

Creative Computing II
Content-based Multimedia Information Retrieval
17th March 2009

This lab sheet aims to illustrate the steps involved in implementing a simple content-based multimedia information retrieval system, in this case for overall colour similarity in images.

1. The first step in building a system is to understand the task that we require it to achieve. For today, that task will be “find the image known to the system whose average colour is perceptually closest to the input (query) colour”. Suggest technical or creative applications corresponding to this task.
2. In order to compute average colour and to evaluate perceptual closeness, we will need some utility colour-handling functions: one to convert from an sRGB colour specification to a CIE XYZ colour (a space in which *colour vision by averaging* takes place); and one for converting from CIE XYZ to CIE L*a*b*. In the programming environment of your choice, write or obtain methods implementing these conversions.
3. The computation of the average colour for each image in the database needs only to be done once. Decide on a data structure in your programming environment of choice which can store all the information about each image that you will need, and implement that data structure.
4. The database of images can be represented as a linear data structure, each of whose elements is itself an instance of the data structure defined in part 3. Implement such a data structure in your programming environment, and write code populate it with a reasonably small set of images and computed metadata.
5. The user of the system will require some means of specifying a query colour: maybe by text input, mouse selection from a set of choices, or some other means. Implement one way of providing a query colour.
6. The act of satisfying the query is to iterate over the data structure from part 4, populated with colour data, computing the distance between the query colour and each of the average image colours in turn, and recording the image with the smallest distance. At the end of the iteration, the image (or other identifying characteristic, such as its filename) should be displayed to the user. Implement this.
7. You should now have a working system. Compare it with those produced by your colleagues. Are any of your design choices particularly different?