ONLINE SKETCH-BASED IMAGE RETRIEVAL USING KEYSHAPE MINING OF GEOMETRICAL OBJECTS

HUDA ABDULAALI ABDULBAQI

A thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

February 2017
To my God, Allah azzawajalla

Then to my beloved mother, spirit of my beloved father, my husband Adel Muhsin, whom without their love and support this research would have never been completed.
ACKNOWLEDGEMENT

Thanks for Allah SWT for everything I was able to achieve and for everything I tried but I was not able to achieve.

To my brilliant supervisor Professor Dr. Ghazali Bin Sulong, you play important role in my studying life, you are truly the supervisor that concern and observe your student. I am greatly appreciative of him for his support and guidance, most importantly, for providing me the freedom to pursue my ideas and find my own path in research. Also, I have gained a wealth of experience and knowledge working under your supervision, which will always be my delight to share along my life’s journey. Thanks also to Assoc. Prof. Dr. Siti Zaiton Mohd Hashim, Assoc. Prof. Dr. Soukaena Hassan Hashim, for their continuous encouragement and support.

Finally and more importantly, I am lucky to have a great family especially my lovely daughter JWD. The love and the support from my brothers and lovely sisters for their support and encourage to overcome the hurdles in my PhD study. Also I would like to express spatial thanks to Prof. Dr. Alaa Hussein Al-Hamami and my friend Assmaa Sadiq for their help when it was most required.
Online image retrieval has become an active information-sharing due to the massive use of the Internet. The key challenging problems are the semantic gap between the low-level visual features and high-semantic perception and interpretation, due to understating complexity of images and the hand-drawn query input representation which is not a regular input in addition to the huge amount of web images. Besides, the state-of-art research is highly desired to combine multiple types of different feature representations to close the semantic gap. This study developed a new schema to retrieve images directly from the web repository. It comprises three major phases. Firstly a new online input representation based on pixel mining to detect sketch shape features and correlate them with the semantic sketch objects meaning was designed. Secondly, training process was developed to obtain common templates using Singular Value Decomposition (SVD) technique to detect common sketch template. The outcome of this step is a sketch of variety templates dictionary. Lastly, the retrieval phase matched and compared the sketch with image repository using metadata annotation to retrieve the most relevant images. The sequence of processes in this schema converts the drawn input sketch to a string form which contains the sketch object elements. Then, the string is matched with the templates dictionary to specify the sketch metadata name. This selected name will be sent to a web repository to match and retrieve the relevant images. A series of experiments was conducted to evaluate the performance of the schema against the state of the art found in literature using the same datasets comprising one million images from FlickerIm and 0.2 million images from ImageNet. There was a significant retrieval in all cases of 100% precision for the first five retrieved images whereas the state of the art only achieved 88.8%. The schema has addressed many low features obstacles to retrieve more accurate images such as imperfect sketches, rotation, transpose and scaling. The schema has solved all these problems by using a high level semantic to retrieve accurate images from large databases and the web.
Dapatan semula imej dalam talian menjadi satu perkongsian maklumat aktif disebabkan penggunaan internet yang berleluasa. Cabaran masalah utama adalah jurang semantik antara ciri-ciri visual peringkat rendah dan persepsi dan tafsiran semantik tinggi, kerana memperkecilkan kerumitan imej dan perwakilan input pertanyaan lukisan tangan yang bukan merupakan input kerap di samping jumlah besar imej web. Selain itu, penyelidikan terkini amat dikehendaki untuk menggabungkan pelbagai jenis perwakilan ciri yang berbeza untuk merapatkan jurang semantik. Kajian ini membangunkan satu skema baharu untuk mendapatkan semula imej secara langsung daripada repositori web. Ia terdiri daripada tiga fasa utama. Pertama, perwakilan input dalam talian baharu berdasarkan perlombongan piksel untuk mengesan ciri bentuk lakaran dan mengaitkannya dengan maksud objek lakaran semantik telah direka bentuk. Kedua, proses latihan dibangunkan untuk mendapatkan templat biasa menggunakan teknik Penguraian Nilai Singular (SVD) untuk mengesan templat lakaran biasa. Hasil daripada langkah ini adalah lakaran kamus pelbagai templat. Akhir sekali, fasa dapatan semula memadankan dan membandingkan lakaran dengan repositori imej menggunakan anotasi metadata untuk mendapatkan semula imej yang paling relevan. Jujukan proses dalam skema ini menukarkan lakaran input yang dilukis dengan bentuk rentetan yang mengandungi unsur-unsur objek lakaran. Kemudian, rentetan dipadankan dengan kamus templat untuk menentukan nama metadata lakaran. Nama yang dipilih ini akan dihantar kepada repositori web untuk padanan dan mendapatkan semula imej yang berkaitan. SATU siri eksperimen dijalankan untuk menilai prestasi skema tersebut terhadap penemuan terkini dalam kesusastraan menggunakan set data sama yang terdiri daripada satu juta imej daripada Flickerlm dan 0.2 juta imej daripada ImageNet. Terdapat dapatan semula yang ketara dalam semua kes dengan 100% ketepatan untuk lima imej pertama yang didapatan semula manakala yang terkini hanya mencapai 88.8%. Skema tersebut telah menyelesaikan banyak halangan ciri-ciri rendah untuk mendapatkan semula lebih banyak imej yang lebih tepat seperti lakaran yang tidak sempurna, putaran, transposisi, dan penskalaan. Skema ini telah menyelesaikan semua masalah tersebut dengan menggunakan semantik tahap tinggi untuk mendapatkan semula imej yang tepat daripada pangkalan data dan web yang besar.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF ALGORITHMS</td>
<td></td>
<td>xix</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td></td>
<td>xvii</td>
</tr>
<tr>
<td>LIST OF SYMBOLS</td>
<td></td>
<td>xviii</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td></td>
<td>xix</td>
</tr>
</tbody>
</table>

