

Image database indexing: Emotional impact assessing

Syntyche Gbèhounou

XLIM Laboratory, UMR CNRS 7252, University of Poitiers, France

Advisors: François Lecellier and Christine Fernandez-Maloigne

PhD thesis defended 3rd October 2013, University of Bristol

Received 19th February 2015; accepted 27th July 2015

1 Abstract

The goal of my PhD was to propose an efficient approach for emotional impact recognition based on CBIR techniques (descriptors, image representation). The main idea relies in classifying images according to their emotion which can be "Negative", "Neutral" or "Positive". Emotion is related to the image content and also to the personal feelings. To achieve our goal we firstly need a correct assessed image database. Our first contribution is about this aspect. We proposed a set of 350 diversified images rated by people around the world. Added to our choice to use CBIR methods, we studied the impact of visual saliency for the subjective evaluations and interest region segmentation for classification. The results are really interesting and prove that the CBIR methods are useful for emotion recognition. The chosen descriptors are complementary and their performance is consistent on the database we have built and on IAPS, reference database for the analysis of the image emotional impact.

2 The new database for emotional impact study: SENSE

The existing databases proposed for emotion study are various in terms of content; from photographs to abstract images but present some weakness. Moreover, on the literature databases, a new evaluation cannot be performed according to confidentiality rules. So we propose a new database with low semantic images in order to ensure the independance of the emotionnal responses during the test sessions.

Our database, called SENSE (Studies of Emotion on Natural image databaSE), is composed of 350 low semantic, natural and diversified images. More information about this image set, free to use for academic research purpose is given in [1]. Using SENSE for new tests or modifications is permitted according the non-attribution rule of the original images.

During our subjective evaluations, we asked observers to indicate the nature of the emotion (Negative, Neutral or Positive) and its power (from low to high). We organized two kinds of evaluations, made several months apart, inducing two image subsets: SENSE1 composed of the whole images and SENSE2 containing regions of interest defined from a visual saliency model proposed by Ferreira Da Silva et al. [2]. The size of the images evaluated during this test varies from 3% to 100% of the size of the original ones. The thumbnails are built with the bounding rectangle of the salient regions. More than 1000 persons rate each of the two databases leading to

Correspondence to: < {syntyche.gbèhounou, francois.lecellier, christine.fernandez} @univ-poitiers.fr >

Recommended for acceptance by Jorge Bernal

DOI <http://dx.doi.org/10.5565/rev/elcvia.709>

ELCVIA ISSN:1577-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

a properly assessed database. The result analysis shows that for the three classes of emotions (Negative, Neutral and Positive), when the viewed area has a size equal to at least 50% of that of the original image, 77% of the images has the same class during the evaluations of SENSE1 and SENSE2. We also notice that 77% of the SENSE1 inconclusive images has a definitive class when using their thumbnails from SENSE2. The visual saliency usage for emotional impact evaluation in our case improves the database assessment. This shows that using this property of our visual system is an interesting compromise for semantic interpretation reduction. We also evaluated the impact of visual saliency using during the computational emotion retrieval based on CBIR techniques.

3 The CBIR approach we used

For this second part of our research we chose CBIR approach for emotion retrieval; from the image features to the visual signatures. Because emotion retrieval is not only content based recognition, we used a machine learning algorithm: SVM. In this document, we only talk about the results considering only two classes: Negative and Positive.

We selected a certain number of local and global features to model emotional information. The more information is given in the [the french version of my thesis](#). An english long summary of my thesis report is also available on request. We used Bag of Visual Words (BoVW) [3] and Vector of Locally Aggregated Descriptors (VLAD) [4] for the image signatures. For the codebook building we used traditional *K-Means* algorithm and a new one we proposed based on Information Gain computing: Iterative Random Visual Word Selection (It-eRaSel) [5]. The proposed emotional impact recognition method based on CBIR techniques is also evaluated on a database considered as a reference for emotion study: IAPS [6].

At first, we noticed that VLAD matches better for the local descriptors and BoVW for global ones. Secondly, we concluded that the visual dictionary has little impact on the behaviour of descriptors for SENSE1 and IAPS. On the other side, the selected descriptors differently perform whether the nature of emotion is positive or negative. However the average of correct classification for each nature of emotion remains higher than the random selection (50%). If we had to make a choice about the best descriptors between those selected, we think that SIFT and its different colour extensions are the more multipurpose. Concerning the images evaluated during SENSE2, we also noted an improvement during the computational recognition for some local features. The usage of the regions of interest obtained with visual saliency model increases the results for positive and negative images especially for SIFT and OpponentSIFT: +10%. Besides the interesting contributions of visual saliency were not shown only for emotion recognition but also in the case of image retrieval.

4 Conclusions and future works

During our thesis researches, the contribution of visual saliency was assessed for image evaluation for emotion recognition and also for image size before computational solution. This functionality of our vision system shows an interest for emotion recognition by improving the evaluation of our image database. The classification results of the regions of interest obtained with saliency model are equivalent or better depending on the descriptors. This work opens up many perspectives for both image retrieval and emotion recognition. One of the more interesting regarding the results presented in this abstract concerns visual saliency usage. For SENSE2, we used a bounding box of the different salient areas, we think that a more precise region definition must be studied: defining different regions of interest by image and determine the emotion of each region. The final emotion of the image could be a combination of the negative and positive areas thereby resuming the idea of the harmony of a multi-coloured image from Solli et al. [7]. The fusion method could be found based on subjective evaluations to determine the correct weighting between negative and positive "patches" to form the final emotional impact.

References

- [1] Gbèhounou, S. and Lecellier, F. and Fernandez-Maloigne, C. and Courboulay, V., "Can Salient Interest Regions Resume Emotional Impact of an Image?", *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, 8047:515-522, 2013.
- [2] Perreira Da Silva, M. and Courboulay, V. and Prigent, A. and Estraillier, P., "Evaluation of preys/predators systems for visual attention simulation", *Proceedings of the International Conference on Computer Vision Theory and Applications*, Paris, 275-282, 2010.
- [3] Sivic, J. and Zisserman, A., "Video Google: A Text Retrieval Approach to Object Matching in Videos", *Proceedings of the International Conference on Computer Vision*, 1470-1477, 2003.
- [4] Jégou, H. and Douze, M. and Schmid, C. and Pérez, P., "Aggregating local descriptors into a compact image representation", *Proceedings of the 23rd IEEE Conference on Computer Vision & Pattern Recognition*, 3304-3311, 2010.
- [5] Urruty, T. and Gbèhounou, S. and Le, H. T. and Martinet, J. and Fernandez, C., "Iterative Random Visual Word Selection", *Proceedings of International Conference on Multimedia Retrieval*, Glasgow, 249-256, 2014.
- [6] Lang, P. J. and Bradley, M. M. and Cuthbert, B. N., "International affective picture system (IAPS): Affective ratings of pictures and instruction manual", *Technical Report A-8*, 2008.
- [7] Solli, M. and Lenz, R., "Color harmony for image indexing", *Proceedings of the 12th International Conference on Computer Vision Workshops*, 1885-1892, 2009.