

C00 Exercise 2: review

Ramana Kumar (rk436)

October 25, 2010

Woodwark Review Form

(Notes to accompany the filled in form.)

The paper is missing a detailed results section, naming all the tested image sets, the method and parameters used on each test, and the measurements made on the resulting synthesized images.

Showing more results, probably as graphs or figures, while keeping the images of the representative examples, is a major revision of content. Minor revisions of presentation include fixing spelling and grammar errors, as detailed on the next page.

1 Summary

A new automatic method for extracting shared static content from a set of images each depicting a similar scene, and producing a synthesized image. The method is repeatedly shifting the pixel at some location towards a weighted mean of the corresponding source pixels, where the weights come from patch neighborhood consistency between a proper subset of the source images. Contributions include the ideas of using spatial coherence and proper subsets, and the mathematical details for an implementation.

2 Major comments

The overall style and structure of the paper is good. The balance is good, too, except for the light results section.

No objective measure of image quality or realism is given. Is the model, defined by the source images, the ultimate target? Perhaps not, since you talk about refining the model. Consider counting ghosts and discontinuities. Also consider manually identifying static content in a source image set that should be preserved in the synthesized image, and measuring whether it is preserved.

You mention testing your method on a variety of image sets. Provide full details: the number and type of sets tested, the parameter values used, and the results. Relevant input data/parameters include n , image size, scene type, starting image, k , and s . Full results would make the following statement precise, for example, “usually many pixels in the template image did converge”.

You assume “similar areas in the images are associated with similar scene elements”. What is the scope of the kinds of images that satisfy this assumption? You mention “stochastic textures” but give the impression that these aren’t your target data sets.

To justify the conclusion that your method compares well to other methods for image synthesis, present a comparison in your results section. Such a comparison would complement the representative results of three other methods in Figure 2.

Do you want to explore your method’s time complexity in detail, with analysis or experimentation? If not, there is no reason to mention it.

Provide a link to the source code of the CImg-based implementation.

3 Minor comments

(A) “as a model for a single photograph [of] this scene.” Do you mean “of this kind of scene”?

(1) “The technical challenges appear similarly” → “Similar technical challenges appear”

(Fig 2) “(middle) [and of] forming large... discontinuous, [even in] ...”

“is rather optimizing spatial... then” → “is optimizing spatial... rather than”

“manifold” → “manifold” or “many”

(2) “we have the input images using [War03]” → “we have aligned...”? (!)

The subscripts in Eq (4) are excessive. Try $\sum_{\mathcal{I} \in S} d_{\mathcal{I}, \mathcal{T}}(x, y)$ where S is the k closest images.

(3) Exponents in Eqs (6) and (10) are hard to read. Consider $\exp(\cdot)$ for e^{\cdot} , or $\mathcal{N}(\cdot, \cdot)$ for Gaussians.

(4) “after few iterations of this non-local” → “after a few...”?

(5) (Fig 9) “and the apply mean shift” → “then applying mean shift”