International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 2, No. 1, 2015, pp.45-50.
ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Comparison of Two Different Distance Functions of Image Retrieval for Detecting Species of Microscopic Fungi in Medical Mycology Laboratory

Zahede Mousavi Kakavand[a] , Abdolah Chalechale[b]

[a]*Department of Computer Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.*

[a]*Department of Computer Engineering, Kermanshah Science and Research Branch, Islamic Azad University, Kermanshah, Iran*.

[b]*Department of Computer Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.*

[b]*Department of Computer Engineering, Kermanshah Science and Research Branch, Islamic Azad University, Kermanshah, Iran*.

Abstract

In medicine, fungi are called as human pathogens. Physiological and immunological characteristics of fungi are not of great importance for detecting fungi species. Rather, the use of structural specifications of fungi in laboratory lam and fungal culture is the best way to detect them. In the method used in this paper, color histogram and horizontal, vertical and diagonal gradients of image edges are used to calculate attribute vector for images of different species of fungi. Then, we estimate the species of unknown fungus through calculating the distance of average attribute vector for different species and attribute vector of the query image and finding the shortest distance. Finaly , with comparison of the obtained results from two distance functions, Canberra & Bray Curtise, it was found that the Canberra dissimilarity is more efficient.

Keywords: image processing, image retrieval, microscopic images, medical mycology, distance function

Introduction:

The fungi were introduced as pathogens for the first time in 1800 and since then numerous fungal diseases and fungal pathogens have been reported (Shadzi, 2008). Between 200,000 fungal species, only about 100 species are pathogenic fungi (Emami, 1997). Although in the past medical mycology was confined to some limited skin infections, today many opportunistic fungal infections have been known with high ability of compatibility with the environment. S0 , the number and dispersion of pathogenic fungi in nature is much less than the opportunistic fungi (Shadzi, 2008).

Fungi can be classified to pathogenic and opportunistic groups. The latter is capable of infecting individuals and opportunistic group is capable of infecting individuals who are susceptible to infection. Despite of advances in medical science in the treatment and prevention of infections caused by pathogenic fungi, the number and variety of opportunistic infections have been intensified; this is due to this fact that the therapeutic advances in medical science in the context of malignant and autoimmune diseases and transplant have increased the lifespan of patients who do not have a healthy immune system, and as a result, opportunistic infectious agents cause dangerous fungal infections in individuals. Fungies that were previously safe and non-infectious for humans today are accounted as one of the major causes of mortality in these patients because these fungi have high compatibility with environmental conditions, and turning into dangerous opportunistic pathogens in the individuals with failure in immune system. For this reason, timely recognition and treatment of fungal infections for rescuing life of these patients, required for providing comprehensive information in this context (Imam, 1999). Cox et al. (1998) in an article entitled "Image analysis of the morphology of filamentous micro-organisms" mentioned image analysis as a strong tool to detect internal morphology of filamentous micro- organisms (Cox et al, 1998).

Blackledge and Barry (2011) in an article entitled "Morphological Analysis from Images of Hyphal Growth using a Fractional Dynamic Model" examined optimization of industrial fermentation processes using filamentous bacteria (Blackledge and Barry, 2011). Papagianni (2014) in an article entitled "Fungal morphology detection using digital image analysis techniques" examined this issue in detection of fungal filaments (Papagianni, 2014).

Since the digital image analysis techniques were introduced, the studies on morphology of filamentous fungi have been wider and image analysis potential has been documented properly in these studies. Medical science is closely related to health of the people and experts in this area require greater utilization of various information sources to increase their academic ability. Medical images assist for compression of large volumes of information and governance of large amounts of data. Medical science largely depends on the images and health experts use medical images for clinical, educational and research purposes (Kayani and Mansourian, 2013) . An important point lies on timely access to proper and relevant information, because most of the questions of experts at this area rise during work or when they face patients; retrieval of information associated to the user's considered image has transformed to a vital issue especially at medical area. Image has a key role as one of the non-textual information sources especially at the area of medicine; since health area has a direct relationship with the quality of life, the use of image is effective in improving performance of medical specialists. On the other hand, image processing has been regarded as one of the most modern emerging technologies, transformed to one of the most used science at all the areas. Robotics, defense systems, medical engineering, industrial control and computer graphics can be mentioned as the major usages of image processing (Aghaei Ata, 2010). The present research intends to compare accuracy of two different distance functions in retrieving similar images to estimate fungus species. In this method, firstly an average attribute vector is estimated for each fungus species based on Morphological features of laboratory lam images and then the distance of the attribute vector of query image with the average vector of each fungus species is acquired through two different distance functions and finaly the species with less distance between its average vector and

