Image Deblurring Methods and Image Quality Evaluation

Vassil Guliashki[1], Dimo Dimov[2]

Abstract – **This paper surveys methods and approaches for digital software image stabilization by algorithmic image deblurring, so that the original form of the image is restored in the best possible way and the image is altered in a sharper, clearer state. Some techniques for image quality evaluation are also considered.**

Keywords – **Software image stabilization, Deblurring methods and approaches, Image quality evaluation.**

I. Introduction

The main sources of blur in a photograph are: motion blur, camera shake and long exposure times. Anyone of these effects leads to worsened quality of the image. Most modern cameras have some built in hardware for image stabilization. This allows the camera to reduce the noise caused by blur in the image without needing to depend on a software level solution. Due to technological limitations, digital cameras are unable to perform software level deblurring on pictures. Advances in image deblurring are important both to the development of modern photography and to the restoration of images and videos that are not as sharp as they can be. Image deblurring has applications in many different real-world problems [8]. The ability to remove noise from images captured in highly technical fields such as astronomy and medicine is critically important for the effective job of related professionals. Deblurring techniques are used also for film restoration [2] and for decoding of bar codes in the supermarkets [3].

Because every picture taken comes out blurry to some degree, image deblurring is fundamental in making pictures sharp and useful [1]. Usually the videos taken from hand held mobile cameras suffer from different undesired and slow motions (handshake, gait wobble, etc.). That's why an improvement of image quality is necessary. The image stabilization techniques can be classified generally as:

- mechanical image stabilization,
- optical image stabilization (OIS), and
- post-processing image stabilization (digital image stabilization).

Mechanical stabilization systems are based on vibration feedback through sensors like gyros accelerometers [20]. The vibrations can be dumped to some extent by gyroscope stabilizers. Unfortunately the same cannot be applied to overcome the forward movement effects. Such problem arises for example when the camera is attached to moving vehicle. Also when the light conditions are bad, the camera needs a long exposure time to gather enough light to form the image and this leads to objectionable blur. A simple way to overcome such problems is to increase the sensitivity of the camera by amplifying the signal from the sensor, which permits faster shutter speed. This may lead to decreasing the image quality because of more noise. Optical image stabilization systems contain either a moving image sensors or an optical element to counteract the camera motion such as a prism or moveable lens assembly that variably adjusts the path length of the light as it travels through the camera's lens system [9], [21]. The use of optical image stabilization allows obtaining a sharp image for shutter speeds 8-16 times slower than without any OIS [10]. This kind of stabilization is not suitable for small camera modules embedded in mobile phones due to lack of compactness and also due to the associated cost, weight and energy consumption. Digital image stabilization tries to smooth and compensate the undesired motion by means of digital video processing. As noted in [22], there are typically three major stages constituting a video stabilization process in the image post processing algorithm: 1) camera motion estimation, 2) motion smoothing or motion compensation, and 3) image restoration.

When deblurring images, a mathematical description of how it was blurred is very important to maximizing the effectiveness of the deblurring process. With real-world photos, we do not have the luxury of knowing the mathematic function by which the image was blurred. However, there exist methods to approximate how blur occurred. Many methods have been developed for image deblurring and there is a big interest in creation of techniques making them more effective (see [10], [22], [23]).

The paper is organized as follows: Section 2 considers some basic and ill-conditioned methods for image deblurring. Section 3 is devoted to the modern deblurring approaches. In section 4 are noted some techniques for image quality evaluation. Finally a summary of main approaches is given.

II. Basic Deblurring Methods

This section considers some basic and ill-conditioned deblurring methods (see [23]).

A. General method for image deblurring

The general method of image deblurring is a direct method for obtaining a blurred or deblurred image from the original or blurred image, respectively. The most basic method is simple matrix multiplication of row and column blurring matrices with the original image. Here is used a very simple linear model for the blurring effects on an image:

$$\mathrm{A}_C \mathrm{X} \mathrm{A}_r^T = \mathrm{B} \qquad (1)$$

[1]Vassil Guliashki is with the Institute of Information and Communication Technologies – BAS, "Acad. G. Bonchev" Str. Bl. 2, 1113 Sofia, Bulgaria, E-mail: vggul@yahoo.com

[2]Dimo Dimov is with the Institute of Information and Communication Technologies – BAS, "Acad. G. Bonchev" Str. Bl. 2, 1113 Sofia, Bulgaria, E-mail: dtdim@iinf.bas.bg

where $\mathbf{A_c}$ and $\mathbf{A_r}$ are the column and row blurring matrices, A_r^T is the transpose, **B** is the blurred image, and **X** is the original image. **B** and **X** are matrices in $\mathbf{R}^{mxn}$ (the set of m-by-n matrices of real numbers), $\mathbf{A_c}$ is in $\mathbf{R}^{mxm}$ and $\mathbf{A_r}$ is in $\mathbf{R}^{nxn}$. It could be expected that the naive solution in this case is $\mathbf{X} = \mathrm{A}_C^{-1}\mathrm{B}(\mathrm{A}_r^T)^{-1}$, but the experiments show that the resulting output does not posses resemblance to the original image.

