

## Supplemental Materials

### 1. Introduction

The supplemental materials contain two parts. In the first part, we give some statistical properties from a large image dataset (Labelme), which will support the effectiveness of our algorithm. In the second part, we show the whole results of separating an image sequence in a crossfade process.

### 2. Statistic properties of natural images

In our paper, we assume three properties of image gradients: the sparsity, the independence between layers and the independence on locations. The first two properties have been well studied and widely used, while the last property is not that well known. This section will give some statistical results about the last property from a large image dataset (Labelme), which can show the reason of the effectiveness of our presented algorithm.

The Labelme dataset [1] contains more than 130,000 pictures and various kinds of scenes, and is suitable to show statistical properties of natural images. We use all pixels in all pictures in Labelme to calculate the correlation between gradients, as:

$$\text{cor}(s) = \sum_{1 \leq k, l \leq 2} \sum_i \sum_x \nabla_k P_i(x) \nabla_l P_i(x + s), \quad (1)$$

where  $P_i$  denotes the  $i$ th pictures in Labelme,  $x$  denotes the location.  $\nabla(\cdot)$  denotes the gradient operator, and  $k$  and  $l$  are the subscripts of gradient vectors, which denote the vertical or the horizontal direction. The results are shown in Fig. 1. In the results, we can know that  $\text{cor}(s)$  ( $s \neq 0$ ) is far smaller than  $\text{cor}(0)$ . It means the gradients on different locations in the same picture are statistically uncorrelated, *i.e.*, are linearly independent. The complete validation of the independence, which needs the estimation of the joint probability function, is computationally unfeasible with a large dataset (it will cost too much time), and hence only the linear independence is shown here. In fact, the motion object function used in the paper is the correlation between gradients of one warped mixture and gradients of the other mixture, and the linear independence on locations is enough to guarantee its effectiveness.

Now we do not use the third hypothesis in the paper, and only suppose that  $\forall i, k, l, x \neq y, \nabla_k L_i(x)$  and  $\nabla_l L_i(y)$  are

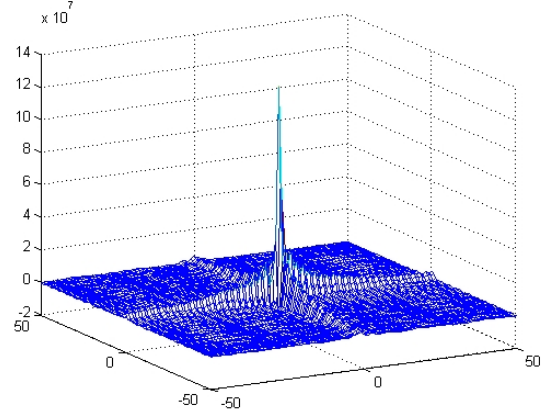


Figure 1. Correlation between gradients in LabelMe.

uncorrelated. As in the paper, by use of the independence between layers, the motion object function (Eq. 4 in the paper) can be expanded as:

$$O(u) = \left| \sum_j a_{2j} G_j(u) \right|, \quad (2)$$

where

$$G_j(u) = \frac{1}{N} \sum_x \nabla^\top L_j(u(x)) \frac{du(x)}{dx} \frac{df_{2j}^\top(x)}{dx} \nabla L_j(f_{2j}(x)). \quad (3)$$

On one hand, when  $u(\cdot) = f_{2j}(\cdot)$ , the summed term is positive determined, and thereby  $G_j(u) > 0$ . On the other hand, When  $u(\cdot) \neq f_{2j}(\cdot)$ ,  $\nabla L_j(u(x))$  and  $\nabla L_j(f_{2j}(x))$  are uncorrelated, and thus  $G_j(u) = 0$ . Consequently, we get the same properties as in the paper: the motion object function  $O(u)$  is significantly larger than zero when and only when the searching motion  $u$  is the same with one correct layer motion.

It is also easy to explain the effectiveness of our algorithm in intuitive perspectives. The gradients can be seen as edges of images, which show clear shapes of every layer on mixtures with little superimposition. With such edges, if we use a correct layer motion to match two mixtures, almost all edges of one layer will be matched and the correlation is large. Otherwise, only a small number of edges can be matched and the correlation is very small.

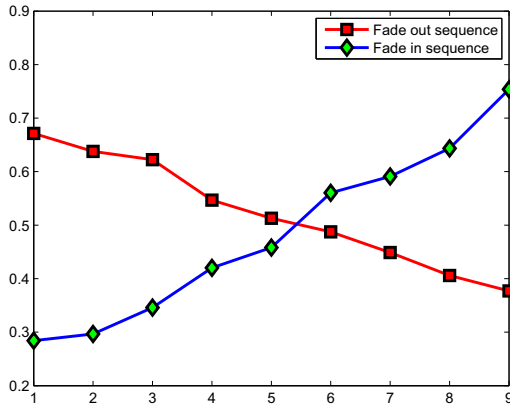


Figure 2. The intensity changes of the fade in and the fade out sequences.

### 3. Separation of a crossfade sequence

In our paper, we have shown the separation of two crossfade images. Indeed, the whole crossfade sequence<sup>1</sup> can be separated. We perform SPBS-M on every three images (the current frame, one in front of it and another one behind it) and use the correlation of layers' gradients to match the orders. Then, the fade in and fade out sequences are obtained, which can be seen in the video "Separation.Of.Crossfade.avi".<sup>2</sup> Fig. 2 shows the intensity changes of the fade in and out sequences, which are approximately linear. Besides the intensity change, the fade in layer of the water dam also has a zoom out effect, which is a nonuniform scaling transformation. Our method can handle both varying mixing coefficients and nonuniform motions, and reconstructs layer images of good quality. To the best of our knowledge, the separation of mixtures with varying mixing coefficients and nonuniform motions has not yet been addressed in open literature.

### References

- [1] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. Labelme: a database and web-based tool for image annotation. *international journal of computer vision. IJCV*, 77(1-3):157–173, May 2008. <sup>1</sup>

<sup>1</sup>The crossfade is from "How Water Won the West" in TREC Video Retrieval Test Collection, at <http://open-video.org/details.php?videoid=489>.

<sup>2</sup>For clear view, the results are played in both normal and inverse orders. The video uses the Xvid codec, which can be downloaded at <http://www.xvid.org/Downloads.15.0.html>.