Content Based Image Retrieval with LIRe

Mathias Lux
Klagenfurt University, Institute for Information Technology
Universitätsstrasse 65-67
Klagenfurt, Austria
mlux@itec.uni-klu.ac.at

ABSTRACT

LIRe (Lucene Image Retrieval) is an open source library for content based image retrieval. Besides providing multiple common and state of the art retrieval mechanisms it allows for easy use on multiple platforms. LIRe is actively used for research, teaching and commercial applications. Due to its modular nature it can be used on process level (e.g. index images and search) as well as on image feature level. Developers and researchers can easily extend and modify LIRe to adapt it to their needs.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Documentation

Keywords

Image retrieval, image features, image search, image indexing

1. INTRODUCTION

Content based image retrieval has been around for some years. Commercial software like IBM QBIC or Cantos Cumulus did not result in a broad application of content based image retrieval. Also research has been consolidated and discussed on a general level several times [7, 21]. Different evaluation organisations and events provide playgrounds for researchers and practitioners to test new and consolidate current approaches [16, 20]. Many research groups and practitioners aim to build on existing tools to avoid reimplementation of existing approaches. LIRe satisfies this need by providing a library of basic and advanced functions needed in the field of visual information retrieval.

MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.
Copyright 2011 ACM 978-1-4503-0616-4/11/11 ...$10.00.

2. ARCHITECTURE

LIRe is provided as Java library to be integrated in existing or yet to be built applications and code. Main approach of LIRe is to hide as much as possible of the underlying complexity of content based image retrieval. LIRe is based on Lucene [15], a text search engine providing inverted indexing, search and fast random access to text indexes.

At a first glance developers encounter in LIRe few parameters to be set and even fewer choices. *DocumentBuilder* classes provide easy access to different low level features and wrap the use of Lucene, which is used as index. *ImageSearcher* classes allow for search and retrieval based on single query images or already indexed documents (see Fig. 1).

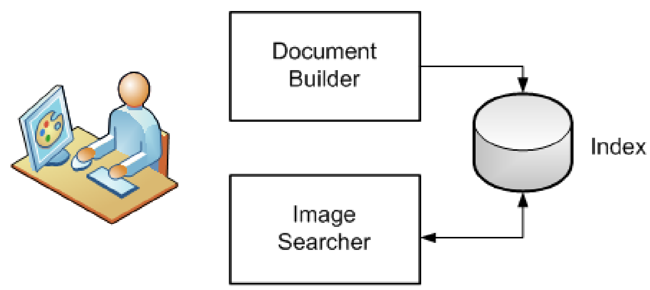

Figure 1: LIRe general architecture.

Developers can access parameters and modules at different levels. Low level features are organized in classes, which implement the LireFeature interface as shown in Fig. 2. Each of the different feature implementations provides means to extract the feature from an image, to compute the distance to a feature of the same class and to serialize and deserialize the histogram data for indexing and storage.

3. GLOBAL FEATURES

LIRe provides a broad range of common and state of the art global image features. Furthermore, extension of LIRe with new global features can be done short time by implementing the LireFeature interface. Indexing and search routines are provided by generalized classes. Features include but are not limited to

1. Color Histograms supporting different color spaces and quantization routines as well as different metrics for distance computation.

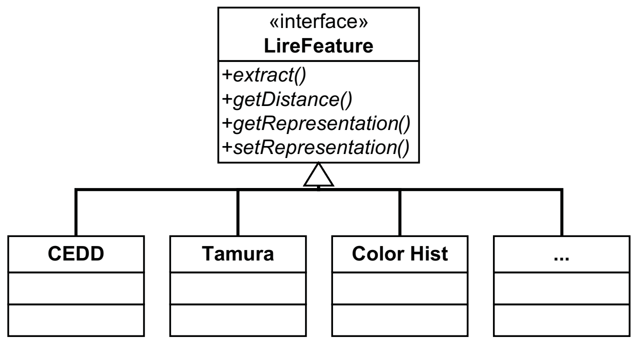

Figure 2: *LireFeature* interface and low level feature classes.

