Abstract

In this work, we are interested in the retrieval of images by the content. This interesting alternative approach which is known more by its English acronym CBIR (*Content Based Image Retrieval*) was proposed in order to remedy to the problems caused by its text-based predecessor (TBIR: *Text based image retrieval*). Mainly, TBIR suffers from two disadvantages. The first one is the difficulty of assigning keywords to images due to its manual appearance. On the other hand, the second disadvantage is the subjectivity of the task. In order to counter these disadvantages, the alternative approach CBIR uses the visual characteristics of the image, such as *color*, *texture*, *shape*… etc., for satisfying the user query. However, relying on such low-level characteristics to represent and express the content of an image, generally very complex, creates a significant gap with the semantics and the meanings of the image content. This gap is commonly known by the *semantic gap*.

Particularly, "texture" is a very widespread and widely used characteristic, since human beings are able to distinguish between different manifestations of this attribute. Nevertheless, on the computational level, this attribute remains difficult to characterize. In the present work, texture represents the focal point. Therefore, a set of texture-based methods were addressed and several contributions were proposed.

Mainly, our first contribution is an improvement of the famous LBP method (*Local Binary Patterns*), where we propose to integrate gradually the local information conveyed by the binary patterns. This is achieved through a new region delimitation scheme, which is different from that based on blocks (classic rectangular regions). Hence, our method is called *GLIBP* (*Gradual Locality Integration of binary patterns*). The success of this instance of the proposed *Framework*, *GLI* (*Gradual Locality Integration*) encouraged us to investigate its feasibility for extracting the characteristics of *color* attribute. In this context, the *GLI-Color* method was proposed as well as two of its variants. The first operates on the HSV color space, hence its name: *GLI-Color* $^{(HSV)}$. As for the second variant, it uses the moments, hence its name *M-GLI-Color*. In the second contribution, we were interested in Wavelets in the context of the characterization of image textural content. From this perspective, the existing literature shows that authors claim that the means of the wavelet sub-bands is equal to (or close to) zero. This assertion seemed to us worthy of interest, since the calculated measurements of the different sub-bands are generally the *statistical moments* and the mean is used when calculating the latter. Thus, as a first step, we experimentally verified the difference between the distribution of the wavelet sub-bands means and that of the wavelet sub-bands energies. The obtained results show an interesting difference between the two distributions. Indeed, to the contrary of the means values, which concentrate around the axis of zero, those of the energies are more dispersed. Based on these results, the calculation of the statistical moments around the energy instead of the mean seemed to us to be an effective modification in order to improve the discriminative capacities of these measures. This hypothesis was the subject of the second step, where we

showed the effect of the consideration of energy in the statistical moment's formulas on the effectiveness. As for our third main contribution, it acts on the comparison step of the CBIR systems. This contribution relies on a *greedy strategy* based on a homogeneity measurement imported from the *biclustering of Gene's Expressions* domain, called *Mean Squared Residue* (*MSR*). The proposed method is distinguished by the fact that it takes advantage of all the images already selected as pertinent for the selection of the following image. This is completely different from the traditional approach of CBIR systems, where only the query is used when searching for images. Indeed, in the proposed approach, the results corresponding to a given query are returned as a cluster, grown incrementally.

All the results obtained by the various proposed algorithms and techniques have shown interesting or even better performances compared to several published works. In particular, the GLIBP method showed performances superior to those of the methods: LBP of Ojala *et al.* (1996), DLEP of Murala *et al.* (2012a), 'Block-based LBP' of Takala *et al.* (2005), among others.

On the other hand, the results obtained encourage us to continue on this field of research and inspire us with perspectives to be explored in the future, particularly the exploitation of other biclustering techniques of genetic expressions in the field of content based image retrieval.

Key words: CBIR, image retrieval by the content, texture, LBP, wavelet, biclustering, features extraction, gradual locality integration