Image registration is a technique that compares two images for purposes of alignment and calibration. A well known problem underlining medical imaging is that the alignment of different exams for the same patient and same anomalies may not match. So we can not make reliable conclusions based on direct comparison of both exams. Generally, we can accomplish alignment of geometric features as well as luminance profile. Besides, there are a vast literature demonstrating their importance and challenges. The luminance alignment is close related to the amount of information between two images. Until the end of 90’s the most known strategy to measure transfer information between two probability distributions of physical systems was through the classical Shannon entropy. A further improvement of this kind of formalism is now the so called Tsallis entropy, which has been proved to be a generalization of Shannon theory. The recent related literature has been shown that Tsallis entropy is suitable for medical imaging with many works proving its performance in several applications. This paper presents an analysis of the probability luminance distribution between images acquired from Cone-Beam and Multi-Slice techniques based on Kullback-Leibler divergence extendable for Tsallis statistics.

1. Introduction

Currently, it is vital to use images for evaluation and monitoring patients in the area of medicine [10], whose methods of acquisition and their technology have been so much improved. A common example is the Computed Tomography (CT), one of the most used and important test today. Different acquisition techniques are developed to improve the quality of the exam with small damage as possible to the patient. Despite their advantages, methods such as Multi-Slice Computerized Tomography (MSCT) have a high dose of radiation. In contrast, studies has shown that the use of techniques such as Cone-Beam Computerized Tomography (CBCT) can be less harmful to the patient [8] [9]. Its importance comes from its frequent use in orthodontics. Then there is a need of evaluating differences between both techniques, taking into consideration the underlining mistakes or difficulties in diagnosis.

In order to minimize the differences between these two types of images, we first perform the registration between them, which consists of luminance and spatial alignment [12]. On the other hand, the spatial alignment can be determined under techniques of geometric transformations, such as in [16], [12],[13]. Once computed these transformation, the differences between the images can be computed by matching their luminance distribution. The most common techniques to accomplish this task is through matching their luminance histograms.

It is well known from Information Theory Area that the probability distribution of an image luminance can carry information related to its semantic content. Other features such as color, texture and spatial relationship between dominant regions also can carry similar information. However, due to the high degree of correlation between pixel features it can be difficult measuring such properties.

The traditional way to measure the amount of information between two images is by computing their relative entropies or Kullback-Leibler divergence. Recently, works of [1],[4], [13], [12] and [16] provided evidences that medical images can be better explained if their physical systems are considered as having non-extensive behavior, which means that they have long-range spatial and temporal interactions. This issue has been little explored in literature for image registration yet.

This paper presents a study of techniques based on non-extensive entropy for comparison of Cone-Beam (CB) and Multi-Slice (MS) CT images used in orthodontics. We carried out our studies under the Kullback-Leibler divergence
extendible for Tsallis entropy. The results show the power of this recent methodology for measuring relation between two probability distributions, and suggest that the pixels interactions of Cone-Beam and Multi-Slice may have sub-extensive behavior.

2. Non-Extensive Entropy

The term entropy appeared first in the field of thermodynamics, where it was used to demonstrate microscopic behaviors under macroscopic physical processes. Initially, it was considered merely a physical property that was thought to be applicable only in the context of thermodynamics. Later, E. Boltzmann and W. Gibbs showed entropy as a statistical measure for various other applications, resulting in the celebrated formula of Boltzmann-Gibbs, where the entropy (S) is the product of the Boltzmann constant (k) by the logarithm of a state W [18][4], given by the following equation:

\[ S = k \log W \] (1)

Later, C. Shannon gave an important contribution using the concept of entropy in the light of a new contexts, where in his work [17] he derived the celebrated Equation (2), known as Boltzmann-Gibbs-Shannon (BGS) entropy. So, \( p_i \) is the probability of finding the system in state \( i \), and \( P = [p_1, \ldots, p_k] \), \( 0 \leq p_i \leq 1 \) and \( \sum_i p_i = 1 \) [13].

\[ S = - \sum_i p_i \ln(p_i) \] (2)

Systems that can be described by the BGS entropy are called extensive systems and have the following important property, called additive property:

\[ S(A + B) = S(A) + S(B) \] (3)

where \( S(A + B) \) is the entropy of a composed system of two independent variables, S(A) and S(B), calculated according to Equation (2). However, this formalism does not describe physical systems that behave as non-extensive systems, requiring a generalization. Then, Constantino Tsallis defined in his work [19], the following equation:

\[ S_q(\{p_1, \ldots, p_k\}) = \frac{1 - \sum_{i=1}^{k} p_i^q}{q-1} \] (4)

Tsallis entropy is a generalized formulation for Boltzmann-Gibbs-Shannon Entropy, where \( k \) is the number of possibilities of the system and \( q \) is the entropy index that characterizes the degree of non-extensivity. Note that, when \( q \to 1 \), Equation (4) meets the traditional Equation (2).