2. MPEG-7 descriptors scalable color, color layout and edge histogram, see [3].
3. The Tamura texture features coarseness, contrast and directionality, see [22]
4. Color and edge directivity descriptor, CEDD, see [4] and Fuzzy color and texture histogram, FCTH, see [5], which are joint histograms combining fuzzy color and edge histograms.
5. Auto color correlation feature defined by Huang et al. [10]

4. LOCAL FEATURES

Besides global features current research focuses also on local features and their application in visual information retrieval [19], the well known *bag of visual words* approach. It has been shown that for many tasks local feature histograms and visual words outperform global features [8]. In LIRe three types of local features are available.

1. Scale-Invariant Feature Transform, SIFT [13] based on the ImageJ[1] implementation,
2. Speeded Up Robust Feature, SURF [2] based on jopensurf[2] and
3. Maximally Stable Extremal Regions, MSER [14], based on the algorithm presented in [17].

All those local features are handled by LIRe like global features, but with a $1 : n$ relation of image to feature. Consequently local features are stored within the Lucene index just like global features. To employ the bag of visual words approach Lire provides *HistogramBuilder* classes. Those classes create a visual words vocabulary using k-means and (i) assign local feature histograms to images stored in the index and (ii) creates a *visual sentence*, like ”v13 v522 v34 . . . ”, from the set of visual words {v0, v1, v2, v3, . . . }. This representation of visual words allows to utilize the fast text based retrieval mechanisms of Lucene for search. The corresponding searcher classes utilize an inverted list with a configurable similarity metric to search through all the visual sentences.

[1]http://rsbweb.nih.gov/ij/
[2]http://code.google.com/p/jopensurf/

5. INDEXING AND SEARCH

In general global as well as local features are stored as either text or byte payload in a Lucene based index. Basic search implementations then use linear search to find the most promising n candidates and return them in a ranked list. For local features a further indexing step allows for the use of an inverted index for the bag of visual words approach (see previous section). However, for several scenarios with millions of images and global features sub linear search in feature spaces is needed. In LIRe two basic mechanisms for faster search in feature spaces are included. *Fastmap* [9] is a rather aged but fast method that allows for dimensionality reduction. With the a reduced number of dimensions features can be organized in data structures for fast retrieval, like for instance R* trees. A newer and more promising approach is the *Metric Spaces* [1] method. This approach utilizes inverted lists to characterize data points in feature spaces by their distance to a set of reference data points. Utilizing a ranked list of nearest neighbours the footrule distance provides an approximation of the original pair wise distance. This approach has been shown to work well on sets of millions of images.

6. APPLICATIONS

Main showcase is the open source application LireDemo, which is provided in the LIRe SVN and on the LIRe web page. Fig. 3 shows the indexing screen of the demo UI. Features include indexing, search, index browsing, selection of different features and an image mosaic option, contributed by students.

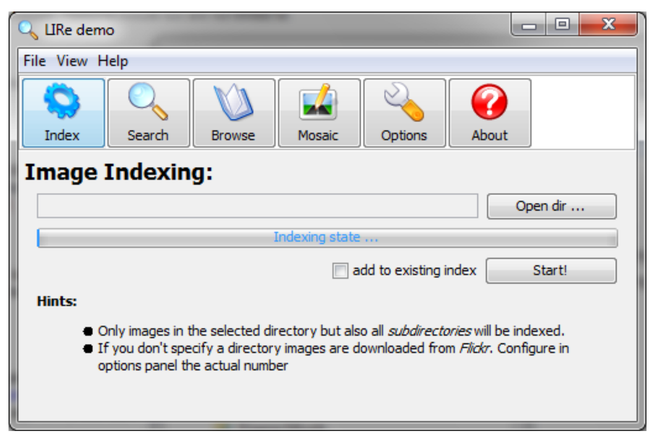

Figure 3: LireDemo user interface screenshot showing the indexing screen.

Within the lifetime of the project multiple uses have been reported. Usage ranges from teaching at university courses, integration in research and use in TRECVid and ImageCLEF, and multimedia retrieval tasks. Besides several others LIRe features are used in [18], [6], and [11]. LIRe has also for intance been integrated in other software like in ZoRa Photo Director[3] for duplicate detection, is used for finding similar photos in asset management application by

[3]http://www.photozora.de

Table 1: **Performance of selected features tested on the Wang Simplicity data set [12] in terms of mean average precicion (map), precision at ten (p@10) and error rate (er).**

Feature	map	p@10	er
Auto color correlogram	0.475	0.725	0.171
CEDD	0.506	0.710	0.178
Color histogram	0.484	0.704	0.205
FCTH	0.498	0.703	0.209
Gabor	0.233	0.248	0.707
JCD	0.510	0.719	0.177
JPEG coefficients histogram	0.446	0.669	0.215
MPEG-7 color layout	0.439	0.610	0.309
MPEG-7 scalable color	0.305	0.470	0.462
MPEG-7 edge histogram	0.333	0.500	0.401
SIFT BoVW	0.183	0.243	0.687
Tamura	0.253	0.359	0.601

mediamid[4] and has been employed for similarity search in mobile applications by R/GA[5].

