Quality Assessment based Denoising to Improve Face Recognition Performance

Samarth Bharadwaj[1], Himanshu Bhatt[1], Mayank Vatsa[1], Richa Singh[1], and Afzel Noore[2]
1 - Indraprastha Institute of Information Technology (IIIT)
New Delhi, India
2 - West Virginia University
Morgantown, WV, USA
1 - {samarthb, himanshub, mayank, rsingh}@iiitd.ac.in, 2 - afzel.noore@mail.wvu.edu

Abstract

A probe face image may contain noise due to environmental conditions, incorrect use of sensors or transmission error. The performance of face recognition severely depletes when the probe image is contaminated with noise. Denoising techniques can improve recognition performance, provided the correct parameters are used. In this paper, a parameter selection framework is presented. In the proposed framework, the optimal parameter set is selected for denoising using quality assessment algorithms with low complexity. Quality score based parameter selection is evaluated on the AR face dataset. A correlation study is discussed to ascertain the relationship between the quality scores and recognition rate. The experiments suggest that the proposed framework improves the performance both in terms of accuracy and computation time.

1. Introduction

Quality of a biometric sample affects the performance of the recognition algorithm. In literature, several research papers exist on analyzing the effects of quality on the performance of different biometric modalities such as iris and fingerprint [2, 3, 4]. Environmental corruption such as noise, blur, adverse illumination and compression rates (in JPEG and other compression techniques) influence the performance of state-of-art recognition algorithms. Several enhancement methods have been proposed in literature to handle these corruptions [1]. However, the parameters chosen for the enhancement algorithms have an adverse effect on the performance of automatic recognition algorithms.

1.1. Noisy Images in Face Recognition

Recent research in face biometrics, in an effort to address covariates such as pose, illumination and expression, have turned towards texture recognition algorithms. Texture algorithms such as Local Binary Patterns (LBP) [7], are known to be more resilient towards these covariates compared to appearance based algorithms such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA). However, the experiments show that the texture algorithms are also susceptible to environmental noise. As shown in Figure 1, noise may be induced due to sensor error, transmission error or due to wrong capturing practices, affect face recognition performance. The experiment performed using data-driven noise and LBP demonstrates the effect of noise on face recognition. Cumulative Match Characteristic (CMC) curves in Figure 2 show a sig-

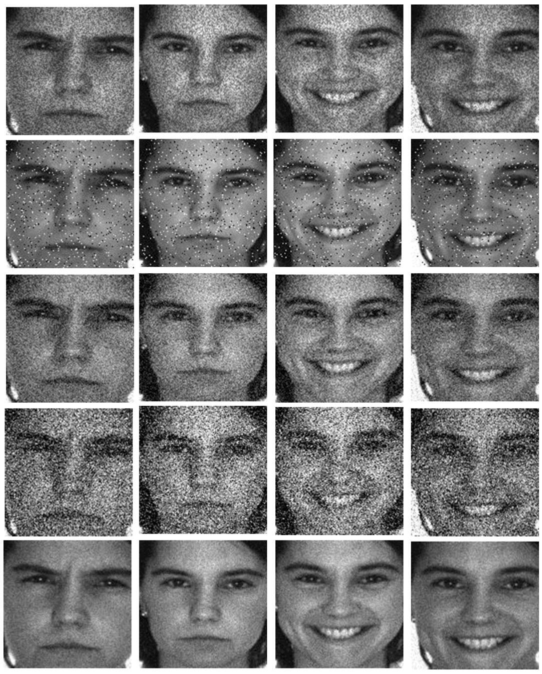

Figure 1. Irregularities due to different types of noise degrade the quality of face images significantly. Sample images from the AR face dataset with synthetic noise.