1 INTRODUCTION 1

1.1 Introduction 1
1.2 Problem Background 4
1.3 Problem Statements 7
1.4 Research Goal 9
1.5 Research Objective 9
1.6 Research Scope 9
1.7 Significance of the Study 10
1.8 Thesis Organization 11
2 LITERATURE REVIEW

2.1 Introduction 13

2.2 Image Retrieval 13
  2.2.1 Classification 14
  2.2.2 Query Modality 14
  2.2.3 Architecture of an Image Retrieval System 15

2.3 Overview of Content Based Image Retrieval 16
  2.3.1 Low-level Features in CBIR 17
    2.3.1.1 Color-based Image Retrieval Methods 17
    2.3.1.2 Texture-based Image Retrieval Methods 18
    2.3.1.3 Shape-based Image Retrieval Methods 18
  2.3.2 Evaluation Techniques 19

2.4 Sketch Based Image Retrieval (SBIR) 22
  2.4.1 Bag-of-Words: An Approach Inspired by Text Retrieval 22
  2.4.2 Feature Extraction 24
    2.4.2.1 Geometrical Keyshape Method 25
    2.4.2.2 Appearance Features 28
  2.4.3 Efficient Index-based 31
  2.4.4 Matching Methods 34

2.5 Mining Technique 37

2.6 SVD Technique 39

2.7 Hybrid based Image Retrieval 39
  2.7.1 Text Annotation or Key Word Based 40
  2.7.2 Label Sketch with Color Based 43

2.8 Gap Analysis of related study 47

2.9 Summary 48
3 METHODOLOGY

3.1 Introduction 49
3.2 Phase A: Problem Formulation 52
3.3 Phase B: Design and Development 52
3.4 Phase C: Implementation 53
3.5 Phase D: Evaluation 56
   3.5.1 Validation 56
3.6 Benchmarking 57
3.7 Software Requirements 58
3.8 Dataset 60
   3.8.1 EitzSKETCH Dataset 61
   3.8.2 Flickr15k Dataset Sketches 62
   3.8.3 MIR Flickr Dataset 64
   3.8.4 ImageNet Dataset 65
3.9 Summary 66

4 ONLINE SKETCH-BASED IMAGE RETRIEVAL 67

4.1 Introduction 67
4.2 Sketch Board Mining 70
   4.2.1 Online sketch board 70
   4.2.2 Proposed Geometric Keyshape 72
   4.2.3 Sketch Interpretation 86
    4.2.3.1 Imperfect Sketch status 91
    4.2.3.2 Scaling status 92
    4.2.3.3 Transpose status 94
4.3 Development a Dictionary of Sketch Templates 96
   4.3.1 Training Process 97
    4.3.1.1 Create Semantic Space of Template Sketch 98
    4.3.1.2 Singular Value Decomposition (SVD) 99
    4.3.1.3 Rotation status 107
   4.3.2 Creating a Semantic Dictionary of Templates 111
<table>
<thead>
<tr>
<th>4.4</th>
<th>Querying by Sketch</th>
<th>115</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1</td>
<td>Drawing Question Sketch Online</td>
<td>116</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Comparing Sketch with Template Dictionary</td>
<td>117</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Matching with Google image Metadata</td>
<td>123</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Retrieving Image</td>
<td>125</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5</th>
<th>EXPERIMENTAL RESULTS AND DISCUSSION</th>
<th>130</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>130</td>
</tr>
<tr>
<td>5.2</td>
<td>Performance Evaluation of Sketch Board Scheme</td>
<td>131</td>
</tr>
<tr>
<td>5.3</td>
<td>Result of Training Stage</td>
<td>134</td>
</tr>
<tr>
<td>5.4</td>
<td>Performance Evaluation of online Retrieval from Google Images Repository</td>
<td>136</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Performance Evaluation for Variety of Sketches</td>
<td>137</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Robustness against Transpose</td>
<td>141</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Robustness against Scaling</td>
<td>145</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Robustness against Rotation</td>
<td>148</td>
</tr>
<tr>
<td>5.4.5</td>
<td>Robustness against Imperfect</td>
<td>150</td>
</tr>
<tr>
<td>5.4.6</td>
<td>Summary of Discussion</td>
<td>153</td>
</tr>
<tr>
<td>5.5</td>
<td>Performance Evaluation versus State-of-the-art SBIR Techniques</td>
<td>156</td>
</tr>
<tr>
<td>5.5.1</td>
<td>Quantitative evaluation</td>
<td>156</td>
</tr>
<tr>
<td>5.6</td>
<td>Summary</td>
<td>158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6</th>
<th>CONCLUSIONS AND FURTHER WORKS</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>160</td>
</tr>
<tr>
<td>6.2</td>
<td>Contributions</td>
<td>162</td>
</tr>
<tr>
<td>6.3</td>
<td>Future works</td>
<td>163</td>
</tr>
</tbody>
</table>

