Brightness correction and stereovision impression based methods of perceived quality improvement of CCTV video sequences

mgr inż. Julian Balcerek, mgr inż. Adam Konieczka,
prof. dr hab. inż. Adam Dąbrowski, inż. Mateusz Stankiewicz,
inż. Agnieszka Krzykowska

Poznań University of Technology, Institute of Control and System Engineering,
Division of Signal Processing and Electronic Systems
Piotrowo 3a Street, 60-965 Poznań, Poland
{julian.balcerek, adam.konieczka, adam.dabrowski}@put.poznan.pl

Abstract. In this paper brightness correction and stereovision impression based methods of the perceived quality improvement of CCTV video sequences are presented. These methods are helpful for the monitoring operator in order to evoke attention during a long time observation. Clearness of picture is increased by using local brightness correction method. Another available option is a 3D visualization that can be used, e.g., to observe important image regions, which require an increased attention. Stereovision impression experiments and a real-time 2D to 3D video conversion tool are described.

Keywords: CCTV, brightness correction, stereovision, anaglyph, 2D to 3D video conversion

1 Introduction

CCTV (*closed-circuit television*) is a system that is commonly used for surveillance in urban areas that may need monitoring. The resulting job of a monitoring system officer (operator), who must observe the monitor screen for many hours, is quite monotonous and exhausting. In this paper we propose some new image processing tools that can serve as optional aids for the monitoring operator, helping him / her to stay actively conscious for a long time. This is because we offer view diversity and improvement of the perceived quality of images. Two mechanisms are proposed:

- increase or decrease of local brightness in order to increase the image clearness and thus the information perception
- real-time 2D to 3D conversion in order to increase the subjective informative content of the observed scenes.

It has to be stressed that the proposed image transformations are optional, thus they will never burden the monitoring operator. They are helpful in special situations only, e.g., if there is an interesting detail, which requires an increased attention. An experimental software has been prepared for verification of the proposed concepts.

Monitored events taking place in evenings and nights, filmed with artificial illumination, may result in low contrast images with poor informative content. The proposed local increase of brightness transforms, as a matter of fact, the view from a real to somehow artificial world but may essentially improve the human perception.

Stereovision impressions in a typical 2D monitoring system are possible, e.g., with the use of special (red / cyan glasses). The scenes are then perceived as if they contained collections of objects at different distances to the viewer. This phenomenon may be used for visualization of important events that have to be noticed and correctly interpreted.

Experiments on human 3D perception have been performed and the proper ways of image processing have been found [9]. Due to 3D impressions, the monitoring operator job is easier as his / her sight may easily be focused on the most important image regions and / or objects such as: people, moving vehicles, etc. This method can also help the monitoring operator to notice and remember important details.

The paper is structured as follows. After an introduction in this section, local brightness correction method is presented in Section 2. Section 3 is devoted to description of the 3D visualization tool together with the stereovision impression experiments and the real-time 2D to 3D video conversion. Conclusions are formulated in the last section.

2 Local brightness correction method

Basic approaches to sharpness improvement, exposure correction, brightness, contrast, and color adjustment, and digital noise removing are presented in [1], [2], [3], and [4]. In this paper a new method for clearness and visibility improvement of CCTV frames captured under difficult circumstances (especially in dark places or at night), is proposed.

In first step of our method for each CCTV frame an RGB to HSV color space conversion is proceeded. H and S components remain not modified. V component image is divided into square blocks starting from the top left corner of the picture. Sizes of all blocks are the same. Number of blocks in the shorter side of the picture is defined by the user as a control parameter. In the next step, for each component block the minimal and the maximal V values are selected. Then all V values are linearly scaled to the whole available range using formula

$$V_{\mathrm{n}} = \left(\frac{V - V_{\min}}{V_{\max} - V_{\min}} \right), \tag{1}$$

