlled according to the predetermined linear function. The end user may have a possibility to conveniently operate on this only one parameter in order to obtain the best perceptible and the most pleasant 3D effect (Fig. 6).

The presented results may be adapted to other visualization methods than the anaglyphs and may be used not only for the 2D to 3D conversion but also as methods of depth range increasing for the real two channel 3D images or sequences of images.

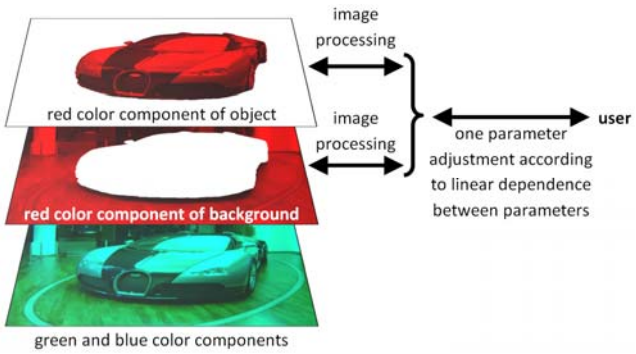


Fig. 6. User adaptable interface for 2D to 3D conversion and 3D visualization

The authors plan to perform new experiments using a higher number of images and sequences of images, more advanced depth maps with higher number of distance levels, and more advanced visualization techniques such as the autostereoscopy. The authors will continue to work on improvement in the conversion schemas. The authors will also try to qualify and quantify influences of all important features, such as environmental conditions, on optimal conversion parameters. The best conversion schemes with the best controlling parameters will be used in a fully automatic 2D to 3D conversion system which can possibly operate in a real time.

There are numerous possible applications for presented conversion approach. It can be applied as an auxiliary 3D tool for operator of visual monitoring system, in 3D displays for modern car interiors, and 3D displays of mobile devices like: smartphones, tablets, or personal navigations.

The paper was prepared within the INDECT and DS projects.

REFERENCES

[1] Tam W. J., Zhang L., 3D-TV Content Generation: 2D-to-3D Conversion, *IEEE International Conference on Multimedia and Expo*, (2006), 1869-1872

[2] Kim D. H., Min D., Sohn K. W., Stereoscopic Video Generation Method Using Motion Analysis, *3DTV Conference*, (2007), 1-4

[3] Li S.; Wang F., Liu W., The overview of 2D to 3D conversion system, *IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD)*, (2010), Vol. 2, Pages: 1388-1392

[4] Chang Y.L., Tsai Y.P.; Chang T. H., Chen Y. R.; Lei S., A depth map refinement algorithm for 2D-to-3D conversion, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (2012), 1437-1440

[5] Dubois E., A projection method to generate anaglyph stereo images, *Proc. of 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'01)*, Volume 3, (2001), 1661-1664

[6] Gallagher A. C., Detecting anaglyph images with channel alignment features, *Proc. of 2010 17th IEEE International Conference on Image Processing (ICIP)*, (2010), 2985-2988

[7] Urey H., Chellappan K.V., Erden E., Surman P., State of the Art in Stereoscopic and Autostereoscopic Displays, *Proceedings of the IEEE*, Volume: 99, Issue: 4, (2011), 540-555

[8] Balcerek J., Dąbrowski A., Konieczka A., Simple efficient techniques for creating effective 3D impressions from 2D original images, *Proc. of New Trends in Audio and Video / Signal Processing: Algorithms, Architectures, Arrangements and Applications NTAV / SPA'2008*, IEEE Poland Section Chapters Signal Processing, Circuits and Systems, Poznań, Poland, (2008), 219-224

[9] Balcerek J., Konieczka A., Dąbrowski A., Stankiewicz M., Krzykowska A,. Brightness Correction and Stereovision Impression Based Methods of Perceived Quality Improvement of CCTV Video Sequences, *Proc. of 4th International Conference on Multimedia Communications, Services and Security (MCSS 2011)*, Springer: *Communications in Computer and Information Science*, (2011), Volume 149, 64-72

[10] Balcerek J., Konieczka A., Dąbrowski A., Marciniak T., Binary depth map generation and color component hole filling for 3D effects in monitoring systems, *Proc. of Signal Processing SPA'2011*, IEEE, Poland Section, Chapters Signal Processing, Circuits and Systems, (2011), 138-143

[11] Balcerek J., Konieczka A., Dąbrowski A., Stankiewicz M., Krzykowska A., Approach to evoking stereovision impressions from images, *Przegląd Elektrotechniczny*, 88 (2012), Issue 6, 17-23

[12] Zhang L., Tam W. J., Stereoscopic image generation based on depth images for 3D TV, *IEEE Transactions on Broadcasting*, 51 (2005), Issue 2, 191-199

[13] Ideses I., Yaroslavsky L., Amit I., Fishbain B., Depth Map Quantization – How Much is Sufficient?, *3DTV Conference*, (2007), 1-4

[14] Chang Y. L., Fang C. Y., Ding L. F., Chen S. Y., Chen L.G., Depth Map Generation for 2D-to-3D Conversion by Short-Term Motion Assisted Color Segmentation, *2007 IEEE International Conference on Multimedia and Expo*, (2007), 1958-1961

[15] Fehn C., Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV, Proc. of SPIE 5291, *Stereoscopic Displays and Virtual Reality Systems* XI, (2004), 93-104

[16] Dąbrowski A., Balcerek J., Konieczka A., An Approach to Adjustment and Reduction of the Number of Controlling Parameters for Simple 2D to 3D Image Conversion Schemes, *Proc. of New Trends in Audio and VIideo/Signal Processing: Algorithms, Architectures, Arrangements and Applications NTAV/SPA'2012*, IEEE Poland Section Chapters Signal Processing, Circuits and Systems, Łódź, Poland, (2012), 227-231

Authors: *MSc Julian Balcerek, e-mail: julian.balcerek@put.poznan.pl, Prof. Adam Dąbrowski, e-mail: adam.dabrowski@put.poznan.pl, MSc Adam Konieczka, e-mail: adam.konieczka@put.poznan.pl* – Poznań University of Technology, Chair of Control and System Engineering, Division of Signal Processing and Electronic Systems

.