International Journal of Control and Automation
Vol. 9, No. 9 (2016), pp.343-352
http//dx.doi.org/10.14257/ijca.2016.9.9.33

Flame detection via Quaternion Discrete Cosine Transform based Spectral Saliency

Haiyan Zhang[1], Shangbing Gao[2], Jianxun Xue[1] and HaoZheng[3]

[1]Faculty of computer and software, Jiangsu Provincial Key Laboratory for advanced manufacturing technology, Huaiyin Institute of Technology, Huai'an, China
[2]The Key Laboratory for traffic and transportation security of Jiangsu Province, Huai'an, China
[3]Key Laboratory of Trusted Cloud Computing and Big Data Analysis, Nanjing Xiaozhuang University, Nanjing, China
zhylhy2008@ 126.com

Abstract

To address the problem of detecting the specific object of flame, this paper proposes quaternion discrete cosine transform based video flame detection algorithm in conjunction with the bottom-up visual attention mechanisms and the flame features. According to the characteristics of the flame to be detected, the brightness, color, texture and motion information of the image are respectively defined and extracted to form a quaternion image, and then the flame saliency map is produced by Quaternion discrete cosine Transform with this quaternion image. Extensive experimental results verify the accuracy and robustness of the proposed method on flame detection.

Keywords*: flame detection, saliency detection, quaternion*

1. Introduction

Video processing techniques for automatic flame have become a hot topic in computer vision during the last decade [1]. The different vision-based fire detectors, as proposed in the literature, led to a large amount of techniques for detecting the presence of fire at an early stage. Current researches, such as the work of [2]-[5], show that these video-based fire detectors promise fast detection and can be a viable alternative or complement to the more traditional techniques. However, due to changing environmental conditions and the variability of shape, motion, transparency and colors, existing video fire detection (VFD) approaches are still vulnerable to missed detections and false alarms.

The key of VFD is to have real-time detection on target region of fire danger in complicated scene, while there is a huge difference between the flame region and non-flame region. The flame regions can be regarded as the visual salient region, when the non-flame region is treated as the background region. However, the VFD methods seldom adopted visual selective attention method. In [6], Stadler et. al investigate five different pixel intensity flickering features to propose a novel flame detection algorithm. Yang and Wang [7] have introduced a kind of new video monitoring model based on hierarchy attention and salient integration framework of multi-source perceptive information, and tried to improve the efficiency and initiative of fire monitoring system with salient feature description and low redundant calculation. However, it has only explained the structure of hierarchy attention without study on the realization technology.

In recent years, starting with the spectrum residual (SR) in 2007 [8], spectral saliency models [9-12] attracted a lot of interest. These approaches manipulate the image's frequency spectrum to highlight sparse salient regions and provide state-of-the-art

ISSN: 2005-4297 IJCA

performance on salient region detection (see [8-12]). Following SR, the Phase spectrum of the Fourier Transform (PFT) [9] was introduced, which achieved nearly the same performance as the SR. Based on PFT, Phase spectrum of Quaternion Fourier Transform (PQFT) [10] was also proposed by combining more features and using the quaternion Fourier Transform. Schauerte et. al [11] put forward a novel face detection algorithm based on the quaternion discrete cosine transform (QDCT). Subsequently, Schauerte et. al [12] also proposed EigenPQFT by using the concepts of eigenaxes and eigenangles.

An interesting development in this area is the use of quaternion as a holistic representation to process color images as a whole [10-12]. The quaternion algebra makes it possible to process color images as a whole without the need to process the image channels separately and, in consequence, tear apart the color information. Therefore, inspired by [8-12], we first consider flame detection as a spectral saliency detection problem based on quaternion. We found that, it can achieve the best flame detection performance using QDCT, while compared to PQFT, EigenPQFT. Thus, a new flame detection algorithm based on spectral saliency is proposed. In this model, intensity, color, texture and motion features are comprised into a quaternion image as individual channel for taking phase spectrum. The main steps of the proposed algorithm are shown in Figure 1.

