A CONTENT-ADAPTIVE MULTI-VIEW VIDEO COLOR CORRECTION ALGORITHM

Feng Shao [1,2], *Gangyi Jiang* [1], *Mei Yu*[1,3], *Ken Chen*[1]

[1]Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211
[2]Department of Information and Electronics Engineering, Zhejiang University, Hangzhou 310027
[3]National Laboratory on Machine Perception, Peking University, Beijing 100871

ABSTRACT

A content-adaptive color correction algorithm for multi-view video is proposed due to variation in lighting or camera parameters. We first establish color correction property between the target image and source image. Then color correction matrix can be obtained by global correction or preferred region matching correction. Finally, video tracking technique is used to correct multi-view video sequences. Experimental results show the proposed algorithm has better correction effect for different multi-view video sequences.

Index Terms— Image processing, color, image segmentation, video signal processing

1. INTRODUCTION

With the rapid development and wide application of digital video, requirements for high quality and variety representation of video become higher and higher. Traditional 2D video is no longer entirely satisfying the requirements. Multi-view image system provides the user with realistic impression due to the high interactivity and photorealistic image quality. It allows the users to change his/her viewing point and viewing direction freely and enjoy more photorealistic 3D images. With these functions, multi-view image processing will have good application in three dimensional television (3DTV)[1] or free viewpoint television (FTV)[2], et al.

Un-consistent color between different viewpoints is a serious problem in multi-view image processing. The reasons include scene illumination, camera calibration, CCD noise, jitter of shutter speed or exposal time, et al. Therefore, Color correction or color compensation is often desirable to compensate these differences as a pre-processing or post-processing step[3]. A variety of color correction technologies have been developed and applied to realize color consistency[4,5]. Usually most of the color correction algorithms are performed independent of the image contents and uniform operation is conveniently applied to all image sequences. While the correction properties for different image sequences are not always consistent. So the different correction operation should be applied for higher color efficiency and better visual quality.

In this paper, we develop a content-adaptive color correction algorithm. This algorithm first establishes color correction property between two viewpoints. Color correction matrix can be obtained by global correction or preferred region matching correction. Then we perform correction on multi-view video using the matrix by video tracking technique.

2. IMAGE SEGMENTATION AND K-L TRANSFORM

In our algorithm, the *RGB* color data are first transformed to *CIELAB* color space and data is processed in *CIELAB* color space[6]. Image segmentation is an important step in our algorithm. Color clustering is the most common method in image segmentation. Let a *CIELAB* space color vector $\mathbf{X}=[L,a,b]^T$ and the mean vector $\boldsymbol{\mu}=[\bar{L},\bar{a},\bar{b}]^T$, Euclidian color distance between a color vector of pixel and the mean vector in class k are defined by

$$\begin{aligned} {}_k d(\text{Euclid}) &= \| \mathbf{X} - {}_k\boldsymbol{\mu} \| \\ &= [(L-{}_k\bar{L})^2 + (a-{}_k\bar{a})^2 + (b-{}_k\bar{b})^2]^{1/2} \end{aligned} \tag{1}$$

After segmentation, covariance matrix for class k is given by

$${}_k\mathbf{C}_x = E[({}_k\mathbf{X} - {}_k\boldsymbol{\mu})({}_k\mathbf{X} - {}_k\boldsymbol{\mu})^T] \tag{2}$$

Color vector ${}_k\mathbf{X}$ in class k is transformed in vector ${}_k\mathbf{Y}$ with K-L transform and projected onto the characteristic space by

$${}_k\mathbf{Y} = {}_k\mathbf{A}({}_k\mathbf{X} - {}_k\boldsymbol{\mu}) \tag{3}$$

The matrix ${}_k\mathbf{A}$ is formed by the eigenvectors of covariance matrix ${}_k\mathbf{C}_x$. The covariance matrix ${}_k\mathbf{C}_y$ of $\{{}_k\mathbf{Y}\}$ is diagonalized whose elements are the eigenvalue of ${}_k\mathbf{C}_x$.