Because of different extra noisy factors except accidental camera movements, knowing the exact blurring matrices for **X** is not sufficient to restore the image. For this reason an error term **E** needs to be included in the equation:

$$\mathrm{B} = \mathrm{E} + \mathrm{A}_C \mathrm{X} \mathrm{A}_r^T \tag{2}$$

Unfortunately, the value of the noise term is unknown, and so the goal of more sophisticated methods is to minimize the influence of the inverted noise $\mathrm{A}_C^{-1}\mathrm{E}(\mathrm{A}_r^T)^{-1}$. The noise causes the deblurred image to be unrecognizable because its high-frequency components are amplified when the small singular values of the blurring matrices are inverted [1]. The simplest way to reduce this effect is by creating rank-k versions of the row and column matrices from their singular value decompositions:

$$\mathrm{A}_k^{\dagger} = \left[\mathrm{v}_1 \ \ldots \ \mathrm{v}_k\right] \begin{bmatrix} \sigma_1 & & \\ & \ldots & \\ & & \sigma_k \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{u}_1^T \\ \ldots \\ \mathrm{u}_k^T \end{bmatrix} \tag{3}$$

While this method provides a better restoration than the previous solution, the noise still has a large effect on the resulting image using vec**(X)** = $\mathbf{A}_k^{-1}$vec**(B)**. Despite the improvement, this simple approach is insufficient to appropriately reconstruct a blurred image.

B. Point spread function and boundary conditions

In order to increase the accuracy of deblurring functions, a more in-depth look at the blurring matrix **A** must be taken. Understanding the method by which the blurring matrix is generated is the first step to take. Since **A** will blur any original image **X** in the same way, we can imagine applying **A** to a black image with a single pixel of light, known as a point source. The resulting 'image' that describes how **A** acts on individual pixels is called the Point Spread Function (PSF) [1]. Depending upon the type of blurring that is being described, we have different functions that model the PSF around a central point of (*k*, *l*). A few PSFs are given below:

PSF for Out-of-Focus Blur has the form:

$$p_{i,j} = \begin{cases} 1/(\pi r^2) & \text{if} \quad (i-k)^2 + (j-l)^2 \le r^2 \\ 0 & \text{elsewhere} \end{cases} \tag{4}$$

PSF for Atmospheric Blur has the form:

$$p_{i,j} = \exp\left(-\frac{1}{2} \begin{bmatrix} i-k \\ j-l \end{bmatrix}^T \begin{bmatrix} s_1^2 & \rho^2 \\ \rho^2 & s_2^2 \end{bmatrix}^{-1} \begin{bmatrix} i-k \\ j-l \end{bmatrix} \right) \tag{5}$$

PSF for Astronomical Telescope Blur has the form:

$$p_{i,j} = \left(1 + \left(\frac{(i-k)}{s_1} \right)^2 + \left(\frac{(j-l)}{s_2} \right)^2 \right)^{-\beta} \tag{6}$$

The function for atmospheric blur is commonly known as the two-dimensional Gaussian function [4], and the function used for astronomical telescopes is called the Moffat function [5]. By convoluting the generated PSF and the original image **X,** we can generate the blurred image without needing to construct the larger blurring matrix. Convolution can also be used to smooth noise or enhance edges in a picture with an appropriate PSF.

Alone, the PSF only allows us to increase the speed at which we reach a poorly deblurred image. In order to improve the accuracy of image reconstructions, we need to take into account that the image we are looking at does not exist in a vacuum, but rather, extends infinitely beyond the captured area. As such, noise can be generated from outside the boundary of the image. In order to counteract this effect, the use of boundary conditions allows us to make a simple assumption about the state of the world outside of the image's frame.

There are three types of boundary conditions commonly used for this purpose, with different situations where they are used. For each boundary condition, the matrix **X** is placed in the center of a matrix three times as large as it. In the *zero boundary condition*, all other entries are left as zero. Zero boundary conditions assume that all space beyond the image is empty, which is a good assumption to make when working with astronomical data. A *periodic boundary condition* assumes that the image repeats infinitely, and so tiles the original image in each of the blocks of the larger matrix. The *reflexive boundary condition* assumes that the image is mirrored outside of the border, and is created by flipping the original image depending on what block it is in, vertically if it is above or below, horizontally if it is to the left or right, and both if it is diagonal of the middle block where the original image was placed.

For certain states of the PSF and assumptions of the boundary condition, we can use fast algorithms in order to calculate the blurred image and a naive solution from it. A PSF P $\in \mathbf{R}^{mxn}$ is separable if there exists $c \in \mathbf{R}^m$ and $r \in \mathbf{R}^n$ such that P = cr^T. When the PSF is separable, under any boundary condition, we can interpret the blurring matrix **A** as a Kronecker product of two smaller matrices. Using the PSF, we are able to calculate the $\mathbf{A_c}$ and $\mathbf{A_r}$ whose Kronecker product form the blurring matrix **A** without ever having to construct **A.** We can then apply singular value decomposition to these matrices as in earlier techniques to generate the naive reconstruction of an image. Because the SVD of a Kronecker product can be represented in terms of the SVDs of the matrices that form it, we can use these singular value decompositions in order to more efficiently compute the naive solution than if we were to decompose the matrix itself [1].