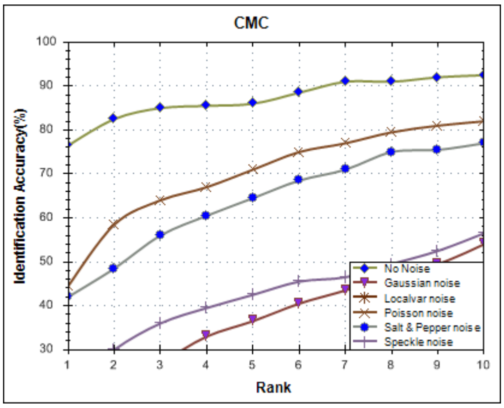

Figure 2. CMC curve of local binary pattern (LBP) + χ^2 matcher: Identification performance decreases when noise is added to the probe images

nificant loss of performance in the identification accuracy due to synthetic addition of noise. Considerable advancements have been made in literature to denoise images, beginning from wavelet based hard thresholding to more elegant soft thresholding techniques. However, the performance of these approaches depend on parameters such as choice of mother wavelet and number of iterations, which are directly dependent on the amount of noise present in the image. It is our hypothesis that the full utility of denoising (or enhancement algorithms) can be realized with a framework that selects the best parameters for each of the given (probe) image.

This paper presents a framework to select the image denoising parameters based on the quality assessment of individual images. The proposed framework utilizes Support Vector Machine (SVM) classification to learn the relationship between image quality assessment scores and the optimal parameters for denoising. The research contribution can be summarized as follows:

1. Understanding the effect of noise artifact on texture based face recognition algorithms.

2. Understanding the relationship between quality enhancement parameters and quality assessment scores.

3. A quality assessment based enhancement framework for parameter selection to improve face recognition performance both in terms of accuracy and computational time.

2. Quantitative Assessment

Several image quality assessment techniques exist in literature that have shown high correlation with the assessment of human subjects [6, 8]. Also, biometric modality specific quality assessment techniques for fingerprint and iris exist in literature. This research focuses on computationally simple quality metrics that possess intuitive relevance and high correlation with face recognition accuracy are considered in this research, namely, *No-reference quality assessment* (Q1) [8] and *Edge spread measure* (Q2) [6].

- **No-reference quality assessment**: Perceptual image quality such as noise, blur, and compression are measured by computing multiple - computationally inexpensive metrics. The no-reference perceptual quality assessment algorithm proposed by Wang *et al.* [8] is a combination of blockiness and activity estimation in both horizontal and vertical direction.

  Blockiness is estimated by average intensity difference between block boundaries of the image x. For an image of size $M \times N$, the blockiness in horizontal direction (B_h) is given by Eq. 1:

$$B_h = \frac{1}{M([N/8]-1} \sum_{i=1}^{M} \sum_{j=1}^{[N/8]-1} |d_h(i,8j)| \quad (1)$$

  where d_h is the differentiating signal in horizontal direction $d_h(m,n) = x(m,n+1) - x(m,n)$ for $n \in [1, N-1]$.

  Activity of the image provides insight on the effects of compression and blur in the image. Activity (A_h) of an image of size $M \times N$ is given by:

$$A_h = \frac{1}{7}\left\{\frac{8}{M(N-1)} \sum_{i=1}^{M} \sum_{j=1}^{N-1} |d_h(i,j)| - B_h\right\} \quad (2)$$

  Activity of an image may also be measured via zero-crossing rate of the image of size $M \times N$ and it is given by:

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^{M} \sum_{j=1}^{N-2} z_h(m,n) \quad (3)$$

  where,

$$z_h(m,n) = \begin{cases} 1 & if\ horizontal\ ZC\ at\ d_h(m,n) \\ 0 & otherwise \end{cases}, \quad (4)$$

  and $n = [1, N-2]$. Similarly, blockiness, activity and zero crossing rate are measured in the vertical direction as B_v, A_v and Z_v. The overall estimation of values B, A, and Z are given by:

$$B = \frac{B_h + B_v}{2}, \quad A = \frac{A_h + A_v}{2}, \quad Z = \frac{Z_h + Z_v}{2} \quad (5)$$

Combining blockiness, activity and zero-crossing rate provides a quality score of the image. These individual estimates are combined as follows,

$$S = \alpha + \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3} \quad (6)$$

where the parameters α, β, γ_1, γ_2 and γ_3 are model parameters that must be estimated for a given data set. Experimentally, for the database used in this research, these values are found to be $\alpha = -245.89$, $\beta = 261.94$, $\gamma_1 = -239.89$, $\gamma_2 = 160.17$ and $\gamma_3 = 64.29$.

