

Face Hallucination: A Review

Jaskiran Kaur¹, Asst. Prof. Manish Mahajan²

¹M-TECH (IT) 4th Sem, Chandigarh Engineering College, Punjab(INDIA)

²Asst. Professor, Department of Information Technology, Chandigarh Engineering College, Punjab(INDIA)

ABSTRACT-For addressing the problem of hallucinating a higher resolution face given a low resolution input face, the goal of face hallucination aim should be to produce high resolution images with fidelity from low-resolution ones. Face hallucination is a technique of domain-specific super-resolution problem having goal to generate high-resolution images from low-resolution inputs, which finds numerous vision applications. Since a LR image can be made from a HR image by a linear convolution process with down sampling, the hallucination problem can be viewed as an inverse task to reconstruct the high-frequency details. In this paper, we study face hallucination which is the process of synthesizing a high-resolution face image from an input LR image, with the help of a large collection of other HR face images.

Keywords-Face hallucination; Super resolution; Non negative matrix factorization;

I. INTRODUCTION

Face hallucination is the synthesis of a high-resolution face picture from an input low-resolution picture, with the help of a large collection of other high-resolution face pictures. The theoretical contribution has a two-step statistical modeling approach that integrates both a global parametric model and a local non parametric paradigm. In the first step, we will get a global linear sample to learn the interrelationship between the higher-resolution face pictures and their smoothed and down-sampled lower resolution ones. At the second step, we will model the residue between an original high-resolution image and the reconstructed high-resolution image after the application of the learned linear model with a patch-based non-parametric Markov network, to get the high-frequency content. With the integration of two global and local models, we can produce photorealistic face pictures. A contribution is a robust algorithm for alignment of the low-resolution face images to obtain very good hallucination consequences. The effectiveness of our approach is given by experiments generating very high-quality hallucinated face images from low-resolution input without the manual alignment.

Any face hallucination algorithm must meet these three constraints as written below:

(a) Data constraint. The resulting image should be close to the original image when it is being smoothed and down sampled.

(b) Global constraint. The result should have all common features of a human face, e.g. mouth, eyes and nose, symmetry,

and many other features. The facial features must be coherent always.

(c) Local constraint. The resulting image must have very specific features of the face image, as with the photorealistic local characteristics.

Many other computer tasks need inferring the missed higher-resolution pictures from the low-resolution input or original image. The particular interest is to infer high-resolution face images from low-resolution face pictures. This effect was brought forward by Baker and Kanade as *face hallucination*. This method has very broad uses in image enhancement method, image compression techniques and face detecting. It can be specifically useful in surveillance system in that the resolution of face image are normally in video clips, but the specific details of facial features which can be found in the potential high-resolution picture are crucial for detection or identification and further inspection.

They have addressed the problem of hallucination of a high resolution face given a low resolution input picture. This problem is processed through sparse coding. To destroy the facial structure, Non-negative Matrix Factorization [1] is first applied to learn a localized part-based subspace.

This subspace is very useful for super-resolving of the incoming low resolution face under reconstruction conditions. For the further enhancement of the specific facial information, they have proposed a local patch method based on sparse coding representation with respect to coupled over complete patch dictionaries, which can be rapid resolved via linear programming. They have shown that this approach can hallucinate high quality super-resolution faces.

II. Face hallucination

At a practical issue, the challenge in face hallucination is the difficulty of alignment faces with below-resolution pictures. Many learning-based image synthesis approaches or methods are required to bring the alignment between the test sample taken and the training examples, e.g. [2]. Even a bit amount of wrong alignment can degrade the synthesized or processed result. The facial features may have very few pixels; in real world images the faces are normally not upright; the scale and position must be evaluated at sub-pixel level. Therefore,

alignment at low-res requires that so accurate measurements can be made from very little bit of data. For addressing this challenge, they have designed a face alignment algorithm to align faces at low-res. The alignment approach finds an affine transform to warp the input image to a template to hike the probability of low-res face image, determined from an Eigen space value representation. For making that alignment step robust, multiple candidate starting points are shown via a stochastic algorithm from which the best alignment result is selected by them itself. Example shown below:



(a) Low-res input (b) Hallucinated by our system (c) Original high-res

Figure 1. Example of face hallucination

Note that the detailed facial features like eyes, eyebrows, nose, mouth and teeth of the hallucinated face (b) are different from the ground truth (c), but perceptually we see it as a valid face image. The processing from (a) to (b) is entirely automatic.