$${}_k\mathbf{C}_y = {}_k\mathbf{A}({}_k\mathbf{C}_y)\,{}_k\mathbf{A}^T = \begin{bmatrix} {}_k\lambda_1 & 0 & 0 \\ 0 & {}_k\lambda_2 & 0 \\ 0 & 0 & {}_k\lambda_3 \end{bmatrix} \tag{4}$$

Letting $\{{}_k\mathbf{Y}_{tar}\}$ be the transformed images for the target image $\{{}_k\mathbf{X}_{tar}\}$ and $\{{}_k\mathbf{Y}_{sou}\}$ for the source image $\{{}_k\mathbf{X}_{sou}\}$. The scaling matrix is applied to $\{{}_k\mathbf{X}_{sou}\}$ resulting in the corrected color vector $\{{}_k\mathbf{Y}_{cor}\}$ to estimate the original color vector as follows:

$$ {}_k\mathbf{Y}_{tar} = {}_k\mathbf{A}_{tar}({}_k\mathbf{X}_{tar} - {}_k\boldsymbol{\mu}_{tar}) \quad (5)$$

$$ {}_k\mathbf{Y}_{sou} = {}_k\mathbf{A}_{sou}({}_k\mathbf{X}_{sou} - {}_k\boldsymbol{\mu}_{sou}) \quad (6)$$

$$ {}_k\mathbf{Y}_{cor} = {}_k\mathbf{S}_k\mathbf{Y}_{sou} = {}_k\hat{\mathbf{Y}}_{tar} = {}_k\mathbf{A}_{tar}({}_k\hat{\mathbf{X}}_{tar} - {}_k\hat{\boldsymbol{\mu}}_{tar}) \quad (7)$$

Here

$$ {}_k\mathbf{S} = \begin{bmatrix} \sqrt{{}_k\lambda_{1\_tar}/{}_k\lambda_{1\_sou}} & 0 & 0 \\ 0 & \sqrt{{}_k\lambda_{2\_tar}/{}_k\lambda_{2\_sou}} & 0 \\ 0 & 0 & \sqrt{{}_k\lambda_{3\_tar}/{}_k\lambda_{3\_sou}} \end{bmatrix}$$

After these operations, we get the following relation between the target and source images.

$$ {}_k\hat{\mathbf{X}}_{tar} - {}_k\hat{\boldsymbol{\mu}}_{tar} = {}_k\mathbf{M}({}_k\mathbf{X}_{sou} - {}_k\boldsymbol{\mu}_{sou}) \quad (8)$$

The correction matrix ${}_k\mathbf{M}$ is given by

$$ {}_k\mathbf{M} = ({}_k\mathbf{A}_{tar})^{-1}({}_k\mathbf{S})({}_k\mathbf{A}_{sou}) \quad (9)$$

3. CONTENT-ADAPTIVE COLOR CORRECTION ALGORITHM

3.1. Color Correction Property Selection

Ideally, if no color or content is changed between images, the correction matrix will be identical with identity matrix. Therefore, we define matrix distance between correction matrix $\mathbf{M}=\{m_{ij},0<i,j\le 3\}$ and identity matrix $\mathbf{I}=\{i_{ij},0<i,j\le 3\}$ as

$$ D_{m,i} = \|\mathbf{M}-\mathbf{I}\| = \frac{1}{9}\sum_{i=1}^{3}\sum_{j=1}^{3}\left|m_{i,j} - i_{i,j}\right| \quad (10)$$

If $D_{m,i}\le T_1$, we regard no correction satisfied for images. If $T_1<D_{m,i}<T_2$, global correction is satisfied for images. Then if $D_{m,i}\ge T_2$, preferred region matching correction is satisfied, and then image segmentation and K-L transform are performed for images.

3.2. Preferred Region Matching Correction

If global correction property is satisfied for image sequence, then we correct source image with global correction matrix $\mathbf{M}$.

