Recognizing Faces with Partial Occlusion using Inpainting

Vijayalakshmi A.
Assistant Professor
Department of Computer Science
Christ University, Bengaluru

ABSTRACT

Face recognition as a biometric is gaining popularity as it is widely being used in surveillance as well as for security. One of the disadvantage of face recognition as a biometric is its low recognition rate when a part of face image is lost due to known or unknown reasons. Occlusion on faces can be caused with the knowledge of a user when the user is covering his face or part of the face on purpose. Occlusion can be categorized as sparse occlusion and dense occlusion. The capability of face recognition system achieves its goal if the occluded part can be recovered for recognition of faces. In this paper, a hybrid inpainting approach is followed to recover the lost region of a face. This approach increases the recognition rate of faces that are occluded. Experimental result on hybrid inpainting proves that the recognition rate on faces increases on comparison with existing methods on occluded faces.

Keywords

Face recognition, Occlusion, Inpainting

1. INTRODUCTION

Biometrics is gaining popularity as security is a major concern in today's world. Among the various biometrics available, face recognition has gained popularity and is widely used for security these days as it is widely accepted among people across the world. Even though face recognition is widely in use, this biometric suffers from reduction in recognition rate due to varied illumination, pause variation and occlusion. Pose variation is a major concern when the input for this biometric is a video sequence. In this case, if the user is unaware of the existence of camera, user is not cooperating with the camera and there is a huge chance of varied poses that makes the recognition task complex. These factors affect the recognition rate of faces in face recognition system.

Occlusion is a major problem that deteriorates the recognition rate of faces. Occlusion on faces or any image leads to the loss of a portion of an image.There can be two different types of occlusion 1) Sparse occlusion and 2) Dense occlusion. Inpainting is a technique that fills in the occluded or lost part of an image using the pixels around the lost region. There are three major categories of inpainting techniques 1) Diffusion based, 2) Exemplar based 3) Global based method.

Image inpainting is a technique wherein the missed pixels or missed data are recovered from the available region and used in reconstructing the missing region. Exemplar-based inpainting which is also known as texture-oriented inpainting. In this method, samples of required texture are used for synthesis. One of the important works in Exemplar-based image inpainting is carried out by the authors in [1]. The algorithm proposed by Criminisi combines the advantage of texture synthesis method and Partial Differential Equation (PDE)-based inpainting method. The method proposed in this work consists of the following major steps for filling the missing region.

i. The image with a missing hole is taken as an input to the algorithm to check the efficiency of occlusion recovery.

ii. The pixels that appear on the edge of the missing region also known as fill front are considered as a first step.

iii. The strength of a pixel on the fill front is calculated to prioritize the pixel to form a patch.

iv. Find the most similar patch from the known region where the pixels are not lost.

v. The prioritised patch is matched against a patch of the same size from the known region.

vi. Once the most similar patch is obtained, the prioritised pixel and the patch surrounding are replaced.

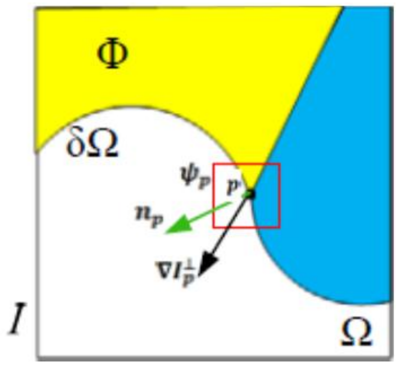

Fig 1: Image Representing Partial Loss of Data.

Figure 1 describes the terms used in the algorithm for occlusion recovery from an image with lost data.

i. I represent the whole image that includes both the known region and the occluded region that is to be filled.

ii. Let Ω represent the target region where the pixels are missing and this is the region that needs to be recovered.

iii. Φ represents the source region from which a most similar patch is selected to fill the target region.

iv. δΩ represent the contour of the region to be filled. This is also being known as fill front.

v. Ψp is the patch surrounding the pixel p.