REFERENCES | 165 |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>178</td>
</tr>
<tr>
<td>Appendix B</td>
<td>181</td>
</tr>
<tr>
<td>Appendix C</td>
<td>187</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary of existing SBIR methods</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of implementation roadmap</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>The state of the art (large data set contains 1.2 million images)</td>
<td>58</td>
</tr>
<tr>
<td>3.3</td>
<td>Topics and subtopics selected for full annotation</td>
<td>64</td>
</tr>
<tr>
<td>3.4</td>
<td>General tree and sub tree</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>Some examples of start and end points for each type of straight line</td>
<td>80</td>
</tr>
<tr>
<td>4.2</td>
<td>Matrix space of drawing in Figure 4.23(a)</td>
<td>108</td>
</tr>
<tr>
<td>4.3</td>
<td>Matrices space of figure 4.23(b)</td>
<td>108</td>
</tr>
<tr>
<td>5.1</td>
<td>Precision of Top 20 images retrieved from Google images repository</td>
<td>155</td>
</tr>
<tr>
<td>5.2</td>
<td>Comparison of % of true positive images retrievals at different ranks P@k of Top (5, 10, 25, 50, 100, and 250)</td>
<td>157</td>
</tr>
<tr>
<td>FIGURE NO</td>
<td>TITLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>What will be the appropriate text query for the above search? Scenario? A sketch query is more expressive in this case.</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>A subset of the sketch queries of EitzSBIR dataset (Eitz et al., 2011)</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Examples of top 10 retrieval results with red Cross unsuccessful retrieval and green circle indicate the object boundaries are surrounded by noisy edges (Wang et al., 2015)</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Various feature extraction methods</td>
<td>25</td>
</tr>
<tr>
<td>2.2</td>
<td>Examples of probabilistic point (McNeill and Vijayakumar, 2006)</td>
<td>38</td>
</tr>
<tr>
<td>2.3</td>
<td>Example of adding text to sketches as a search query (Hu and Collomosse, 2013)</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>Semantic sketch for one specific image to be searched (Liu et al., 2010)</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>Block diagram of research design</td>
<td>51</td>
</tr>
<tr>
<td>3.2</td>
<td>A diagram representation of the stages involved in the proposed SBIR approach</td>
<td>53</td>
</tr>
<tr>
<td>3.3</td>
<td>Execution of the applet in Client Side Machine</td>
<td>59</td>
</tr>
<tr>
<td>3.4</td>
<td>Execution of Java Servlets at Server Side and return of results to the client side</td>
<td>60</td>
</tr>
</tbody>
</table>
3.5 An excerpt of the sketch queries of EitzSKETCH dataset (Eitz et al., 2011)

3.6 Sample of landmark from Flicker 15 dataset sketches (Hu and Collomosse, 2013)

3.7 Some sketch queries from ETIZ Sketch (Hu and Collomosse, 2013)

4.1 General framework of the proposed methodology

4.2 The sketch board editor

4.3 A block diagram of data acquisitions by the sketch board mining during a sketch drawing session

4.4 The eleven element types (a) horizontal line, (b) vertical line, (c) northward line, (d) southward line, (e) northward curve, (f) southward curve, (g) eastward curve, (h) westward curve, (i) circle, (j) x-axis ellipse, (k) y-axis ellipse

4.5 General frameworks to determine geometrical keyshape

4.6 The important points of each drawn element

4.7 Various types of straight lines, (a) vertical line, (b) horizontal line, (c) northward-east line, (d) northward-west line, (e) southward-east line and (f) southward-west line

4.8 Four types of curve (a) westward curve, (b) eastward curve, (c) southward curve, and (d) northward curve

4.9 Three types of Keyshape (a) circle, (b) x-axis ellipse, (c) y-axis ellipse

4.10 Angles of elements

4.11 Interpretation of sketch element from spatial relationships to concept

4.12 Example of the sketch watch to determine spatial relationship correlated with direction

4.13 Specifying the direction of each element based on an imaginary correlated with centroid point
4.14 Example of Six teapot sketches (a) different types of
drawing sketch teapot and (b) imperfect sketch