attribute vector of the query image is selected as the suitable species, whereby accuracy of two distance functions is compared and the results are represented in tables and diagrams.

The research method

In this study, two different distance functions in image retrieval for detecting type of Microscopic fungi in Laboratory of Medical Mycology compared. In this method, firstly an average attribute vector based on morphological features of laboratory lam images is calculated, and then the distance of the attribute vector of the query image with the vector of average images of fungal species has been obtained through two different distance functions and the species with less distance are selected as the suitable species, whereby accuracy of two distance functions is compared and the results are represented in tables and diagrams. In the method of the present research, regarding importance of image color feature, to fetch the required characteristics of image, color histogram of image is acquired, firstly with transforming RGB image to HSV and then by quantizing histogram of H, S and V components to 24 levels. At the next stage, firstly the image noise is removed via Gaussian elimination and then Laplacian sharpening filter is applied on the image. Then, horizontal, vertical and diagonal gradients of the image edges are calculated via Sobel and Roberts operators. Now, the corresponding vectors are acquired by transforming these values to degree and quantizing the frequency diagram of these values to 24 levels. Considering the vectors of different images among the existing images, we calculate the average vector for different species of fungus in the dataset. To find the fungus species in the query image, the corresponding vectors of this image have been provided and then species of unknown fungus is estimated through calculating distance of these vectors with the average vectors of the images of different species of fungi in the dataset and considering the least distance. Distance function is a positive real value function, so that it gives a non-negative value for both sample vectors when two samples differ from each other and/or it gives zero value when two samples are same. This function is symmetrical, that is, the distance of the first sample from the second sample is equal to the distance of the second sample from the first sample. The less value of distance function, will be the less distance of two vectors and the more similarity of samples. To evaluate the proposed method, firstly the average attribute vectors for each group of images and the attribute vectors of the query image must be prepared (Tahmasebi & Amouzegar, 2011). In this article, accuracy in two distance function of Canberra and Bray Curtis is compared.

Canberra Dissimilarity

Canberra dissimilarity is a numerical criterion for the distance between paired values in the vector space introduced by G. N. Lance and W. T. Wiliams during 1966 and 11967, found as a weight version of distance L_1(Manhattan). This distance is used as a metric for the arranged lists for the purpose of discovering the penetration in computer security. Canberra distance between vectors p and q in a real vector space is defined in formula 1(Tahmasebi & Amouzegar, 2011). In formula 1, pi represents i^{th} component of vector p, qi represents i^{th} component of vector q and d represents Canberra distance.

$$d(p,q) = \sum_{i=1}^{n} \frac{|p_i - q_i|}{|p_i + q_i|}$$ Formula 1

Bray Curtise dissimilarity

This distance function as been defined based on the distance between two attribute vectors. This distance function has been detected as the best and most optimum classification method for k-Nearest Neighbors algorithm in medical data mining systems (Tahmasebi & Amouzegar, 2011). This distance function is calculated by adding and subtracting absolute value of elements in two attribute vectors that calculated and then sum of all the adding absolute value of elements and sum of all the subtracting absolute value of elements are calculated and finaly sum of the subtractions are divided into the sum of additions.

$$d(x,y) = \frac{\sum_{i=1}^{n} |x_i - y_i|}{\sum_{i=1}^{n} (x_i + y_i)}$$ Formula 2

Findings

In this section, the results form distance between attribute vectors of the query images and the average attribute vectors of different species of fungus existing in database are mentioned via Canberra and Bray Curtis distance functions to estimate species of unknown fungus. In tables below, the correct results and their percent for each of vectors acquired through the images have been proposed for each of the mentioned methods. In these tables, the mentioned variables are equivalent to the vectors below:

C: this represents the correct results from comparison of the vector from quantizing histogram of the color of the query image with the average vector for the images of any fungus species for the existing images in the dataset.

