

Full-Quaternion Color Correction in Images for Person Re-identification

Reynolds León Guerra^(✉), Edel B. García Reyes, and Francisco J. Silva Mata

Advanced Technologies Application Center (CENATAV), Havana, Cuba
rleon@cenatav.co.cu

Abstract. Nowadays, video surveillance systems are very used to safeguard airport areas, train stations, public places, among others. Using these systems, the person re-identification is an automated task. However, there are many problems that affect the good performance of person re-identification algorithms. For example, illumination changes in the scenes is one of the essential problems. It increases the false colors on the person appearance. Moreover, in the extraction of low level features (color) there is a need to obtain reliable colors of the image or person appearance. To this end, we propose a new algorithm using full-quaternions for color image representation. The quaternionic trigonometry and the Quaternion Fast Fourier Transform are used to improve person image in the frequency domain. Finally, to see the transformed image, an adaptive gamma function is developed. Experiment results on two datasets (VIPeR and GRID) show consistent improvements over the state-of-the-art approaches.

Keywords: Full-quaternion · Person image · Color
Fourier transforms

1 Introduction

Automation of tasks linked with video surveillance have received a lot of attention recently by computer vision researchers [1]. This is because cameras located in avenue or public places allow person tracking, anomaly detection or person re-identification (Re-Id). The aim of person re-identification is identifying the same person in different cameras with non-overlapping camera views [2]. The task of Re-Id is a challenging problem caused by illumination changes, occlusion, pedestrian poses, background clutters, viewpoints and so on [3].

All illumination changes generate variance in the person appearance [4], for instance, false colors in person's clothes. It is clear because it is necessary to get true colors to decrease the error using descriptors as color histograms. Moreover, a stage in Re-Id is the feature extraction [5] that has a lot of influence on the image comparison (in a second stage). As explained above, color is an important feature for person re-identification.

The problem of looking for robust colors to illumination change is called color constancy (also color invariant, photometric normalization or color correction).

To address this problem, several alternative have been proposed for Re-Id algorithms. Ma et al. [6] propose using the Gray World algorithm to improve the parts of the person body. Kviatkovsky et al. [7] make groups of colors in a log-chromaticity space of the upper/lower parts of human body assuming that color clouds are invariant under some conditions. Jung et al. [8] use Shades of Gray algorithm to improve the person images. Eisenbach et al. [9] improve the image, taking as reference the person's clothes and the variations of illumination when the person passes through the scene in a learning process. Allouch [10] applies Retinex algorithm to improve images. Monari [11] developed an algorithm using shadow information of the person to estimate local illumination. Liao et al. [12] developed a version the multiscale Retinex algorithm to get vivid colors in person images. However, previous works for color correction in Re-Id do not take into consideration the connections among each color channel in the person image.

Person images have edges and isolated colors that arise from the self shadow or some anomaly (these isolated colors and edge are considered as false colors, because they have no relation to the pixel colors nearby), and are ignored by the algorithms, for instance, Gray World, which was designed for generic images (as landscapes). We propose a new algorithm for color correction in person images using full-quaternion (to preserve the relationships among colors) and the Quaternion Fast Fourier Transforms (QFFT) for the removal of the false colors previously mentioned.

Contributions of this paper are summarized as follows: First, a full-quaternion is designed using local and neighborhood information of each pixel. Second, in the frequency domain the person image is modified through its spectral module. Third, the image is displayed with an adaptive gamma function.

This paper is organized as follows. In Sect. 2, the algebra of the quaternion is explained to understand the following sections. The proposed approach, the full-quaternion, the QFFT and gamma function are explained in Sect. 3. Experimental results are presented and analyzed in Sect. 4. Finally, the conclusions are set out in Sect. 5.