4.15 (a) sketch of a small teapot and (b) the same object
drawn but larger size

4.16 Example of Transpose (a) represent the location
northward-west (b) sketch location southward-east

4.17 The steps of developing sketch template dictionary

4.18 Example of five sunflower sketches drawn

4.19 The matrix of eigenvector

4.20 Bicycle sketch matrix

4.21 The U matrix of SVD decomposition

4.22 The S matrix of SVD decomposition

4.23 The V matrix of SVD decomposition

4.24 The result of elimination

4.25 (a) sketch of sailboat (b) the Reflection of sketch

4.26 Example of SVD for sketch (a) in Figure 4.25

4.27 Example of SVD for sketch (a) in Figure 4.25

4.28 The Singular values of components for both, (a)
original sketch and (b) mirror sketch

4.29 Example of templates related to two sketches, moon, and pyramid

4.30 Dictionary Scheme

4.31 Summary of process of how to retrieve online image
sketches

4.32 The life cycle of sketch query

4.33 Sequence of operations for similarity matching

4.34 Example of matching query template with dictionary

4.35 Process of send metadata to URL Google by (WS)

4.36 Example of sunflower image metadata annotation

4.37 Example of metadata annotation of sunflower images

4.38 Example of matching query with annotation and
retrieved images

5.1 The GUI with Board Mining
5.2 Example of Drawing glasses sketch (a) Drawing space (b) Board mining

5.3 Utility part

5.4 Sample of Training Iteration of Flicke Object Sketch

5.5 Sample of Training Iteration of EitzSKETCH Object Sketch

5.6 Sydney Opera sketch images retrieved

5.7 Sydney Opera sketch images retrieved

5.8 KLCC sketch images retrieved

5.9 Burj Alarab sketch images retrieve

5.10 Sailboat sketch images retrieved

5.11 Bicycle sketch images retrieved

5.12 Sunflower sketch images retrieved

5.13 Scissors sketch images retrieve

5.14 Sketch pair (a) and (b) exhibiting transpose with the symmetry concept retrieved

5.15 Sketch pair (a) and (b) exhibiting transpose with the symmetry concept retrieved

5.16 Sketch pair (a) and (b) exhibiting scaling symmetry concept retrieved

5.17 Sketch pair (a) and (b) exhibiting scaling symmetry concept retrieved

5.18 Sketch pair (a) and (b) exhibiting with the reflection symmetry concept retrieved

5.19 Sketch pair (a) and (b) exhibiting with the reflection symmetry concept retrieved

5.20 Sample1 of drawing pyramid Sketch

5.21 Sample 2 of drawing a pyramid

5.22 Sample 3 of drawing a pyramid

5.23 Retrieval performance of the proposed method in comparison to the state-of-art in the case of Best (B) image retrieved
# LIST OF ALGORITHMS

<table>
<thead>
<tr>
<th>ALGORITHM NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Determination of line or curve keyshapes</td>
<td>78</td>
</tr>
<tr>
<td>4.2</td>
<td>Determination of various line type keyshapes</td>
<td>79</td>
</tr>
<tr>
<td>4.3</td>
<td>Computation of curve types</td>
<td>82</td>
</tr>
<tr>
<td>4.4</td>
<td>Determine circle or ellipse</td>
<td>84</td>
</tr>
<tr>
<td>4.5</td>
<td>Identify spatial relationship between keyshape type</td>
<td>90</td>
</tr>
<tr>
<td>4.6</td>
<td>Proposed SVD-based training process</td>
<td>101</td>
</tr>
<tr>
<td>4.7</td>
<td>Similarity matching based on Dot product</td>
<td>123</td>
</tr>
<tr>
<td>4.8</td>
<td>Image Annotation and retrieved images</td>
<td>127</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>AROP</td>
<td>Angular Radial Orientation Partitioning</td>
</tr>
<tr>
<td>ARP</td>
<td>Angular Radial Partitioning</td>
</tr>
<tr>
<td>BoVW</td>
<td>Bag Of Visual Words</td>
</tr>
<tr>
<td>BoW</td>
<td>Bag-Of-Words</td>
</tr>
<tr>
<td>CBIR</td>
<td>Content Based Image Retrieval</td>
</tr>
<tr>
<td>COM</td>
<td>Click On Mouse</td>
</tr>
<tr>
<td>CS</td>
<td>Chain Score</td>
</tr>
<tr>
<td>DT</td>
<td>Distance Transform</td>
</tr>
<tr>
<td>EHD</td>
<td>Edge Histogram Descriptor</td>
</tr>
<tr>
<td>EI</td>
<td>Edge Indexing</td>
</tr>
<tr>
<td>EM</td>
<td>Expectation–Maximization algorithm</td>
</tr>
<tr>
<td>Fm</td>
<td>F-measure</td>
</tr>
<tr>
<td>Fp</td>
<td>False Positive</td>
</tr>
<tr>
<td>GC</td>
<td>Geometric Consistency</td>
</tr>
<tr>
<td>GF-HOG</td>
<td>Gradient Field HOG</td>
</tr>
<tr>
<td>HELO</td>
<td>Histogram of Edge Local Orientations</td>
</tr>
<tr>
<td>HLR</td>
<td>Histogram of Line Relationsh</td>
</tr>
<tr>
<td>HOG</td>
<td>Histogram of Oriented Gradients</td>
</tr>
<tr>
<td>HTD</td>
<td>Homogeneous Texture Descriptor</td>
</tr>
<tr>
<td>IR</td>
<td>Image retrieval</td>
</tr>
<tr>
<td>MAP</td>
<td>Mean Average Precision</td>
</tr>
<tr>
<td>OCM</td>
<td>Oriented Chamfer Matching</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>QBIC</td>
<td>Query By Image Content</td>
</tr>
<tr>
<td>RGB</td>
<td>Read, Green, Blue</td>
</tr>
<tr>
<td>ROI</td>
<td>Region Of Interest</td>
</tr>
<tr>
<td>SBIR</td>
<td>Sketch Based Image Retrieval</td>
</tr>
<tr>
<td>S-HELO</td>
<td>Soft-Histogram of Edge Local Orientations</td>
</tr>
<tr>
<td>SIFT</td>
<td>Scale Invariant Feature Transform</td>
</tr>
<tr>
<td>STD</td>
<td>Structure Tensor Descriptor</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>TBD</td>
<td>Texture Browsing Descriptor</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>VSW</td>
<td>Visual Saliency Weighing</td>
</tr>
<tr>
<td>WS</td>
<td>WebSockets</td>
</tr>
<tr>
<td>WWW</td>
<td>World Wide Web</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS

P - Precision
R - Recall
n - total number of image retrieval
m - total number of relevant images in database
P - Precision
Pq - precision of each query
Q - Number of queries in the benchmark
B - Best
W - Worst
A - Average
x₁,y₁ - start point of element
xₙ,yₙ - End point of element
xᵢ,yᵢ - Chin point
xₘ,yₘ - Mid-point
xₑ, yₑ - Imaginary centre-point
D₁ - distance between imaginary centre point and mid-point
D₂ - distance between imaginary centre point and one of either start or end point
ΔX - The difference in x-axis
ΔY - The difference in y-axis
θ - angle
X̄ - Centroid x
Ȳ - Centroid y
λ₁ - eigenvalue
V - vector which represents the sketch components
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Solving Sketch variances</td>
<td>178</td>
</tr>
<tr>
<td>B</td>
<td>Sketch Based Image Retrieval Results</td>
<td>181</td>
</tr>
<tr>
<td>C</td>
<td>Visual Results</td>
<td>187</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, the explosive use of the Internet has supported the study of information retrieval systems, which has become a very active research area among computer scientists. The classical approach for information retrieval is based on textual information, where a user writes a textual query composed of target concepts and the retrieval system returns a ranking with the most relevant documents stored in a database or on the web. From the early days of the Internet to present, there has been a great effort in the research community to develop efficient algorithms capable of searching and indexing text documents (Amanda and Jansen, 2004; Rondo and Manuel, 2013). The result of this effort is reflected in the fact that the present day usage of search engines is found everywhere with results characterized by efficiency and effectiveness. Because of the increase of media devices, the web content is no longer based only on textual documents, but also on multimedia content like images, videos, etc. On-going research on information retrieval is focused on multimedia information retrieval, and in particular, image retrieval systems have attracted a great number of researchers coming from different communities like computer vision, multimedia retrieval, data mining, among others, leading to a vast number of related publications. This is due to the fact that images are ubiquitous and easy to capture by
users (Lee and Chen, 2009). An image retrieval system returns a set of images which are ranked in a certain order under a similarity function in response to a query given by a user.

Content based image retrieval (CBIR) requires another image as query as opposed to the concept based query (Singh and Sinha, 2012; Datta and Wang, 2005). The retrieval system has to extract some relevant information from the input image and from those stored in the database (test images). The extracted information is commonly represented as vectors of numerical features that will be compared by using a similarity function. Traditionally, a CBIR system receives a regular image showing colour and texture as the input. In this way, in order to start the search, the user must have an example image resembling what the user is looking for.

Although content based image retrieval seems to overcome the aforementioned drawbacks belonging to the concept of image retrieval, the fact of containing a regular image as query can represent a serious problem. The image retrieval system used commonly by many people because they do not depend on the desired image, thereby, having such a query of any image may not be possible, reduce the usability of image retrieval system. Frequently, users are looking for an image without having any related images available. Therefore, they need a natural way to express their query. That is, making a sketch of what the user expects as an answer could overcome the absence of a regular example image. Sketch Based Image retrieval (SBIR), a kind of query method also supported by the fact that emerging touch screen based technology is becoming more popular, allows the user to make a sketch directly on the screen. Consider the example in Figure 1.1; the users have a particular image in their mind that cannot be easily expressed with text. If they attempt to make a text-based image search using the Keyword Mountains, the results will be very generic word and time must be spent browsing the collection for desired images (Bozas et al., 2014).
Figure 1.1 What will be the appropriate text query for the above search? Scenario? A sketch query is more expressive in this case.