The additive property can explain better extensive systems, but fails in explains non-extensive ones, where Tsallis has proposed the following generalization, called pseudo-additive property:

\[ S_q(A \ast B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B) \] (5)

Both additive and pseudo-additive equations are widely used in image thresholding methodologies for background extraction. However, achieving an optimal threshold is still a challenge. In [6] was proposed a threshold method using the BGS entropy. The work considers two probability distributions, one for the foreground and one for the background. Then, the optimal threshold is computed using the Equation(3). In [1] the methodology proposed by [6] was improved by using non-extensive formalism. Rodrigues in [16] proposed an algorithm for classification of ultra-sound images of breast cancers. His strategy has used a non-extensive entropy. Also in [16] was proposed the first algorithm for automatic computation of the q index, optimizing results.

In [4] are discussed and presented results based on non-extensive relative entropy, or kullback-Leibler Divergence, which is detailed in the following section.

3. Kullback-Leibler Divergence

The Kullback-Leibler divergence or relative entropy, as it is also known, is similar to BGS entropy. Furthermore it considers two probability distributions and computes the divergence between them. With Kullback-Leibler divergence we can measure the gain of information between two regions from the same images. The traditional Kullback-Leibler divergence is defined by the following equation:

\[ D_{KL}(P : P') = \sum_{i} p_i \cdot \log \frac{p_i}{p_i'} \] (6)

where \( P \) and \( P' \) are two probability distributions: \( P = \{p_1, p_2, \ldots, p_n\} \) and \( P' = \{p_1', p_2', \ldots, p_n'\} \).

The Kullback-Leibler divergence can be computed using the non-extensive formalism proposed by Tsallis, that has been defined as [16]:

\[ D_{KL}(P : P') = \sum_{i} \frac{p_i^q}{q-1} \cdot (p_i^{1-q} - p_i'^{1-q}) \] (7)

This equation measures the divergence between two probability distributions for non-extensive systems.

So similar to Equation (6), which measures the divergence of extensive systems, Equation (7) measures the divergence of those non-extensive. In the limit when \( q \to 1 \) it can be shown that Equation (7) meets Equation (6), then Equation (7) is a generalization of Equation (6). The works of [1], [5], [14] and [15] have suggested that Tsallis statistics is a powerful tool for medical image segmentation and the works of [3], [7], [2] and [11] have shown
that the Kullback-Leibler Divergence for non-extensive systems has promising results when applied to the problem of image classification.

Usually one cannot know a priori the behavior of systems as being extensive or non-extensive. The usual solution is then to compute the divergence for various $q$ values including $q = 1$ and choosing $q$ values with best performance. When the value is less than 1.0 it is said that the system has a sub-extensive behavior; when $q = 1$ is observed, it is said that the system is extensive; and when $q > 1$ occurs, it is said that the system is super-extensive [20].

In our work, we propose the use of extensible Kullback-Leibler divergence and investigate the $q$ value that minimizes the amount of information between the two types of images of CT: Cone-Beam and Multislice, both taken for orthodontics.

4. The Proposed Methodology

As we are interested in investigate the power of Kullback-Leibler divergence as a tool for measure luminance information only, the comparison proposed here is carried out only under a range of gray scale $L = [0 : 255]$. Then, the complete Hounsfield scale for CT images was mapped for this range. Figure 1(left) shows an example of CB image and Figure 1(right) shows an example of its corresponding MS image.

Figure 3 shows the diagram of the proposed methodology. In step 1 we have non-normalized input images. The manually normalization is accomplished in step 2, where the captured images by the Cone-Beam method were just cut and maintained without resize or rotations. On the other hand, the Multi-Slice images were cut but rotated in order to meet the CB images as well as possible. Since the probability distributions are invariant to rotation as well as translation, these geometric tasks do not have any influence over our results and are accomplished in order to better visualization purposes only. For the same reason, there is no need to match their respective geometric centers. Figure 2 shows the corresponding normalized images (step 3 in the diagram).

In step 4, the histograms are computed in order to get the probability distributions. Both histograms are shown in Figure 4. We can observe that they are visually similar as also show Figures 2(left) and 2(right), respectively.

With the histograms of each image computed, we calculate the Kullback-Leibler divergence, accomplished in step 6.

5. Experimental Results

In our experiments, first we show the behavior of Tsallis entropy in comparing the information content for 8 both CB and MS images. Table 1 shows the entropy difference between CB and MS images under increasing $q$ values (0.1 to 0.9). Column 2 shows the entropy for CB, column 3 for MS and column 4 shows the corresponding relative difference. The lower the $q$ value the larger is the non-extensive entropy intensity, enhancing their corresponding differences. That’s because the smaller the $q$ value the greater is the corresponding histogram powering, increasing the small histograms entries associated with low probability values.

In order to evaluate the robustness of Tsallis Entropy under changes of Signal-to-noise ratio (SNR), we have applied increasing gaussian noise with standard deviation varying from 2 to 14 over BC and MS images. These strategy simulates the cases when we have scattered histograms.

