**Abstract**

Since low-resolution(LR) image captured by the surveillance system can be noisy and even occluded by masks or sunglasses, directly using traditional methods to reconstruct high-resolution(HR) image may lead to unnatural effects which are not tolerable. In this paper, we proposed a face hallucination method that could handle those occlusions with no occluded training data or prior knowledge of occlusions needed in advance. This is achieved by using adaptive robust sparse coding to distinguish the occluded regions and then using position-patch based face hallucination method to synthesize interested non-occluded regions. Experimental result demonstrates that our HR image has not only better overall PSNR and SSIM but also better image quality in local non-occluded areas. Further, by ignoring occluded areas, we can achieve a better performance in the extended long distance recognition problem than some other state-of-the-art methods.

**Introduction**

Face hallucination refers to a learning-based super-resolution technique that synthesizes a HR face image by its LR input. With the help of this technique, many skills and applications like image compression or face recognition could be improved due to the better image quality it provides. Since surveillance systems are widely used in the modern society, these kinds of image processing technique attract much attention of researchers in recent years in order to deal with the uncontrollable image quality caused by the rather poor surveillance device. Recently, one trend of face hallucination research turns to study on multi-view face image synthesis*\cite{ A Simple Approach to Multiview Face Hallucination } \cite{ Multi-view face hallucination based on sparse representation }* which synthesizes a frontal view of face image form face images with various poses. In this paper, still, we focus our attention on super-resolution of face images and extend traditional method to cope with challenges of occlusions on input images.

Several methods are proposed for face hallucination these years. By using a training set of HR and LR image pairs, most of the state-of-art methods learn a co-occurrence model and use the details of HR training images to predict the missing structure of the LR input. In short, the basic idea is to fit the income LR input with the LR training images and then recover the HR output by using the corresponding HR training images.

The ways of fitting LR input can be divided into different categories. A global method (fitting the input by whole face images) *\cite{ Hallucinating face by eigen-transformation }\cite{ An example-based face hallucination method*

*for single-frame, low-resolution facial images}* is an efficient one while the result might be unsatisfactory due to the neglect of local details. Local patch method (fitting the input by image patches)has been widely used these days owing to its promising performance but, nevertheless, the key factor of the hallucination performance lies in how to fit the LR input patches more properly and get a better reconstruct weight. Chang *\cite{ Super-resolution through neighbor embedding }*proposed a super-resolution method by selecting K patches as candidates and finding the reconstruct weight for target input patch through solving a least squares regression(LSR). Ma*\cite{ Hallucinating face by position-patch }* improved Chang's method by taking advantages of positions of the face structure and proposed the position-patch method that selects the patches in same face position as reconstruct candidates. Jung *\cite{ Position-patch based face hallucination using convex optimization }* followed the idea of position-patch but assumed that the image patches can be sparsely represented(SR) by training imaged. Hence a L1 regularizer is added to get a more accurate result. Other methods like Jiang's *\cite{ Position-Patch Based Face Hallucination via Locality-Constrained Representation }* whichintroduced a locality constraint to improve SR and *\cite{ Coupled-Layer Neighbor Embedding For Surveillance Face Hallucination* *}* which uses couple-layer information to correct the hallucination result are more complex extensions of the position-patch method that are shown to achieve a better performance than the original one.

However, the performance of the method above could be greatly degraded when the input image has unwanted occlusions on it. There are several reasons for this. First of all, since there is no occlusion pattern pre-collected in the training set, face hallucination method above would use neutral, clean face images to fit the input occluded areas and somewhat reconstruct those areas unnaturally. In addition, the reconstruction of non-occluded areas would also be interfered when those methods are trying to fit the input image patch containing both occluded and non-occluded areas. In consequence, without taking any measures to prevent such problem, the result of face hallucination is not reliable. In the following paper, we will propose a robust face hallucination method that adaptively detects the areas of occlusions and deal with the hallucination problem separately on both occluded and non-occluded areas.