The pixels at the fill front are prioritised and appropriate patch is chosen around the prioritised pixel. Once the patch is found, pixels are diffused inwards. To prioritise the pixels on the fill front, the confidence term of the pixel and the data term of the pixel are used. Figure 2 explain the detailed procedure followed in filling pixels using the above method.

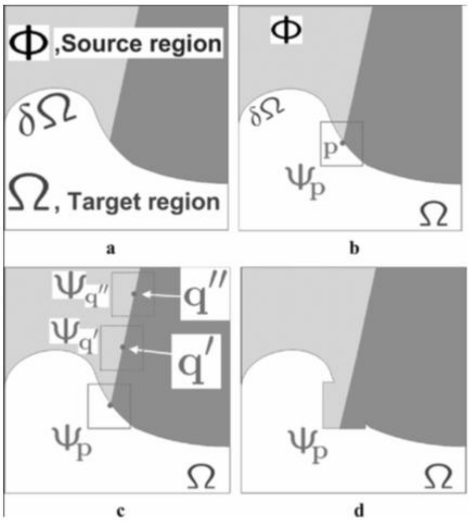

Fig 1: Procedure of Inpainting.

The square template Ψp is the patch centred at a pixel, p. This pixel p belongs to δΩ, which is the fill front. The patch Ψp has at its centre, the pixel p. If Ψp lies in the continuation of an edge, the most similar match Ψq ϵ Φ also lies in the same edge.

2. EXISTING WORK

Face detection is an important step in face recognition technique. For real-time face detection, viola Jones algorithm gives a good detection rate. The work proposed by[2] describes a machine learning approach for detecting objects which give high detection rate. According to the work carried out, it is a composition of three key features: (i) New image representation which authors represented as an Integral image. This helps in quickly computing the features used by the detector. (ii) Learning algorithm based on AdaBoost. This selects features from and yields efficient classifiers. (iii) Combines complex classifiers which discard the background region and does the computation of object-like regions. Based on this work, extended face detection for multi-view faces is proposed in [2]. In this work, authors used building different detectors for different views, [3], and [4]. For this work, the authors extended their previous work for rotated and profile faces, as well as new sets of rectangle features, are defined for rotated face detection. The authors in their work used classifiers based on the weak classifier set from the Adaboost algorithm by [5].

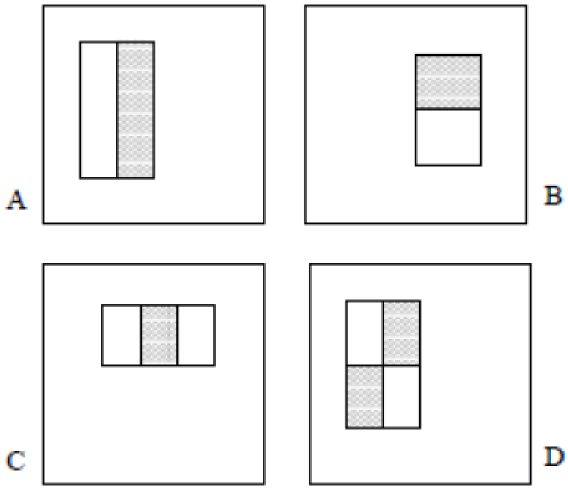

Fig 3: Haar features considered for Adaboost.

The features shown in Figure 3, are the important Haar features used to detect a face region.

