

Image Summaries using Database Saliency

Radhakrishna Achanta*

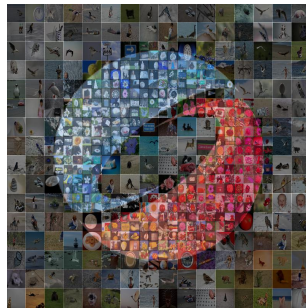
Appu Shaji

Pascal Fua

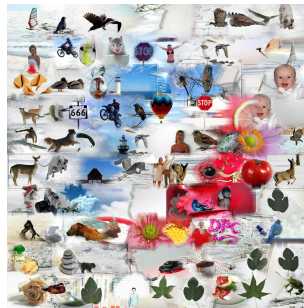
Sabine Süsstrunk



(A)



(B)



(C)



(D)

Figure 1: Database saliency automatically finds the most salient images from thousands of images for creating interesting mosaics (eg. A and B) or collages (eg. C and D). More results available at: http://ivrg.epfl.ch/supplementary_material/RK_SIGGRAPH_Asia09/

Database saliency

It is useful to have a small set of representative images from a database of thousands of images to summarize its content. There are two key aspects of such image summaries: how to generate them and how to present them. We address both issues. We extend the idea of image saliency to databases and introduce the notion of *database saliency*. We argue that in image databases, there are certain images that are more uncommon or *salient* than others and therefore are more interesting. We compute the database saliency value of an image as its total distance from all the pre-defined cluster centers of the database. We demonstrate the use of database saliency in two visualization applications: creating image collages and mosaics using automatically chosen salient images.

We cluster the images using the *k-means* algorithm. For this we need an image signature and a distance measure to compute the similarity between two image signatures. We propose the use of *saliency weighted color coherence vectors* (SWCCV) as the image signature and a weighted χ^2 similarity measure. Simple color histograms are poor abstractions of images as two visually dissimilar images can have the same histogram. Color coherence vectors (CCV) alleviate this problem by labeling pixels as coherent or incoherent pixels. A pixel p_i is incoherent with a given neighborhood $\mathcal{N}(p_i)$ if and only if $\exists p_j \in \mathcal{N}(p_i), \max(p_i - p_j) > T_n$, where T_n is a threshold on the number of neighbors that should be similar to a pixel to determine its coherence. To obtain an SWCCV, we use an image saliency map [2009] to weigh each bin of the CCV with the total pixel saliency of the pixels falling within that bin. Thus, more weight is given to the salient pixels than to others. The degree of similarity between two images is computed as a weighted χ^2 distance between their corresponding SWCCV's.

If M images are needed for an application, we create M clusters of the database. We rank images in decreasing order of their total distance from the M cluster centers. In this way, we select those images that are at the boundaries of clusters owing to their distinctness with respect to other images. We can now choose images for an application, such as a collage, either by simply choosing the top M images from a database saliency ranked list, or by choosing M images from this list that satisfy a weighted cost function C_w that best matches a region of a user-provided template image (e.g. the SIGGRAPH logo of Fig. 1) with an image in the database.

Variable tile size image mosaics

Conventional image mosaics have same sized tiles that are occupied by the “best matching” image from a database. If the image tiles are big, the underlying image is blocky, while if the tiles are small, the image tiles are hardly discernible. We create an image mosaic that assigns variable size image tiles based on a novel energy map decomposition technique.

We use the image saliency maps proposed by Achanta et al. [2009] as our energy maps. We recursively subdivide the input image in four equal quadrants, subject to the energy contained within each quadrant exceeding a threshold T_e . Each quadrant is then overlaid by the database saliency ranked image whose average *CIELAB* vector has the lowest Euclidean distance from the average *CIELAB* vector of the quadrant according to the cost function C_w . The original image and the chosen tile images are averaged to get the final mosaic.

Automatic image collages

Creating a collage from images is an interesting consumer application [2006]. In order to create a collage as in Fig. 1(C and D) we first stitch together rows of salient images chosen by our technique, and then vertically stitch the created rows. To join any two adjacent images, we choose the best seam cut in their overlap region. This is done using an improved technique of finding a seam cut that uses image saliency maps [2009] instead of conventional gradient energy maps, and new forward edge energy terms that take into account both color and intensity information. Thus, in the resulting mosaic *image saliency* helps join images without altering salient content while *database saliency* helps in choosing the most interesting images of the database.

References

- ACHANTA, R., HEMAMI, S., ESTRADA, F., AND SÜSTRUNK, S. 2009. Frequency-tuned salient region detection. In *IEEE CVPR*, 1597–1604.
- ROTHER, C., BORDEAUX, L., HAMADI, Y., AND BLAKE, A. 2006. Autocollage. *ACM SIGGRAPH* 25, 3, 847–852.

*[radhakrishna.achanta;appu.shaji;pascal.fua;sabine.sustrunk]@epfl.ch