If global correction property is not satisfied, image segmentation and K-L transform are performed, color mapping relations between regions can be automatically established. Then correct each segmentation region by the mapping relations respectively. However, the classification errors are the problem we cannot avoid. Post-filtering processes is necessary to resolve the error problem which will bring large calculation burden. Therefore, preferred region matching is applied to establish the mapping relation between target and source images. The principle of preferred region selection is the best eigenvalue similarity between target and source images, defined by

$$ \arg\min_{0<k\le K}\left(\sum_{i=1}^{3}(1 - {}_k\lambda_{i\_tar}/{}_k\lambda_{i\_sou})^2\right) \quad (11)$$

3.3. Video Tracking

Because of high temporal correlation existed between two consecutive frames, and imaging condition always unchanged for a fixed viewpoint, video tracking technique is used to correct multi-view video. First we establish color correction matrix between the target viewpoint s_1 and the source viewpoint s_2 in time t_0. For subsequent frames in viewpoint s_2, if the similarity with source image in t_0 is larger than a threshold T_h, correct current frame using the color correction matrix. Otherwise, update the color correction matrix. Generally, hue information is used as similarity comparison criterion. For *CIELAB*, hue angle is $arctan(b^*/a^*)$. The similarity is defined as:

$$ Sim(h_t^h, h_{t+1}^h) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\left|h_t^h[i] - h_{t+1}^h[i]\right|}{\max(h_t^h[i], h_{t+1}^h[i])}\right) \quad (12)$$

Here h_t^h and h_{t+1}^h are hue histograms for different frames.

3.4. The Proposed Algorithm and Performance Evaluation

With the above analyses, the proposed content-adaptive color correction algorithm is briefly summarized as follows:

(1) Get $D_{m,i}$ of the two viewpoint images to establish color correction property. If no correction is satisfied, turn to step (5), if global correction is satisfied, turn to step (3), otherwise, turn to step (2).

(2) Perform image segmentation and K-L transform for target and source images, and select preferred region by formula (11). Then calculate the preferred region correction matrix.

(3) Correct the source image with the correction matrix, then consistent color appearance with target image.

(4) Track temporal frames using the correction matrix. If similarity is larger than a certain threshold, update the correction matrix by back to step (1).

(5) Accomplish all the operation.

In order to objectively evaluate the correction quality, the similarity of global or preferred region between target and source images is calculated. The similarity defined by

$$ Sim(\lambda_{tar}, \lambda_{cor}) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\left|\lambda_{i\_cor} - \lambda_{i\_tar}\right|}{\max(\lambda_{i\_cor}, \lambda_{i\_tar})}\right) \quad (13)$$

Here λ_{tar} and λ_{cor} are global or preferred region eigenvalue for the target and source images, respectively. N is color dimension. The maximum similarity is 1.

4. EXPERIMENTAL RESULTS

In the experiments, multi-view video, 'object3', 'golf1' and 'flamenco1', provided by KDDI Corp[7], are used as test sets, in which the size of images is 320×240. The multi-view images were taken by a horizontal parallel camera configuration with eight viewpoints and 200mm camera interval. The color consistency among these original viewpoints images is poor because of lighting or imaging device. Thus, the color correction is necessary if these multi-view images will be used to render new virtual arbitrary viewpoint image.

Fig.1-2 (a) and (b) show the target and source images in t=0, as 'object3' and 'golf1'. Global correction properties are satisfied. Fig.1-2(c) show the corrected images. Fig.1-2(d) show the video tracking correction results in t=100. The global correction matrix is

$$\begin{bmatrix} 0.979 & -0.015 & -0.049 \\ 0.021 & 0.988 & 0.070 \\ 0.058 & -0.087 & 0.941 \end{bmatrix}, \begin{bmatrix} 1.012 & 0.085 & -0.094 \\ -0.011 & 0.784 & 0.049 \\ 0.074 & 0.068 & 1.010 \end{bmatrix}.$$

For 'flamenco1' sequence, preferred region matching correction is used. Fig.3(a) and (b) show the target and source images in t=530, and Fig.3(c) and (d) show the global correction image and preferred region matching correction image, respectively. Because global correction property is not satisfied for 'flamenco1', better visual quality can be achieved by preferred region matching correction. Fig.3(e) shows the tracking correction result in t=589. The preferred region matching correction matrix is

$$\begin{bmatrix} 0.990 & 0.092 & -0.064 \\ -0.027 & 1.005 & -0.142 \\ 0.018 & 0.193 & 1.044 \end{bmatrix}.$$

In Fig.7 (a)-(c), similarity between the target and source images and similarity between the target and corrected images are compared. The corrected image's similarities are larger than that of the original images obviously.