- **Edge spread measure**: Marziliano *et al.* [6] proposed edge spread as a measure to estimate irregularities based on edges and their adjacent regions. Specifically, it computes the effect of irregularity in an image based on the difference in image intensity with respect to the local maxima and minima of pixel intensity at every row of the image. Edge spread can be computed in horizontal as well as vertical directions. However, the experiments in [6] show that either of the two directions suffices for quality assessment.

In the proposed framework, both quality scores are provided to the SVM classifier as a 2D quality vector.

3. Image Denoising

Intuitively, denoising a noisy face image improves the face recognition performance, provided the right set of parameters are used. In the proposed quality assessment based denoising framework, wavelet based soft thresholding technique is used for denoising, also known as BayesShrink [1].

3.1. BayesShrink

BayesShrink [1] is an adaptive, data-driven wavelet thresholding approach for image denoising. The wavelet thresholds are derived from the Bayesian approach assuming that the data follows a generalized Gaussian distribution (GGD). This assumption is based on empirical findings that any natural image can be summarized by a GGD. From this assumption the mean square error (MSE) for each wavelet sub-band is modeled as a Bayesian squared error with known priors for each distribution applied independently and identically. Here the idea is to find soft-thresholds that minimize the Bayesian risk. Formally, given an uncorrupt image $f_{i,j}$ of size $M \times N$, the noisy image $g_{i,j}$ can be written as

$$g_{i,j} = f_{i,j} + \epsilon_{i,j} \quad (7)$$

where $i = 1, \cdots, M$, $j = 1, \cdots, N$, $\epsilon_{i,j}$ is independent and identically distributed (iid) noise assumed as normal N(0, σ^2) and independent of image signal $f_{i,j}$. The purpose is to find an estimate $\hat{f_{i,j}}$ of the image $f_{i,j}$ that minimizes

$$MSE(\hat{\mathbf{f}}) = \frac{1}{N^2} \sum_{i,j=1}^{N} (\hat{f_{i,j}} - f_{i,j})^2 \quad (8)$$

Further, eq 7 in matrix form is given by $Y = X + V$, where X and V are independent of each other, hence

$$\sigma_Y^2 = \sigma_X^2 + \sigma^2 \quad (9)$$

Here σ^2 is the actual variance of noise distribution. From the detail sub-bands of wavelet transform, a threshold T is estimated as

$$\hat{\mathbf{T}}(\hat{\sigma}_X) = \hat{\sigma}^2 / \hat{\sigma}_X \quad (10)$$

where, $\hat{\sigma}^2$ is estimated variance of noise obtained from the wavelet transform of Y. Note that soft thresholding keeps the overall Bayesian risk small as compared to hard-thresholding techniques[1]. The denoising algorithm is governed by two parameters: first is choice of mother wavelet and second is number of iterations required to denoise the image.

3.2. Candidate Parameter Set

Several combination of parameters are possible for denoising using BayesShrink algorithm. For a given probe image, it is computationally expensive to sweep the entire parameter space, however it is more feasible to first assess the level of quality and select the most suitable parameter. In this research, it is observed that the performance of BayesShrink, in terms of improving face recognition accuracy, is related to the type of wavelet used and in some cases, the number of iterations. Hence the following subset of parameters are considered as candidate parameter set: Haar Wavelet (P1), Symmlet Wavelet (P2), Duabechies Wavelet (P3), Beylkin Wavelet (P4) (all single iteration) and Symmlet Wavelet with two iterations (P5) that is termed *Symmlet2*. As we will analyze the performance in Section 5.2, while P5 gives the best performance, P1 is computationally least expensive. Hence, correct use of these parameters can improve accuracy as well as computational time.

4. Proposed Quality Assessment based Denoising Framework

The proposed a quality assessment based optimal parameter selection framework uses SVM classification. The training and testing phases of the framework are discussed.

4.1. Training

Training phase of the proposed framework is shown in Figure 3. The training labels for the parameter selection are

[1] For further details of BayesShrink, refer to Chang *et al.*[1].