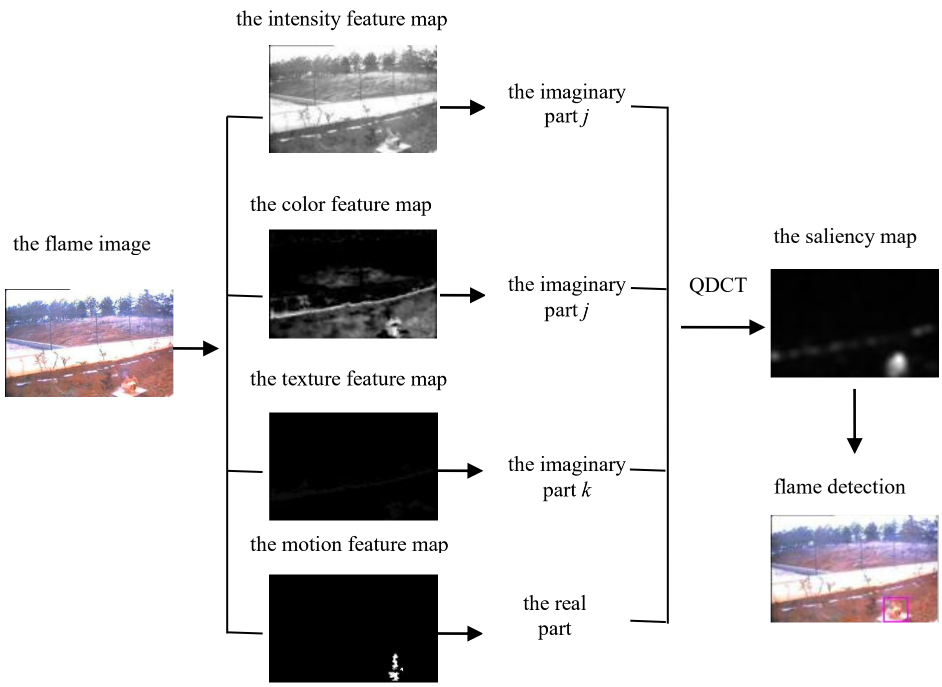

Figure 1. Main Steps of the Proposed Flame Detection Algorithm

2. The QDCT Model

The QDCT model can be divided into two stages: a quaternion representation of an image; secondly, discrete cosine transform of quaternion is used to deal with this quaternion image, and the spatiotemporal saliency map is obtained.

2.1. Quaternion Representation of an Image

A quaternion has four components (one real part and three imaginary parts) and can be represented in a hypercomplex form as [13]：

$$q = x_1 + x_2 \cdot i + x_3 \cdot j + x_4 \cdot k \tag{1}$$

where $x_1, x_2, x_3, x_4 \in R$, and i, j, k obey the following multiplication rules: $i^2 = j^2 = k^2 = -1$, $i \cdot j = -j \cdot i = k$, $j \cdot k = -k \cdot j = i$, $k \cdot i = -i \cdot k = j$ .This demonstrates that the multiplication of quaternions is not commutative.

The conjugate and modulus of a quaternion q follow the definitions for complex numbers.

2.2. The Quaternion Discrete Cosine Transform (QDCT)

According to the definition of quaternion, the multiplication of quaternion is not commutative. Thus, the form of Forward Quaternion Discrete Cosine Transform (FQDCT) and Inverse Quaternion Discrete Cosine Transform (IQDCT) have two categories, left-handed form and right-handed form [13]. In this paper, we will use the left-handed FQDCT ^{L} and IQDCT ^{L} as default QDCT and IQDCT, respectively to demonstrate the process of saliency detection.

