

A Global Precision View for Information Retrieval Evaluation Adapted to Image Retrieval Systems

Tahar Gherbi¹, Ahmed Zeggari^{1*}, Zianou Ahmed Seghir²

¹Math and Computer Sciences Dept. Echahid Cheikh Larbi Tebessi University, Tebessa, Algeria

²Faculty ST, ICOSI Lab, University of Khenchela, Khenchela, Algeria

*Corresponding author's e-mail: ahmed.zeggari@univ-tebessa.dz

ABSTRACT

This paper presents a comprehensive analysis of the prevalent evaluation metrics employed in content-based image retrieval. Initially, these widely used metrics are inspired in and influenced by general information retrieval principles, which primarily focused on textual data rather than visual content. In addition, collecting together all or the most of relevant results is not considered by the standard evaluation measures. However, this characteristic is crucial in the context of visual information retrieval. This paper underscores the need for a novel evaluation metric that addresses this particular characteristic.

Keywords: CBIR evaluation, Grouping relevant, Precision

1 Introduction

The essential goal of this work is to observe and highly evaluate systems with more dispersed irrelevant groups while still taking into account how the images that are relevant to the user query are arranged. The literature cites and uses a range of metrics for evaluating image retrieval systems [1]- [3], most of them are initially adapted to the general domain of information retrieval [4], [5]. Instead of displaying the results in a list like in text retrieval results, image retrieval systems often rank the entire collection using a distance metric, they display the results as a screenbased table of images rather than a linear list. People can quickly determine whether a photographic result is relevant with just a brief check, which is not always possible with text retrieval mm. It is necessary to explore and incorporate these differences between information retrieval and image retrieval into the evaluation metrics.

2 Limits of the standard metrics

The conventional quantitative measurements do not take into consideration the following characteristics, which are very critical for a proper quantitative estimation of content-based image retrieval systems:

- Distinction power between the closest results. Therefore, from our perspective, the following query is of some interest: Are two retrieval results, having the same precision value, similar? Can their equivalent systems therefore be assessed as being identical systems?
- The returned window has a significant density of pertinent results in which the pertinent images are grouped in either a small or large collection area. This characteristic was neglected by the actual evaluation metrics.

3 Evaluation requirements

The quantitative assessment metrics for information retrieval (IR) are made to adhere to certain standards, such as how closely they correspond with a user satisfaction criterion, how well they can distinguish between different retrieval results, and how easy it is to understand and to apply them. A good performance metric for a typical evaluation scenario must meet the following standards:

- The quantity of pertinent images that the system returned.



- how is the size of the returned list.
- the final order of the pertinent images in the returned list.
- The interrelations between the relevant images that were retrieved must be grouped and clustered together as much as possible in a nearest area of a retrieved window.

The last characteristic, in our opinion, represents the singularity of CBIR system evaluation. A decent metric must therefore include all of the aforementioned characteristics.

4 Grouping of relevant results

If the results presented to a user are grouped together, by considering images that are extremely similar regarding the user's query, they might be extremely helpful for visual attention by the user. Either from user's perspective, it is challenging to find the most relevant returns if they are scattered. The user in this case, must verify sequentially the returned results in order to identify the pertinent ones. Due to the dispersion of the pertinent images that such a system returned in this instance, the user cannot benefit from the natural quick visual inspection.

5 Conclusion

In summary, our research findings are highly persuasive, and it is evident that the presented study has shown a critical view of the actual evaluation metrics, which are inspired from the general information retrieval domain; as well as their inaccuracy as CBIR evaluation measures. Specifically, the ability to capture the grouping characteristic of relevant images, crucial for a user expectations and natural inspection of results, has been emphasized. Our work recommends the incorporation of this characteristic for a more robust evaluation of content-based image retrieval results, thereby facilitating a more accurate comparison among various image retrieval systems.

How to Cite

T. Gherbi, A. Zeggari, Z. A. Seghir, "A Global Precision View for Information Retrieval Evaluation Adapted to Image Retrieval Systems", *AIJR Abstracts*, pp. 147–148, Feb. 2024.

References

- [1] H. Müller, W. Müller, D. Squire, S. Marchand-Maillet, T. Pun: "Performance evaluation in content-based image retrieval: overview and proposals". *Pattern Recognit. Lett.* 22(5): 593-601 2001
- [2] M. Sanderson, "Performance Measures Used in Image Information Retrieval". In: Müller, H., Clough, P., Deselaers, T., Caputo, B. (eds) *ImageCLEF. The Information Retrieval Series*, vol 32. pp.81-94, 2010 Springer, Berlin, Heidelberg.
- [3] Savvas A. Chatzichristofis, Chryssanthi Iakovidou, Yiannis S. Boutalis and Elli Angelopoulou Angelopoulou, "Mean Normalized Retrieval Order (MNRO): a new content-based image retrieval performance measure". *Multimed Tools Appl.* 70, 1767–1798 (2014).
- [4] R. W. P. Luk, "Why is information retrieval a scientific discipline?". *Foundations of Science.* 27 (2): 427–453, 2022.
- [5] C. D. Manning, P. Raghavan and H. Schütze, "Introduction to Information Retrieval", Cambridge University Press; Cambridge, UK. 2008