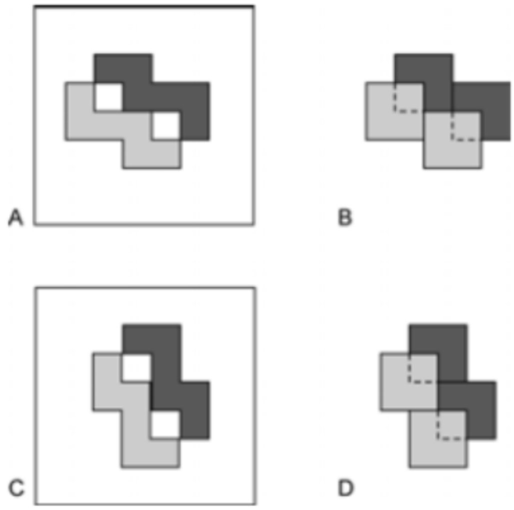

Fig 4: Diagonal Haar Features Designed for Face Detection

Figure 4 shows the diagonal filters used by the authors that are used to detect non-upright faces and non-frontal faces giving highly accurate results. Authors used a sequence of the complex classifier to improve the efficiency in computation. Input window is evaluated on the first classifier and if it returns false then computation ends on that window if it returns true, next classifier in the cascade evaluates the window. In a similar way, window passes through all the classifiers and if true, the detector will return a true value for that window. In order to detect the face with different poses, the poses are divided into various classes and different detectors are trained for each class. The authors in [6] described the algorithm for AdaBoost in their work in detail and is summarised below. The working of AdaBoost algorithm includes one weak classifier is selected at each step. A weight is assigned to the data at each step and a weak learner is constructed based on the weight, [7]. To estimate how the data is classified, the weak learner produces an empirical error. The weak classifier with the best classifier with the lowest error is selected and weights are updated accordingly. The iteration loop stops if the empirical error calculated is 0 or >= ½.

Once the face is detected from the sequence of video, next is to check if this face image suffers from occlusion. The authors in [8] proposed a method using stereo algorithm that gives in disparity maps that can detect occlusion. In their approach, a 3D array is constructed with elements (r,c,d) with (r,c) for each pixel in the reference image and d signifying the range of disparity. Authors have assumed that the disparity maps possess a unique value per pixel and are continuous everywhere. In this method, normalized correlation or squared differences are used to set the intensity values. With each pixel in the image, (r,c,d) element with the maximum match value is found if the match value is found to be lower than a threshold, the image is classified as occluded. Authors in [9] proposed a method of dividing the face into patches and matching each patch with the gallery images to detect occlusion. In [10] authors proposed a method to track objects based on particle filter. To execute this method the pixels in an object were classified as foreground and background pixels using background subtraction method. Upon extracting the object from background, the object is considered to be the region of interest. Ellipse is used to represent the image and occlusion is detected by merging and splitting the image. SVM based approach is used for detecting occlusion in face.

2.1 Support vector machine

Support Vector Machine (SVM) is a pattern classification algorithm developed by [11]. The idea used in SVM classification is structural risk minimisation. Training a SVM is equivalent in training a quadratic programming problem and hence it is difficult when the number of data is large.

Support Vector machine can be used as a regression technique and classification problem. According to [11], SVM in its simplest linear form is a hyper plane that separates a set of positive examples from a set of negative examples with maximum margin. This is depicted in Figure 5. The figure shows the positive data in squares and the negative data in circles. There can be 'n' number of lines that can differentiate the positive and negative data. In SVM the line that separates the data with maximum distance is chosen.

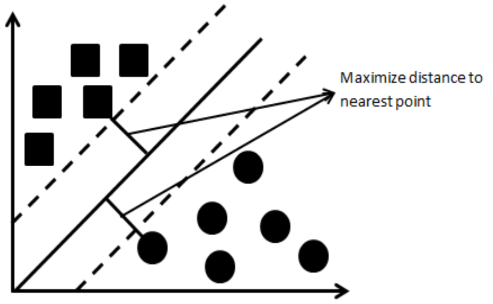

Fig 5: Classification of data using support vector machine.

Support vector machine (SVM) works well with variety of problems under different domains like hand written character recognition, face detection, pedestrian detection and text categorization. Even then, there exist minor disadvantage with SVM as the training algorithm for SVM is slow for large problems. SVM training algorithms suffer with a minor disadvantage because they are complex, subtle and difficult to implement. SVM in its linear form, the maximum margin is defined by the distance of the hyper plane to the nearest of the positive and negative examples, [12].