Face hallucination is higher-resolution of face pictures, or verification of the details of features of a face from a low-resolution picture. The method or technique of sparse coding can be used. The various basis decomposition methods can be used such as nonnegative matrix factorization and Principal Component Analysis. Because of the significance of face images in facial recognition systems and other applications, face hallucination has become a very successful area of research.

Face hallucination is a domain-specific super-resolution problem having the goal to generate high-resolution images from low-resolution images or input image, which will find the number of vision applications. Since a Low Resolution image can be modeled from a HR image with the linear convolution process with down sampling process, the hallucination problem can also be viewed as an inverse task to reconstruct the high-frequency specific details. As in the recent work they have focused on the generic super-resolution problem, considerable less attention is being paid to face hallucination.

III. Related work

In the paper [10] addresses the problem of hallucinating a high resolution face given a low resolution input face. The problem is approached through sparse coding. To exploit the facial structure, Non-negative Matrix Factorization (NMF) is first employed to learn a localized part-based subspace. This subspace is effective for super-resolving the incoming low resolution face under reconstruction constraints. To further enhance the detailed facial information, we propose a local patch method based on sparse representation with respect to coupled over complete patch dictionaries, which can be fast solved through linear programming. Experiments demonstrate that their approach can hallucinate high quality super-resolution faces. Their experiments were conducted on the face database FRGC Ver. 1.0. All these face images were aligned by an automatic alignment algorithm using the eye positions, and then cropped to the size of 100×100 pixels. To obtain the sparse subspace! Spanned by W , we selected 540 face images, covering genders, different races and facial expressions. To prepare the coupled dictionaries needed by our sparse representation algorithm, we sampled approximately 100,000 patch pairs from the training images. The patches are of size 5×5 pixels. 30 outside faces were chosen as our testing cases, which were blurred and down sampled to the size of 25-by-25 pixels.

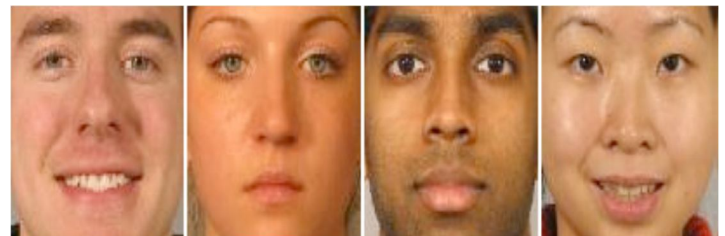


Figure 2. Example training faces in the algorithm.

In the algorithm, there is only one parameter λ that they need to determine. Experimentally, they found that setting $\lambda = 0.005$ generally offers satisfactory results. They compared their algorithm with bi cubic interpolation and back projection. The results indicate that their method can generate much higher resolution faces.

In the paper [2] studied face hallucination or synthesizing a high-resolution face image from a low-resolution input, with the help of a large collection of other high-resolution face images. They develop a two-step statistical modeling approach that integrates both a global parametric model and a local nonparametric model. First, they derive a global linear model to

learn the relationship between the high-resolution face images and their smoothed and down-sampled lower resolution ones. Second, the residual between an original high-resolution image and the reconstructed high-resolution image by learned linear model is modeled by a patch-based nonparametric Markov network, to capture the high-frequency content of faces. By integrating both global and local models, they can generate photo-realistic face images. Their approach is demonstrated by extensive experiments with high-quality hallucinated faces. This paper described the followings.

What resolution of face is needed to do face hallucination? Obviously there are two extreme cases. When the input is only one pixel, then face hallucination becomes a problem random face synthesis constrained by that the average intensity is given. When the input has very high resolution, e.g. 128×96 , then there is no need to do hallucination, either. Therefore, there exists a range of resolutions in which face hallucination makes sense. In this paper they have chosen low-res at 32×24 for face hallucination. Most face detection systems have been designed using a 20×20 or 24×24 templates.

