Color correction method for multi-view video based on background segmentation and dominant color extraction

Feng Shao[1], Gangyi Jiang[1,2], Mei Yu[1,2], and Yo-Sung Ho[3]
[1] Faculty of Information Science and Engineering, Ningbo University, Ningbo, 315211, China
[2] National Laboratory on Machine Perception, Peking University, Beijing, 100871, China
[3] Dept. of Inform. & Comm., Kwangju Institute of Science and Technology, Kwangju, 500-712, Korea

Abstract: Free viewpoint video (FVV) system is one of the most attracting three-dimensional video applications in the future, which can offer depth impression of the observed scenery and allow interactive selection of viewpoint and direction. A color correction method is proposed in this paper to eliminate un-consistent color appearance between viewpoint images. Experimental results show that the proposed method is quite effective.

1. Introduction

It has been recognized that multi-view video processing, including multi-view video color correction, multi-view video coding, virtual view rendering, etc, is the key technology for FVV[1]. Because multi-view video is taken from different viewpoints with multiple cameras, many factors, including light source, shadows, camera aperture, exposure time, non-linearity of transducer, and camera processing (such as auto-gain-control and color balancing) will affect the final image[2]. The un-consistent color between different views has a significant impact on ultimate 3D display and virtual viewpoint rendering.

In this paper, a color correction method for multi-view video based on background segmentation and dominant color extraction is proposed. First, background is segmented from image by edge detection and disparity estimation. Then, with the theory of basic color categories (BCCs)[3], dominant colors[4] from the categories are extracted from background for reference image and input image, and then the corresponding color mapping relationships are built. Finally, by defining multi-view key frames, color correction is achieved for multi-view video.

2. The Proposed Method for Multi-view Video

In multi-view imaging, foreground objects seem to be easier to be influenced by external factors than background. Therefore, background is first segmented from image. First, Susan edge detection is performed to extract object contour. By horizontally scanning edge blocks from left to right, disparity estimation is performed for edge blocks. Then background is extracted based on the disparity information since background usually is far away from cameras compared with foreground. Hadamard transform is used as similarity rule during block matching because it is insensitive to luminance/ chroma change.

2. 1 Dominant Color Extraction

After background segmentation, by using first color labeling and refined color labeling, each pixel (x,y) in background is described by one of the 11 BCCs and the corresponding probabilities ${}_iP_{xy}$ belonging to ith BCC. The mean μ_i and standard deviation σ_i belonging to the ith BCCs can be computed by using the ${}_iP_{xy}$ and pixel intensity $I(x,y)$. The percentage that the ith BCC occupying in the whole 11 BCCs can be described by

$$p_i = \sum_{x,y} {}_iP_{xy} \Big/ \left(\sum_i \sum_{x,y} {}_iP_{xy} \right) \qquad (1)$$

Then $\{p_i\}$ is sorted in descending order. If the accumulative percentage $\left(\sum_{i=1}^{M} p_i \right) < T_1$, regard the colors with highest p_i as dominant colors. M is the dominant color number. Here, T_1 is a threshold controlling dominant color number. The dominant colors is represented as

$$F = \{c_i, p_i, \mu_i, \sigma_i\} \qquad (2)$$

where c_i is the ith dominant color; p_i represents its percentage value; μ_i and σ_i are its color mean and color standard variation.

Finally, for the reference image and input image, match their dominant colors, and perform color correction for the pixel $I^{in}(x, y)$

$$I^{out}(x,y) = \sum_{i=1}^{M} \left(\frac{w_i \sigma_i^{ref}}{w_i \sigma_i^{in}} \left(I^{in}(x,y) - w_i \mu_i^{in} \right) + w_i \mu_i^{ref} \right) \qquad (3)$$

Here w_i is a weighting parameter, and $w_i = p_i \Big/ \left(\sum_{i=1}^{M} p_i \right)$.