The formulas for deblurring with Kronecker decomposition and SVD are given as follows:

$$\mathrm{P} = \mathrm{cr}^{\mathrm{T}} \tag{7}$$

$$\mathrm{A} = \mathrm{A_r} \otimes \mathrm{A_c} \tag{8}$$

$$A_r = U_r \Sigma_r V_r^T \quad (9)$$

$$A_c = U_c \Sigma_c V_c^T \quad (10)$$

$$X = V_c \Sigma_c^{-1} U_c^T B U_r \Sigma_r^{-1} V_r^T \quad (11)$$

For PSFs that are doubly symmetric, under the reflexive boundary condition we can apply the two-dimensional (2D) Discrete Cosine Transform (DCT), which computes the eigenvalues of the blurring matrix **A** from the PSF and blurs/deblurs the image without constructing **A** [1]. The 2D DCT is (usually) defined by two 1D DCT, by rows:

$$\hat{x}_k = \omega_k \sum_{j=1}^{n} x_j \cos\frac{\pi(2j-1)(k-1)}{2n}, \quad k=1,2,...n; \quad (12)$$

and similarly by columns, by $n=m$.

The 2D Complex Fourier Transform (CFT) can be applied whenever there is a periodic boundary condition. The respective 1D CFT, by rows has the form

$$\hat{x}_k = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} x_j e^{-\frac{2\pi i(j-1)(k-1)}{n}}, \quad k=1,2,...n; \quad (13)$$

and the similar form by columns, by $n=m$, where $i^2=-1$.

For both the 1D FTs, CFT and DCT, the well known FFT (Fast Fourier Transform) algorithm is usually applied to reduce processing time. The composition of results over both dimensions is commutative and obtains the resulting matrix [4], [6], [7]. These approaches make possible to compute efficiently the blur and naive restoration of large images efficiently, but are still incredibly vulnerable to noise in the data corrupting the image.

C. Spectral analysis

In each of these efficient algorithms for deblurring images, small eigenvalues of the blurring matrix cause a large buildup of inverted noise that corrupts the restoration of the original image. In order to damp noise from the blurred image, a process called spectral filtering in introduced to the deblurring process [1]. The method used here involves generating an m-by-n matrix **Φ,** and element-wise dividing it by the eigenvalue matrix **S.** Truncated Singular Value Decomposition (TSVD) from *Subsection A.* is an example of spectral filtering. An implementation method of this is to only keep singular values that are greater than or equal to a specified tolerance value. The formula of TSVD method for Spectral Filtering has the form:

$$\Phi_i \equiv \begin{cases} 1 & i = 1,\dots,k \\ 0 & i = k+1, \dots, N \end{cases} \quad (14)$$

The method where the small singular (spectral) values are removed is known as Spectral Filtering. The Spectral Filtering is used to eliminate the noise. In the Truncated Singular Value Decomposition (TSVD) based method all singular values below a certain tolerance were removed from the matrix **S.**

Another method for generating the filter factors is the Tikhonov method (see [23], [27]). The Tikhonov method for Spectral Filtering is based on the following equation:

$$\Phi_i \equiv \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \quad (15)$$

The α in the Tikhonov method equation is called the regularization parameter; this parameter behaves in much the same way as the choice of k does in TSVD in damping small eigenvalues that would increase the inverted noise.

The Spectral Filtering methods allow a much better reconstruction of the original image. Unfortunately the quality of the reconstructed image is not good enough, so that more sophisticated methods of image deblurring are required in order to generate more accurate deblurred images.

Some experimental results about the performance of the above listed methods are reported in [23]. All PSFs in the tests were generated using Gaussian (atmospheric) blur with $p = 0$ and $s_1 = s_2$, and all images were manipulated under periodic boundary conditions with the FFT method, which is the fastest for images of any size. The used images are all based on a sharp original image that is blurred through a given method so that the PSF of the blur is known. A picture of fruit bowl on a table was taken with an aperture 5,5 using a tripod for stabilization. The image was tested on blurring and deblurring using a number of different PSF *s*-variation. For the experiments is used the psfGauss function for MATLAB (see http://www2.imm.dtu.dk/~pcha/HNO/), where s=2 by default.

Initially the blurring and deblurring was performed without any noise. It was determined that as the radius of the PSF increased, the noise that was observed in the restored image would increase as well. The observation is done, that even with a Gaussian blur using s=10, which leaves the blurred image such that the fruit bowl is simply an unidentifiable blur, there is very little graininess in the restored image. What noise is noticeable from this deblurring is concentrated around edges in the photo (such as the edges of the table). When the blur is increased to use s=15, the blurred image becomes even less recognizable, and in the full-sized image can be seen a noticeable amount of grain. By convoluting the grainy image with a low pass filter, the graininess from the restored image is smoothed; however, a slightly blurry image is received in exchange. It is noted that an image blurred without any random noise added in can be blurred to a high degree without significant loss or noise in the restored image.

The same approach was used to create the blurred image and then deblur the image with noise, in order to test the sensitivity of the eigenvalue matrix. The fruit bowl image has been blurred with PSF with s = 2. When noise is added the following equation is used:

$$B = B + err \,\|B\|_F\, E, \quad (16)$$

where $\boldsymbol{E}$ is a matrix containing normalized random noise of the same size as $\boldsymbol{B}$, and $\|\boldsymbol{B}\|_F$ is the Frobenius norm of $\boldsymbol{B}$ (here $\boldsymbol{B}$ is the blurred image). The condition number (i.e. $\|\boldsymbol{S}\|.\|\boldsymbol{S}^{-1}\|$) of the eigenvalue matrix **S** in a case of 1% noise reaches $2.68e^{+18}$. This means that even a very small amount of noise is magnified to massive degrees when inverted in the deblurring process. When **S** element-wise divides the Fourier transform of the noisy blurred image, the noise is magnified by an enormous factor, dominating the image. By testing this out with different degrees of error, it has been found that the amount of error should be reduced to $1e^{-12}$ in order for less than half of the image to be dominated by noisy patches. When the noise was reduced by a factor of ten the effect of inverted noise was reduced to a level where the observer must look closer at any section of the restored image to notice the corruption due to inverted noise. Further, an experiment was done with an error factor of $1e^{-14}$ and it has been found that the inverted noise can be purged more or les from the image. In conclusion this deblurring method is insufficient to successfully remove any amount of noise that would occur in a real world situation, so that a technique that allows cancelation of the small singular values that are exploding the inverted noise is necessary.