$$FQDCT^L(p,q) = \alpha_p^M \alpha_s^N \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u_q I_q(m,n) N(p,s,m,n) \tag{2}$$

where $I_q(m,n)$ is a two-dimensional M×N quaternion matrix. u_q is a pure quaternion named a quaternionzation factor which meets the constraint that $u_q{}^2 = -1$. Similar to the traditional DCT, the values of α_p^M ,α_s^N and $N(p,q,m,n)$ are taken as

$$\alpha_p^M = \begin{cases} \sqrt{\dfrac{2}{M}} & for \quad p \neq 0 \\ \sqrt{\dfrac{1}{M}} & for \quad p = 0 \end{cases}, \quad \alpha_s^N = \begin{cases} \sqrt{\dfrac{2}{N}} & for \quad s \neq 0 \\ \sqrt{\dfrac{1}{N}} & for \quad s = 0 \end{cases} \tag{3}$$

$$N(p,s,m,n) = \cos\left[\frac{\pi(2m+1)p}{2M}\right]\cos\left[\frac{\pi(2n+1)s}{2N}\right] \tag{4}$$

Accordingly, the IQDCTL is defined as follows:

$$IQDCT^L(m,n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p^M \alpha_s^N u_q C_q(p,s) N(p,s,m,n) \tag{5}$$

where $C_q(p,s)$ is a two-dimensional M × N quaternion matrix. Besides u_q, α_p^M, α_s^N and $N(p,s,m,n)$ are the same as those used in the FQDCT.

3. The Fire Detection Method Based on QDCT Model

To detect the specific object of fire, the flame characteristics should be taken into account to represent the quaternion image. The intensity, color, texture, motion are the main characteristics of the flame, thus we consider the information of these four channels.

3.1. The Information of the Four Channels

(1)The information of intensity and color channel

In the RGB color space of fire regions, the value of red channel $r(t)$ is generally the maximum, while the value of blue channel $b(t)$ is minimum. Thus, the red channel compete with the green channel. The information of color channel can be represented as the difference between the pixels value of the red and green channel in the current frame:

$$C(t) = r(t) - b(t) \tag{6}$$

The feature maps of intensity and color are shown as Figure 1.

(2)The information of texture channel

Firstly, the flame regions are extracted by the following two rules [4]:

$$R_{i,j} > R_t \tag{7}$$

$$S_{i,j} > (S_t(255 - R_{i,j}))/R_t \qquad i \in \{1,2,...,n\}, j \in \{1,2,...,m\} \tag{8}$$

where S_{ij} and R_{ij} represents the saturation and the R component of the ijth pixel respectively. R_t denotes the threshold of the R component, while S_t is the threshold of the saturation, $R_t \in [115,135]$, $S_t \in [55,65]$, the size of the image is $n \times m$.

(a) the key frame of flame video


(b) the candidate flame regions

Figure 2. The Extracted Candidate Flame Regions

Then, we compute the texture feature by using the Local Ternary Patterns (LTP) [14], a generalization of the Local Binary Pattern (LBP) [15] local texture descriptor that is more discriminant and less sensitive to noise in uniform regions. In LTP, graylevels in a zone of width $\pm t$ around i_c are quantized to zero, ones above this are quantized to +1 and ones below it to -1, *i.e*. the indicator $s(u)$ is replaced by a 3-valued function:

$$s'(u, i_c, t) = \begin{cases} 1, & u \geq g_c + t \\ 0, & |u - g_c| < t \\ -1, & u \leq g_c - t \end{cases} \tag{9}$$

where t is a user-specified threshold. The computational formula of LTP encoding is equation (10).

$$LTP_{P.R}(x_c, y_c) = \sum_{i=0}^{P-1} 3^i S(u) \tag{10}$$

where P denotes the total number of neighbor pixels, R represents the radius of the neighborhood.