Why global and local modeling?

Theoretically they can solve face hallucination by Bayesian MAP in one step then why do we bother to decompose $p(IH)$ into two steps of global and local modeling? The resolution of the high-res face image is $128 \times 96 = 12288$, requiring too many training examples to estimate a reasonable probability distribution $p(IH)$ even using advanced machine learning algorithms, e.g. they designed a hybrid model for $p(IH)$, i.e. a global face model by eigen faces to capture the global facial features, and a local face model by Markov network to capture the local facial details.

In the paper [4] they have studied that face hallucination, or synthesis a high-resolution face picture from an original low-resolution image, with the help of a huge collection of other high-resolution face pictures. The theory contribution is a two-step statistical modeling process which integrates both a global parametric model and a local nonparametric model. In the first step, they will derive a global linear model to learn the relationship between the high-resolution face images and their smoothed and down-sampling lower resolution ones. At the second step, they model the residue between an original high-resolution image and the reconstructed high-resolution image after applying the learned linear model by a patch-based non-

parametric Markov network, to capture the high-frequency content. By integrating both global and local models, they can generate photorealistic face images. A contribution is a robust warping algorithm for alignment of the low-resolution face images to achieve good hallucination results. The effectiveness of our approach is shown by extensive experiments generating high-quality hallucinated face pictures from low-resolution original image without manual alignment.

In the paper [5] which is describing that the faces often appear very small in size in surveillance imagery just because of the broad fields of view that are used and the relatively very large distance between the cameras picture and the scene. In applications like face recognition or identification, face detection etc. resolution enhancement techniques are generally very essential for it. Super resolution is the process of determining and addition of the missing high frequency information in the image to improve the resolution. It is highly applicable in the areas of recognition, identification, compression, etc. Face hallucination is a subpart of super resolution. This work is intended for enhancing the visual picture quality and resolution of a facial image. It mainly focuses on the Eigen transform face super-resolution techniques in transform domain. Pros of Eigen transformation based technique is that, it does not need any repetitive optimization techniques and so comparatively faster. Eigen transform is performed in wavelet transform and discrete cosines transform domains and the results are presented. The results establish the fact that the Eigen transform is very efficient in transform domain also and thus it is directly applied with slight amendments on the compressed images.

In this paper [6], they presented a new face hallucination method based on similarity constraints to produce a high-resolution (HR) face image from an input low-resolution (LR) face image. This method is modeled as a local linear filtering process by incorporating four constraint functions at patch level. The first two constraints focus on checking if the training images are similar to the input face image. The third is defined in the HR face image, which is to impose the smoothness constraint between neighboring hallucinated patches. The final constraint computes the spatial distance to reduce the effect of patches that are far from the hallucinating patch. Experimental evaluation on a number of face images demonstrates the good performance of the proposed method on the face hallucination task.