2 Algebra of the Quaternions

In 1843, Hamilton presents the quaternions [13], denoted with letter \mathcal{H} . If $q \in \mathcal{H}$, then it is represented as follows:

$$\{q = t + xi + yj + zk | (t, x, y, z) \in \mathcal{R}\} \quad (1)$$

Where the complex operators $\mathbf{i}, \mathbf{j}, \mathbf{k}$ have the next rules $\{i^2 = j^2 = k^2 = ijk = -1, ij = k = -ji, ki = j = -ik, jk = i = -kj\}$. This shows that the multiplication of quaternions is not commutative. Other form to represent the quaternion is:

$$q = Sq + Vq \quad (2)$$

Where $(Sq = t)$ denotes the real or scalar part and Vq denotes the vector or imaginary part.

$$Vq = xi + yj + zk \in \mathcal{H}(Vq) \quad (3)$$

If $Sq = 0$, the quaternion is a pure quaternion and $q = Vq$, if $Sq \neq 0$ the quaternion is a full-quaternion. Other properties and operations are:

$$|q| = (t^2 + x^2 + y^2 + z^2)^{1/2} \tag{4}$$

$$q = Sq - Vq \tag{5}$$

Where the expression (4) and (5) are the module and conjugate respectively. If $|q| = 1$ and $q \in \mathcal{H}(Vq)$, then it is called unit pure quaternion.

Its representation in polar form is:

$$q = |q|e^{\tau\phi} = |q|(\cos\phi + \tau\sin\phi) \tag{6}$$

Where:

$$\phi = \tan^{-1}\left(\frac{|Vq|}{Sq}\right) \tag{7}$$

$$\tau = \frac{Vq}{|Vq|} \tag{8}$$

are respectively eigenangle and eigenaxis. The sine, hyperbolic cosine and inverse hyperbolic cosine [14] are represented as follows:

$$\sin(q) = ((\sin(Sq) \cosh(|Vq|)) + (\cos(Sq) \sinh(|Vq|)\tau)) \tag{9}$$

$$\cosh(q) = (\exp(q) + 1/\exp(q))/2 \tag{10}$$

$$\operatorname{acosh}(q) = \log(q + \sqrt{q^2 - 1}) \tag{11}$$

If $\hat{q} = a + bi + cj + dk$ and $\tilde{q} = w + xi + yj + zk$ and $(\hat{q}, \tilde{q}) \in \mathcal{H}$, then add, subtraction and multiplication are as follows:

$$\hat{q} \pm \tilde{q} = (a \pm w) + (b \pm x)i + (c \pm y)j + (d \pm z)k \tag{12}$$

$$\hat{q}\tilde{q} = (S\hat{q}S\tilde{q} - V\hat{q} \cdot V\tilde{q}) + (S\hat{q}V\tilde{q} + S\tilde{q}V\hat{q} + V\hat{q} \times V\tilde{q}) \tag{13}$$

Where, $\langle \cdot \rangle$ is the scalar product and \times is the vector cross product.

3 The Proposed Approach

Person image is represented to the domain $q \in \mathcal{H}$ (full-quaternion) where the real part is a value that represents the clear-dark contrast effect in the image. After, the image is transformed using hyperbolic cosine and with the Quaternion Fast Fourier Transform are modified the coefficients (module). Finally the person image is retrieved through the inverse QIFFT, the inverse hyperbolic cosine and the gamma function.

3.1 Full-Quaternion

An image has three color channels (Red, Green, Blue) and this can be represented by $f_{(x,y)}$. Nevertheless, the information among color channels is not associated and this is a constraint to the image analysis (the algorithms work only with the intensities of each channel). Also, all the colors depend of their neighbor colors to be analyzed. For example, in a patch with size 3×3 the neighbor pixels are gray and the central pixel is black, then the black color (noise) is generally omitted and the patch is classified as a gray homogeneous region. To solve this difference among colors, we start from the expression (9), as $f_{(x,y)} \in \mathcal{H}(Vq)$ then $Sq = 0$ implying that vector part can only be used in the expression (9). The expression (8) is to normalize a quaternion, getting a unit pure quaternion. It is applied to all pixels, where take the value $|q| = 1$, and each pixel is located in the surface of a unitary sphere. It is deductible that $\sinh(|Vq|)$ acts as a scale factor on τ increasing or decreasing the size of the quaternion. We called clear-dark contrast effect (CDCE) to the increment or decrement of τ , because the dark colors are close to zero and light colors are close to higher values (taking the RGB cube as reference), for example the pink or white.