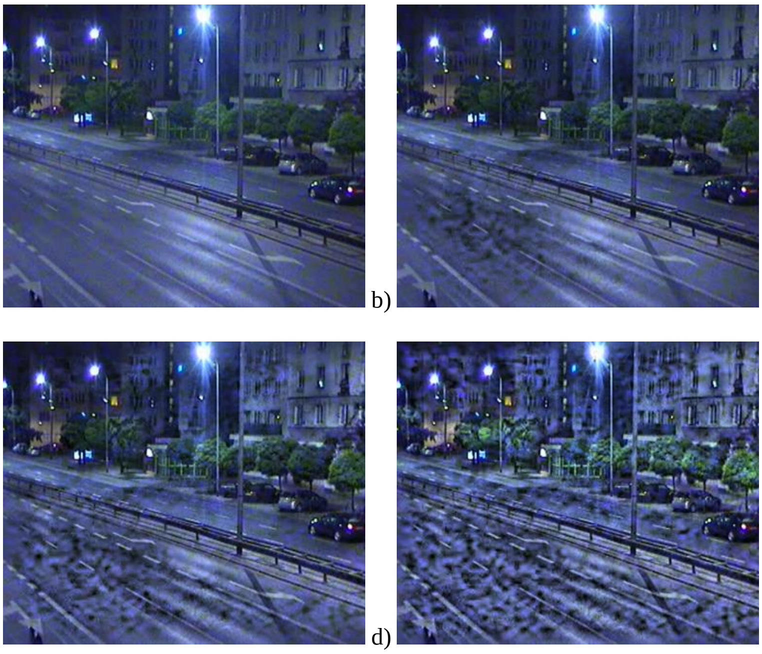
where:

V_{n} – new V component value of a pixel,
V – V component value of this pixel in the original picture,
$V_{\max}$ – maximal V component value in the block,
$V_{\min}$ – minimal V component value in the block.

In order to avoid creation pseudo-edges in the final image, this operation must be repeated. Number of iterations is the same as the one dimensional pixel size of the block. Computations in each iteration (excluding the first one) start one pixel below and one pixel to the right in relation to the previous iteration. In the next stage, new V component values must be averaged for each pixel. Near the picture edges original pixel values are left.

The proposed algorithm facilitates the viewer to perceive differences in brightness of dark objects. It helps in person identification by inspection, in reading inscriptions (i.e. license plates), or in distinguishing colors of very dark picture elements in CCTV recordings. Unfortunately, the proposed local brightness enhancement is a cause of digital noise enhancement in the picture. However, despite this inconvenience, recognition of noisy objects is much easier. This is an important advantage of the proposed method in comparison to classic manual or automatic brightness / contrast adjustment. Furthermore, our method is fully automatic.

As already mentioned, the only parameter, which user can modify, is the number of picture blocks *n*. If *n* is small, the visual effectiveness is low and the computational complexity is high (many iterations), however if *n* is too large, the final picture is not clear enough. It was experimentally proved, that the optimal value for *n* is about 10. The results of the proposed method for various values of *n* are presented in Fig. 1.



a) b) c) d)

Fig. 1. Local brightness correction: a) original picture, b) picture processed with *n*=5, c) picture processed with *n*=10, d) picture processed with *n*=20

Compression loss of the original picture does not have any considerable influence on the quality of the final picture. In extreme cases, i.e., if the compression factor is greater than 100:1, compression noise artifacts may be visible.

3 3D visualization tool

3.1 Stereovision impression experiments

Differences in viewpoints between two two-dimensional (2D) images observed separately by the left and the right eye simultaneously are interpreted by human brain as a three-dimensional (3D) scene [5]. A strict 3D reproduction is influenced by the choice of the interaxial distance, focal length of the reference camera, and the convergence distance [6]. An interesting feature of human 3D perception is that in spite of some incomplete or even inconsistent 3D information, evoking the three-dimensional illusion is still possible.

Almost entire visible color space can be modeled with three color components as RGB. Even if one eye (e.g., the left one) reaches only a single (e.g., the red) component of the left eye image, and the right eye sees the two remaining components (i.e., green and blue) of the right eye image, our brain perceives not only a 3D scene but sees it in quasi-true colors [7], [8]. This 3D effect can be realized using classic 2D equipment by observing specially prepared flat images referred to as the *anaglyph images* with red and cyan filter glasses. The red glass removes blue and green components from the left eye view while the cyan glass filters out the red component from the right eye view.

We have experimentally observed that even if merely the red color component is simply horizontally shifted, a 3D effect appears. This is the way to transform 2D images into their suggestive, perhaps not exact but plausible 3D anaglyph versions. We have conducted a series of experiments. The first and the second experiments were only preliminary experiments in small group of viewers and the third experiment was the main experiment. PUT workers and students were examined.