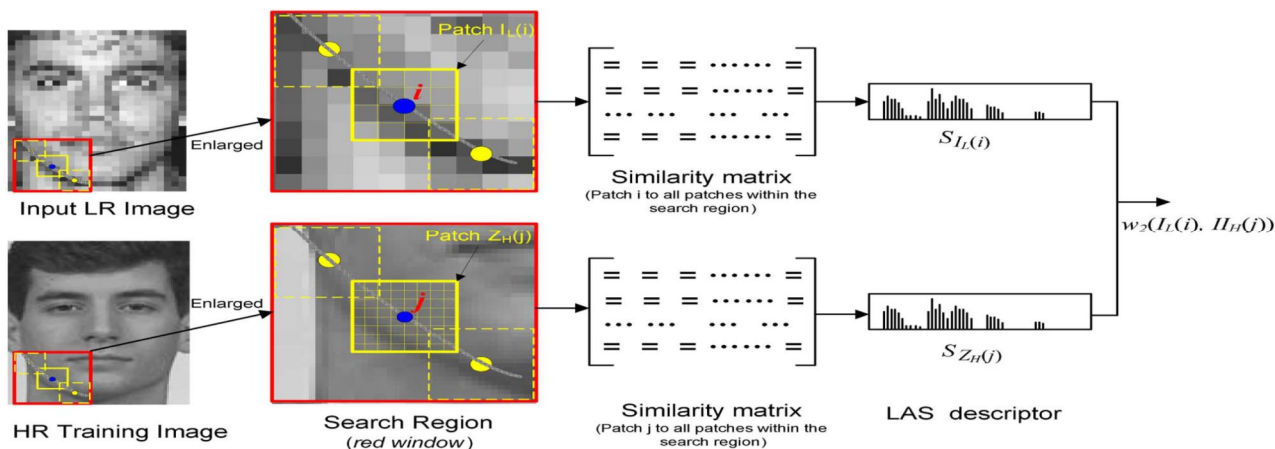


Figure3. In this Illustration of computation, where red window refers to the search region, and yellow windows with solid lines denote the LR and HR patches. Top row: LAS descriptor computation in the input LR image. Bottom row: LAS descriptor computation in a HR training image.

In their work, all the face images are first aligned based on three marked positions: the centers of the two eyes and the center of the mouth. Then the original HR image is cropped to a size of 128 96, where the distance between the eye centers is fixed at 40 pixels, and the vertical distance between the eyes and mouth is fixed at 48 pixels. Its corresponding 32 24 LR images can be obtained by the down sampling process. Given an input LR face image, then divide it into a number of overlapping patches with the size of 4*4. The overlapping pixel is set to 3, which corresponds to 12 pixels in the HR face image. In this letter, they employ laplacian cost function, i.e., to compute the similarity constraints. Four control parameters define are set and respectively, which shows good performance from their empirical study. To handle the possible alignment errors, they simply set to treat the spatial effect of the neighboring patches uniformly.

In the paper [7] explained about Limits on Super-Resolution and How to Break Them. Nearly all super-resolution algorithms are based on the fundamental constraints that the super-resolution image should generate the low resolution input images when appropriately warped and down-sampled to model the image formation process. (These reconstruction constraints are normally combined with some form of smoothness prior to regularize their solution.) In the first part of this paper, we derive a sequence of analytical results which show that the reconstruction constraints provide less and less useful information as the magnification factor increases. We also validate these results empirically and show that for large enough magnification factors any smoothness prior leads to overly smooth results with very little high-frequency content (however many low resolution input images are used.) In the second part of this paper, we propose a super-resolution

algorithm that uses a different kind of constraint, in addition to the reconstruction constraints. The algorithm attempts to recognize local features in the low resolution images and then enhances their resolution in an appropriate manner. We call such a super-resolution algorithm a *hallucination* or *reconstruction* algorithm. We tried our hallucination algorithm on two different datasets, frontal images of faces and printed Roman text. We obtained significantly better results than existing reconstruction-based algorithms, both qualitatively and in terms of RMS pixel error.

In the first half of this paper they have showed that the super-resolution reconstruction constraints provide less and less useful information as the magnification factor increases. The major cause of this phenomenon is the spatial averaging over the photosensitive area; i.e. the fact that S is non-zero. The underlying reason that there are limits on reconstruction-based super-resolution is therefore the simple fact that CCD sensors must have a non-zero photosensitive area in order to be able to capture a non-zero number of photons of light. The analysis assumes quantized noiseless images; i.e. the intensities are 8-bit values, created. The results of enhancing the resolution of a piece of text by a factor of 2. (Just a single input image is used.) The hallucination algorithm produces a clear, crisp image using no explicit knowledge that the input contains text. The RMS pixel intensity error is also almost a factor of 2 improvements over the other algorithms.

In the paper [8] has presented the work of higher-resolution restoring of all of images, they identify the application of learnt picture models within a generative Bayesian framework. They demonstrated that restoration of much higher quality rather than that determination of classical maximum likelihood estimation

can be achieved by either constraining the solution to lie on a very restricted sub-space, or by using the sub-space to describe a spatially varying prior. This sub-space can be learned from image instances. The methods are used on both real and artificial images of text content and faces, and results are compared with R.R. Schultz and R.L. Stevenson's (1996) MAP estimator. They have considered in particular images of scenes for which the point-to-point mapping is a plane projective transformation which has eight degrees. As real image instances, registration is obtained from the images using automatic paradigms.