The transformation $f_{(x,y)} \longrightarrow f_{(x,y)}^H \in \mathcal{H}$, is performed as follows. A sliding window is built and the value of the function ($\sinh(|Vq|)$) of the pixels neighboring the central pixel is obtained in each position. The value of the real part of the full-quaternion is the average value of the neighborhood, as explained above.

$$\xi = \left\{ \frac{1}{\eta} \sum_{\nu=1}^{\eta} \log(\sinh(|Vq|) + 1)_{\nu} \right\} + ri + gj + bk \tag{14}$$

Where, $\xi \in f_{(x,y)}^H$, $\eta = 8$, $\{r, g, b\}$ are color channels.

3.2 Hyperbolic Cosine and Fourier Transforms

Differently from the other works, we consider that a modification of the coefficients of the spectral module in the person image allows obtaining more reliable colors (considering the false colors as noise in the signal). The quaternion hyperbolic cosine (see expression 10) is used to integrate the value of the real part with the vector part of the quaternion where a new full quaternion is obtained before applying the QFFT [15]. The QFFT allows transforming the frequency coefficients using filters in the spectral module where is possible to eliminate noises present in the spatial domain.

$$F(p, s) = S \sum_{m=0}^{m-1} \sum_{n=0}^{n-1} e^{-\mu 2\pi \left(\frac{pm}{M} \right) + \left(\frac{sn}{N} \right)} f(m, n) \tag{15}$$

$$f(m, n) = S \sum_{p=0}^{p-1} \sum_{s=0}^{s-1} e^{\mu 2\pi \left(\frac{pm}{M} \right) + \left(\frac{sn}{N} \right)} F(p, s) \tag{16}$$

Where, $S = \sqrt{\frac{1}{MN}}$, the expression (15) and (16) are the direct and inverse Quaternion Fourier Transforms respectively, μ is unit pure quaternion, p and s are frequency coefficients, m and n are spatial coordinates of the image. The filter used is a Gaussian filter applying a convolution on \mathcal{T} (Module):

$$\zeta = \mathcal{T} * \Gamma \quad (17)$$

Where, $\{*\}$ is convolution operator and Γ is the Gaussian filter. Returning the image to the spatial domain is done as follows:

$$f(m, n) = exp(S \sum_{p=0}^{p-1} \sum_{s=0}^{s-1} e^{\mu 2\pi(\frac{pm}{M}) + (\frac{sn}{N})} ((\zeta + F(p, s))\beta)) \quad (18)$$

Where, β is a scale factor. After the image obtained in the spatial domain $f(m, n)$ is transformed by the expression (11).

3.3 Visualization

A problem with QIFFT is that $f(m, n) \in \mathcal{H}$, to see the image in $\mathcal{H}(Vq)$ is developed a gamma function as follows:

$$f(x, y) = |C_i|_{(x, y)}^{1/|t|} \quad (19)$$

Where, C is color channel, $i \in \{r, g, b\}$, (x, y) are spatial coordinates of the image. The gamma function is adaptive because each channel is affected by the real part t of full-quaternion.

4 Experimental Results

Our experiments are to evaluate the best performance of the color correction algorithms in person images used for person re-identification. However, to the best of our knowledge there are no datasets with ground truth for true colors in person image (for person re-identification). Therefore, to evaluate the algorithm, we use as metric the Area Under Curve (AUC) for Cumulated Matching Characteristics (CMC) curve [16].

The datasets used are VIPeR and GRID. The VIPeR [17] has the image of the same person when passing in front of two different cameras; it contains 1264 images of people divided into 632 images for camera A and camera B. It has strong variations in views and changes in lighting. The characteristics in GRID [18, 19] are: it contains 250 probe images (camera A) captured in one view and 250 images of the same person (in camera B), also there are 775 additional images that are not in the camera A, that is why we reject them and only work with 250 images. The images have strong variations in color and pose.

The parameters used in the experiment for our algorithm called **FqCC** (Full-quaternion Color Correction) are the following: $\mu = ((i + j + k)/\sqrt{3})$, $\beta = 1.5$.