In the scenario of Figure 1.1, a sketch query consists of different types of lines and curves and an intuitive way to describe the users' thoughts to the machine. Obviously, a detailed rendition of the query requires artistic skill. A more convenient way, which this work adopts, is to sketch the main feature lines of a shape or scene.

Of course, a sketch may be enriched by adding colour; however we claim that making a sketch only by strokes is the easiest and most natural way for querying. While many exist approaches for detecting sketch-based object of image retrieval using small datasets that involves less than 1 million images (Saavedra and Bustos, 2014; Eitz et al., 2012; Hu et al., 2010), relatively a little work has been done on huge (web)-scale image retrieval that include more than 1 million images (Parui and Mittal, 2014; Zeng et al., 2014; Cao et al., 2011). It represents the natural search development due to the high technology progress in the field of Internet use. The easy way to search for required images from Web is by a sketch that can be drawn using the mouse of a personal computer or the touch screen of a modern mobile phone. In this work, our focus is online sketch based search on web images.
1.2 Problem Background

There are currently billions of web pages available on the Internet using hundreds of millions of images. It is not allowed to access or make use of the information in these huge image collections unless they are organized so as to accept efficient browsing, searching, and then retrieval of all image data. Sketch Based Image Retrieval SBIR has been identified as a vital form of CBIR, and a simple way to manage the user query is using a hand-drawing line-based, a hand drawn sketch consisting of a set of elements that leaked texture, color, and shadow filling. Figure 1.2 depicts examples of sketches.

Figure 1.2  A subset of the sketch queries of EitzSBIR dataset (Eitz et al., 2011)

The researchers applied different techniques to retrieve images based on sketches. The low level features approach represents the most interesting direction. The matching method in this direction focuses on a small-scale dataset which obtain a good result in matching and retrieved images (Saavedra and Bustos, 2014; Eitz et al., 2010; Eitz et al., 2009) by using the low level features with general descriptor features such as Histogram of Oriented Gradients (HOG), Edge Histogram Descriptor (EHD), Scale Invariant Feature Transform (SIFT) and edge detection
methods like canny and Soubel on retrieval, but with high False Positive (Fp) retrieved images in large dataset. In practice, the huge database is in high demand to be certain that the system can always find desired matches for any sketch query. Only a few attempts exist to deal with large dataset or online to retrieve images with high accuracy. In general, the matching techniques used low level feature extraction, edges, salient corners, points or object boundaries. To determine similar images based on certain sketch, a linear scan over all database images are prepared selectively, such as object boundary selection that uses sketch edge description as a tool to reduce the input of noisy edges. Some researchers (Wang et al., 2015) improve this method using spatial coherent constraint that treats the false matches that degrade retrieval performance. The observed results successfully indicate that boundary of the object are surrounded by some noisy edges, but the top 10 retrieved images were not satisfactory, as shown in Figure 1.3.

<table>
<thead>
<tr>
<th>Sketches</th>
<th>Image Retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="example.png" alt="Sketch 1" /></td>
<td><img src="example.png" alt="Image 1" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 2" /></td>
<td><img src="example.png" alt="Image 2" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 3" /></td>
<td><img src="example.png" alt="Image 3" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 4" /></td>
<td><img src="example.png" alt="Image 4" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 5" /></td>
<td><img src="example.png" alt="Image 5" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 6" /></td>
<td><img src="example.png" alt="Image 6" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 7" /></td>
<td><img src="example.png" alt="Image 7" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 8" /></td>
<td><img src="example.png" alt="Image 8" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 9" /></td>
<td><img src="example.png" alt="Image 9" /></td>
</tr>
<tr>
<td><img src="example.png" alt="Sketch 10" /></td>
<td><img src="example.png" alt="Image 10" /></td>
</tr>
</tbody>
</table>

**Figure 1.3** Examples of top 10 retrieval results with red Cross unsuccessful retrieval and green circle indicate the object boundaries are surrounded by noisy edges (Wang et al., 2015)

Most related works use salient contours or salient detection to efficiently achieve retrieval from 4.5 million images with a query sketch by determining the salient part of an image and retrieving images with objects similar to the query. However, determining saliency considers as a hard problem and the result shows better precision with (72%) retrieval accuracy (Li et al., 2014; Zhou et al., 2012). This system filters out a large number of irrelevant images quickly. Saliency
detection using the hierarchical saliency detection method that is robust on images with complex structures, extracts the significant regions of an image by extracting primary object contours using edge detection on a saliency map. Many obstacles affect the local operation such as missing the primary object contours. Therefore, the high-contrast of the small-scale textures and also the complex structures in images lower the performance (Zhang et al., 2016; Hu and Collomosse, 2013).

