

# Instantaneous mental image search with range queries on multiple region descriptors

Julien Fauqueur\*

Signal Processing Group, Department of Engineering, University of Cambridge, UK  
<http://www.eng.cam.ac.uk/~jf330/>

**Abstract.** The Mental Image Search paradigm allows the user to retrieve images which match the target image s/he has in mind without a starting example. We present a novel approach for this paradigm which enables multiple descriptor range-query, which is necessary to match the more or less precise idea of the user's mental image. Sophisticated query can be formulated in a simple and intuitive way through the query interface which maps directly to the region multifeature space using a grid file index. Images are retrieved instantaneously on a large database.

## 1 Introduction

In Content Based Image Retrieval, when a starting image example is not available, the Mental Image Search paradigm allows the user to retrieve images corresponding to the image s/he has in mind. A system implementing this paradigm should satisfy both following challenging aspects:

- *rich user expression* : in the absence of image example, the user must have the possibility to specify a wide range of visual content with the query interface.
- *simple interaction* : the system must be intuitive, simple and fast to allow the user to perform iterative search if necessary.

Existing systems do not meet both points simultaneously. The popular Query-by-Example paradigm requires simple interaction but provides poor user expression. On the other hand, Query by Sketch approach enables rich user expression but usually requires complex drawing interaction.

The new framework we propose attempts to meet both challenges. Indeed, to find a mental image, we allow the user to query images containing regions which have a precise (or vague) combination of visual descriptors. A query requires very

---

\* This work has been carried out with the support of the UK Data and Information Fusion Defence Technology Centre.

few checkbox clicks and matching images are retrieved instantaneously. This is made possible by the use of a multidimensional index based on grid files [2] and a simple query interface which directly maps to the index space and enables easy range-query formulation. The index structure gives directly access to the list of regions (and images) which have any given combination of descriptors. Range query on multiple descriptors is achieved by accessing contiguous buckets in the indexing space. A more detailed version of this paper is available [1].

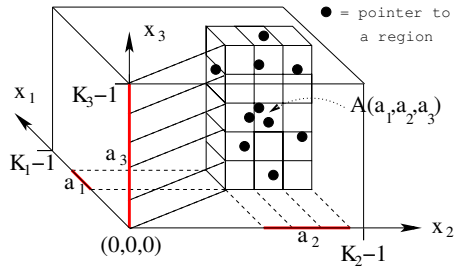
## 2 Building the multi-descriptor index

Images are first segmented into visually salient components which may constitute relevant search keys for the mental image. To perform mental image search, we want to use various region descriptors which are simple and intuitive for the user : average color, position, size and color heterogeneity (as a simple texture descriptor). It defines a  $D$ -dimensional region feature space  $\mathcal{F}$  ( $D = 7$ , here). A uniform quantisation of  $\mathcal{F}$  is achieved by quantising each dimension into  $\{K_1, \dots, K_D\}$  levels. In the new quantised region feature space  $\mathcal{F}' \subset \mathbb{N}^D$ , each region falls into a unique bucket. Each bucket contains regions which have similar features. The indexing structure is a multidimensional array  $A$  such that for all  $(a_1, \dots, a_D)$  in  $\mathcal{F}'$ ,  $A(a_1, \dots, a_D)$  gives the list of regions which fall into bucket  $(a_1, \dots, a_D)$ . Since we only store non-empty buckets, the size of the index file will virtually not be affected by the sparsity of array  $A$ . Thanks to this index, retrieving regions which have a particular set of descriptors  $(a_1, \dots, a_D)$  (e.g. a given color and a given position ...) is as simple as accessing the bucket  $A(a_1, \dots, a_D)$ . Note no loop nor complex computation is involved.

