

Contents

1. Introduction	1
1.1. Motivation	1
1.2. Objectives	2
1.2.1. General Objective	2
1.2.2. Specific Objectives	2
1.3. Methodology	2
1.4. Structure of this Work	3
2. Theoretical Framework	4
2.1. Image Analysis	4
2.2. Visual Descriptors	5
2.2.1. Gray Histogram Descriptor	6
2.2.2. Color Layout Descriptor	7
2.2.3. Histogram of Oriented Gradients	8
2.2.4. DeCAF7	9
2.3. Content-based Image Retrieval	9
2.3.1. Similarity Search	9
2.3.2. Indexing	10
2.4. The Semantic Web and the Multimedia Content	13
2.4.1. The Web of Data	13
2.4.2. Multimedia on the Web of Data	19
2.4.3. Similarity Search with SPARQL	21
3. System Overview	24
3.1. Images of IMGPEdia	25
3.2. Visual Descriptors of IMGPEdia	26
3.3. IMGPEdia Knowledge-base	27
4. Image Processing	29
4.1. Reference Implementations for Visual Descriptors	29
4.2. Descriptor Extraction Process	30
4.3. Similarity Search	31
4.4. Similarity Search Results	34
5. Data Modelling and Publication	38
5.1. IMGPEdia Ontology	38
5.2. Knowledge-base Structure	40

5.3. Links to DBPEDIA Resources	42
5.4. SPARQL Endpoint	43
6. Results	45
7. Conclusions	53
7.1. Relevance and Contribution	54
7.2. Future Work	55
Bibliography	57
A. IMGPEdia ontology in Turtle	61

List of Tables

2.1.	5 Star Scheme for Linked Open Data Evaluation	19
3.1.	Count of encodings of the images from WIKIMEDIA COMMONS.	25
4.1.	Descriptor calculation benchmark.	31
6.1.	Number of resources of IMG _{PEDIA}	45
6.2.	Results of Listing 6.1 query	46
6.3.	Results of Listing 6.2 query	47
6.4.	Results of Listing 6.3 query	48
6.5.	Results of Listing 6.5 query	50
6.6.	Results of Listing 6.4 query	52

List of Figures

2.1. Color spaces.	5
2.2. GHD calculation process	6
2.3. Icon generation example	7
2.4. HOG computation steps	8
2.5. R-tree example	10
2.6. K-D tree space division and tree representation	11
2.7. Boundaries for distance using pivots	11
2.8. Exclusion area in a pivot-based search.	12
2.9. Example of bad search results on the Web.	13
2.10. Semantic Web Stack	14
2.11. Example of an RDF graph.	15
2.12. Ontology example.	17
2.13. Linking Open Data cloud diagram	20
3.1. WIKIMEDIA COMMONS image sample.	25
3.2. The data to be merged in the IMGPEdia ontology.	27
4.1. Pipeline of computation of the nearest neighbors using FLANN	33
4.2. Distribution of the in-degree of the images on the similarity graph of IMGPEdia.	34
4.3. 10-NN of Hopsten Marktplatz picture using HOG descriptor.	35
4.4. 10-NN of Hopsten Marktplatz picture using GHD.	36
4.5. 5-NN of Hopsten Marktplatz picture using CLD.	36
4.6. Average color icon for Hopsten Marktplatz and its first two nearest neighbors	37
4.7. 10-NN of a picture of a building in Dreifertstraße street using CLD	37
5.1. IMGPEdia ontology core.	40
5.2. IMGPEdia ontology for similarity links.	40
5.3. Virtuoso endpoint query interface	44
6.1. Results of Listing 6.1 query	46
6.2. Results of Listing 6.2 query	47
6.3. Some results of Listing 6.4 query	49
6.4. Results of Listing 6.5 query	51