At the end the spectral filtering is applied to eliminate the noise. Primarily, the method that was used was the Truncated Singular Value Decomposition (TSVD) based method, where all singular values below a certain tolerance were removed from the matrix **S**. It has been established that with the TSVD filtering approach the tolerance needs to be between 0.5 and 0.5 times the error in order to produce a reasonable image. The range is smaller as the amount of blur induced in the image increases; for s=4, a range of 0.05 to the error amount produces a good reconstruction. This implies that the range of tolerances that will allow for the deblurred image to be of good quality will shrink as the blurred image gets blurrier. It is shown, that applying this filtering technique good results (with reasonable image quality) can be obtained with error noise as high as 0.1%. Applying the Tikhonov method to filter the singular values generates results that have less noise than their TSVD counterparts, but are blurrier than them as well. This is due to the different way the singular values are trimmed by the different filters. In conclusion this approach is insufficient to create a truly sharp restored image from the blurry, noise added image. The same is the situation when a naturally blurry image is deblurred using spectral filtering. For image with low noise the tolerance and PSF parameters chosen in the deblurring attempt were 0.1 and 12, respectively. The reconstructed image is sharper than the started image, but its quality is not good enough.

III. Modern Deblurring Approaches

A. Blur Model

The homogenous blurring can be described by *convolution* (see [10]):

$$\mathbf{z} = \mathrm{u} * h[x,y] = \int \mathrm{u}(x-s, y-t)h(s,t)\,ds dt\,, \qquad (17)$$

where **u** is an original image, ***h*** is called the *convolution kernel* or *point-spread function* (PSF) and **z** is the blurred image. In our case of camera motion blur the PSF is a plane curve given by an apparent motion of each pixel during the exposure.

If the focal length of the lens is short or camera motion contains a significant rotational component about the optical axis, this simple model is not valid. The blur is then different in different parts of the image and is a complex function of camera motion and depth of scene [29].

Nevertheless, this spatially varying blur can be described by a more general linear operation:

$$\mathbf{z} = \mathrm{u} *_v h[x,y] = \int \mathrm{u}(x-s, y-t)h(x-s, y-t; s,t)\,ds dt\,, \quad (18)$$

where *h*, again called *point-spread function* is the (variable) kernel of this *space-variant convolution.* The subscript *v* distinguishes from ordinary convolution, denoted by asterisk.

Because the rotational component of camera motion is usually dominant, the blur is independent of depth and the PSF changes in a continuous gradual way. Therefore the blur can be considered locally constant and can be locally approximated by convolution. As discussed in [10], this property can be used to efficiently estimate even the space-variant PSF.

B. Deblurring approaches

Various techniques have been proposed for stabilizing videos taken under different environment from different camera systems. Below are listed some modern deblurring approaches:

- Multiple underexposed / noisy images

The simplest way to avoid camera motion blur is to take a sequence of underexposed images so that the exposure time is short enough to prevent blurring. After registration, the whole sequence can be summed to get the original sharp image with a reasonable noise level. Unfortunately this idea turns out to be impractical for more than a few images because of the time needed for sensor read-out (see [10]).

- Blind restoration from single blurred image (deconvolution)

The blur is usually assumed to be homogenous in the whole image for simplicity. In this case the blur can be modeled by convolution. That is why the reverse problem to find the sharp image is called *deconvolution.* If the PSF is not known, which is the case in most real situations, the problem is called *blind deconvolution*.

The blind deconvolution problems from a single image are very hard solvable in contrast to the non-blind deconvolution problems, which can be easily solved. To find a stable solution some additional knowledge is required. The most common approach is regularization, applied both on the image and blur. Regularization terms mathematically describe *a priori* knowledge and play the same role as prior distributions in stochastic models. Good published blind deconvolution methods are those of Fergus et al [25], Shan et al [26], as well as Mignotte [24].

- Multiple blurred images (deconvolution)

This approach is extensively studied at the present time. The idea is to use multiple images capturing the same scene but blurred in a different way. The camera takes two or more successive images and each exhibits different blurring due to the basically random motion of the photographer's hand or, for example, aircraft vibrations. Multiple images permit estimation of the blurs without any prior knowledge of their shape, which is hardly possible in single image blind deconvolution [28].

- One correctly exposed but blurred and one underexposed image

This is a particular case of multi-image setup. This approach is most advantageous and attracted considerable attention only recently. Taking images with two different exposure times (long and short) results in a pair of images, in which one is sharp but underexposed and another is correctly exposed but blurred. Instead of the underexposed image we can equivalently take an image with high increase the sensitivity of a camera (ISO). Both can be easily achieved in continuous shooting mode by exposure and ISO bracketing functions of DSLR (Digital Single-Lens Reflex) cameras. As noted in [10] for Canon compact cameras these functions can be written in the scripting language implemented within the scope of the CHDK project (http://chdk.wikia.com/wiki/CHDK).