SS: this represents the correct results from comparison of the vector from quantizing the values obtained from calculation of the Sobel gradient in the query image with the average vector for the images of any fungus species for the existing images in the dataset.

RS: this represents the correct results from comparison of the vector from quantizing the values obtained from calculation of the Roberts gradient in the query image with the average vector for the images of any fungus species for the existing images in the dataset.

Main: this represents the correct results for each repetition on image testing and comparison of the attribute vectors of query image with the average vectors of the images of different fungus species in the database.

Laplacian/Canberra	C	SS	RS	Main
Right number	48	45	48	62
Right type in 5 first order	29	23	24	23
percent	55.17%	51.72%	55.17%	71.26%

Table 1-6- the correct results from calculation of Canberra distance between attribute vectors of images after applying Laplacian filter

Laplacian/Courtise	C	SS	RS	Main
Right number	33	16	17	38
Right type in 5 first order	36	37	40	37
percent	%37.93	% 18.39	% 19.54	% 43.67

Table 7-6- the correct results from calculation of Bray Curtise distance between attribute vectors of images after applying Laplacian filter

As seen in Figure 1 representing number of correct image from comparison of average attribute vectors from the images in the database with the attribute vectors of query image, it can be implied that Canberra distance function has had more correct result than Bray Curtis distance function.