5. COLCLUSION

The color non-consistency of multi-view images, synchronously captured from a scene, is an important problem to be solved in multi-view image systems, such as free viewpoint television and 3DTV. In this paper, we have proposed a content-adaptive color correction algorithm of multi-view video. Well visual quality is achieved. In the future work, we will go deep into researching how to increase correction accuracy and discussing noise influence for color correction.

ACKNOWLEDGEMENT

This work was supported by the Natural Science Foundation of China (grant 60472100,60672073), Natural Science Foundation of Zhejiang Province (grant RC01057, Y105577), Zhejiang Science and Technology Project of China (grant 2004C31105), and the Key Project of Chinese Ministry of Education (grant 206059).

6. REFERENCES

[1] Matusik W, Pfister H, "3DTV: A Scalable System for Real-time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes", *In Proceeding of SIGGRAPH*, ACM, Los Angeles, 2004, 23(3): 814-824.

[2] Masayuki Tanimoto, "Free viewpoint television – FTV", *Picture Coding Symposium*, San Francisco, USA, 2004,

[3] ISO/IEC JTC1/SC29/WG11, "Luminance and Chrominance Compensation for Multi-view Sequences using Histogram Matching", MPEG2005/M12487, Nice, France, Oct. 2005.

[4] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, et al, "Color transfer between images", *IEEE Computer Graphics and Applications*, IEEE, 2001, 21(4):34-41.

[5] U. Fecker, M. Barkowsky, A. Kaup, "Improving the Prediction Efficiency for Multi-View Video Coding Using Histogram", Matching, *Picture Coding Symposium*, Beijing, China, 2006.

[6] G. Cui, M.R.Luo, B.Rigg, et al, "Uniform Color Spaces Based on the Din99 Colour Difference Formula", *Color Research & Application*, 2002, 27(4): 282-290.

[7] ISO/IEC JTC1/SC29/WG11, "KDDI Multi-View Video Sequences for MPEG 3DAV Use", MPEG2004/M10533, Munich, Mar. 2004.

(a) Target image (b) Source image (d) Corrected image (d) Tracking image

Fig.1 Correction result for 'object3'

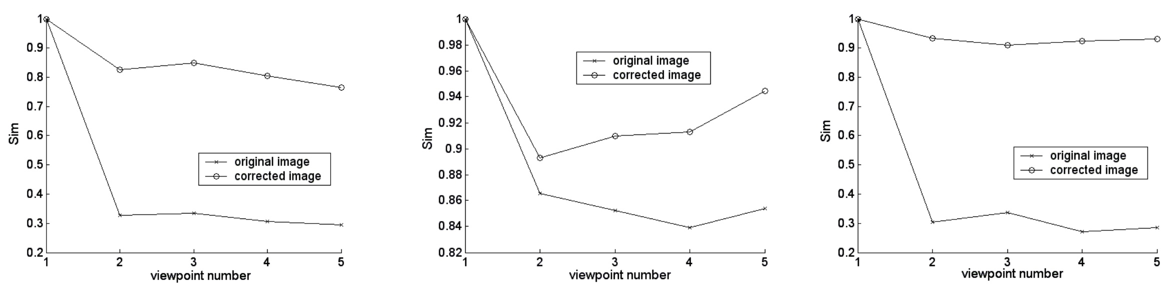
(a) Target image (b) Source image (d) Corrected image (d) Tracking image

Fig.2 Correction result for ‘golf1’

(a) Target image, (b) Source image, (c) Global correction image, (d) Preferred region matching correction

(f) Tracking image

Fig.3 Correction result for ‘flamenco1’



(a) object3 (b) golf1 (c) flamenco1

Fig.4 Similarity comparison