3. PROPOSED WORK

This work concentrates on an improvement in recognition rate when human faces are partially occluded in a video succession. Inpainting is one of the method to recover the lost region of the image/face. In the proposed work, a preparation set is made with reproduced blocked appearances. The recreation of occluded region is performed using Inpainting strategy. Inpainting is a technique of recovering lost region in an image. There are different methods to recover lost pixels in an image. Exemplar inpainting is one of the best methods to recover the lost region. In this work a hybrid technique is used where in a PDE method and Modified Exemplar Inpainting is utilized as a part of remaking the occluded face region.

The Modified exemplar inpainting is described in the Table 1.

Table 1 . Modified Exemplar Inpainitng

I: The whole image
Ω: The portion of the image I that is occluded
Ψ_p: the patch formed for a search

Step 1: *If* p ϵ I
C(p)=0;
else
C(p)=1;
Step 2:*For* all pixels on the fill front
{

$$C(p) = \frac{\sum q \in \Psi_p \cap \emptyset \, C(q)}{|\Psi_p|}, 0 \leq C \leq 1$$

Calculate gradient magnitude $D(p) = \sqrt{{G_x}^2 + {G_y}^2}$

Calculate C(p).D(p)
} *Continue*

Step 3: Form a patch Ψ_p with p as the pixel at center.
Step 4: *For* all p ϵ I- Ω
Find the most similar patch $\Psi_p{}'$
Step 5: Replace the similar patch with Ψ_p
Step 6: Repeat step2 to step 6 until Ω is filled.

The above algorithm is explained in detail below.

Step 1: The target region initialization. In this step, the missing areas in the image are identified.

Step 2: Computation of patch priority.

The priority is given to each pixel on the fill front based on if the pixel is a continuation of the strong edge and if they are surrounded by high confidence pixels.

For a patch Ψp with center at p for any p ϵ δΩ, the priority P(p) of that pixel is set with the equation,

$$P(p) = C(p)D(p) \quad (4.10)$$

Where $C(p)$ is the confidence term and $D(p)$ is defined as the data term.

During initialisation the values of $C(p)$ is set as:

$$C(p) = 0 \; for \; all \; p \in \Omega$$

(4.11)

$$C(p) = 1 \; for \; all \;\; p \in I – \Omega$$

(4.12)

C (p) is a confidence term that gives information of the surrounding pixels around p. Those patches that have more of the pixels around it available are the patches that are prioritized for filling.

C (p) is calculated as $C(p) = \frac{\sum q \in \Psi_p \cap \emptyset \; C(q)}{|\Psi_p|}$, $0 \leq C \leq 1$

(4.13)

To calculate the data term, sobel edge detector is used to find the strength of the pixel on the edge. As larger mask reduce the error rate, in this case a mask of 2 is used with c=2, [13].

This leads to a sobel operator,

$$Mx = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \quad \text{and}$$

$$My = \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{matrix}$$

The major steps in finding the magnitude of the gradient are as given in Table below.

Once the input image is introduced to the framework, the occluded area is recognized if any. Features from the face are extracted using Curvelet Transform.

4. RESULTS AND DISCUSSIONS

To check the performance of the proposed algorithm, various images are given as the input to the system to analyse the result after inpainting is performed on them. Figure 6 depicts the image by name Lena on which a part of pixels at various locations are covered with green colour. This image is subjected as input to the algorithm which could recover the lost region. The recovered image is shown in Figure 7.