The window and standard deviation of the Gaussian filter are respectively 3 and 0.8. The person image is segmented with a parsing person algorithm [20] to get the trunk and legs (see Fig. 1(i) and (j)). The feature vector is composed of the histograms of the color channels -red, green and blue- of the trunk and legs (because there are pixels that belong to clothing), also the hue and contrast histograms of the whole image using the $\sinh(|Vq|)$ for contrast and the hue channel of the HSV color space. The Bathacharyya distance [21] is used three times for the person marching for 16, 32 and 64 bins in the feature vector histograms (see Table 1). The algorithms used in the comparison are applied in the pre-processing step for person re-identification such as: multiscale retinex (MSR) adapted in the work of Liao [12], Shades of Grey (SofG) and Grey-World (GW) the framework [22]. Others algorithms as maxRGB (Mrgb), Grey-Edge (GE), Weighted Grey-Edge (WGE) [23] are used and original images (OI) without transformations, too.

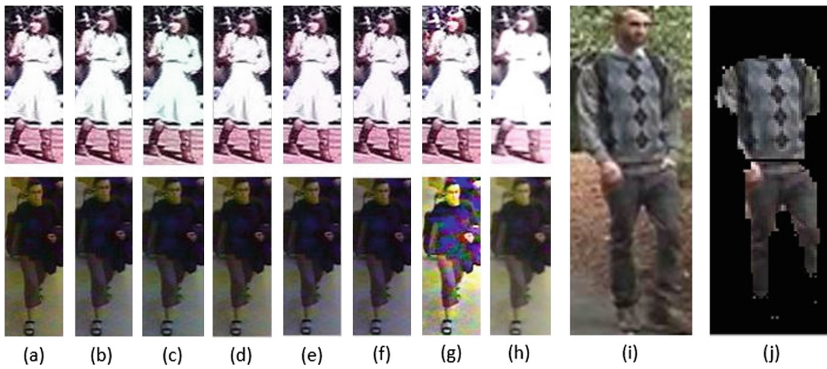


Fig. 1. Shows different results of the algorithms used for color correction (row above are VIPeR images and row below are GRID images), (a) OI, (b) SofG, (c) GW, (d) Mrgb, (e) GE, (f) WGE (g) MSR, (h) **FqCC**. The image (j) is the trunk and legs of the image (i). (Color figure online)

In Fig. 1 (h - above) it can be seen that our algorithm removes the noise (black small regions in the dress) by white color in correspondence with true color of the dress and the arm edges are smoothed unlike the other algorithms. Also in Fig. 1 (h - below) image contrast is improved and yellow color decreases (considered as noise). The results obtained are shown in Table 1 where our algorithm gets the best results for a value of 0.7308 (to 16 bins) in the VIPeR dataset and 0.5876 (to 32 bins) in the GRID dataset. The values obtained in the GRID dataset are lower than those of VIPeR dataset due to the high noise level of the person images in the GRID dataset. The results show that not always the algorithms used to improve the colors in the images achieve their goals, as it is the case of SofG, GW, GE and WGE (with VIPeR) that obtained values smaller than the experiment with the image without transforming it and MSR, SofG, GW,

Table 1. Experiments with Bathacharyya distance.

Methods	VIPeR - 632			GRID - 250		
	16 bins	32 bins	64 bins	16 bins	32 bins	64 bins
FqCC	0.7308	0.7281	0.7251	0.5812	0.5876	0.5775
Original image	0.7076	0.7035	0.7021	0.5674	0.5726	0.5779
Multiscale retinex	0.7153	0.7116	0.7107	0.5526	0.5400	0.5339
Shades of grey	0.6834	0.6787	0.6766	0.5435	0.5404	0.5468
Grey-world	0.6759	0.6725	0.6713	0.5578	0.5642	0.5708
MaxRGB	0.7140	0.7091	0.7065	0.5607	0.5668	0.5686
Grey-edge	0.6906	0.6854	0.6828	0.5731	0.5785	0.5859
Weighted grey-edge	0.6931	0.6882	0.6857	0.5665	0.5731	0.5806

Mrgb and WGE in the GRID dataset. However, the algorithm herein proposed obtains the best results by contributing to the elimination of false colors (edges and isolated colors). This results are also, by the transformations of the image in the frequency domain and the incorporation of (CDCE) as an element of the full-quaternion.

5 Conclusions

Obtaining a value (real part), based on the clear-dark contrast effect, allows the construction of a full-quaternion and applying transformations in the frequency domain make it possible to correct the color and to improve the image. It was possible to establish an adaptive gamma function to display the image without losing information of the real part. For future works, we are going to design a local approach to color correction and design a metric of image quality (without ground truth).