Due to the high dimension of the flame samples, we reduce dimension using PCA in order to improve the efficiency. To this end, the 328 flame regions of 243 images are trained, the LTP characteristic matrix A with the size of 105×328 is obtained, then A is reduced dimension by PCA to acquire the characteristic matrix B with the size of $N \times 328$ (N =4), the average eigenvector U is obtained by taking the mean value by row on B. The principal elements can be computed through the contribution rate of the eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n, (n=105)$ of the matrix A. The contribution rate of each eigenvalue is defined as:

$$C_i = \lambda_i / \sum_{j=1}^{n} \lambda_j, (i=1,2,...,n) \tag{11}$$

The sum of the contribution rates of the first $m(m \leq n)$ eigenvalues is:

$$SC_m = \sum_{i=1}^{m} \lambda_i / \sum_{j=1}^{n} \lambda_j \tag{12}$$

SC_m denotes the ratio of the information of compression image and the original image. In the image compression and restoration process, the original image can be generally greater than 80% used to restore, which automatically determine the number of the principal elements.

Then the distance d between the LTP eigenvectors of all candidate flame regions V_1, V_2, …, V_n and the average eigenvectors V is computed, and finally the reciprocal of d is taken to obtain the texture feature map of all the candidate flame regions. Figure 3 shows the corresponding texture feature maps.

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

(a) the key frame of flame video; (b)the corresponding flame texture feature map; (c) the key frame of flame video; (d)the corresponding flame texture feature map.

Figure 3. The Flame Texture Feature Map

(3)The information of motion channel

The motion saliency of flame is indicated by the accumulated difference. In order to obtain the superior flame motion region detection result, the appropriate cumulative number of frames should be selected. Considering the flames flicker frequency, frame rate and the running time, we choose adjacent 10 frames of the video to carry out the accumulated difference operation. Suppose the tth frame of the flame video f_t be the current frame, f_{t-1}, f_{t-2}, …, f_{t-9} is the former 9 frames, respectively. The steps are as follows:

Step1: Extracting the candidate flame regions $f_{ca,t}$, …, $f_{ca,t-9}$ in f_{t-1}, f_{t-2}, …, f_{t-9} according to the equations (7), (8);

Step2: Calculating the absolute value of the difference between two adjacent frames frame $f_{di,t-1}$, …, $f_{di,t-9}$, where $f_{di,t-1} = \left| f_{ca,t} - f_{ca,t-1} \right|$;

Step3: Obtaining all the candidate flame regions $f_{ac,t}$ of f_t after calculating the cumulative difference, $f_{ac,t} = \sum_{j=t-9}^{t-1} f_{di,j}$;

Step4: Extracting all the candidate flame regions $c_{a,1}$, …, $c_{a,n}$ in $f_{ac,t}$;

Step5: Computing the mean values of the n flame candidate regions respectively, denoted by a_1, …, a_n, and computing the total mean value $a = \left(\sum_{i=1}^{n} a_i \right) / n$;

Step6: Selecting the candidate flame regions which satisfy a condition of $a_i > a, (i \in \{1,2,...,n\})$ to eliminate the background area with sudden luminance change or the color of the fires. These regions are denoted by $a_{i,j}, (i \in \{1,2,...,n\}, j = 1,2,...,m)$, respectively.

Step7: To avoid the influence of the sudden luminance change, supposing the number of pixels in the jth candidate flame region to be N. The number of times which the luminance change of these N pixels between the two adjacent frames is greater than twice of the mean value a' are represented by $p_1, p_2,..., p_N$, respectively. Then, selecting the pixels which satisfy the condition of $p_k > 1, (k \in \{1,2,..., N\})$, the number of these pixels is supposed to be M. Thus, these pixels are represented by $p_{k,l}, (k \in \{1,2,..., N\}, l = 1,2,..., M)$, then Computing the average $T_h = \left(\sum_{l=1}^{M} p_{k,l} \right) / M$, $(k \in \{1,2,..., N\})$.

As the dynamic characteristic value of the connected domain, T_h is assigned to the value of all these pixels in the same connected domain. In this way, the final motion feature map is obtained, as shown in Figure 1.