- Other methods for image stabilization

Various 2D stabilization algorithms are presented in [11 and 12]. Hansen *et al.* [11] describe the implementation of an image stabilization system based on a mosaic-based registration technique. Burt and Adelson [13] propose a multi-resolution spline technique for combining two or more images into a larger image mosaic. They describe a system which uses a multi-resolution, iterative process that estimates affine motion parameters between levels of Laplacian pyramid of images. From course to fine levels, the optical flow of local patches of the image is computed using a cross-correlation scheme. The motion parameters are then computed by fitting an affine motion model to the flow.

Matsushita *et.al* [14], proposed in 2006 the direct pixel based full frame video stabilization approach using hierarchical differential motion estimation with Gauss Newton minimization. The Gauss error functions are minimized iteratively to find the optimized motion parameters. After motion estimation, motion inpainting is used to generate full frame video. This method performed well in most videos except in those cases when large portion of video frame is covered by a moving object, because this large motion makes the global motion estimation unstable. R. Szeliski, [15] presented in 2006 a survey on image alignment to explain the various motion models, and also presented a good comparison of pixel based direct and feature based methods of motion estimation. The efficiency of the feature based methods depends upon the feature point's selection [16]. Rong Hu, *et al* [17] proposed in 2007 an algorithm to estimate the global camera motion with SIFT (Scale Invariant Feature Transform). These SIFT features have been proved to be affine invariant and used to remove the intentional camera motions. Derek Pang *et al* [18] proposed in 2010 the video stabilization using Dual-Tree complex wavelet transform (DT-CWT). This method uses the relationship between the phase changes of DT- CWT and the shift invariant feature displacement in spatial domain to perform the motion estimation. Optimal Gaussian kernel filtering is used to smoothen out the motion jitters. This phase based method is immune to illumination changes between images, but this algorithm is computationally complex.

The feature-based approaches are even faster than the direct pixel based approaches, but they are more prone to local effects and their efficiency depends upon the selection of feature points. The direct pixel based approaches use optimally the information available in motion estimation and image alignment, since they measure the contribution of every pixel in the video frame. Hence, the direct pixel based approaches can be used for aligning the sequence of the frames in a video. Hierarchical motion estimation can be used to further improve the stabilization efficiency [9].

IV. Image Quality Evaluation

To estimate the sharp of an image, two different ideas were proposed in the literature. The first one adjusts the contrast of the underexposed image to match the histogram of the blurred one [30]. However, this technique is applicable only if the difference between exposure times is small. The second way [31], [32] uses the image pair to estimate the blur and then deconvolves the blurred image. This path was followed by [33], where the authors show an effective way to suppress ringing artefacts produced by Richardson-Lucy deconvolution. An algorithm of this type is proposed in [10]. It is designed for space-variant blur and may be applied even for wide angle lenses.

In [19] Wang and Li consider perceptual image quality assessment (IQA) algorithms. These algorithms have a common two-stage structure: 1) local quality/distortion measurement, and 2) pooling. In the first stage, image quality/distortion is evaluated locally, where the locality may be defined in space, scale (or spatial frequency) and orientation. For example, spatial domain methods such as the mean squared error (MSE) and the structural similarity (SSIM) index [37], [38] compute pixel- or patch-wise distortion/quality measures in space, while block-discrete cosine transform [39] and wavelet-based [40]-[44] approaches define localized quality/distortion measures across scale, space and orientation. Such localized measurement approaches are consistent with our current understanding about the human visual system (HVS), where it has been found that the responses of many neurons in the primary visual cortex are highly tuned to the stimuli that are "narrow-band" in frequency, space and orientation [45]. The local measurement process typically results in a quality/distortion map defined either in the spatial domain or in the transform domain (e.g., wavelet subbands).

It is expected that IQA methods can automatically predict human behaviours in evaluating image quality [34]-[36]. The researchers have achieved significant progress in measuring local image quality/distortion, but the pooling stage remains not good understood. The potential of spatial pooling has been demonstrated by experimenting with different pooling strategies [46] or optimizing spatially varying weights to maximize the correlation between objective and subjective image quality ratings [47]. A common hypothesis underlying nearly all existing schemes is that the pooling strategy should be correlated with human visual fixation or visual region-of-interest detection. This is supported by a number of interesting recent studies [47]-[49], where it has been shown that sizable performance gain can be obtained by combining objective local quality measures with subjective human fixation or region-of-interest detection data. In practice, however, the subjective data is not available, and the pooling stage is often done in simplistic or ad-hoc ways, lacking theoretical principles as the basis for the development of reliable computational models.

The existing pooling approaches can be categorized generally in four groups [19]. They are briefly discussed below:

- *Minkowski pooling*

Let q_i be the local quality/distortion value at the i-th location in the quality/distortion map. The Minkowski summation is given by:

$$Q = \frac{1}{N}\sum_{i=1}^{N} q_i^p \quad (19)$$

where N is the total number of samples in the map, and p is the Minkowski exponent. To give a specific example, let q_i represent the absolute error, then (19) is directly related to the l_p norm (subject to a monotonic nonlinearity). As special cases, p = 1 corresponds to the mean absolute error (MAE), and p = 2 to the mean squared error (MSE). As p increases, more emphasis is shifted to the high distortion regions. Intuitively, this makes sense because when most distortions in an image is concentrated in a small region of an image, humans tend to pay more attentions to this low quality region and give an overall quality score lower than direct average of the quality map [36]. In the extreme case $p = \infty$, it converges to $\max_i\{p_i\}$, i.e., the measure is completely determined by the highest distortion point. In practice, the value of p typically ranges from 1 to 4 [38]-[43]. In [36], it was shown that Minkowski pooling can help improve the performance of IQA algorithms, but the best p value depends upon the underlying local metric q_i and there is no simple method to derive it.