The weakness in the indexing methods represents another challenge for image retrieval online. To solve this an indexing method is proposed for image retrieval by decomposing the sketch or the image into different spatial regions and then measuring the correlation of the gradients direction in the image and the strokes direction in the sketch by using two descriptors; Tensor and Edge Histogram (Eitz et al., 2009). Cao et al. (2011) and Bozas and Izquierdo (2012) also used an index friendly raw contour based method. This method compares the edges and their directions to proximate sketch-to-image similarity. This comparison uses the edges and their directions at approximately the same location in the sketch and the image after scale normalization.

Both present indexing methods above fail to address the different scale, transpose and rotation changes as the query was not identical to the stored images determinant (Sun et al., 2013) proposed a new system to support query-by-sketch that uses a Compact visual word with sketch to find desired images based on a label with a particular shape in order to build real-time search indexing useful for billions of images. This method improved the image retrieval in terms of object recognition but lacked of position-invariant matching. These methods yield a highly scalable schema, yet the quality of the retrieval results is questionable, as evaluation of popular SBIR databases is not provided.

The **dimensional** representation of sketch reflects one of the barriers in matching consideration with the image retrieval that belong to scale, transformation, and Rotation variations. The joint information technique was used to measure the relative angles among all pairs of sampled points along a contour. Riemenschneider
*et al.* (2011) used descriptors to apply the partial matching mechanism between two such angle-based, with the range of consecutive points and the help of the integral image-based approach. But in general, the computational complexity in dense descriptors can be considered very high for large datasets. In addition, the performed experiments of very small datasets which include less than three hundred images and therefore the applicability of this approach in large scale is again not very promising. Other approaches represented object boundaries as chains of connected segments and the images database are pr-processed to gain such chains that contain a high possibility of containing the object. Parui and Mittal (2014) used two chain extraction strategies to capture the object shape information: a) long chains overlapping from contour segment networks and b) segmented boundaries of object proposals. To make online matching more accurate and fast retrieval, each database image is pre-processed to find sequences of contour segments (chains) that capture adequate shape information that are represented by specific variable length descriptors. This approach achieves percentage of True Positives precision in the top 5 ranking with a retrieved rate of 80.8%.

### 1.3 Problem Statements

Sketch Based Image Retrieval (SBIR) can be considered an important method in image retrieval due to the widespread use of the internet. However, obstacles appear in retrieval accuracy when accessing large-scale databases because of lacking of suitable features, effective indexing, and also verities of geometric shape. Therefore, based on the background of this study, the SBIR still suffers from the following issues.

**First**, poor performance in retrieving images from large database due to low level feature utilization in the edge detection (Canny and Soubel) and boundary methods like contours and salience that bound the object in order to match with
sketched queries (Zhang et al., 2016; Wang et al., 2015; Miguelenaet al., 2014). Furthermore, to date, no work is found on feature determination in real time for image extraction. Hence, this research proposed new feature extraction.

Second, in spite of the accuracy of image retrieval, indexing remains one of the goals of the SBIR that exhibit good performance in large repositories. Only a few studies focus on building an index that can deal with a real-time sketch-based search by building an inverted index structure depending on the edge pixel locations, or compact visual word with sketch in order to build real-time search indexing useful for 2 million images (Sun et al., 2013; Cao et al., 2011). However, all these approaches are based on an essential assumption, which is that the user aims to search image results based on spatial consistency. Therefore, the problem of web-scale sketch-based image retrieval remains very challenging in terms of building a real-time sketch-based image search engine for large-scale and variant database due to the lack of efficient and effective indexing solutions. This work proposed a new method to index the template of each sketch which able to interact both web domain and query.

Third, at present the performance of matching technique is not satisfactory in large scale database, especially in multivariate sketch drawing concepts (scale, transpose, and rotation). Many researchers apply different techniques to retrieve images based on a large database, such as “salient” object descriptor or joint information method, but these techniques demonstrated low value when used with large database due to scale and transform issues (Zhou et al., 2012; Riemenschneide et al., 2011). Also, Parui and Mittal (2014) improved a large scale sketch-based image retrieval method to solve the problem by specifying long sequences of contour segment chains from each image and support them using variable length descriptors based on similarity preservation. However, the retrieved accuracy evaluation as indicated by the true positive (TP) low value achieved a top 25 ranking from 1.5 million images database. This research looks into solving the mentioned issues by proposing two-tier matching approach the first, the similarity in the corresponding sketch vectors and second, the sketch semantic concept correlated with metadata.
1.4 Research Goal

To propose a new online sketch-based image retrieval schema based on mining objects key geometrical shapes with greater accuracy than current state-of-the-art.

1.5 Research Objective

1. To develop an online sketch board schema that captures and interprets the key geometrical shapes of a sketch drawn of a simple object.
2. To develop a dictionary for common sketch templates to represent variation in style of drawing of different age groups.
3. To propose a new interpretation technique based on spatial relationship using (SVD) to tackle special sketches that is exposed to transpose, scaling and rotation.
4. To propose a new two-tier matching approaches that retrieves the drawn sketch's closest matches, from image repositories.

