

Research on Medical image retrieval based on HSI color model

Wang Jinguo^{1, a}, Wang Na^{2, b*} (corresponding author), Wang rui^{3, c}

¹Department of Urology, the First Hospital of Jilin University, China

²Department of Anesthesiology, the First Hospital of Jilin University, China

³Department of Anesthesiology, the First Hospital of Jilin University, China

^awangjinguolily@163.com, ^blilyly12345@163.com, ^cXiaoben6666@126.com

Keywords: Medical image. Retrieval algorithm. HSI color model.

Abstract. The ideal image retrieval should be both automatic extraction of keywords and the bottom of the content-based retrieval. For medical image retrieval, not only to take into account the content of the image itself, but also related to the image of other medical information, such as the text of the diagnostic report of consciousness. With the development of image processing technology and database technology, medical information retrieval technology and advanced semantic retrieval technology must be in the content based medical image retrieval plays an important role, which is the inevitable direction of development in the future, but also the only way for medical database application.

Introduction

With the progress of digital medical imaging equipment, after CT, pet, SPECT, DSA, MRI, digital ultrasonic, digital gastrointestinal radiography and digital mammography, the advent of new technology of computer X-ray photography and digital X-ray photography, medical imaging comprehensive digital new era has arrived. In the past ten years, China has introduced the advanced medical image equipment, which has been imported from China, and has played a positive role in improving the diagnostic level and strengthening the management of the hospital level. Because of the limitation of the financial distress and the design level of instrument, most medical image equipments do not consider the image acquisition, processing, storage and transmission function, at best, a printer or X optical film for image recording is configured. The diagnosis of the doctor's diagnosis is to observe the instrument's screen image or film, with the experience of the individual, the diagnosis is more subjective. With the continuous improvement of the degree of modernization of our country, in the hospital to solve the above problems have been put on the agenda[1].

However, in the implementation of the system, how to realize automatic, intelligent retrieval, query and management of massive medical image data can be easily and quickly and accurately find the required information, so that the management can be relieved from a large number of monotonous and ineffective manual management. And content-based image retrieval technology, especially in the field of medical image aided diagnosis, more and more attention. Because of the current diagnostic techniques, especially those who can not immediately come to the diagnosis of the case, the use of similar cases to assist in the diagnosis, this time using content-based image retrieval technology can easily get the case with the current case images, thereby saving the manual search and the need for a large number of human and time, to facilitate and improve the accurate diagnosis of patients[2].

But medical image because of the data volume of large and on the accuracy of the high demand, the building has a greater complexity, and the speed of retrieval, query accuracy high, small feature

amount of storage space. So I introduced the content based image retrieval method to medical image database to launch a large number of research work[3].

Local accumulation histogram retrieval based on HSI color model

Choice of color model. Color is a very important visual feature of the image, relative to other features (such as texture, shape, etc.), color is very stable, with the characteristics of the rotation, translation is not sensitive to change, showing strong robustness. It is an important method to simplify the extraction and classification of targets, so it is important to choose the right expression way. The expression of the color feature depends on the selected color model, and not all of the color models are consistent with human feelings. From the point of view of physics, the perception of color originates from the spectrum of electromagnetic radiation of the retina. These energy distributions can be recorded as $E(I)$. It distributions are visible at the wavelength of 350~780nm.

$E(I)$ is closely related to the distribution of the $S(I)$ and the reflection properties of the illuminated object illuminated by the light source. This relationship can be expressed as:

$$E(I) = S(I)R(I) \quad (1)$$

But the feeling of human beings is related to many physical phenomena, psychological effects and physiological actions, and the relationship is complex. At present, the proposed color model can be divided into four types according to the different basis:

Colorimetric color model: Physical measurement based on spectral reflectance.

Psychophysical color model: Perception of color based on human.

Physiological color model: There are 3 basic color perception cone cells in the human retina.

Opponent color model: Perception test.

Considering the use of color model, color model is divided into two categories, namely, the color model for hardware device and the color model for visual perception. Through the introduction and research of different color models, we use the HSI color model, RGB is the most commonly used and the most basic color space, the HSI color model can be obtained by the linear or nonlinear transformation of the basic RGB color model. So the RGB color model is introduced before the introduction of the HSI color model.