In our first experiment we observed that during increasing the red component shift value, the 3D effect was stronger and stronger but the image quality was decreasing, because some annoying artifacts near to the image contours appeared to be stronger and stronger visible. Lastly, starting from a specific value of the shift, two separate images without any 3D effect were perceived. In 83% of cases the 3D effect occurred. Among viewers, who observed the 3D effect when the red component was shifted to the left, in 92% of cases viewers noticed that the distance to the overall image surface was increased. When, however, the red component was shifted to the right, in 67% of the examined cases, the viewers noticed that the distance to the overall image surface was decreased. The others viewers perceived the exactly opposite effect [9]. However, the subjective informative content of the image seemed in both cases to be increased.

In our second experiment among viewers who observed a 3D effect when the red component was shifted to the left, in 71% of the cases the viewers noticed the increasing distance to the overall image surface. When the red component was shifted to the right, in 66% of the cases they also noticed that the distance was increasing [9].

In our third experiment test images were observed by 30 viewers. Resolutions of original images were 1280×1024 pixels. In 4 test pictures the red color component was shifted from 0 to 20% (0–255 pixels) of the original image width, both to the left and to the right side. Shifting step was 20 pixels because in case of a smaller step differences between pictures were too weak and experiment could be too tiring for the viewers. For each original picture there were 26 red component shifted pictures, 13 to the left and 13 to the right.

Application implemented in *C++* programming language and *QT* library was used (Fig. 2). Four different images were displayed on a 15.4 inches diagonal wide-screen with about 0.5 m distance from the viewer to the screen. Images were rescaled to the 960×768 pixels resolution in order to fit the viewing area for tests.

Viewers rated images by answering to given questions. The following questions and scoring criteria have been used. The first question was: “Can you see any 3D effect in the picture?” and there were two possible answers “yes” or “no” according to “1” or “0” value. The second question was: “Are there any visible artifacts?” and there were also two answers “yes” or “no” according to “1” and “0” value. The third question was “How does the distance between the viewer and the displayed objects change?” and two possible answers “increases” and “decreases” according to “1” and “0” value. If the viewer did not see any 3D effect, the button for this question was deactivated. The last question was “What is the perceptible quality of the picture?” and there were three possible answers “good”, “medium” and “bad” according to the values “3”, “2” and “1”.

Fig. 2. Application window of for image testing

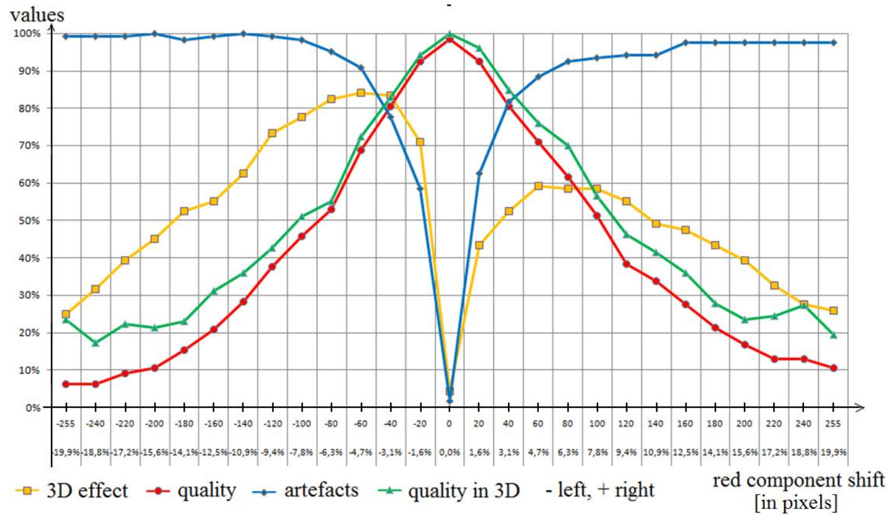


Fig. 3. Results of experiments with red component shifts

Results of our experiments explain how the human brain interprets and perceives quality of the image, occurrence of the 3D effect, and the level of artifacts after simple red component shifts (Fig. 3). They are the basis for development of simple methods for real-time 2D to 3D image or video conversion in CCTV systems.

3. 2 Real-time 2D to 3D conversion

An example of a 3-dimensional CCTV application is a 3D motion detection and assessment system described in [10]. Among possible advantages of using 3D instead of a 2D classic system is reduction of false alarms in case of difficult environment conditions like shadows, reflections, rain, or snow. Another advantage of using the 3D approach, i.e., a system with more than one camera, is reliability. Indeed, in case of a camera failure the system can still operate with the remaining camera(s).