Figure 5 shows two images of CB, one with no noise (left) and other (right) which was applied a noise with standard deviation 14. Also, we have applied five values for $q$ (0.1, 0.3, 0.5, 0.7, 0.9). In the graphic of Figure 6 and 7 we can see that for $q$ lesser than 0.5 the value of entropy increases significantly; and for $q$ greater than 0.5, approaching to 0.9, the entropy value decreases under increasing gaussian noise. It suggests that even under large gaussian noise,
as that of Figure 5, the Tsallis entropy can generate observable values. Combining this information with that from Table 1, we can conclude that the larger difference between two images occur for values around $q = 0.5$, suggesting a non-extensive system for the images used here. When we see the physical system as a traditional extensive system under severe gaussian noise, it is quite impossible to distinguish the differences between them.

The differences highlighted by Shannon and Tsallis entropies can be projected over their respective Kullback-Leibler divergencies. To show it we have applied Equation (7) to measure the relative information between CB and MS by interchanging the images. Table 2 shows the Kullback-Leibler divergence for CB under MS (column 1) and for MS under CB (column 2) showing its Euclidian distance in col-

Figure 4. Histogram of normalized images. (left) Cone Beam histogram; (right) Multi-Slice histogram.

Figure 5. (left) Original image from Cone Beam; (right) image with gaussian noise with standard deviation 14.

Figure 6. Tsallis entropy under increasing Gaussian noise for CB.
Table 1. Table showing the differences in Tsallis Entropy between CB and MS images

<table>
<thead>
<tr>
<th>(q)</th>
<th>(s(mt))</th>
<th>(s(cb))</th>
<th>(\frac{[s(mt)-s(cb)]}{\max(s(mt), s(cb))})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>128.8339</td>
<td>91.3420</td>
<td>0.2910</td>
</tr>
<tr>
<td>0.2</td>
<td>68.3955</td>
<td>49.6353</td>
<td>0.2743</td>
</tr>
<tr>
<td>0.3</td>
<td>37.1733</td>
<td>27.6408</td>
<td>0.2564</td>
</tr>
<tr>
<td>0.4</td>
<td>20.7459</td>
<td>15.8134</td>
<td>0.2377</td>
</tr>
<tr>
<td>0.5</td>
<td>11.9378</td>
<td>9.3262</td>
<td>0.2188</td>
</tr>
<tr>
<td>0.6</td>
<td>7.1189</td>
<td>5.6942</td>
<td>0.2001</td>
</tr>
<tr>
<td>0.7</td>
<td>4.4240</td>
<td>3.6159</td>
<td>0.1827</td>
</tr>
<tr>
<td>0.8</td>
<td>2.8799</td>
<td>2.3984</td>
<td>0.1672</td>
</tr>
<tr>
<td>0.9</td>
<td>1.9711</td>
<td>1.6669</td>
<td>0.1543</td>
</tr>
</tbody>
</table>

Table 2. Kullback-Leibler Divergence with interchanged images. Column 4 shows the Euclidian distance between Columns 2 and 3 for different \(q\) values.

<table>
<thead>
<tr>
<th>(q)</th>
<th>(K(cb: mt))</th>
<th>(K(mt: cb))</th>
<th>Euclidian Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.3597</td>
<td>0.2862</td>
<td>0.4597</td>
</tr>
<tr>
<td>0.3</td>
<td>0.5642</td>
<td>0.5474</td>
<td>0.7861</td>
</tr>
<tr>
<td>0.5</td>
<td>0.9799</td>
<td>0.9508</td>
<td>1.3654</td>
</tr>
<tr>
<td>0.9</td>
<td>6.4464</td>
<td>6.2551</td>
<td>8.9823</td>
</tr>
</tbody>
</table>

6. Conclusions

This paper shows the sensitiveness of non-extensive Tsallis entropy in a comparative study between Cone-Beam and Multi-Slice CT images. By simple histogram matching it is possible to investigate the contribution of each bin related to each probability of the whole physical system but it is not possible to clearly see the histogram differences. Nevertheless, as the use of Cone-Beam exams has becoming popular and can substitute the Multi-Slice images for various exams, there is a growing need to investigate their differences. Then, the simple euclidean distance or traditional relative entropy may not enhance the main differences, at least when the important pixels are those related with the small probabilities.

When we need to investigate the contribution of small probabilities the extendable Kullback-Leibler divergence may be the appropriate choice. The Kullback-Leibler divergence allows the fine-tuning of the entropic parameter \(q\) for matching two images as probability distribution histograms. This matching can computes the amount of relative information between samples and also allows the sensitiveness under noise and small probabilities when it has some demanded significance. In our results we show that the extendable Kullback-Leibler may be a powerful tool for investigate the differences of information under such situations, which is accomplished in step 6 of our methodology (Section 4).

We can note that, when the \(q\) parameter moves way from 1.0 towards 0 (zero) the entropy value becomes more observable. This behavior remained less invariant even under the strong presence of noise.

This behavior suggest that the Kullback-Leibler divergence is better used by considering the image as non-extensive systems; which means that the statistical states (or luminance probabilities) may have long-range spatial and temporal interactions, even for the small probabilities. It is interesting for systems where the small probabilities are semantically important.

The results presented here only reinforce those found in literature for other types of applications, requiring however, a larger investigation to prove its effectiveness in medical images.

Figure 7. Tsallis entropy under increasing Gaussian noise for MS.
References