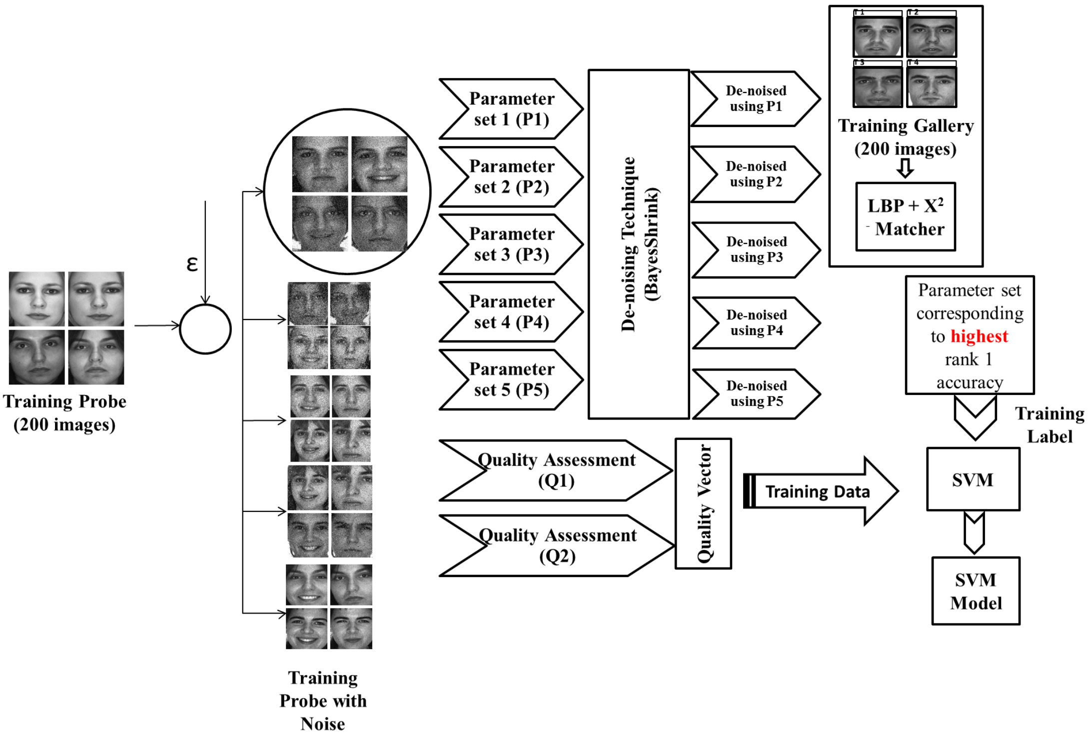

Figure 3. The training scheme of the proposed assessment based denoising framework. The process presented here is for a single noised training probe set; this process is repeated for all the five data driven noises.

generated using the training data which is partitioned into gallery and probe set.

- The images in the training-probe-set are each corrupted systematically by data driven noises, namely, Gaussian(white) noise, local variant white noise, Poisson noise, salt & pepper noise and speckle noise.
- Each of these corrupt training-probe-set are denoised with the wavelet based BayesShrink denoising algorithm[1] with each of the i candidate parameters $P_{1..i}$.
- The quality vector for each image with quality scores $[Q_1, Q_2]$ is computed and used as the training sample for a multi-class SVM classifier. The best parameters of the SVM classifier are converged upon by minimizing the training error via a 10-fold cross validation.
- The class label corresponding to each of the quality vector is the parameter $P_{1..i}$ which results in the best rank-1 efficiency with the training-gallery-set using local binary pattern (LBP) as the face recognition algorithm.
- While Figure 3 illustrates the process for a single noisy training probe set, the process is repeated for all the five data driven noises.

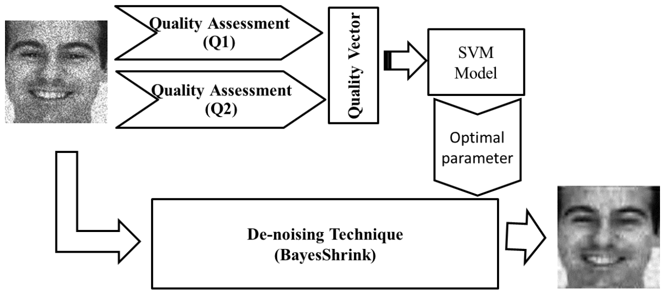

Figure 4. The testing scheme of the proposed assessment based denoising framework

Figure 7 shows a scatter plot of the quality scores of the training data. The illustrated class labels correspond to the best parameters selected for denoising. The classes are well separated, confirming the initial hypothesis that images with a certain set of quality scores require a specific parameter for the best denoising.

