

Semantic Improvements to MPEG-7 Descriptors for Content-Based Image Retrieval

K. Salvador Aguilar Domínez

Tecnológico Nacional de México / Centro Nacional de Investigación y
Desarrollo Tecnológico,
Departamento de Ciencias de la Computación,
Mexico

kevin.aguilar17ca@cenidet.edu.mx

Abstract. This abstract shows a brief description of the principal parts of research project called: “Improvements to Recommended Descriptors by the MPEG-7 standard for content-based image retrieval”, who is carried out for obtain the degree of doctor, in the department of computational sciences in CENIDET.

Keywords: MPEG-7, CBIR, semantic, descriptors.

1 Introduction

The file multimedia storage is increasing; therefore, the need arises to efficient filter, search and identify similar visual information [2]. The Moving Picture Experts Group (MPEG) developed the MPEG-7 standard (Multimedia content description interface [2]) to solve this problem and its applications are varied, like [3, 8]. One of the applications that are being given to the visual descriptors of this standard is in Content Based Image Retrieval (CBIR) systems [4, 6], but the descriptors present some problems in image recovery among millions of pictures [5, 9, 10].

Attempts have been made to propose modifications to these descriptors [10, 11], however, the methods are very mathematical and still present problems in recovery: the results only are similar in one of their characteristics such as color or shape [5]. These descriptors could be improved by introducing semantics to avoid those errors that are only similar in some characteristics and do not consider the semantic content of the image.

2 Previous Works in the Area

Currently, several MPEG-7 standard descriptors are still being implemented for images [12], as well as for video [13] and audio [14]. Similarly, the CBIR systems are still applied to solve many problems and work has been done to improve them as shown in [15], where is proposed a novel CBIR technique based on the visual words fusion of Speeded-Up Robust Features (SURF) and Fast REtina Keypoint (FREAK) feature descriptors. Other example is [16], this work focuses on a uniform partitioning scheme

which is applied in the Hue, Saturation and Value (HSV) color space to extract Dominant Color Descriptor (DCD) features; in the proposed CBIR scheme, the DCD features were initially extracted as the color features, and then an appropriate similarity measure was applied. Among other works that solve other problems [17–21]. Modifications and proposals of different descriptors have been made to solve some problems as mentioned in [22], which proposed a novel feature descriptor named “Correlated MicroStructure Descriptor (CMSD)” for image retrieval; this descriptor represents high level semantics by identifying microstructures via establishing correlations between texture orientation, color, and intensity features. Other work is [23], where propose an accurate edge detector using Richer Convolutional Features (RCF), that encapsulates all convolutional features into more discriminative representation, which makes good usage of rich feature hierarchies, and is tractable to training via backpropagation. Among others works [24–26]. One example of semantic used is the semantic Web [27].

3 Research Objectives

General objective: Investigate, propose and evaluate an improved CBIR descriptor that overcomes one deficient or some deficiencies found in current ones, by means of a semantic descriptor.

Specific objectives: Study the current descriptors; Detect deficiencies; Propose improvements (semantic descriptor); and Apply it to a domain accepted by the international community as a test platform.

4 Methodology

We are considering the methodology in two parts: a) the collection of data and b) the process.

a) Data collection. It will seek to obtain collections of problematic digital images, that is, those where other descriptors have problems and fail; for example, extreme cases where, if semantics are not used, the descriptor would fail, such as the retrieval of images of specialized materials or dangerous activities, in these the semantic description is required not only form, texture and color.

In this sense we plan to have collections created by ourselves, as well as those commonly used in specialized literature; This may mean having synthetic digital image collections (artificially created to evaluate a certain aspect), as well as real-world digital images. With this, we hope to cover the entire possible spectrum of practical interest.

b) Process. We assume two basic stages: i) learning, where the proposed algorithm-descriptor is trained and its parameters are automatically tuned for a set of particular images; At the same time, an expert human has the possibility of making an additional subsequent adjustment, to include the human expert, and ii) recognition, where new images are fed to observe the behavior of the previously created model. At the heart of the model, there will be not only color or shape descriptors, but semantic descriptors,

similar to how Frames [28] or other forms of representation of semantic knowledge [29, 30].

5 Preliminary Results

It has been done the experimentation with the descriptors using the datasets in the literature whit the evaluation metric ANMRR (Averaged Normalized Modified Retrieval Rate), which is the recovery rate proposed by MPEG-7 [31]. The characterization time of the entire dataset is also measured in the case of Edge Histogram Descriptor (EHD) and the similarity calculation time in Color Layout Descriptor (CLD). similarly, we have presented proposals for the calculation of similarity using the CLD descriptor. This preliminary results were supervised by M. Lux, professor at Klagenfurt University, Austria [32] and Dr M. Mejía-Lavalle, researcher at TecNM/ CENIDET, Mexico.