## 3 Multi-descriptor range query

Depending on the query, the user may want to be specific on one visual descriptor, be vague on another one and ignore the others. For instance, to search skin parts, color may be relevant but not size and position. So the user may also want to be able to query from a partially specified set of descriptors, to express that a descriptor is somewhat relevant (“somewhat centered within the image”) or completely irrelevant (“any position within the image”). This means searching the partial range or the full range of any descriptor. We write a user query  $\mathcal{Q}$  as a list of subsets of quantised values for all dimensions:  $\mathcal{Q} = \{S_1, \dots, S_D\}$ , where  $S_i$  is any subset whose elements belong to  $\{0, \dots, K_i - 1\}$ . For example, the range query illustrated in figure 1 is written as:  $\mathcal{Q} = \{\{a_1\}, \{a_2 - 1, a_2, a_2 +$

$1\}, \{0, \dots, K_3-1\}$ . In general, solving a query  $\mathcal{Q}$  comes down to determining the bucket unions:  $\mathcal{R} = \bigcup_{(a_1, \dots, a_D) \in \mathcal{Q}} A(a_1, \dots, a_D)$ , where subsets  $S_i$  are provided by the user through the query interface and where  $A$  is the index array.



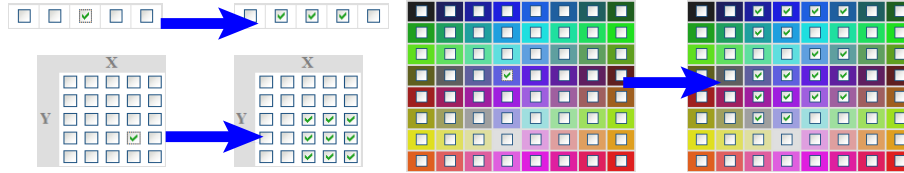
**Fig. 1.** Example of range query in a 3D feature space around bucket  $A(a_1, a_2, a_3)$ . It is decomposed into a point query around  $a_1$ , a 1-nearest-neighbor around  $a_2$ , and a full range query along  $x_3$  axis. Matching regions are represented within the solid lines.

## 4 User interface and range query interaction

In the mental image search context, the query interface is an essential aspect of the system, as it should easily let the user choose the relevant features which best describe the region in the target image. The interface comprises a *feature palette* for each feature which provides a perceptual way to select quantised feature values. Figure 2 shows the palettes for the size (same as for texture), position and color features. Each checkbox corresponds to a quantised feature value in  $\mathcal{F}'$ . Below each palette, three buttons (see [1]) implement the range query: “none”, “more”, “all”. Clicking button “none” deselects all values of the corresponding feature. “all” checks all the values. It is used to perform a full range query on the corresponding feature. Finally, “more” implements an “automatic range query selection” by expanding the query by 1 nearest neighbor in the quantised feature space around the currently selected value. Once a button is pressed, the selection of corresponding checkboxes is automatic and instantaneous. Color space was quantised into 4 bins per dimension, and color, size and texture into 5 per dimension. These quantisation levels were found to be a good trade-off between the feature palette usability and the search precision.

## 5 Results

We used the Corel database of 9995 segmented images from which 50220 regions are detected. The content of this database is heterogeneous : landscapes, portraits, objects, flowers, cars, animals, kitchens, food, etc. To assess the viability of



**Fig. 2.** Three feature palettes and the result of automatic range query selection with the “more”. In each case, one box is checked by the user and further clicking the “more” button makes the automatic checking of the contiguous neighbor in the feature space.

our approach various mental image search scenarios were performed : “mediterranean skies”, “british skies”, “sea/swimming pool”, “white backgrounds”, “fireworks”. Figure 3 shows the top results for the “objects on white background” along with the check boxes for the corresponding query. For all scenarios, many semantically relevant images could be retrieved after one or two queries. This is sufficient in the mental image search context, since finer query by example search engines may be used in a second step, if a search refinement is required. On average on a 700MHz PC, retrieving the top 200 images takes 0.01 second at most, while a search is considered as instantaneous when below 0.1 second [2].



**Fig. 3.** Query and results for “object on white background” (showing only first results).

## 6 Conclusion

We presented a new approach for mental image search which has the major advantage to allow the user to express range queries on multiple region descriptors. Both specific and vague queries can be expressed on the visual content of the mental image with a minimal user interaction. The mental image can be retrieved efficiently and instantaneously and, if necessary, constitute the starting image for a finer query-by-example search.

## References

1. J. Fauqueur. Instantaneous mental image search with range queries on multiple region descriptors. *Technical Report TR 512, University of Cambridge*, 2005.
2. J. Nievergelt, H. Hinterberger, and K. C. Sevcik. The grid file: An adaptable, symmetric multikey file structure. *ACM Transactions on Database Systems*, 1984.