Both classic and stereovision systems can be used for obstacle and vehicle detection in traffic [11], [12]. In case of a single-camera system the vehicle localization and tracking should be based on a 2D image sequence [11]. However, the precise vehicle distance can only be computed using a stereovision system based on the offset, measured between the left and the right images. Pairs of the corresponding points in the left and the right images have to be found and mapped into a calibrated 3D world model using the stereo geometry. Then moving objects can detected and tracked using a quite involved algorithm [12].

We suggest another approach, which, instead of the described automatic detection, can help people to extract information from the classic 2D monitoring system. Thus in our case a 3D effect is realized with 2D to 3D conversion methods, which must be simple enough to operate in real-time on a video sequence. We assume that the anaglyph images are generated and are observed with red and cyan filter glasses. The proposed real-time 2D to 3D conversion schema is illustrated in Fig. 4.

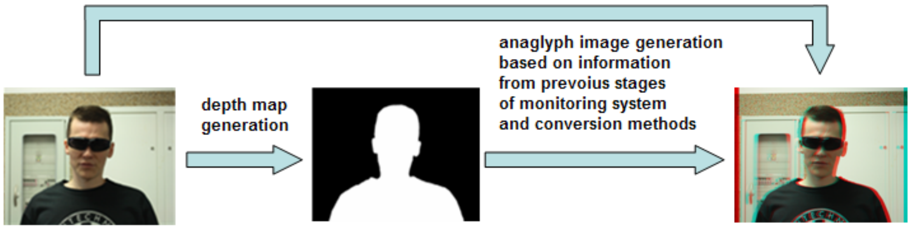


Fig. 4. A real-time 2D to 3D conversion schema for monitoring system

An *OpenCV* computer vision library is used. Examples of windows of a real-time 2D to 3D conversion application are presented in Fig. 5.

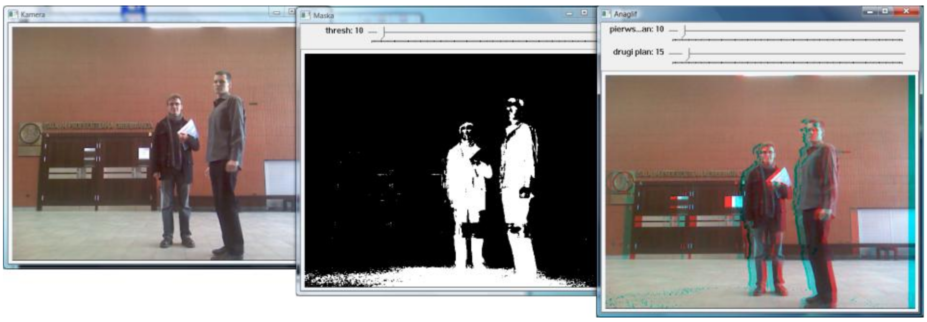


Fig. 5. Examples of windows of a real-time 2D to 3D conversion application using a direct shift method

We proposed simple techniques for creating effective 3D impressions (stereovision illusion of depth) from 2D original images. They are based on the simplest, i.e., binary depth map [9]. The following algorithms were suggested: direct shift, direct shift with interpolation, segment scaling, and segment shifting (Figs. 6 and 7).

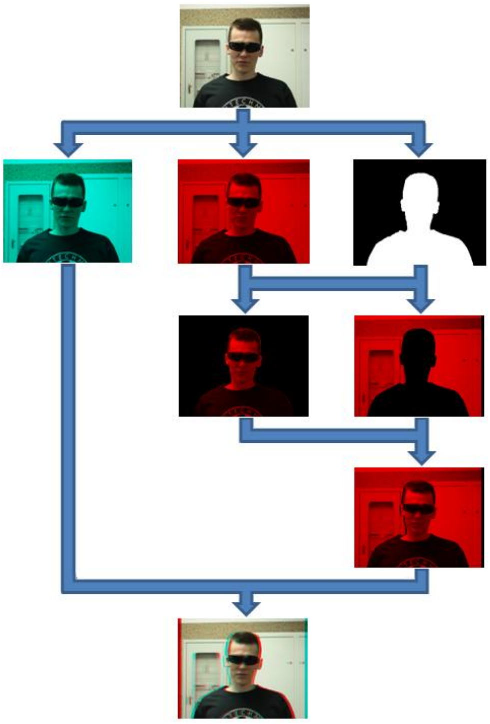


Fig. 6. Anaglyph generation for direct shift method [9]