**Proposed Method**

**2.1. Robust Sparse Coding**

Since the LR input images may be partially occluded, we need to find the occluded region in advance before performing further processing. In this work, we utilize the robust sparse coding method proposed in [?] to neglect the occluded pixels, and focuses only on the clean part of the image to calculate the reconstruction

error. In this way, the non-occluded part can be separated and left unaffected by the occluded region. The minimization problem in RSC is stated below:

...

**2.2. Adaptive Robust Coding and Mask Learning**

A deeper examination of RSC would reveal that the parameter τ can be interpreted as the proportion of images we want to reserve. For example, if we set τ to be 0.6, the worst 40% of pixels in reconstruction would be marked and neglected while the cleanest 60% of pixels would be reserved and used for further recognition purpose.

In section 2.1, we demonstrated that RSC is capable of masking out the corrupted pixels and producing a clean image for further hallucination tasks. However, the proportion of the unwanted pixels are predetermined by τ ( which controls

the threshold that w(ek) starts to drop rapidly). Since the occluded proportion in every images varies tremendously, it doesn’t make sense to fix τ at a single value for all patches. In consequence, the biggest problem of occlusion-robust face hallucination becomes to adaptively determine how many pixels we are going to drop for each input image.

To solve this problem, we take some assumptions into account.

1. The reconstruct error of the occluded region would be a lot greater than non-occluded region.

2. The reconstruct error in all occluded region would be at the same level while the reconstruct error in non-occluded region would be at another level which is lower than above.

Therefore, if we adjust the parameter τ dynamically, when τ is too large(the reserved region still contains occlusions), the reconstruct error of the reserved region would be high, otherwise(τ is too small and too many pixels including clean ones are dropped) the error would be low. Further, although the reconstruct error would fall as τ goes down, there should be two trend in the relation curve between τ and the reconstruct error. By the second assumption, if τ falls into the range that the reserved region contains occlusions, the reconstruct error would fall linearly and sharply when we drop a few more pixels. On the contrary, if τ falls into the range that too many pixels have already been dropped, the reconstruct error would fall mildly when dropping more pixels. The curve we mention above is plotted in figure(?).

(figure)

Consequently, by finding the intersection point of the two straight line shown above, we can determine the value of τ. As we sample through the relation between τ and the reconstruct error, the curve does indicate the two trends of relation but, however, cannot easily be fit by exact two straight lines. To approximate our assumption, we use an exponential curve to fit the sampled points in practice. An exponential fit is chosen not only because it can reflect the two trends mentioned above better than other curves but also for the convenience that a smooth curve could neglect the effect of some outlier sample points. The approximation curve is shown in figure(?).

(figure)

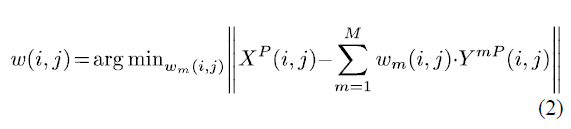
In order to find the target value of τ in the new exponential curve we approximate, we use a point of a fixed slope in the exponential curve to simulate the intersection point of the two lines in our assumption. Since we have no pre-knowledge of the occlusion, we still need to divide some training images as validations to select this slope. Nonetheless, with the properly selected slope value, we could adaptively find τ for input images without knowing the proportion of occluded region.

At last, we get a mask of occluded region for every input images. We further assume that the occlusions are contiguous and use a median filter to remove the discrete dots of the learned mask. Figure(?) shows the effectiveness of using adaptive robust coding to find different value of τ and the learned masks for corresponding τ.

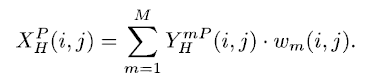
**2.3. Position-patch Based Face Hallucination**

As mentioned in the introduction, position-patch based method is one of the most popular face hallucination method in recent years. In this paper, we choose one adaptation of the position-patch based method which uses sparse representation to find the reconstruct weight for each position-patch. This method is first proposed by Jung*\cite{ Position-patch based face hallucination using convex optimization }* who thought that the LR input could be linearly represented by the training LR images sparsely. We combine this method with our mask learned above to perform face hallucination and the process is briefly stated below.