In teaching LIRe allows for instance for easy construction of image search engines or analysis of image and video sets. Students can employ established methods on different data sets, extend existing methods or change parameters and learn about the impact.

7. PERFORMANCE

Retrieval performance is of course a matter of test data set and choice of parameters like number of bins, quantization routine and weighting scheme, but tests indicate that LIRe features performance similar to well known evaluations. Features tested include the standard RGB color histogram with 1024 bins and L1 distance function. The SIFT based bag of visual words implementation employs standard TF*IDF weighting from Lucene. Overall results are shown in Tab. 1. A closer look on the performance in different categories is presented in Fig. 4.

8. CONCLUSIONS

LIRe has been around for some time now after the inital release in 2006. It has been downloaded from sourceforge.net 18.521 times[6]. While this would not be an impressive number for an end user application it is a great success for a content based image retrieval software library, also in the light of the fact that this does not include downloads via SVN and beta releases not released on sourceforge.net. While LIRe has no regular stable release cycle its code base grows constantly. New features are added and tested and bugs are fixed as soon they are found and reported. All additions are submitted to the SVN and are available as open source for the public. All in all LIRe provides (i) a great base line for further research in content based image retrieval, (ii) a valuable resource for teaching the implementation of content based search, (iii) a robust and easy-to-use library for easy integratrion of standard content based image retrieval functions in applications.

[4]http://www.mediamid.com
[5]http://www.rga.com/
[6]as stated on May 9th, 2011, on the sourceforge.net homepage

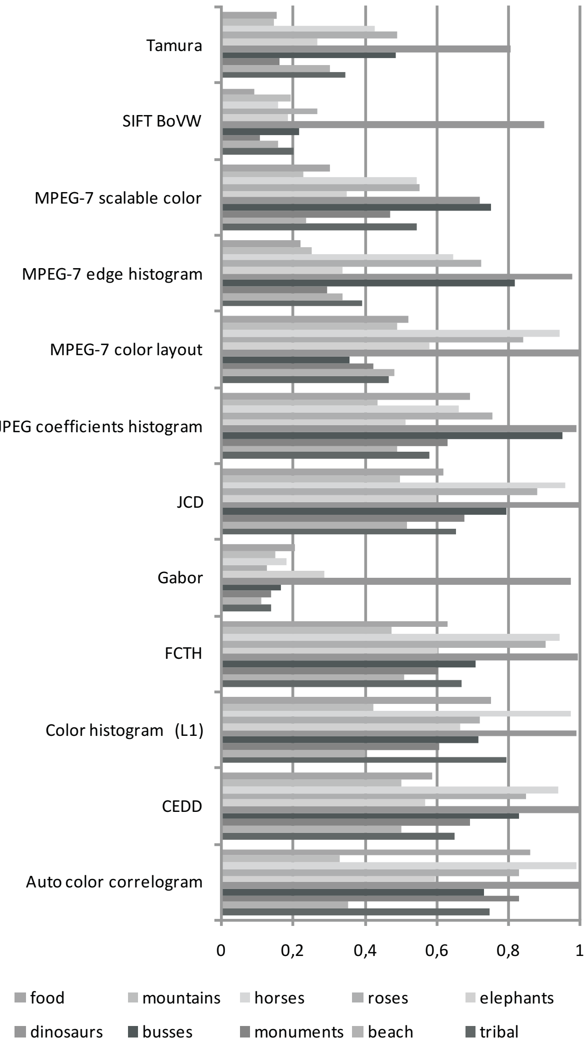

Figure 4: MAP per category in the Wang Simplicity data set.