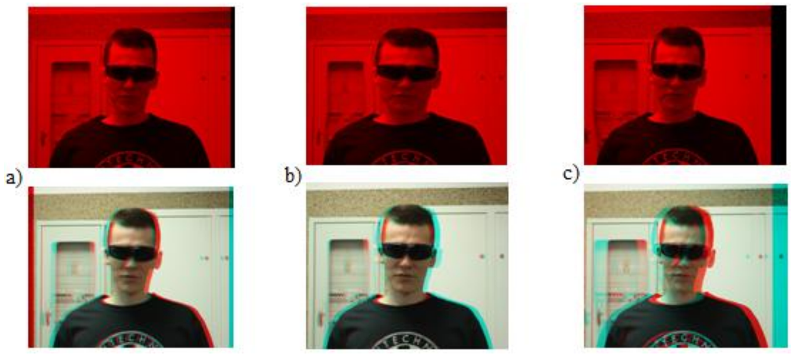


Fig. 7. Red color component and anaglyph generation using: a) direct shift with interpolation, b) segment scaling, c) segment shifting [9]

The anaglyphs in Figs. 6 and 7 are of astonishingly good quality due to strong and tolerant human abilities of perception the 3D effects despite various kinds of distortions. Our two simple methods (segment scaling and segment shifting) based on a binary depth map and specific image deformations allow to obtain quite natural 3D effects with almost imperceptible artifacts [9].

4 Conclusions

The described approach for image clarity improvement and 2D to 3D transformations can be used in CCTV systems for processing both video and for still images.

Current state of experiments with the proposed operator aids proved their effectiveness and encourages for further work on this subject. In future more involved depth maps and more advanced image enhancement methods might be used.

5 References

1. Yibin, Y., Boroczky, L.: A new enhancement method for digital video applications. IEEE Transactions on Consumer Electronics, Volume 48, Issue 3, 435–443 (2002)
2. Cuizhu, S., Keman, Y., Jiang, L.: Shipeng Li Automatic image quality improvement for videoconferencing. Proceedings of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), Volume 3, 701–704 (2004)
3. Zicheng, L., Cha, Z., Zhengyou, Z.: Learning-based perceptual image quality improvement for video conferencing. 2007 IEEE International Conference on Multimedia and Expo, 1035–1038 (2007)
4. Hough, K., Marshall, S.: Soft Morphological Filters Applied to the Removal of Noise From CCTV Footage. The IEE International Symposium on Imaging for Crime Detection and Prevention, ICDP, 61–66 (2005)
5. Onural, L.: An Overview of Research In 3DTV. Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services. 14th International Workshop on 27–30 June 2007, p. 3 (2007)
6. Fehn, C.: Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV. Proc. of the SPIE Stereoscopic Displays and Virtual Reality Systems XI San Jose, CA, USA, 93–104 (2004)
7. Dumbreck, A.A., Smith, C.W.: 3-D TV displays for industrial applications. Stereoscopic Television, IEE Colloquium on, 15 Oct 1992, 7/1–7/4 (1992)
8. Dubois, E.: A projection method to generate anaglyph stereo images. Acoustics, Speech, and Signal Processing, 2001. Proc. (ICASSP ’01) 2001 IEEE International Conference on Volume 3, 7–11 May 2001, 1661–1664 vol. 3 (2001)
9. Balcerek, J., Dąbrowski, A., Konieczka, A.: Simple efficient techniques for creating effective 3D impressions from 2D original images. Proc. of Signal Processing NTAV/SPA’2008, Poland Section, Chapter Circuits and Systems IEEE, 219–224, Poznań, Poland, 25–27th September 2008 (2008)
10. Nelson, C.L.: 3-dimensional video motion detection and assessment. Proc. of IEEE 38th Annual 2004 International Carnahan Conference on Security Technology, 11–14 October 2004 (2004)
11. Bertozzi, M., Broggi, M., Fascioli, A., Nichele, S.: Stereo Vision-based Vehicle Detection. Proc. of the IEEE Intelligent Vehicles Symposium 2000, Dearbon (MI), USA, 3–5 October, 2000 (2000)
12. Nedevschi, S., Danescu, R., Frentiu, D., Marita, T., Oniga, F., Pocol, C., Schmidt, R., Graf, T.: High Accuracy Stereo Vision System for Far Distance Obstacle Detection. 2004 IEEE Intelligent Vehicles Symposium, 292–297, University of Parma, Parma, Italy, 14–17 June, 2004 (2004)

The paper was prepared within the INDECT project.