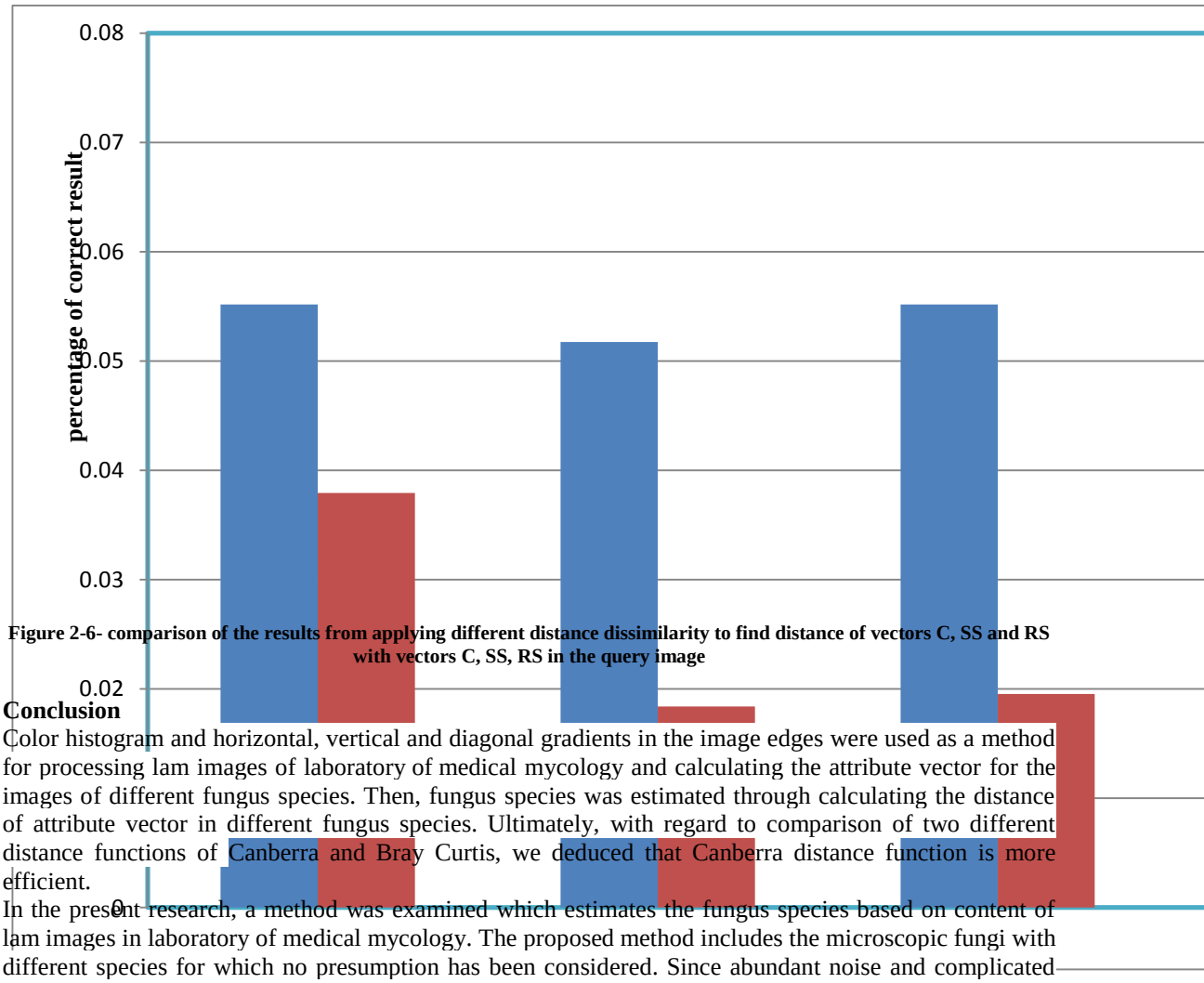


Figure 2-6- comparison of the results from applying different distance dissimilarity to find distance of vectors C, SS and RS with vectors C, SS, RS in the query image

Conclusion

Color histogram and horizontal, vertical and diagonal gradients in the image edges were used as a method for processing lam images of laboratory of medical mycology and calculating the attribute vector for the images of different fungus species. Then, fungus species was estimated through calculating the distance of attribute vector in different fungus species. Ultimately, with regard to comparison of two different distance functions of Canberra and Bray Curtis, we deduced that Canberra distance function is more efficient.

In the present research, a method was examined which estimates the fungus species based on content of lam images in laboratory of medical mycology. The proposed method includes the microscopic fungi with different species for which no presumption has been considered. Since abundant noise and complicated morphology are the attributes of microscopic images, common and effective methods for recovering images on microscopic images do not have efficiency. Therefore, such attributes must be considered in designing image retrieval system for these images.

We must consider more suitable criteria and more efficient distance functions in future research so as to increase yield of image retrieval method. Since the fungi have passed different stages during their lifetime with apparent changes in them, learning methods such as support vector machine seem more suitable at this area.

References:

Aghaee Ata, Hussein. “Specifying the coordinates of the center of pieces with automatic image processing”. Construction and production.. 33.

Emami, M. Zeini, Faride. “Comprehensive Medical Mycology”. Tehran University Press.

J. M. Blackledge and D. J. Barry: Morphological Analysis from Images of Hyphal Growth using a Fractional Dynamic Model, EG UK Theory and Practice of Computer Graphics (2011), pp. 1–8

Kayani, M. Karami, Yazdan. “Image Information retrieval in health on the Web from the perspective of medical science professionals: a qualitative study”. Health Management. 53.

Maria Papagianni : Characterization of Fungal Morphology using Digital Image Analysis Techniques, Papagianni, J Microb Biochem Technol 2014, 6:4

P. W. Cox, G. C. Paul and C. R. Thomas:Image analysis of the morphology of filamentous micro-organisms, Microbiology (1998), 144,817-827

Shadzi, Shahla. “Medical Mycology, Laboratory Diagnosis”. SID.

Tahmasebi, Hamid Reza. Amoozegar, Malihe. “Introducing measures that are appropriate for classification k closer neighbors in medical data mining”. National Conference on Computer and Information Technology.