HSI color model. HSI is the most commonly used color model in color processing, in which H is expressed in hue, S indicates saturation, and I is expressed in density (corresponding to luminance or gray). People distinguish between 3 basic characteristics of color: brightness, hue and saturation. The luminance is proportional to the reflectivity of the object, and if no color is only the change of the intensity of a dimension. For color, the more white in the color, the more bright, the smaller the number of black and white. The tone is the main light wave in the mixed spectrum. Saturation is related to a certain degree of color purity, pure spectral color is a complete package and the degree of saturation of the white light is gradually reduced. Hue and saturation are called color. So the color can be expressed in both luminance and color.

HSI model consists of two important features. Firstly, the luminance component and the chroma component are separated, and the I component is independent of the color information of the image. The second is that the H and S components are closely connected with the human feeling color. These features are the HSI model which is very suitable for image processing analysis based on human visual system for color perception characteristics.

Color feature expression

There are many kinds of color features, such as histogram, histogram, local accumulation histogram, color layout, central moment method, and so on. Because of the large amount of color image data, so a variety of methods common to use more effective and compact method to express color information.

Statistical histogram. In order to describe the image features, the image can be characterized by statistical histogram. The statistical histogram of the image features is actually a 1-D of the discrete function:

$$H(k) = \frac{n_k}{N} \quad k = 0,1,2,\dots,L-1 \quad (2)$$

On the K representation of the image features, L is the number of features available, n_k is the number of pixels in the image with a feature value of N, the total number of K type image pixels.

When characters cannot image in all of the value, the statistical histogram appears to be zero. The emergence of these zero to compute histogram intersection bring great influence, so that the calculated matching values and can not correctly reflect the color difference between the two graphs. To some extent, this problem can be overcome by increasing the number of image features, which is related to the visual characteristics of human eyes. However, there is a problem in this simple quantization. In the case of hue, the color of the two is close to the same distance. This problem occurs in the vicinity of the quantization interval, which is caused by the quantization itself.

In addition, the color feature similarity distribution is not uniform, such as blue color of the distribution range is wider than the yellow, so the simple and so on the same interval can not accurately reflect the color difference between the two. In order to solve the problem of zero, the cumulative histogram. Cumulative histogram can greatly reduce the number zero in the original histogram, the two colors on the axis of the distance will maintain the similar degree between them and is proportional.

Cumulative histogram

The histogram of image feature statistics is also a discrete function of 1-D:

$$I(k) = \sum_{i=0}^k \frac{n_i}{N} \quad k = 0,1,2,\dots,L-1 \quad (3)$$

The meaning of each parameter in the formula is the same as in the previous one, and the color image can be accumulated by 3 components.

Local accumulation histogram

In the Euclidean distance as the similarity measure, the cumulative histogram can reflect the correlation of the signal in the distribution of the sampling points, that is, the distribution of the relative distance of a sample point, in a sense, the more similar to the sample point. But cumulative histogram can reflect the premise of this nature is: the signal itself to have such characteristics, that is, the distance between the two points on the axis of the small two points more similar to the distance, it is worth noting that this condition has not been attached importance to. The following example of a color signal is discussed:

In the HSI space, $H=0$ is corresponding red. $H=60^\circ$ is corresponding yellow. $H=120^\circ$ is

corresponding green.

So the distance between the yellow and red in the H is not equal to the distance between the red and the green, the yellow is more close to the red. But from human visual perception, yellow and red, green and red between the two more similar. In other words, although the H shaft yellow chroma value than green chroma values close to the value of the red color, but in the visual effect of yellow are not necessarily better than green and red are similar. This shows that the distribution of the color signal from the visual sense, does not meet the premise of cumulative histogram. Natural color image of the color is more complex, the general accumulation of the histogram will be mixed with different color signal.

References

- [1] CHEN T W,,CHEN Y L,CHIEN S Y.Fast image segmentation based on K-means clustering with histograms in HSI colorspace. IEEE 10th Workshop on Multimedia Signal Processing(MMSP08) . 2008.
- [2] David G Goodenough,Andrew D,et al.Processing Hyperion andALI for Forest Classification. IEEE Transaction on Geoscienceand Remote Sensing . 2003/
- [3] Boardman J W,Kruse F A.Automated Spectral Analysis:AGeolog-ical Example Using AVIRIS Data. Proceedings of 10th The-matic Conference on Geologic Remote Sensing . 1994.