Referencias

1. Baresi, L., Colazzo, S., Mainetti, L., Morasca, S.: W2000: A modelling notation for complex web applications. In: Mendes, E., Mosley, N. (Eds): Web Engineering, Springer, pp. 335–364 (2006)
2. Koch, N., Knapp, A., Zhang, G., Baumeister, H.: UML-based web engineering: An approach based on Standards. In: Rossi, G., Pastor, O., Schwabe, D., Olsina, L. (Eds.): Web engineering: Modelling and implementing web applications, Springer, pp. 157–191 (2008)
3. Sikora, T.: The MPEG-7 visual standard for content description - An overview. In: IEEE Transactions on Circuits and Systems for Video Technology, 11(6), pp. 696–702 (2001)
4. Pattanaik, S., Bhalke, D.G.: Efficient content based image retrieval system using Mpeg-7 Features. In: International Journal of Computer Applications, 53(5), pp. 19– 24 (2012)
5. Mejía-Lavalle, M., Pérez-Lara, C., Ruíz-Ascencio, J.: The MPEG-7 visual descriptors: A basic survey. In: International Conference on Mechatronics, Electronics and Automotive Engineering, pp. 115–120 (2013)
6. Hyun, J., Kim, H.K., Oh, W.G.: Study on performance of MPEG-7 visual descriptors for deformable object retrieval. In: 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) (2015)
7. Tyagi, V.: MPEG-7: Multimedia content description standard. Content-Based Image Retrieval, pp. 85–100 (2017)
8. Vertan, C., Badea, M.S., Florea, C., Florea, L., Bădoiu, S.: MPEG-7 visual descriptors se-lection for burn characterization by multidimensional scaling match. E-Health Bioeng, In: E-Health and Bioengineering Conference (EHB), pp. 253–256 (2017)
9. Zinzuvadia, K.M., Tanawala, P.B.A., Brahmhatt, P.K.N.: Feature based image retrieval using classification and relevance feedback technique. In: International

- Journal Advance Engineering and Research Development, 2(5), pp. 508–513 (2015)
10. Reta, C., Solis-Moreno, I., Cantoral-Ceballos, J.A., Alvarez-Vargas, R., Townend, P.: Improving content-based image retrieval for heterogeneous datasets using histogram-based descriptors. *Multimedia Tools and Applications*, 77(7), pp. 8163–8193 (2018)
11. Nazir, A., Ashraf, R., Hamdani, T., Ali, N.: Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor. In: *iCoMET International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development*, pp. 1–6 (2018)
12. Georgescu, F.A., Răducanu, D., Datcu, M.: New MPEG-7 scalable color descriptor based on polar coordinates for multispectral earth observation image analysis. In: *IEEE Geoscience Remote Sensing Letters*, 14(7), pp. 987–991 (2017)
13. Ling-Yu Duan, Vijay Chandrasekhar, Shiqi Wang, Yihang Lou, Jie Lin, Yan Bai, Tiejun Huang, Alex Chichung Kot, Wen Gao: Compact descriptors for video analysis: The emerging MPEG standard. *IEEE Computer Vision and Pattern Recognition* (2017)
14. Lee, J., Lee, J.S.: Music popularity: Metrics, characteristics, and audio-based prediction. In: *IEEE Transactions on Multimedia*, 20(11), pp. 3173–3182 (2018)
15. Jabeen, S., Mehmood, Z., Mahmood, T., Saba, T., Rehman, A., Mahmood, M.T.: An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model. *PLoS One*, 13(4), pp. 1–24 (2018)
16. Fadaei, S., Amirfattahi, R., Ahmadzadeh, M.R.: A new content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features. *IET Image Process*, 11(2), pp. 89–98 (2017)
17. Xia, Z., Zhu, Y., Sun, X., Qin, Z., Ren, K.: Towards privacy-preserving content-based image retrieval in cloud computing. In: *IEEE Transactions on Cloud Computing*, 6(1), pp. 276–286 (2018)
18. Huang, F., Jin, C., Zhang, Y., Weng, K., Zhang, T., Fan, W.: Sketch-based image retrieval with deep visual semantic descriptor. *Pattern Recognition*, 76, pp. 537–548 (2018)
19. Wu, M., Xiao, W., Hong, Z.: Similar image retrieval in large-scale trademark databases based on regional and boundary fusion feature. *PLoS One*, 13(11), pp. 1–25 (2018)
20. Dash, J.K.: Content-based image retrieval using fuzzy class membership and rules based on classifier confidence. *IET Image Process*, 9(9), pp. 836–848 (2015)
21. Mohamadzadeh, S., Farsi, H.: Content-based image retrieval system via sparse representation. *IET Computer Vision*, 10(1), pp. 95–102 (2016)
22. Dawood, H., Alkinani, M.H., Raza, A., Dawood, H., Mehboob, R., Shabbir, S.: Correlated microstructure descriptor for image retrieval. In: *IEEE Access*, 7, pp. 55206–55228 (2019)
23. Liu, Y., Cheng, M.M., Hu, X., Wang, K., Bai, X.: Richer convolutional features for edge detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(8), pp. 1939–1946 (2019)

24. Zhang, J., Xia, Y., Xie, Y., Fulham, M., Feng, D.D.: Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features. In: *IEEE Journal of Biomedical and Health Informatics*, 22(5), pp. 1521–1530 (2018)
25. Napoletano, P.: Visual descriptors for content-based retrieval of remote-sensing images. In: *International Journal of Remote Sensing*, 39(5), pp. 1343–1376 (2018)
26. Fernández, L., Payá, L., Reinoso, O., Jiménez, L.M., Ballesta, M.: A study of visual descriptors for outdoor navigation using Google street view images. *Journal of Sensors* (2016)
27. Villagrán, A., Hernandez-Gress, N., Gonzalez-Mendoza, M.: *Introducción a las tecnologías de la web semántica*. Research in Computing Science (2012)
28. Minsky, M.: *A framework for representing knowledge* (1974)
29. Douglas, B.L.: CYC: A large-scale investment in knowledge infrastructure. In: *Communications of ACM*, 38(11), pp. 33–38 (1995)
30. Chen, L., Lambon-Ralph, M.A., Rogers, T.T.: A unified model of human semantic knowledge and its disorders. *Nature Human Behaviour*, 1(3), pp. 39 (2017)
31. Manjunath, B.S., Sikora, T., Salembier, P.: *Introduction to MPEG-7 : Multimedia content description interface*. John Wiley & Sons (2002)
32. Lux, M., Marques, O.: *Visual information retrieval using java and lire*. Morgan & Claypool Publishers (2013)