First, we apply the learned mask of the LR input to mask out the region of occlusions for both LR input and LR training set. Also, we interpolate this LR mask to become HR and then mask out the corresponding occluded region in HR training set. We then divide the masked LR input, (\* means the image is masked), into overlapping position-patches denoted by , where (i,j) is the position index and N is the number of patches. The mth masked LR training image, , is also divided into patches in the same way, denoted by , where m=1, 2, ...,M and M is the number of training images. The reconstruct weight in the position (i,j) could be determined by the following minimization problem:

 <-裡面還要加入L1的regularizer和表示mask的\*

The masked HR output in position (i,j), is then obtained by

 <-裡面還要加入表示mask的\*

After combining the HR patches and averaging the overlapping area, we obtain the hallucinated HR output with the mask, . At last, we perform bicubic interpolation on the LR input to get a HR image on the masked region, and then add it to to get our final result, . The flow chart of our method is summarized in fig. 1.

**Experimental Result**

**3.1. AR Database**

To demonstrate the superiority of our hallucination method under conditions of occlusion, we perform the experiment on AR database. The AR database contains 126 individuals with more than 4000 frontal face images. There are two sessions in it both consist of images with different variations like expression, illumination and facial occlusions. In our experiment, we consider a subset of AR by randomly choosing 100 individuals with 50 men and 50 women. Two sessions are used for training and testing purpose separately while images with occlusions are removed in the training set. We also remove images with variation of illumination since we focus only on occlusions in this work. In short, there are 4 clean images and 2 occluded images for a individual in testing set. All images are converted to grayscale and cropped to 165x120 pixels. After that, images are down-sampled to 96x96 and 24x24 to be HR and LR images respectively. In the following, we divide the experiment into two parts and we would, first, demonstrate the better quality of our hallucinated images and, then, show these HR images with better quality can help face recognition in reality.

**3.2. Hallucination Result**

Since there is scarcely any method handling occlusions in the face hallucination problem, we only need to compare our result with a few representative hallucination methods. In our experiment, we select Jung's method, which is a widely used position-patch method and the bicubic interpolation as base lines for comparison. The reason for choosing Jung's method is that since we, in our work, use Jung's method to hallucinate the clean region of the face images, this comparison is able to show the effectiveness of the learned mask proposed in our work.

There are two parts of the comparison. In part A, which is a traditional comparison setting, we use the whole HR image of the LR input as ground truth to compute PSNR and SSIM. In part B, for fairness, we use the mask learned in our method to mask out the occluded regions on the result of every hallucination method as well as the HR ground truth. And then we compute PSNR and SSIM for the left clean region. The superiority of our method in both part A and B proves that for occluded images, the performance of common hallucination method would degrade on both occluded and clean region if no protective step like the method we propose is taken. The experimental result is shown in fig(?) and table(1).

**3.3 Recognition Result**

One of the most important applications for hallucination is to improve the performance of face recognition when the quality of the input query image is poor. In reality, occlusions may also exist and enhance the difficulty of face recognition. In this work, we deal with the problem mention above and we would show that our HR result could help it.

We use RSC, which is proposed in [?] to perform face recognition due to the robustness it provide to handle images with occlusions. We fine tune the parameter in RSC and perform it on the following 4 settings.

1.LR training set and LR testing set

2.HR training set and HR testing set hallucinated by bicubic interpolation

3.HR training set and HR testing set hallucinated by Jung's method

4.HR training set and HR testing set hallucinated by our method (occluded region is masked out by the mask learned in our method)

The experimental result is shown in table(2) and it demonstrates that our hallucination method helps RSC to do better while other methods only lower the performance.