Europeana DSI 2– Access to Digital Resources of European Heritage

MILESTONE

MS6.3: Advanced image discovery development plan

Revision	Version 0.1
Date of submission	
Author(s)	Sergiu Gordea, AIT - Austrian Institute of Technology David Haskiya, Europeana Foundation
Dissemination Level	Public

Co-financed by the European Union
Connecting Europe Facility

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
1		Sergiu Gordea	AIT - Austrian Institute of Technology	

Statement of originality:

This milestone contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

Table of Contents

Introduction

This Milestone Document has the goal to prepare an activity plan for the work to be carried out within the scope of *Subtask 6.3.4 Develop image discovery services*, which has the following formal description in Description Of Action document:

"Subtask 6.3.4. Develop content-based discovery services (infrastructure, search index, APIs) to support image search by similarity and to support improving and extending the existing browse by colour functionality. This will support not only image based discovery and browse in Europeana Collections but will also be made available for 3rd party use via APIs documented on Europeana Labs."

Within the next sections we present the goals and the requirements that guide the development of the advanced image similarity service (see Goals & Requirements) and the requirements for the image dataset used by the service (see Dataset). The state of the art report (see State of the Art Report) presents a set of alternative approaches to be evaluated for enhancing the existing search services. The final section summarizes the activities required for implementing the service and the development plan (see Implementation of advanced image search services).

Goals & Requirements

The goal of the Subtask 6.3.4 is to enhance the existing algorithms used for image discovery purpose in Europeana and Europeana related projects, with the aim to enhance their scalability and improving their retrieval accuracy. For improving the scalability and maintenance of the services, we plan to migrate existing implementations to use the mainstream technologies used in Europeana APIs (e.g. solr-based solutions). In order to enhance the accuracy of search algorithms, we plan to integrate state of the art technologies, by evaluating state of the art algorithms developed in the past years, which make use of local image descriptors. However, the fine tuning of advanced algorithms requires experimentation with the concrete/targeted Europeana image datasets.

The advanced image similarity search service is meant to be implemented as a core service that will serve different applications and user scenarios. Consequently, the development of the service will be guided by the particular functional needs expressed in terms of user scenarios and information needs expressed in terms image datasets.

User stories

The following user stories are the main scenarios aimed to be supported by the service.
Note that while these user stories are phrased from the perspective of a user of Europeana Collections they indirectly set the requirements on the Image similarity API and backend. For every user story, the user might as well be a developer that integrates this service in third party applications.

Find similar to existing Europeana images

Short description: As a user I want to find in Europeana similar images to the one I can see in Europeana portal.

This is the core and central use case for image similarity in Europeana. The user should be able to trigger this option in multiple ways:

1. When a user is accessing a search result page. Each item in the search results that is part of the similarity index should have an action/option where the user can activate a search for similar images with one click.
2. When user is looking at an item page. If the item is part of the similarity index it should have an action/option where the user can activate a search for similar images similar with one click.
3. When user is looking at an item page. If the item is part of the similarity index it should include a gallery of the most (visually) similar 4-12 items, which the user access directly when clicking on to view (note: this is like the current similar items, but based on visual similarity not metadata)
4. When a user is looking at an enlarged image and marks an area of the image to search for similar images.[1]

Find similar to user provided images

Short description: As a user I want to find images in Europeana similar to the one I upload or the one I can reference with a web link.
The user should be able to trigger this option in multiple ways:

1. By dragging and dropping a locally stored image into the Europeana Collections search box
2. By dragging and dropping an image from another website into the Europeana Collections search box
3. By taking a photo with their mobile device (e.g. CultureCam[2])
4. By providing the URL to an image available in the Web
5. By uploading an image file

Filtering search results

Short description: As a user I want to be able to filter the set of returned images by using facets. This user story is orthogonal to the previous two user stories, in the sense that any image similarity search result must be possible to be reduced to a subset that matches the list of facets selected by user. The following facets[3] must be supported and available for the user to choose from (i.e. one or more can be selected):

- Theme: Art, Fashion, Music
- Content type/motif: A more complete list of image content types/motifs will be created which includes the ones available within the Art, Fashion and Music themes. (see also Section Specific themes and content types /motifs)
- Copyright: Can I use it? and specific rights statements (i.e. same functionality as in search api)

[1] A good comparison here is Pinterest: https://www.pinterest.com/
[2] http://culturecam.eu/
[3] All these facets already exist in Europeana Collections with the exception of Category/Motif.