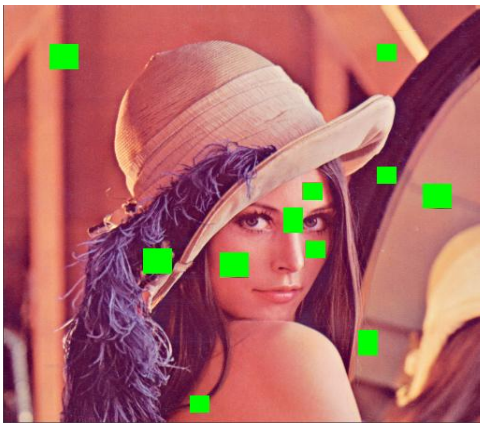

Fig 6: Input image1 with regions of the image covered

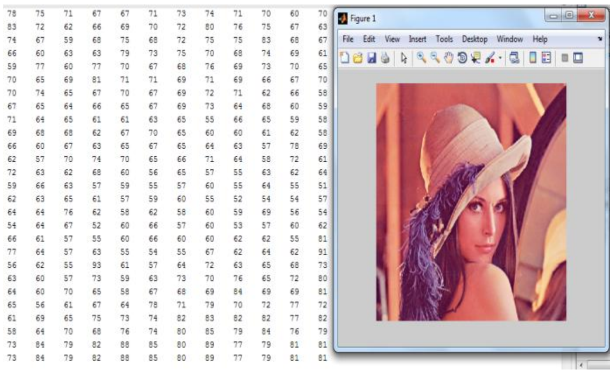

Fig 7: Recovering the lost region of image1

Table 2. Recognition rate with respect to different patches

	P1	P2	P3	P4	P5
Proposed method	86%	95%	98%	96%	98%

Table 2 describes recognition rate of faces with respect to different rectangular patches applied on the face. This proves that the recovery of lost region can be carried out using inpainting technique so that recognition rate of faces can be improved.

5. ACKNOWLEDGMENTS

Thanks to Department of Computer Science, Christ University for the support given in completion of this work.

6. REFERENCES

[1] Criminisi A., Pérez P. and Toyama K., “Region filling and object removal by exemplar-based image inpainting”, IEEE Transactions on Image Processing, vol.13, no. 9, pp.1200-1212, 2004.

[2] Viola P. and Jones M., “Rapid object detection using a boosted cascade of simple features”, in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,vol. 1, pp. 511-518, 2001.

[3] Rowley H., Baluja S. and Kanade T., “Rotation invariant neural network-based face detection”, In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 38-44, 1998.

[4] Schneiderman H., and Kanade T. “A statistical method for 3D object detection applied to faces and cars”, in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 746-751, 2000.

[5] Freund Y. and Schapire R.E., “A desicion-theoretic generalization of on-line learning and an application to boosting”, *European Conference on Computational learning theory, Springer: Berlin Heidelberg*, pp. 23-37, 1995.

[6] Meynet J., Popovici V. and Thiran J. “Fast face detection using adaboost”, *No. EPFL-STUDENT-86954*, 2003.

[7] Schapire R. E., “Explaining Adaboost”, in *Empirical Inference, Springer Berlin Heidelberg*, pp. 37-52, 2013.

[8] Zitnick C. L. and Kanade T., “A cooperative algorithm for stereo matching and occlusion detection”, *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 7, pp.675-684, 2000.

[9] Chen Z., Xu T. and Han Z., "Occluded face recognition based on the improved SVM and block weighted LBP", *in Proceedings of IEEE International Conference on Image Analysis and Signal Processing*, pp. 118-122, 2011.

[10] Pan X., Chen X. and Men A., "Occlusion Handling Based on Particle Filter inSurveillance System", in *Proceedings of International Conference on Computer Modeling and Simulation* IEEE, vol 1, pp.179-183, 2010.

[11] Wang Y., "An Analysis of the Viola-Jones face detection algorithm", *Image Processing on Line*, vol. 4, pp.128-148, 2014.

[12] Cortes C. and Vapnik V., "Support-vector networks", *Machine learning*, vol.20, no.3, pp.273-297,1995.

[13] Vincent O. R. and Folorunso O., "A descriptive algorithm for sobel image edge detection", in *Proceedings of Informing Science & IT Education Conference*, vol. 40, pp. 97-107, 2009.