4.2. Testing

The trained SVM model is used to select the parameters for denoising, as shown in Figure 4. Given an input (probe) image, the quality vector $[Q_1, Q_2]$ is calculated. Using this quality vector as input, the parameter class obtained from the trained SVM is used to denoise the input image.

Table 1. Different noise artifacts used in the experiments with their associated parameters.

Corruption	Parameters
Gaussian noise	$\sigma = 0.01$
Localvar noise	Dependent on local intensity
Poisson noise	λ=1
Salt and Pepper noise	d = 0.05 or 5%
Speckle noise	v = 0.05

5. Experimental Results

The experiments are conducted on the AR face database [5] containing 756 frontal face images pertaining to 126 subjects (i.e. six images per subject). From this dataset, images corresponding to 50 subjects are chosen for training and the remaining are in the testing set. For experimental purposes, different noise artifacts are synthesized for each image. The kernel parameters used to introduce these artifacts in the images are indicated in Table 1. Before evaluating the proposed framework, a correlation study is performed to establish that the combination of the two quality assessment scores utilized are indeed indicative of the performance of the face recognition algorithm.

5.1. Correlation Analysis

A correlation experiment is performed to analyze the effect of image quality assessment algorithms on face recognition algorithms.

- The training set which is divided into gallery and probe set, with two images per subject in the gallery and four in the probe is used.
- Each artifact kernel is applied on a new copy of the probe set and the average of no-reference quality score [8] and blur assessment score [6] are computed individually. The recognition accuracy is also computed for each set using LBP based face recognition algorithm [7]. The features are extracted and χ^2 -matcher is used to generate match scores.
- The noise artifacts are applied on the existing probe sets and then again quality assessment scores and recognition rate are computed. This process is repeated 10 times until the images are visually incomprehensible. The noise parameters are chosen by visual inspection starting from barely perceptible to visibly incomprehensible corruption. Note that the gallery set here is uncorrupted.
- Pearson's correlation is computed on the recognition rate and mean quality score pairs. The results are shown in Table 2. This experiments suggest that the quality assessment algorithms are, generally, correlated with the performance of the facial feature extractors. Specifically, if the quality score indicates a degradation in the quality of the sample, there is also degradation in the recognition accuracy.

Table 2. Pearson's correlation between recognition accuracy and mean quality score given by no-reference quality assessment algorithm [8] and edge spread assessment algorithm [6].

	No-reference [8]	Edge spread [6]
Gaussian noise	0.97169	0.6773
Localvar noise	0.67448	0.0835
Poisson noise	0.98479	-0.5926
Speckle noise	0.95135	-0.1864
Salt and Pepper noise	0.98872	0.2594

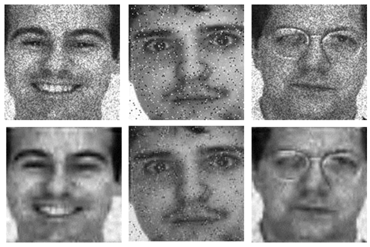

Figure 5. Sample noisy (top) and denoised (bottom) images obtained using the proposed framework.

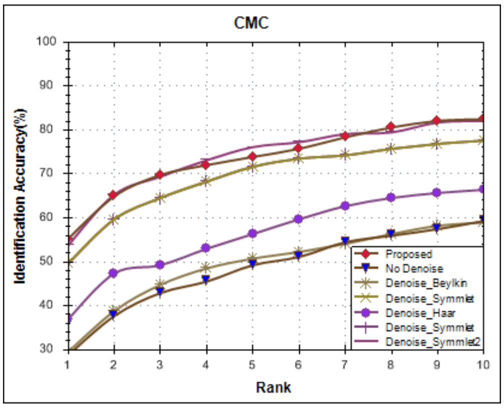

Figure 6. CMC of LBP based face recognition using proposed parameter selection framework and each parameters without selection. The selection framework slightly improves performance and reduces computation time.