1.6 Research Scope

The scope of this study covers the following:

1. The study utilizes the standard dataset of sketches in order to evaluate performance,
   a. EitzSKETCH dataset (Eitz et al., 2011).
b. Flickr15k dataset (Hu and Collomosse, 2013).

2. For validation of the retrieval images and benchmarking, two datasets are used as a standard large datasets provided by
   a. 1 million Flickr images (Huiskes and Lew, 2008).
   b. 0.2 million images from Imagenet database (Deng et al., 2009).

3. Canvas drawing is used as a tool to input the sketch online.

4. SVD technique is selected to create an effective semantic dictionary.

5. The Drawing sketch collected from four different age groups between 5 to 50 years old and consisting of 50 kids from primary school (5-11 years old), 50 from secondary school (12-18 years old), 50 college students (19-22 years old) and 25 others (30-50 years old).

6. The independent language used in building the system is JDK under the Net Bens 8.1 environment supported by Apache Tomcat 7 Software.

1.7 Significance of the Study

It is hoped that the proposed online sketch-based image retrieval approach will overcome the challenges existing in retrieving images from large scale (web) database. This approach tries to return images similar to a user-provided image query-based sketch.

The existing sketch based image retrieval approaches show some encouraging results. However, these approaches were implemented using offline images retrieved. The motivation of conducting this study is to propose innovation technique to input sketch query and to retrieve images from web repository in Real-time. This work mimics in-depth learning techniques in parallel to the efforts in the Google brain system.

Furthermore, the new approach is currently used in a Laptop or PC under Windows and its ability to be extended and used in Android application to be used in mobile phones. In this way, a SBIR system would provide great benefit to people with hand-
motility issues who cannot interact with a computer, write queries or take pictures to use as queries. Moreover, current approaches on gaze interaction have allowed disabled people to interact with computer to perform any kind of task like Digital library, Crime prevention, photo sharing sites and search engines. The emerging touch-screen based technology that allows any kind of users to make a sketch query drawing directly on the screen

In practical terms, this approach can be used as commerce in the field of Digital school for learning online, which can be used in web applications by allowing users to make a sketch query drawing directly on the screen. The proposed sketch based image retrieval technique can innovate the drawing software to be used by users that limited skill in drawing capability to enhance them learning.

1.8 Thesis Organization

This thesis is organized as follows. The rest of chapter 1 provides a brief description highlighting the aims of each chapter and ends with a short summary. Each chapter in this thesis is written to be self-contained, but there exists cohesion among the thesis structure in all chapters in order to ensure the understanding and free flow of presentation of the thesis content. It should also be put in mind that mathematical definitions and notations are introduced at various positions in the thesis to render consistency and better understanding of the presentation.

Chapter 2 provides an in-depth overview of relevant literatures on Content based image retrieval and focuses on the Sketch-based image retrieval which incorporates an analysis of the existing literature in relation to the study object. The review covers most of the methods of feature extraction used previously by using many methods of edge detection, using many descriptors to find the similarity between the query and the image in datasets. This study also presents the literature
of small dataset and large dataset, as well as the technique used to reduce the false matching.

Chapter 3 presents a clear roadmap of this study to guide the reader to achieve a quick grasp of the detailed research framework, including the details of datasets related to the two sketch datasets and the huge dataset to images retrieved based on the sketch. The layout of the entire research framework, strategies, and procedures is highlighted.

Chapter 4 discusses the detailed design and development of the proposed approaches of online Sketch Based Image Retrieval, which includes sketch board to input sketch online, feature extraction using mining algorithms, interpretation of the sketch, building Dictionary based on SVD technique, matching in conceptual domain and retrieved images from the Google repository.

Chapter 5 provides the experimental result, detailed analysis, and discussion related to explaining the quantitative measurements that are carried out for the performance evaluation and implementation of the use of standard measurement approaches. To do so, a series of experiments were conducted using the online Google repository images, where the experimental result is interpreted and discussed in detail with various types of queries and simple and complex sketches. In addition, the performance of the proposed method was benchmarked against the best up-to-date techniques on sketch based image retrieval found in the literature. Chapter 6 concludes by emphasizing the major contributions, significant findings, and recommended future directions of the present study.
REFERENCES


Chang, R. (2014). *Effective Graph-Based Content--Based Image Retrieval Systems for Large-Scale and Small-Scale Image Databases*. Doctor Philosophy, University Logan.

Chang, Ran, “Effective Graph-Based Content--Based Image Retrieval Systems for Large-Scale and Small-Scale Image Databases” (2013). All Graduate Theses and Dissertations. Paper 2123


in Proceeding of the 2005 IEEE International Conference on Computer Vision (ICCV'05, 1604-610).


Kumar, G., and Bhatia, P. K. (2014,). A detailed review of feature extraction in image processing systems. Proceeding of the 2014 IEEE.Fourth International Conference on Advanced Computing and Communication Technologies. 5-12


International Journal of Multimedia and Ubiquitous Engineering, 4 (2), 115-124