2. 3 Extend to Multi-view Video

The *CIEL*a*b** color space was intended to be a perceptually uniform color space, so that equal distance in the color space represent equal perceived differences in appearance. Color difference is defined as the Euclidean distance between two images in this color space

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \qquad (4)$$

To extend to multi-view video, by defining multi-view key frames, the color difference ΔE_k between input image and reference image can be obtained. For other multi-view frames, dominant color information of key frame is used for correction with color difference ΔE_{k+i}. If $|\Delta E_{k+i} - \Delta E_k| > T_2$, update current frame as key frame and implement color correction again.

3. Experimental Results

Experiments are performed on multi-view image sequence 'objects2', 'flamenco1' and 'golf1'[5]. Figs.1-3(a) and (b) show the reference image and input image of 'objects2', 'flamenco1' and 'golf1' in the first and second viewpoints. Figs.1-3(c) show the corrected image with foreground, and Figs.1-3(d) show the corrected image with background. For indoor scene, such as 'objects2' and 'flamenco1', the foreground influence is very clear, while for 'golf1', the influence can almost be neglected. Fig.4 shows color differences between the reference image and the original input image, the input image corrected frame by frame, and the input image corrected with dominant color tracking. From the figure, it is obvious that the proposed method achieves almost smallest color difference, as well as smallest key frames number.

Acknowledgement

This work was supported by the Natural Science Foundation of China (grant 60672073), NSFC/KOSEF Joint Research Project, Natural Science Foundation of Zhejiang Province (grant Y105577), and the Key Project of Chinese Ministry of Education (grant 206059).

References

[1] W. Matusik, H. Pfister, "3DTV: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes," In *Proceeding of SIGGRAPH*, Vol.23, no.3, pp.814-824, 2004.

[2] E. Reinhard, M. Ashikhmin, B. Gooch, et al, "Color transfer between images," *IEEE Computer Graphics and Applications*, Vol.21, no.4, pp.34-41, 2001.

[3] Y. Chang, S. Saito, M. Nakajima, "Example-based color transformation of image and video using basic color categories," *IEEE Trans on IP*, Vol.16, pp.329-336, 2007.

[4] S. Chang, T. Sikora, A. Puri, "Overview of the MPEG-7 standard," *IEEE Transactions on Circuits and System for Video Technology*, Vol.11, no.6, pp.688-695, 2001.

[5] ISO/IEC JTC1/SC29/WG11, KDDI multi-view video sequences for MPEG 2DAV use, M10533, 2004.

(a) (b) (c) (d)

Fig.1 'flamenco1' correction results

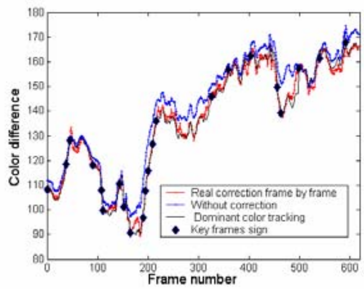
(a) (b) (c) (d)

Fig.2 'objects2' correction results

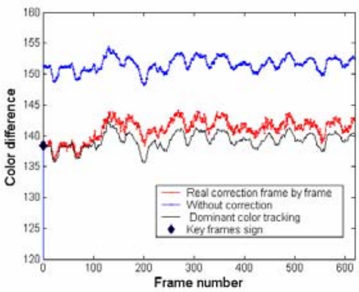
(a) (b) (c) (d)

Fig.3 'golf1' correction results

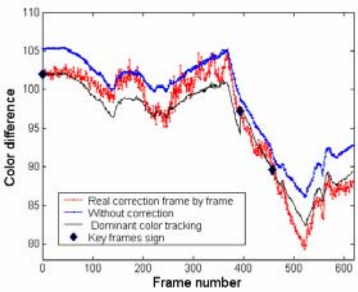
(a): Reference image; (b): Input image; (c) Corrected image with foreground; (d) Corrected image with background



(a) flamenco1



(b) objects2



(c) golf1

Fig.4 Color difference comparison