5.2. Recognition Experiment

As discussed in Section 4, SVM model is learned using the training labels from the data driven approach on the training set of 50 individuals from the AR face dataset[5]. Figure 5 illustrates denoised output of the proposed algorithm. The scatter plot of the training set is shown in Figure

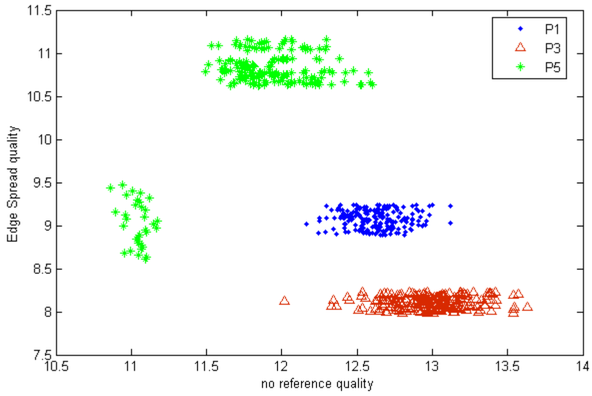

Figure 7. Scatter plot of the quality score of training data. The illustrated class labels correspond to the best parameters selected for denoising. This indicates that images with a certain set of quality scores require a specific parameter for best denoising.

Table 3. Average computation time for denoising an image with each image enhancement parameter and Rank-1 efficiency with testing-gallery-set. All values computed in Matlab with a Dual-core CPU and 2 GB RAM

Parameters	Computation Time	Rank-1 Identification
Haar Wavelet	0.01sec	36.94%
Daubechies	2.77sec	50%
Symmlet	3.87sec	49.62%
Symmlet2	4.41sec	53.73%
Proposed	**3.68sec**	**55.22%**

7. The best performance label corresponded to three of the five parameters.

To evaluate the performance of the framework, random noise is added to probe images in the testing data set. As shown in Figure 6 (CMC curves) and Table 3, the performance of the face recognition algorithm on the images denoised using the proposed approach is comparable to/better than the other denoising approaches. These results show that the proposed framework not only improves the accuracy but also reduces the computational time.

6. Conclusion

This research presents a quality assessment based denoising framework to improve the results of denoising by selecting optimal parameters. The results discussed in this work suggest that no-reference quality and edge spread quality assessment techniques correlate with the recognition performance of automated face recognition systems. Further, noisy images in different parts of the quality space require completely different parameters to result in the best possible denoising process. The proposed framework can easily be extended to a larger set of assessment and enhancement techniques. One limitation of the framework is the large computation time for training, however, we found that the generated prediction model is quite versatile. Large computational hardware will allow for a more comprehensive sweep of the parameter space.

7. Acknowledgment

The authors are thankful to Dr. A. Martinez for providing the AR face database. This research is partially supported by the Department of Defense, United States of America.

References

[1] S. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9(9):1532 –1546, 2000. 169, 171, 172

[2] Y. Chen, S. Dass, and A. Jain. Fingerprint quality indices for predicting authentication performance. In *Audio- and Video-Based Biometric Person Authentication*, volume 3546 of *Lecture Notes in Computer Science*, pages 160–170. Springer Berlin / Heidelberg, 2005. 169

[3] H. Fronthaler, K. Kollreider, and J. Bigun. Automatic image quality assessment with application in biometrics. *Computer Vision and Pattern Recognition Workshop*, 2006. 169

[4] N. Kalka, J. Zuo, N. Schmid, and B. Cukic. Estimating and fusing quality factors for iris biometric images. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 40(3):509 –524, May 2010. 169

[5] A. Martinez and R. Benevento. The ar face database. *CVC Technical Report #24*, 1998. 172, 173

[6] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. A no-reference perceptual blur metric. *International Conference on Image Processing*, pages 57–60, 2004. 170, 171, 173

[7] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002. 169, 173

[8] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600 –612, april 2004. 170, 173