FELIX STÜRMER

SKETCH-BASED IMAGE RETRIEVAL USING CURVELETS

[July 10, 2012 at 13:02 – classicthesis version 0.1]

SKETCH-BASED IMAGE RETRIEVAL USING CURVELETS

FELIX STÜRMER

An Evaluation of Curvlet-Based Cross-Domain Descriptors for Sketch-Based Image Retrieval

January 2012 – version 0.1

[July 10, 2012 at 13:02 – `classicthesis` version 0.1]

ABSTRACT

Short summary of the contents...

ACKNOWLEDGMENTS

acknowledgments go here...

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

1 INTRODUCTION

1.1 MOTIVATION

Paragraph about increase in visual data, mobile cameras, medicine, etc...

At the core of the research into content-based image retrieval (CBIR) lies the need to be able to access the growing repositories of visual data in a convenient and efficient manner. In this context "convenient" describes the ability for the user to express the query without a complex reformulation of the intent to make it accessible to the query processor. At the same time the computational efficiency becomes more important as the amount of data to search grows. This issue becomes even more critical as the use of mobile, power-limited devices increases across many areas of application, such as autonomous vehicles or handheld augmented reality devices.

Research into text-based information retrieval has brought into existence many statistical methods to query a potentially large body of text using text as the query input. This preserves the close mapping of the intent of the user to the expression of the query and thereby makes the process accessible to users without knowledge about the internal workings of the retrieval system. Providing the means to access a large amount of visual data using a system with similar properties has turned out not to be an easy problem to solve. Using text-based querying for that purpose depends on the ability to reliably label visual data, which would require solving the general object recognition problem first [2]. To avoid that obstacle and to free the retrieval system from the requirement of translating between textual and visual information, many methods to search an image database using visual similarity have been developed.

While the goals of those systems are very similar, they differ considerably in many aspects of the processing pipeline. The query input ranges from example images over drawings to predicate describing color and shape distribution. Similarly, the structure and content of the databases and the means by which the systems query and rank the results vary significantly. This thesis focuses on evaluating a system that uses hand-drawn sketches as inputs to query databases of either full-color images or contour images. The Fast Discrete Curvelet Transform [1] is used to analyse image segments.

1.2 OUTLINE

Chapter 2 presents the structure of the problem and prior solutions. The following Chapter 3 proposes several variations of a particular solution using the Fast Discrete Curvelet Transform [1]. The experimental setup and its results are documented in Chapter 4 and analysed in Chapter 5. In Chapter 6 several possible conclusions are drawn and pointers towards future research are given.

2

BACKGROUND & RELATED WORK

Illustration of basic structure of CBIR using local/global features.

Most approaches can be characterized by looking at three stages in their processing pipeline:

INPUT FORMAT The structure of the input data determines the amount of information available to the subsequent processing steps. Possible preprocessing steps include color space conversion, scaling and edge extraction.

EXTRACTED FEATURES Many algorithms produce a large number of coefficients that can be reduced to a set of feature coefficients using by techniques such as vector quantization or principal component analysis (PCA).

DISTANCE METRIC In order to rank the images according to similarity a metric is used to calculate the distance in feature space between two sets of feature coefficients. The selection of a metric is often closely coupled with the feature extraction algorithm.

2.1 INPUT FORMAT

Complete vs incomplete sketches, intra-/cross-domain

2.2 FEATURES

- bag of features from k-means clustered visual words [video google]
- histogram of oriented gradients [chalechale + refs]

2.3 METRIC

- after ranking using euclidean distance, rank by spatial similarity [video google]
- Earth Mover's distance? [rubnerljcv00]

3 PROPOSED SOLUTION

Proposed solution goes here...

3.1 INPUT FORMAT

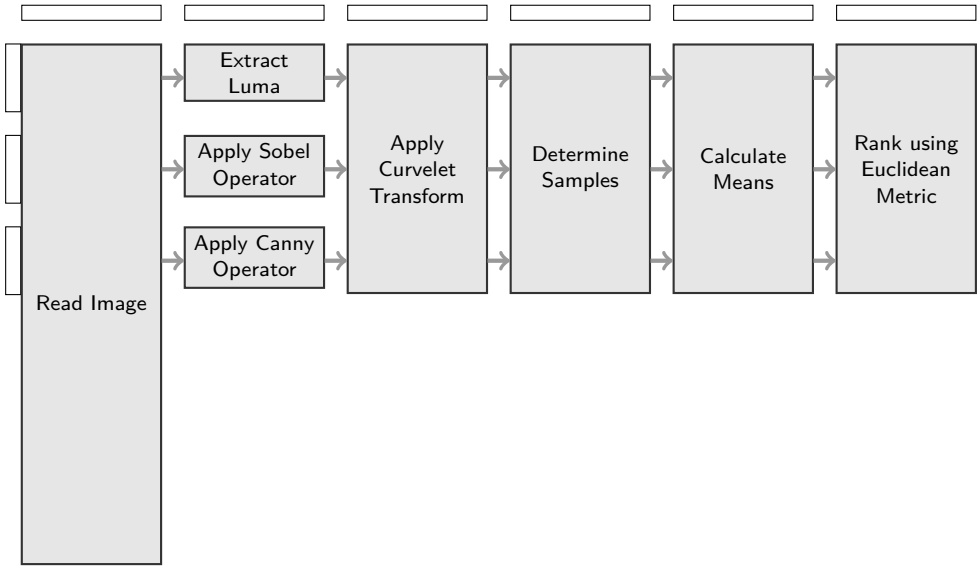
- Luma component (Y') of Y'UV representation
- Gradient magnitude of Sobel operator of luma component
- Canny edge map of luma component
- gPb

3.2 FEATURE EXTRACTION

- Global features: mean and standard deviation
- Local features: visual words via k-means clustering
- great comparison of sampling for k-means clustered vws [nowak06]

3.3 DISTANCE METRIC

- Euclidean Distance
- cosine distance?
- EMD?


Read Image
Extract Luma
Apply Sobel Operator
Apply Canny Operator
Apply Curvelet Transform
Determine Samples
Calculate Means
Rank using Euclidean Metric

4

EXPERIMENTAL RESULTS

Experimental results go here...

5 ANALYSIS

Analysis goes here...

6 CONCLUSION

Conclusion goes here...

BIBLIOGRAPHY

[1] E. Candes et al. “Fast discrete curvelet transforms.” In: *Multiscale modeling and simulation* 5.3 (2006), 861–899.

[2] A. W.M Smeulders et al. “Content-based image retrieval at the end of the early years.” In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22.12 (2000), 1349–1380.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

DECLARATION

Put your declaration here.

Berlin, January 2012

Felix Stürmer