9. ACKNOWLEDGMENTS

We'd like to thank at least some of the numerous people having contributed code and ideas to Lire or having published their work as open source allowing integration into LIRe: Anna-Maria Pasterk, Arthur Li, Arthur Pitman, Bastian Hösch, Benjamin Sznajder, Christian Penz, Christine Keim, Christoph Kofler, Dan Hanley, Daniel Pötzinger, Fabrizio Falchi, Giuseppe Amato, Janine Lachner, Katharina Tomanec, Lukas Esterle, Manuel Oraze, Marian Kogler, Marko Keuschnig, Rodrigo Carvalho Rezende, Roman Divotkey, Roman Kern and Savvas Chatzichristofis.

This work was partially supported by the Lakeside Labs GmbH, Klagenfurt, Austria and funding from the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant 20214/17097/24774.

10. REFERENCES

[1] G. Amato and P. Savino. Approximate similarity search in metric spaces using inverted files. In

Proceedings of the 3rd international conference on Scalable information systems, InfoScale '08, pages 28:1–28:10, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[2] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. In *Computer Vision - ECCV 2006*, volume 3951 of *Lecture Notes in Computer Science*, pages 404–417. Springer, 2006.

[3] S.-F. Chang, T. Sikora, and A. Puri. Overview of the mpeg-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):688–695, June 2001.

[4] S. A. Chatzichristofis and Y. S. Boutalis. CEDD: Color and Edge Directivity Descriptor. A Compact Descriptor for Image Indexing and Retrieval. In A. Gasteratos, M. Vincze, and J. Tsotsos, editors, *Proceedings of the 6th International Conference on Computer Vision Systems, ICVS 2008*, volume 5008 of *LNCS*, pages 312–322, Santorini, Greece, May 2008. Springer.

[5] S. A. Chatzichristofis and Y. S. Boutalis. FCTH: Fuzzy Color And Texture Histogram A Low Level Feature For Accurate Image Retrieval. In *Proceedings of the 9th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2008*, pages 191–196, Klagenfurt, Austria, May 2008. IEEE.

[6] P. Daras, T. Semertzidis, L. Makris, and M. G. Strintzis. Similarity content search in content centric networks. In *Proceedings of the international conference on Multimedia*, MM '10, pages 775–778, New York, NY, USA, 2010. ACM.

[7] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40(2):1–60, 2008.

[8] T. Deselaers, D. Keysers, and H. Ney. Features for image retrieval: an experimental comparison. *Inf. Retr.*, 11(2):77–107, 2008.

[9] C. Faloutsos and K.-I. Lin. Fastmap: A fast algorithm for indexing, data miningand visualization of traditional and multimedia datasets. In *ACM SIGMOG*, pages 163–174, CA, USA, 1995.

[10] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih. Image indexing using color correlograms. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition, CVPR '97*, volume 00, pages 762–768, San Juan, Puerto Rico, June 1997. IEEE.

[11] D. Inkpen, M. Stogaitis, F. DeGuire, and M. Alzghool. Clustering for photo retrieval at image clef 2008. In *Proceedings of the 9th Cross-language evaluation forum conference on Evaluating systems for multilingual and multimodal information access*, CLEF'08, pages 685–690, Berlin, Heidelberg, 2009. Springer-Verlag.

[12] J. Li and J. Z. Wang. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25:1075–1088, September 2003.

[13] D. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150 –1157 vol.2, 1999.

[14] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761 – 767, 2004. British Machine Vision Computing 2002.

[15] M. McCandless, E. Hatcher, and O. Gospodnetic. *Lucene in Action, Second Edition*. Manning, 2010.

[16] H. MÃijller, P. Clough, T. Deselaers, and B. Caputo, editors. *ImageCLEF*. Springer, 2010.

[17] D. Nister and H. Stewenius. Linear time maximally stable extremal regions. In *Proceedings of the 10th European Conference on Computer Vision: Part II*, ECCV '08, pages 183–196, Berlin, Heidelberg, 2008. Springer-Verlag.

[18] M. Rahman, S. Antani, and G. Thoma. A biomedical image retrieval framework based on classification-driven image filtering and similarity fusion. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 1905 –1908, 30 2011-april 2 2011.

[19] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. *Computer Vision, IEEE International Conference on*, 2:1470, 2003.

[20] A. F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and trecvid. In *MIR '06: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, pages 321–330, New York, NY, USA, 2006. ACM Press.

[21] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380, December 2000.

[22] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6):460–472, June 1978.