- *Local quality/distortion-based pooling*

The intuitive idea that more emphasis should be put at high distortion regions can be implemented in a more straight-forward way by local quality/distortion-based pooling. This can be done by using a non-uniform weighting approach, where the weight may be determined by an error visibility detection map [54]. It may also be computed using the local quality/distortion measure itself [36], such that the overall quality/distortion measure is given by:

$$Q = \left(\sum_{i=1}^{N} w(q_i) q_i\right) \Big/ \left(\sum_{i=1}^{N} w(q_i)\right) \quad (20)$$

where the weighting function $w(\cdot)$ is monotonically increasing when q_i is a distortion measure (i.e., larger value indicates higher distortion), and monotonically decreasing when q_i is a quality measure (i.e., larger value indicates higher quality). Another method to assign more weights to low quality regions is to sort all q_i values and use a small percentile of them that correspond to the lowest quality regions. For example, in [55] and [56], the worst 5% or 6% distortion values were employed in computing the overall quality scores. Local quality/distortion-based pooling has been shown to be effective in improving IQA performance, as reported in [36], [56], though the implementations are often heuristic (for example, in the selection of the weighting function $w(\cdot)$ and the percentile), without theoretical guiding principles.

- *Saliency-based pooling*

Here the "saliency" is used as a general term that represents low-level local image features that are of perceptual significance (as opposed to high-level components such as human faces). The motivation behind saliency-based pooling approaches is that visual attention is attracted to distinctive saliency features and, thus, more importance should be given to the associated regions in the image. A saliency map $\{w_i\}$, created by computing saliency at each image location, can be used as a visual attention predictor, as well as a weighting function for IQA pooling as follows:

$$Q = \left(\sum_{i=1}^{N} w_i q_i\right) \Big/ \left(\sum_{i=1}^{N} w_i\right) \quad (21)$$

Given an infinite number of possible saliency features, the question is what saliency should be used to create w_i. This can range from simple features such as local variance [36] or contrast [57] to sophisticated computational models based upon automatic point of gaze predictions from low-level vision features [56], [58]-[61]. It has also been found that motion information is another useful feature to use in the pooling stage of video quality assessment algorithms (see [50], [62], [63]).

- *Object-based pooling*

Different from low-level vision based saliency approaches, object-based pooling methods resort to high-level cognitive vision based image understanding algorithms that help detect and/or segment significant regions from the image. A similar weighting approach as in (21) may be employed, just that the weight map w_i is generated from object detection or segmentation algorithms. More weights can be assigned to segmented foreground objects [64] or on human faces [63], [65]-[67]. Although object-based weighting has demonstrated improved performance for specific scenarios (e.g., when the image contains distinguishable human faces), they may not be easily applied to general situations where it may not always be an easy task to find distinctive objects that attract visual attention.

In summary, all of the previous pooling strategies are well motivated and have achieved certain levels of success. Combinations of different strategies have also shown to be a useful approach [56], [62], [63], [67]. However, the existing pooling algorithms tend to be ad-hoc, and model parameters are often set by experimenting with subject-rated image databases. What are lacking are not heuristic tricks but general theoretical principles that are not only qualitative sensible but also quantitative manageable, so that reliable computational models for pooling can be derived.

Wang and Li [19] have proposed an information theoretic pooling method. The approach they use is saliency-based. The resulting weighting function has interesting connections with the proposed pooling method. The same authors have tested in their work the hypothesis that when viewing natural images, the optimal perceptual weights for pooling should be proportional to local information content, which can be estimated in units of bit using advanced statistical models of natural images. It was found first, that the information content weighting leads to consistent improvement in the performance of IQA algorithms. Second, with information content weighting, even the widely criticized peak signal-to-noise-ratio can be converted to a competitive perceptual quality measure when compared with state-of-the-art algorithms. Third, the best overall performance is achieved by combining information content weighting with multiscale structural similarity measures. There is a general belief, that the human vision system is an optimal information extractor, and this is a widely assumed conception in the computational vision science [53]. Information theoretic methods are relatively new for IQA. The visual information fidelity (VIF) method [51] is one such successful method, but VIF was not originally proposed for pooling purpose. In [50], based upon statistical models of Bayesian motion perception [52], motion information content and perceptual uncertainty were computed for video quality assessment. In [36], simple local information-based weighting demonstrated promising results for improving IQA performance. The information content weighting method proposed in [19] is built upon advanced statistical image models and combines them with multiscale IQA methods. The result is superior performance in extensive tests using six independent databases, so that the general hypothesis mentioned above obtains strong support.

V. SUMMARY

In this paper are reviewed methods for and approaches to software stabilization of still images in sense of removing blur, caused by camera motion during exposure time. Basic deblurring methods as well as modern deblurring approaches are considered. Finally some techniques for image quality evaluation are surveyed.

The modern deblurring approaches are the following:

- Avoiding blur from the beginning by taking a sequence of underexposed images. This idea is impractical because of the time needed for sensor read-out.

- Deblurring from a single image. The disadvantages are speed and difficulties with the segmentation of moving objects.

- Deconvolution from a sequence of blurred images. The main disadvantage of this kind of methods is speed.