In the paper [9] they presented an example-based facial sketch system. Their system generates a sketch from an original image, by learning from example sketches drawn with a particular style by an artist. There are two main elements in our system: a non-parametric sampling process and a flexible sketch model. First the original image i.e. input image pixel and its neighborhood; the c distribution of a sketch point is computed by querying the instances and finding all same neighborhoods of that under some conditions. An "expected sketch picture" is then drawn from the distribution to reflect the drawing pattern. Finally, facial sketches are obtained by incorporating the sketch model. Experimental results demonstrate the better quality of their techniques.

In this paper [10] addressed the problem of hallucination of a high resolution face given a low resolution input face. The problem is approached using sparse coding repetition. To destroy the facial structure, non-negative matrix factorization is the first employed to memorize a localized part-based subspace. This subspace is very effective for super-resolving the incoming below resolution face under reconstruction constraints. For further enhancing the too detailed facial information, they proposed a local patch method based on sparse representation with respect to coupled over complete patch dictionaries, which can be effectively solved through linear programming. Experiments demonstrate that their approach is able to hallucinate higher quality super-resolution faces.

In this paper[11] addressed that because of number of very important applications of face images, like long distances surveillance imagery and verification of identity, face hallucination have been one of the good research topics in the last decade. This paper made a survey of different techniques to higher quality face hallucination by looking at previous theoretical backgrounds and consequences. The strengths and weaknesses of these techniques are identified to form a base for a current sparsity based method for super- resolving mis aligned face images.

In this paper[12] In this paper, they studied that the face hallucination, or synthesis of a high-resolution(HR) face image from an input low-resolution image, by taking help from a huge collection of other high-resolution face images. Their contribution is a two-step statistical modeling approach that is integrating both a global parametric model and a local non parametric model. In the first step, they derived a global linear model to learn the relationship between the high-resolution (HR) face images and their smoothed and down-sampled lower resolution (LR) images. In second step, they modeled the residue between an original high-resolution image and the reconstructed high-resolution image after applying the learned linear model by a patch-based non-parametric Markov network, for capturing the high-frequency content. By integrating both global and local models, they generated photorealistic face images. A practical contribution is a robust warping algorithm to align the low-resolution face images to obtain better hallucination results. The effectiveness of our methodology is shown by extensive experiments generating high-quality hallucinated face images from low-resolution input with no manual alignment.

In this paper [13], they studied face hallucination, or synthesizing a high-resolution face image from low-resolution (LR) input, by taking the help of a large collection of high-resolution face pictures. They developed a two-step statistical modeling approach that integrates both a global parameter model and a local non parametric model. First, they will derive a global linear model to learn the relationship between the high-resolution face images and their smoothed and down-sampled lower resolution ones. Second, the residual between an original high-resolution image and the reconstructed high-resolution (HR) image by a learned linear model is defined by a patch-based nonparametric Markov network, for capturing the high-frequency content of faces. By integrating both global and local models, they can generate photorealistic face images. Their approach is demonstrated by extensive experiments with high-quality hallucinated faces.

In the paper [14] they studied that call methods for achieving high-resolution enlargements of pixel-based images super-resolution algorithms. Many applications in graphics or image processing could benefit from such resolution independence, including image-based rendering (IBR), texture mapping, enlarging consumer photographs, and converting NTSC video content to high-definition television. They built on another training-based super-resolution algorithm and developed a faster and simpler algorithm for one-pass super-resolution. Their algorithm requires only a nearest-neighbor search in the training set for a vector derived from each patch of local image data. This one-pass super-resolution algorithm is a step toward achieving resolution independence in image-based

representations. They did not expect perfect resolution independence—even the polygon representation is not having that—but increasing the resolution independence of pixel-based representations is an important task for IBR.