- Data provider
- Pixel dimensions/size
- File format/MIME-type
- Colour/Greyscales
- Orientation

Dataset

The first step for implementation of advanced image search services is the selection of the dataset to be used by the service. The quality of the search results is directly dependent by the technical quality of the images. Typically, the image processing and retrieval services require a certain size of the image content. The technical requirements specified by image search libraries recommend a minimal image size of 500-640 pixels on the smallest dimension (height or width). The used algorithms are able to work with smaller images, however for optimal retrieval is recommended to have comparable sizes with the previously named values. Using high resolution images increases considerable the processing time without providing a relevant increase of the search accuracy.

Consequently, we aim at using the 400px thumbnails available in Europeana.Additionally, the color information has an important impact on the quality of most image retrieval algorithms. With respect to this requirement, the image content from Art History and Fashion collection are best suited to be used within this service.

Specific themes and content types /motifs

The guiding principles for the selection of image dataset include:

1. Image content types (e.g. subject categorizations) must be possible to align with one or more themes and thematic collections (Art, Fashion, Music are top priority, other themes may be added later)
2. Image content types must be specific enough and provide meaningful information to the end user (i.e. Clothing, not Fashion)
3. Image content types should not be too narrowly defined as that would be very difficult to scale (so Music instruments, not Flutes)

Note that for all suggested links/selections there will be some noise (e.g. in the sense of false positives associated to an image content type) given that these are selected using free text search as no such categorization schema is used in the metadata (e.g. in subject field for example). This can only be avoided through a manually curation process, which is time consuming approach and is not scalable for the targeted dataset sizes. A mixed approach may be used in which the selections represent handpicked training sets for automatic classification using machine learning algorithms.

Art subsets

Paintings

The paintings contained within the thematic art collections.

Link to search result with a selection of paintings.

Sculpture

The sculptures contained within the thematic art collections.

Link to search result with a selection of sculptures.

Note: There may well be drawings and other indirect images of sculptures in this result set. It may be difficult to narrow it down to e.g. only photographs of sculptures.

Posters

The posters contained within the thematic art collections.

Link to a search result with a good selection of posters. Note: NOT all posters in Europeana.

Illuminated manuscripts

The illuminated manuscripts contained within the thematic art collections.

Link to search results for illuminated manuscripts.

Tapestries

The tapestries contained within the thematic art collections.

Link to search results with tapestries

Porcelain ceramics

Link to a search results for porcelain ceramics (API-query)

Fashion subsets

Shoes

Link to search results for shoes (API-query)

Headgear

Link to search results for headgear (API-query)

Fashion shows

The fashion photography contained within the thematic fashion collections.

what:"fashion show" OR "fashion photograph"

Fashion sketches, drawings and illustrations

Link to search results for fashion sketches, drawings and illustrations (API-query)

Jewellery

The jewellery contained within the thematic fashion and art collections.

A good set within the Fashion Collections.

A set of of jewellery within the Art Collections

Objects and tools

Music instruments

The music instruments contained within the thematic music collections.

Link to images of music instruments

Evaluation of alternative solutions

The similarity search algorithms typically need to make a tradeoff between the server response time and the precision of retrieval. A higher accuracy of the search is obtained when combining different features, using both global and local descriptors. The server response time is mainly depending on the complexity of the similarity measures, the complexity of feature descriptors and the method used for indexing the extracted features.

The advanced image search service will enhance the solutions implemented in Europeana Creative project, by including the state of the art solutions that use local feature descriptors and additional faceting/filtering functionality.
However, it is required to evaluate which of the possible enhancements provide better results on the given dataset by running automatic evaluations of alternative solutions. The evaluation can be carried out only when clear criteria for assessing the quality of search results are defined. This can be achieved when a representative subset of items in the dataset are properly categorized according to the objects and/or subjects of the image content. This will be achieved by analyzing the description and categorization (i.e. subject) fields existing in the metadata.

For this step, the existing third party software implementations will be used, without integrating them in the Europeana platform.