3.2. The Quaternion Flame Image

The information of motion channel and texture channel is more important for the fire detection compared to the other channels, thus they should be given greater weight. It has been found through experiment that the optimum weights of $M(t)$、$C(t)$、$I(t)$、$T(t)$ are chosen $\frac{1}{3}$、$\frac{1}{6}$、$\frac{1}{6}$、$\frac{1}{3}$. As thus, the quaternion images is defined as:

$$q(t) = \frac{1}{3}M(t) + \frac{1}{6}C(t)\mu_1 + \frac{1}{6}I(t)\mu_2 + \frac{1}{3}T(t)\mu_3 \tag{13}$$

4. The Experimental Results and Analysis

In our work, computer simulations are conducted in Matlab on an Intel Core i5 processor running at 2.30 GHz with 8 GB real memory. The proposed fire detection algorithm has been tests on different fire detection bench mark videos (http://signal.ee. bilkent.edu.tr/ VisiFire/index.html) and non-flame video test sets (from the internet) with the resolution of 320×240. The following tests 6 video clips, among them there are 4 fire scenarios and 2 complex interference scenarios. The 6 video clips were numbered from a to f, as listed in Table 1.

Table 1. The Details of Testing Video Clips

Video No.	Scene descriptions
a	Fire burns fiercely; cars with strong lights on the far motor road; the grass is close to the flame in color; the wall is in white, which is close to the color of flame core; people add fuels several times.
b	The floor and wall are close to the flame in color; people strolled in the monitor scene; sun light is strong.
c	Fire burns fiercely in the pan against a wall; the wall color is close to that of flame; flame sways obviously under windy weather; Sun light is weak.
d	Forest fire
e	Road scene in the night; passing cars and street lamps with weak light reflection.
f	Road scene in the night; passing cars street lamps with strong light reflection.

The fire detection algorithm was compared with the two algorithms developed by Celik [3] , Chen [4] and Rong [5] using on the 6 video clips. Figure 4 presents the detection results and the corresponding saliency maps of 6 different fire scenes. The fire regions have been successfully marked by red rectangles although against the background of various interference sources, such as red lawn, walking man, moving car lights and the ground reflection. Table 2 compares the fire detection results using the proposed algorithm and the three algorithms by Celik [3], Chen [4] and Rong [5] for the 6 testing videos. From Table 2, it can be seen that the fire detection rate of the proposed algorithm is higher than the other three algorithms.

Table 2. Fire Detection Rates for Different Algorithms in 6 Scenes

Video No.	a	b	c	d	e	f
Video frames	400	707	438	244	43	35
Fire frames	400	707	438	244	0	0
Celik detection frames	233	579	329	228	24	19
Chen detection frames	362	661	429	234	33	29
Rong detection frames	366	558	400	241	13	11
Our detection frames	386	671	423	244	5	3
Celik detection rate(%)	58.3%	81.9%	75.1%	93.4%	44.2%	45.7%
Chen detection rate(%)	90.5%	93.5%	97.9%	95.9%	23.3%	17.1%
Rong detection rate(%)	91.7%	78.9%	91.3%	98.8%	68.2%	68.6%
Our detection rate(%)	96.5%	94.9%	96.6%	100%	88.4%	91.4%

Figure 4. The Test Images In 6 Scenes and Their Corresponding Saliency Maps

5. Conclusions

In this study, a new fire detection algorithm via the spectral saliency in terms of color, intensity, texture and motion characteristics for fire video clips has been developed and successfully validated in different fire scenes. In the future, more studies are needed to eliminate the non-fire objects which resemble the color, texture, flicker of the fires.

Acknowledgments

This work is sponsored by the National Natural Science Foundation of China (NSFC) #61402192, JiangSu Qing Lan Project, Jiangsu 333 Project, six talent peaks project in Jiangsu Province, the open fund for the Key Laboratory for traffic and transportation security of Jiangsu Province (TTS2015-05), the open fund of Jiangsu Provincial Key Laboratory for advanced manufacturing technology (HGAMTL-1401).