In the paper [15] addressed that the objective presented in the work is the super-resolution restoration of a set of images, and they investigated the use of learnt image models within a generative Bayesian framework. They have shown that restoration of far higher quality than that determined by classical maximum likelihood estimation can be achieved by either constraining the solution to stay on a restricted sub-space, or by using the sub-space to define a spatially changing prior and sub-space can be learnt from image examples. The methods are then applied to both real and artificial images of plain text and faces, and results are compared to R.R. Schultz and R.L. Stevenson's MAP estimator. They considered in specific images of scenes for which the point-to-point mapping is a plane projective transformation which has 8 degrees of freedom. In the real image instances, registration got from the images using automatic methods.

IV. CONCLUSION

In this paper, we studied the face recognition performance over different resolutions. There are various applications of face appearance modeling, e.g. random face Synthesis and the hallucination method based on Eigen transformation could extract the maximum facial information from the low-resolution face images and render some high frequency facial feature to make the face image better. It is making the automatic face recognition easier. Many successful experimental results prove that face hallucination can be applied in real applications to enhance the resolution of face for both face recognition and face image editing. We have focused on other applications of our face appearance modeling methods.

References

[1] Lee, D.D., And Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 6755 (October1999), 788-791.
[2] Wei Liu, Dahua Lin, and Xiaoou Tang. Face Hallucinating Faces: TensorPatch Super-Resolution and Coupled Residue Compensation. *Proc. CVPR,2006*.

[3] S. Baker and T. Kanade. Hallucinating Faces. *IEEE International Conference on Automatic Face and Gesture Recognition*, March 2000.
[4] D. Capel and A. Zisserman. "Super-resolution. Ph.D. Thesis, Department of Eng. Science, University of Oxford, 2001.
[5] H. Chang, D.-Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. *CVPR*, 2004.
[6] Hongliang Li, Linfeng Xu, Guanghui Liu "Face Hallucination via Similarity Constraints" *IEEE SIGNAL PROCESSING LETTERS*, VOL. 20, NO. 1, JANUARY 2013.
[7]. S. Baker and T. Kanade, "Limits on super-resolution and how to break them" *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2000.
[8]. D. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models" *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 627–634, 2001
[9]. H. Chen, Y.Q. Xu, H.Y. Shum, S.C. Zhu, and N.N. Zheng, "Example-based facial sketch generation with non-parametric sampling", *Proc. IEEE Int'l Conf. Computer Vision*, pages 433–438, 2000.
[10]. Jianchao Yang, Hao Tang, Yi Ma, Thomas Huang, "Face Hallucination via Sparse Coding".
[11]. Yan-liang jian-huang-lai, wei-shi zheng "A survey on face hallucination" Springer-W. S. Zheng CCBRLNCS 7701.
[12]. Yan-liang jian huang-lai, wei-shi zheng and Zemin Cai "A survey on face hallucination" 7th Chinese Conference, CCBRLNCS 2012, Guangzhou, China, December 4-5, Proceedings, pp. 83-93, 2012.
[13]. Ce Liu, Heung-Yeung Shumy and Chang-Shui Zhang, "A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model", http://research.microsoft.com/enus/um/people/ce/lu/pdfs/facehallucination_cvpr01.pdf
[14]. W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*,
[15]. D. Capel and A. Zisserman. Super-resolution from multiple views using learnt image models. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 627–634, 2001.
[16] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal ℓ_1 -norm solution is also the sparsest solution. *Comm. on Pure and Applied Math*, Vol. 59, No.6, 2006
[17] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal ℓ_1 -norm near-solution approximates the sparsest near-solution. Preprint, accessed at <http://www-stat.stanford.edu/~donoho/>. 2004.
[18] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *IJCV*, 2000.
[19] Ce Liu, Heung-Yeung Shum, and William T. Freeman, "Face Hallucination: Theory and Practice" Accepted by *International Journal of Computer Vision*. Vol. 22, Issue 2, 2002.
[20] R.C. Hardie, K.J. Barnard, and E.A. Armstrong. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE TIP*, 1997.
[21] S. Baker and I. Matthews. Lucas-kanade 20 years on: A unifying framework. *International Journal on Computer Vision*, 56(3):221–255, March 2004
[22] A. Blake, B. Basclé, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *Proc. European Conference on Computer Vision*, pages 312–320, 1996.