Large Scale Sketch Based Image Retrieval Using Patch Hashing*

Konstantinos Bozas and Ebroul Izquierdo

School of Electronic Engineering and Computer Science
Queen Mary University of London, Mile End Campus, UK, E1 4NS
{k.bozas,ebroul.izquierdo}@eecs.qmul.ac.uk

Abstract. This paper introduces a hashing based framework that facilitates sketch based image retrieval in large image databases. Instead of exporting a single visual descriptor for every image, an overlapping spatial grid is utilised to generate a pool of patches. We rank similarities between a hand drawn sketch and the natural images in a database through a voting process where near duplicate in terms of shape and structure patches arbitrate for the result. Patch similarity is efficiently estimated with a hashing algorithm. A reverse index structure built on the hashing keys ensures the scalability of our scheme and at the same time allows for real time reranking on query updates. Experiments in a publicly available benchmark dataset demonstrate the superiority of our approach.

1 Introduction

The exponential growth of publicly available digital media during the last two decades highlighted the need for efficient and user friendly techniques to index and retrieve images and videos from large scale multimedia databases. Despite the considerable progress of content-based image retrieval (CBIR) [1], where the goal is to return images similar to a user provided image query, most of the multimedia searches are text-based (e.g. Google Images, YouTube). The latter requires user intervention to tag all the available data and has two main drawbacks: (i) it is time consuming and (ii) most importantly user subjective. It is a common belief that images cannot be succinctly communicated based on words; humans would probably use different words to describe a scene based on their cultural background and experience. It follows from the above that in specific scenarios searching images by text query will return frustrating and dubious results to the user. Sketch based image retrieval (SBIR) emerged as a more expressive and interactive way to perform image search; here the query is formed as a hand-drawn sketch of the imaginary picture. Obviously, a shaded rendition of the query requires great artistic skill [2] and most users will refrain from exercising considerable effort and time to draw it. A more intuitive way is

* This work is partially supported by EU project CUBRIK under grant agreement FP7 287704.

G. Bebis et al. (Eds.): ISVC 2012, Part I, LNCS 7431, pp. 210–219, 2012.

Fig. 1. Top 5 results returned from our system in some queries of [5]

to sketch the main feature lines of a shape or scene, a design choice we follow in this paper, and is supported by recent studies where it has been demonstrated that lines are drawn along contours [3] and line drawings can encode certain shapes almost as well as shaded images [4]. Binary sketches can be easily drawn using the mouse in a personal computer or the touch screen of a modern touch screen device.

Finding similarities between a binary drawing and a database of colored pictures taken under arbitrary conditions consists a challenging problem. Images and binary sketches do not share many common modalities. Images contain rich information in domains such as color and texture, while sketches can be described only by their shape and spatial configuration, therefore traditional CBIR methods relying on texture and color cannot be inherited in SBIR. Furthermore, the vast amount of photos uploaded to social media websites signify the crucial role of *scalability*. We need to be able to search large collections of images in reasonable query times as well as update databases and index files with minimum computational cost.

This paper focuses both on scalability and retrieval quality. We tackle these challenges by retrieving *near duplicate in terms of shape* images to a binary sketch query using a hashing technique. Our approach assumes that users are looking for images spatially consistent with their query, for instance if they draw a sunset scene and the sun has been placed at the top right of the canvas that indicates that images displaying the sun in (approximately) the same spot will be preferred, therefore our system is designed to retrieve near duplicate images