- The most perspective approach is to use a pair of images, one blurred and one underexposed. Its main advantages are good speed, reliability, ability to deal with space-variant blur and the potential to segment moving objects.

ACKNOWLEDGEMENT

This study is partly supported by the project № BG161PO003-1.1.06-0038 to the EU operative program „Development of Bulgarian economy competitiveness” and by the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).

REFERENCES

[1] P. C. Hansen, J. Nagy, and D. O'Leary, *Deblurring Images: Matrices, Spectra and Filtering*, SIAM, Philadelphia, 2008.

[2] L. Joyeux, S. Boukir, B. Besserer and O. Buisson. Reconstruction of degraded image sequences. Application to film restoration. *Image and Vision Computing*, 19:504-516, 2000.

[3] T. Wittman. Lost in the supermarket: Decoding blurry barcodes. *SIAM News*, 37(7): 16, September 2004.

[4] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1989.

[5] A. F. J. Moffat. A theoretical investigation of focal stellar images in the photographic emulsion and applications to photographic photometry. *Astronom. Astrophys.* 3:455-461, 200.1969.

[6] P. J. Davis. *CirculantMatrices.* Wiley, New York, 1979.

[7] C. F. Van Loan. *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia, 1992.

[8] S. Reeves, Image Deblurring - Introduction. 2007, http://blogs.mathworks.com/steve/2007/08/13/image-deblurring-introduction/

[9] M. Sorel, Recent advances in deblurring and image stabilization. Czech Academy of Sciences, 2008, http://zoi.utia.cas.cz/files/ASI2008slides.pdf.

[10] M. Sorel, F. Sroubek and J. Flusser, Recent advances in space-variant deblurring and image stabilization. *Unexploded ordnance detection and mitigation*, (Editor: J. Byrnes), Springer Science + Business Media B.V., 2009, pp. 259-272.

[11] M. Hansen, P. Anandan, K. Dana, G. van der Wal, and P.J. Burt. "Real time scene stabilization and mosaic construction". *Int. Proc. DARPA Image understanding Workshop,* pages 457-465, Monterey, CA, November 1994.

[12] F. Liu, M. Gleicher, H. Jin, A. Agarwala, "Content-Preserving Warps for 3D Video Stabilization," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2009)*, vol. 28 No. 3, 2009, pages. 44:1-44:9.

[13] P. J. Burt and E. H. Adelson, “A multiresolution splin with application to image mosaics”, *ACM Transactions on Graphics*, Vol. 2, No. 4, October 1983, pp.217-236.

[14] Y. Matsushita, E. Ofek, W.Ge, X.Tang, and H.Y.Shum, "Full frame video Stabilization with motion inpainting", *IEEE Transactions on Pattern Analysis and Machine Intellig.*, vol. 28 No. 7, July 2006, pages 1150-1163.

[15] R. Szeliski, "Image Alignment and Stitching: A Tutorial," Technical Report MSR-TR- 2004-92, Microsoft Corp., 2004.

[16] O. Adda, N. Cottineau, M. Kadoura, "A Tool for Global Motion Estimation and Compensation for Video Processing" LEC/COEN 490, Concordia University, May 5, 2003.

[17] R. Hu, R. Shi, I. Shen, W. Chen2 "Video Stabilization Using Scale Invariant Features". 11thInt. Conf. Information Visualization IV'07 IEEE 2007.

[18] D. Pang, H. Chen and S. Halawa, "Efficient Video Stabilization with Dual-Tree Complex Wavelet Transform", EE368 Project Report, Spring 2010.

[19] Z. Wang and Q. Li, “Information Content Weighting for Perceptual Image Quality Assessment”, *IEEE Transactions on image processing*, Vol. 20, Issue: 5, May 2011, pp. 1185-1198

[20] Multimedia, "Use Image Stabiliz. for Gyroscopic Stabilizer", URL: http://www.websiteoptimization.com/speed/tweak/stabilizer

[21] C. Morimoto, R. Chellappa, "Fast electronic digital image stabilization" Proc. 13^{-th} International Conference on Pattern Recognition, 25-29 August 1996, vol. 3, pp. 284-288

[22] P. Rawat, J. Singhai, "Review of motion estimation and video stabilization techniques for hand held mobile video", *Signal & Image Processing*, 2011, Vol. 2, № 2, ISSN / EISSN: 222939220976710X (Editor: Academy & Industry Research Collaboration Center (AIRCC)) pp.159-168.

[23] S. Jaffe, “Image Deblurring”, Major Qualifying Project Report submitted to the Faculty of the Worcester Polytechnic Institute, Prof. Homer Walker – major Advisor, 29.12.2013, https://www.wpi.edu/Pubs/E-project/Available/E-project-123013-150749/unrestricted/Image_Deblurring.pdf

[24] M. Mignotte, “Fusion of regularization terms for image reconstruction”, *Journal of Electronic Imaging*, Vol. 19 (3), 033004, (Jul-Sep 2010). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.186.684&rep=rep1&type=pdf

[25] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graphics, SIGGRAPH 2006 Conf. Proc., Boston, MA*, 25:787-794, 2006.

[26] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics (SIGGRAPH)*,v. 27 (3), August 2008, doi: 10.1145/1360612.1360672

[27] A. Tikhonov and V. Arsenin. *Solution of Ill-Posed Problems.* Wiley, New York, 1977.

[28] F. Sroubek and J. Flusser. Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Process.*, 14(7):874-883, July 2005.