References

[1] J. Chen, Y.P. He and J.Wang, "Multi-feature fusion based fast video flame detection", Building and Environment, vol. 45, no. 5, **(2010)**, pp. 1113—1122.
[2] D.C. Wang, X .Cui, E. Park, C.L. Jin and H. Kim, "Adaptive flame detection using randomness testing and robust features", Fire Safety Journal, vol. 55, **(2013)**, pp. 116—125.
[3] T. Celik and H. Demirel, "Fire detection in video sequences using a generic color model", Fire Safety Journal, vol. 44, no. 2, **(2009)**, pp. 147—158.
[4] T.H. Chen, P.H. Wu and Y.C. Chiou, "An early fire-detection method based on image processing", Proceedings of IEEE International Conference on Image Processing, vol. 3, **(2004)**, pp. 1707—1710.
[5] Y. Li, D. Wei, Z. Mu, Z. Xiong, Y. Wang and W. Yin, "Study of the Time-collocation of Signal Lamp at Intersection", vol. 2, no. 1, **(2015)**, pp. 13-16.
[6] Stadler, A. Windisch, Tjark and K. Diepold, "Comparison of intensity flickering features for video based flame detection algorithms", Fire Safety Journal, vol. 66, **(2014)**, pp. 1—7.
[7] J. Yang and R.Wang, "Video flame detection technology based on computer vision", China Journal of image and graphics, vol. 13, pp. 1221—1234, **(2008)**, (in Chinese).
[8] X. Houand, L. Zhang, "Saliency detection: A spectral residual approach", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **(2007)**, pp. 1—8.
[9] M. Zhao, D. Zhao, Z. Jiang, D. Cui, J. Li and X. Shi, "The gray prediction GM (1, 1) model in traffic forecast applicast application", Mathematical Modeling of Engineering Problems, vol. 2, no. 1, **(2015)**, pp. 17-22.
[10] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression", IEEE Trans. Image Processing, vol. 19, no. 1, **(2010)**, pp. 185—198.
[11] B. Schauerte and R .Stiefelhagen, "Predicting human gaze using quaternion DCT image signature saliency and face detection", Proceedings of IEEE International Conference on Applications of Computer Vision, **(2012)**, pp. 137—144.
[12] W. Zhu, C. Li, X. Xu and X. Liang, "Analysis on the Structure Transformation of Landing Craet", Mathematical Modeling of Engineering Problems, vol. 2, no. 1, **(2016)**, pp. 7-10.
[13]W. Feng and B. Hu, "Quaternion discrete cosine transform and its application in color template matching", In: Congress on Image and Signal Processing (CISP'08), vol. 2, **(2008)**, pp 252–256.
[14] J. Xia, L. Xiao and L. Wan, "Application of Random-fuzzy Probability Statistics Method", Mathematical Modeling of Engineering Problems, vol. 3, no. 1, **(2016)**, pp. 19-24.
[15] T. Ojala, M. Pietikainen andT. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, **(2002)**, pp. 971-987.

Authors

Haiyan Zhang, she was born in 1980, received the BS degree in Computer science and technology from the Changchun University in 2003. She received the MS degree in computer application from Nanjing University of Science and Technology in 2010. Her current research interests include pattern recognition and computer vision.

Shangbing Gao, he was born in 1981, received the BS degree in mathematics from the Northwestern Polytechnical University in 2003. He received the MS degree in applied mathematics from the Nanjing University of Information and Science and Technology in 2006. He received the Ph.D. degree with School of Computer Science and Technology, Nanjing University of Science and Technology (NUST). Since 2014 he has been an associate professor at Huaiyin Institute of Technology, China. His current research interests include pattern recognition and computer vision.

Jianxun Xue, he was born in 1978, received his M.S degree in aerospace manufacturing from Nanjing University of Aeronautics and Astronautics, Nanjing, China in 2003.He is a lecturer at Huaiyin institute of technology, Huai'an, China. His current research interests include image processing and software. terests include software and network.

Hao Zheng, he was born in 1978, received the his BS degree from South East University in 1998, the MS degree from Nanjing University Posts and Telecommunications in 2005, and the PhD degree in pattern recognition and intelligence system from Nanjing University of Science and Technology in 2013. He is currently a postdoctor of South East University. His research interests include pattern recognition, image processing, face recognition, computer vision.