References

1. Wu, Z.: Human Re-Identification. Springer, Heidelberg (2016). <https://doi.org/10.1007/978-3-319-40991-7>
2. Deng, X., Ma, B., Chang, H., Shan, S., Chen, X.: Deep second-order Siamese network for pedestrian re-identification. In: Lai, S.-H., Lepetit, V., Nishino, K., Sato, Y. (eds.) ACCV 2016. LNCS, vol. 10112, pp. 321–337. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-54184-6_20
3. Wang, H., Gong, S., Xiang, T.: Highly efficient regression for scalable person re-identification. arXiv preprint [arXiv:1612.01341](https://arxiv.org/abs/1612.01341) (2016)
4. Schumann, A., Gong, S., Schuchert, T.: Deep learning prototype domains for person re-identification. arXiv preprint [arXiv:1610.05047](https://arxiv.org/abs/1610.05047) (2016)
5. de Carvalho Prates, R.F., Schwartz, W.R.: Appearance-based person re-identification by intra-camera discriminative models and rank aggregation. In: 2015 International Conference on Biometrics (ICB), pp. 65–72. IEEE (2015)

6. Ma, B., Li, Q., Chang, H.: Gaussian descriptor based on local features for person re-identification. In: Jawahar, C.V., Shan, S. (eds.) ACCV 2014. LNCS, vol. 9010, pp. 505–518. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16634-6_37
7. Kviatkovsky, I., Adam, A., Rivlin, E.: Color invariants for person reidentification. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(7), 1622–1634 (2013)
8. Jung, J., et al.: Normalized metadata generation for human retrieval using multiple video surveillance cameras. *Sensors* **16**(7), 963 (2016)
9. Eisenbach, M., et al.: Learning illumination maps for color constancy in person reidentification (2013)
10. Allouch, A.: People re-identification across a camera network. Master of Science Thesis, KTH Royal Institute of Technology, Stockholm, Sweden (2010)
11. Monari, E.: Color constancy using shadow-based illumination maps for appearance-based person re-identification. In: 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance (AVSS), pp. 197–202. IEEE (2012)
12. Liao, S., et al.: Person re-identification by local maximal occurrence representation and metric learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2197–2206 (2015)
13. Morais, J.P., Georgiev, S., Sprig, W.: *Real Quaternionic Calculus Handbook*. Springer, Basel (2014). <https://doi.org/10.1007/978-3-0348-0622-0>
14. Turner, J.D.: Quaternion analysis tools for engineering and scientific applications. *J. Guidance Control Dyn.* **32**(2), 686–693 (2009)
15. Ell, T.A., Sangwine, S.J.: Hypercomplex fourier transforms of color images. *IEEE Trans. Image Process.* **16**(1), 22–35 (2007)
16. Bolle, R.M., et al.: The relation between the ROC curve and the CMC. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies, vol. 2005, pp. 15–20. IEEE (2005)
17. Gray, D., Brennan, S., Tao, H.: Evaluating appearance models for recognition, reacquisition, and tracking. In: Proceedings of the IEEE International Workshop on Performance Evaluation for Tracking and Surveillance (PETS) (2007)
18. Liu, C., Gong, S., Loy, C.C., Lin, X.: Person re-identification: what features are important? In: Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012. LNCS, vol. 7583, pp. 391–401. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33863-2_39
19. Loy, C.C., Xiang, T., Gong, S.: Time-delayed correlation analysis for multi-camera activity understanding. *Int. J. Comput. Vis.* **90**(1), 106–129 (2010)
20. Luo, P., Wang, X., Tang, X.: Pedestrian parsing via deep decompositional network. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2648–2655 (2013)
21. Naik, N., Patil, S., Joshi, M.: A scale adaptive tracker using hybrid color histogram matching scheme. In: 2009 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET), pp. 279–284. IEEE (2009)
22. Van De Weijer, J., Gevers, T., Gijzenij, A.: Edge-based color constancy. *IEEE Trans. Image Process.* **16**(9), 2207–2214 (2007)
23. Gijzenij, A., Gevers, T., Van De Weijer, J.: Improving color constancy by photometric edge weighting. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(5), 918–929 (2012)