State of the Art Report

Europeana Image Search. An image similarity search algorithm, using global image descriptors was developed within the scope of Assets4Europeana project [Amato 2011] and was further developed within the Europeana Creative project, in which was also integrated with the CultureCam frontend [Gordea 2015]. In this approach, a nearest neighborhood algorithm is used to reduce the search space and implement and efficient indexing solution, that ensures optimal execution performance at retrieval time. This is achieved by selecting a pivot set and computing the distances between each indexed image and the pivots. At runtime, the search results are ordered by their similarity relative to the pivot set. A detailed description of the feature extraction and indexing process is presented in [Amato 2011].

While the global feature descriptors, are not able to catch the particularities of objects available in the image, often the local descriptors are not appropriate for assessing the overall similarity of the images. As the amount of higher quality images is available in Europeana was increasing considerably in the last years and standard thumbnails are available in 400X pixels, it is expected that multimodal search algorithms will improve the search accuracy for the heterogeneous content available in Europeana.

In the past years, the MPEG standardization expert group was focusing on creating standardized representations for image descriptions in order to support development of efficient image retrieval systems[Sikora 2001]. While the early standardization work was concentrating on specification of the so called global descriptors, the latest activities focus on local descriptors, which are more appropriate for classification and pattern recognition purposes [Bianco 2015].

There are many different state of the art algorithms available which implement the extraction of local visual features that are used in image recognition and retrieval scenarios. However, in order to support this functionality in client-server environments, supporting various client devices it is important to use standardized solutions. Starting with 2010 the Moving Picture Experts Group (MPEG) worked on creating an ISO standard that specifies the extraction and compression of Compact Descriptors for Visual Search (CDVS) [Duan 2015].

A comparative evaluation of the CDVS approach applied on six different datasets using various variations of the search algorithms is presented in [Bianco 2015]. The search algorithms are based on CDVS specifications but use different algorithms used for detection of interest points. The reported results show very good performance for near-duplicate image search, however, these datasets are much smaller and less heterogeneous than the Europeana dataset.

The interest point detection is a core component used in the extraction of local features. SIFT and SURF interest point detectors were developed in the past decades and proved to be very effective for structure and object detection algorithms. However, these are patented solutions and never detectors like KAZE and ORB were proposed later on to reduce the computational effort and robustness to noise of detection algorithms [Rublee 2011, Alcantarilla 2012]. An open source implementation for various local feature extraction algorithms (i.e. including SIFT, ORB, KAZE) is available in OPENCV library[4]. Pastec[5] is an open source library that implements near duplicate image detection based on indexing of ORB feature descriptors.

[4] http://docs.opencv.org/3.0-beta/modules/features2d/doc/feature_detection_and_description.html
[5] http://pastec.io/

For the implementation of advance image similarity search service, not patented and open source solutions will be taken in consideration.

Lire is implemented as a Java Library that offers an open source implementation of common feature extraction (i.e. mainly global features) and search algorithms for content based image retrieval[6] [Lux 2008]. The search engine is implemented on the top of Lucene library which is well known as the main open source text retrieval library. In 2015 the beta version of the Lire Solr plugin was released, which exploits the engineering work invested by the solr community with the aim to offer a better scalability and configurability for Lucene based search indexes by making them cloud compliant. Lire offers support for extraction of various feature descriptors and an implementation for various similarity measures that can be used in combination with specific features. However, the retrieval accuracy of image retrieval algorithms is highly dependent on the characteristics of the given datasets and by appropriate combination of feature descriptors and weighting schemes. A comparative evaluation of various search algorithms applied on 4 different datasets is presented in [ref]. However, these experiments present the performance of available feature descriptors, but not their combinations.

In the context of advanced image retrieval service, the feature extractors and the solr plugin are the most relevant functionality offered by Lire. A solr based implementation of the search algorithms will ensure the scalability of the search service. Moreover, this will provide support for integrating similarity search with filtering and faceting functionality. An open question is related to possibility of porting the different indexing solutions to a solr/lucene based implementation.

Implementation of advanced image search services

The advanced image search service has to be integrated within the Europeana APIs, consequently this has to be implemented using the technology used in the APIs stack (i.e. using Spring and Solr technologies) and must provide a REST interface powered by Springfox/Swagger libraries. The service will exploit the scalability and performance advantages built in Solr system.
The application architecture that shows the main building blocks of the Advanced Image Search Service and dataflows exchanged with the dataset aggregation tools, administration console and Frontend applications is presented in Figure 1.

[6] http://www.semanticmetadata.net/wiki/

Application Architecture

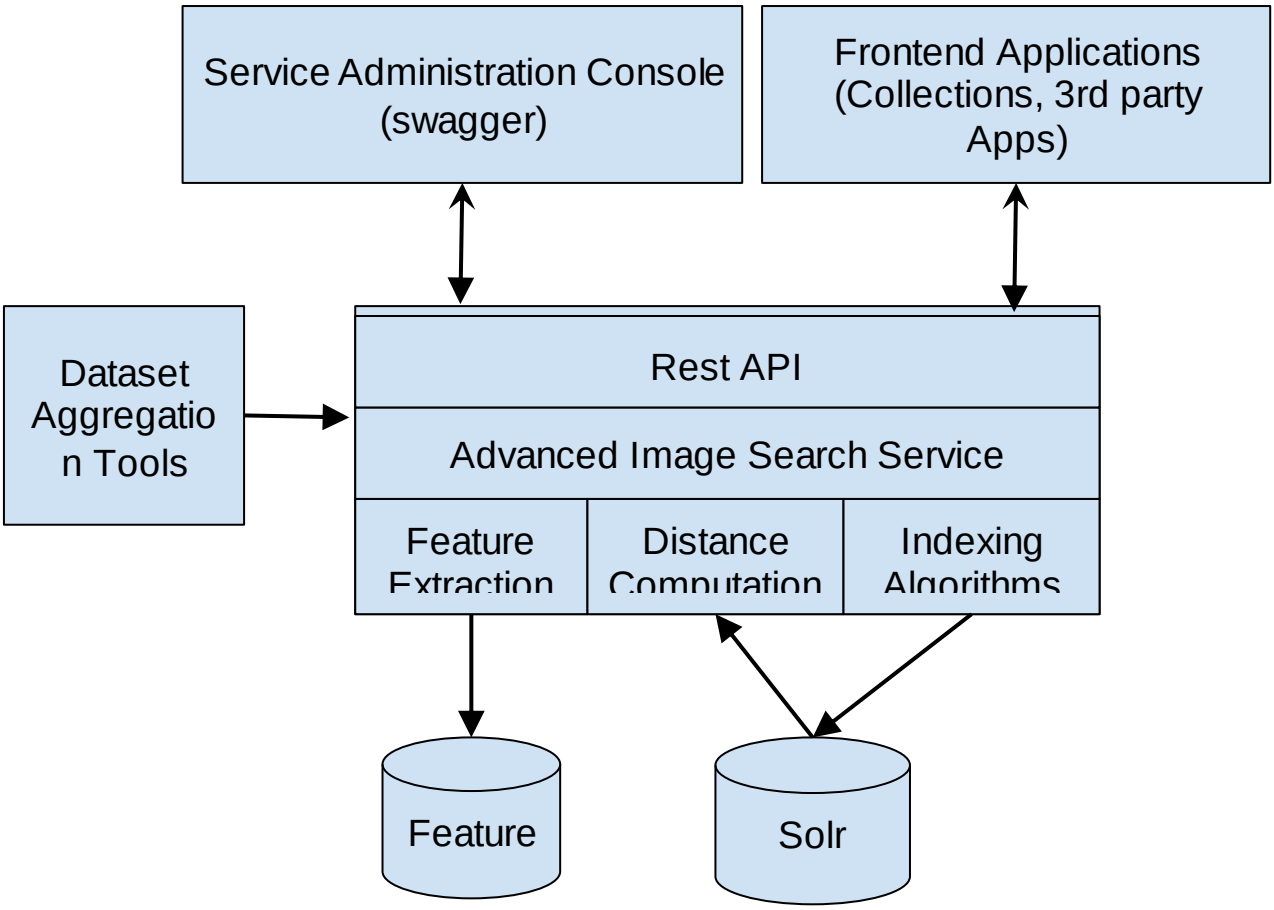

Figure 1 Application Architecture

In order to implement these goals the following activities need to be carried out:

- Dataset Aggregation. By using the selection criteria and examples provided in Section Dataset, the image content that will be served by the service will be aggregated in a dataset. This Dataset must include the ids of Europeana records and the Urls made available for downloading the content.
- Migration of existing image search service. The image search services developed within the Europeana Creative project uses an index based on Lucene. In order to ensure a better scalability and management of the image index, the implementation of this algorithm will be ported to a solr based solution, possibly by reusing the Solr plugin solution implemented in Lire. This solution will be used as baseline for the experimental evaluation.
- Implementation of the image search API. The service must implement a Restful API using the same technologies as the other Europeana technologies (i.e. based on spring mvc and swagger). Additionally the metadata fields used for faceting and filtering functionality needs to be stored in the image index as well.
- Experimental evaluation. An experimental evaluation will be carried out to assess the performance of state of the art algorithms when applied on the selected dataset (e.g.

CDVS, Amato, Pastec). One of the goals of the comparative evaluation is to evaluate the best performing algorithms for different categories of content (i.e. tailored by subject categories, and technical properties like grayscale vs color images). Consequently it will be evaluated if an automatic algorithms selection at runtime may enhance the user experience.

- Migration of novel image search algorithms. It is anticipated that different algorithms will provide better accuracy for particular search scenarios or content categories. Consequently, a multi-modal search approach will better serve the user scenarios defined in Section Goals & Requirements. The novel algorithms need to be migrated on solr based implementations in order to be integrated with the baseline solution and with the filtering/faceting functionality.
- Integration in Europeana Collections. The graphical user elements and controls for integrating the functionality of the advanced search in the Europeana Collections need to be created for supporting the user stories presented in Section User stories.

Activity Plan

Activity	Short Description	Due date
Dataset Aggregation	Selection of image content used by the search service	M7
Migration of existing image search service	Implementation of existing image search service using solr (including feature extraction)	M8
Implementation of the image search API	Implementation of Restful API for existing image search service (Search API)	M8
Filtering and faceting	Implementing filtering and faceting functionality in the search API	M9
Experimental evaluation	Evaluate effectiveness of novel search algorithms on the given dataset	M10
Migration of novel image search algorithms	Implement novel algorithms using solr based solution (including feature extraction)	M12
Integration in Europeana Collections	Develop GUI components for integrating image search in Europeana Collections	M12
Final version of the search	The final deployment of the	M13

API	image search service.	

References

[Amato 2011] Giuseppe Amato, Paolo Bolettieri, Fabrizio Falchi, Claudio Gennaro, and Fausto Rabitti, “Combining local and global visual feature similarity using a text search engine” in CBMI, 2011, pp. 49–54.

[Gordea 2015] Sergiu Gordea and Michela Vignoli, “CultureCam: Visual exploration of cultural heritage content by professional designers” in 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2015, pp. 1-6

[Sikora 2001] T. Sikora, “The mpeg-7 visual standard for content description-an overview,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 11, no. 6, pp. 696–702, Jun 2001.

[Duan 2015] L. Duan, T. Huang, W. Gao, "Overview of the MPEG CDVS standard", *Proc. IEEE Data Compression Conf.*, pp. 323-332, 2015-Apr.

[Bianco 2015] S. Bianco, D. Mazzini, D.P. Pau, R. Schettini, Local detectors and compact descriptors for visual search: A quantitative comparison, Digital Signal Processing, Volume 44, September 2015, Pages 1-13

[Lux 2008] Lux Mathias, Savvas A. Chatzichristofis. Lire: Lucene Image Retrieval – An Extensible Java CBIR Library. In proceedings of the 16th ACM International Conference on Multimedia, pp. 1085-1088, Vancouver, Canada, 2008

[Rublee 2011] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. 2011. ORB: An efficient alternative to SIFT or SURF. In *Proceedings of the 2011 International Conference on Computer Vision* (ICCV '11). IEEE Computer Society, Washington, DC, USA, 2564-2571.

[Alcantarilla 2012] Alcantarilla, P. F., Bartoli, A., Davison, A. J.: KAZE features. Proceedings of the 12th European conference on Computer Vision, 214{227 (2012)