[29] M. Sorel and J. Flusser. Space-variant restoration of images degraded by camera motion blur. *IEEE Trans. Image Process.*, 17(2):105-116, February 2008.

[30] Q. R. Razligh and N. Kehtarnavaz. Image blur reduction for cell-phone cameras via adaptive tonal correction. pages I: 113-116, 2007.

[31] S. H. Lim and D. A. Silverstein. Method for deblurring an image. *US Patent Application*, Pub. No. US2006/0187308 A1, August 24 2006.

[32] M. Tico, M. Trimeche, and M. Vehvilainen. Motion blur identification based on differently exposed images. In Proc. IEEE Int. Conf. Image Process., pp. 2021-2024, 2006.

[33] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, p. 1, ACM, New York, 2007.

[34] T. N. Pappas, R. J. Safranek, and J. Chen,, A. Bovik, Ed., "Perceptual criteria for image quality evaluation," in *Handbook of Image and Video Proc.*, 2nd ed. New York: Academic, 2005.

[35] H. R. Wu and K. R. Rao, Eds., *Digital Video Image Quality and Perceptual Coding*, Boca Raton, FL: CRC Press, 2005.

[36] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment*. San Rafael, CA: Morgan & Claypool Publishers, March 2006.

[37] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett*., vol. 9, no. 3, pp. 81-84, March 2002

[38] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[39] A. B. Watson, "DCTune: A technique for visual optimization of DCT quantization matrices for individual images," in *Proc. Soc. Inf. Display Dig. Tech. Papers*, 1993, vol. XXIV, pp. 946-949.

[40] R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., May 1989, pp. 1945-1948.

[41] S. Daly, A. B. Watson, Ed., "The visible difference predictor: An algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, Cambridge, MA: MIT Press, 1993, pp. 179-206.

[42] J. Lubin, , A. B. Watson, Ed., "The use of psychophysical data and models in the analysis of display system performance," in *Digital Images and Human Vision*, Cambridge, MA: MIT Press, 1993, pp. 163-178.

[43] P. C. Teo and D. J. Heeger, "Perceptual image distortion," in Proc. IEEE Int. Conf. Image Process., 1994, pp. 982-986.

[44] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Trans. Image Process.*, vol. 6, no. 8, Aug. 1997, pp. 1164-1175.

[45] B. A. Wandell, Sinauer Associates, Inc., *Foundations of Vision*, 1995.

[46] Z. Wang and X. Shang, "Spatial pooling strategies for perceptual image quality assessment," in Proc. IEEE Int. Conf. Image Process., Atlanta, GA, Oct. 2006, pp. 2945-2948.

[47] E. C. Larson and D. M. Chandler, "Unveiling relationships between regions of interest and image fidelity metrics," *Proc. SPIE Vis. Commun. Image Process.*, vol. 6822, pp. 6822A1-16, Jan. 2008.

[48] E. C. Larson, C. T. Vu, and D. M. Chandler, "Can visual fixation patterns improve image fidelity assessment?," in Proc. IEEE Int. Conf. Image Process., Oct. 2008, vol. 3, pp. 2572-2575.

[49] U. Engelke, V. X. Nguyen, and H.-J. Zepernick, "Regional attention to structural degradations for perceptual image quality metric design," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Mar.-Apr. 2008, pp. 869-872.

[50] Z. Wang and Q. Li, "Video quality assessment using a statistical model of human visual speed perception," *J. Opt. Soc. Amer. A*, vol. 24, pp. B61-B69, Dec. 2007.

[51] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Process., vol. 15, no. 2, pp. 430-444, Feb. 2006.

[52] A. A. Stocker and E. P. Simoncelli, "Noise characteristics and prior expectations in human visual speed perception," *Nature Neurosci.*, vol. 9, pp. 578-585, 2006.

[53] E. P. Simoncelli and B. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, pp. 1193-1216, May 2001.

[54] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag*, vol. 19, pp. 011006:1-011006:21, Jan.-Mar. 2010.

[55] S. Wolf and M. H. Pinson, "Spatio-temporal distortion metrics for in-service quality monitoring of any digital video system," *Proc. SPIE*, vol. 3845, pp. 266-277, 1999.

[56] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 193-201, Apr. 2009.

[57] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer.*, vol. 7, pp. 2032-2040, Oct. 1990.

[58] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* , vol. 20, no. 11, pp. 1254-1259, Nov. 1998.

[59] E. Peli, "Feature detection algorithm based on a visual system model," *Proc. IEEE*, vol. 90, no. 1, pp. 78-93, Jan. 2001.

[60] A. Ninassi, O. L. Meur, P. L. Callet, and D. Barbba, "Does where you gaze on an image affect your preception of quality? Applying visual attention to image quality metric" in Proc. IEEE Int. Conf. Image Process., Apr. 2007, vol. 2, pp. 169-172.

[61] U. Rajashekar, A. C. Bovik, and L. K. Cormack, "Gaffe: A gaze-attention fixation finding engine," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 564-573, Apr. 2008.

[62] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process.: Image Commun.*, vol. 19, Special Issue on Objective Video Quality Metrics, pp. 121-132, Feb. 2004.

[63] Z. K. Lu, W. Lin, X. K. Yang, E. P. Ong, and S. S. Yao, "Modeling visual attention's modulatory afterreffects on visual sensitivity and quality evaluation," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1928-1942, Nov. 2005.

[64] W. Osberger, N. Bergmann, and A. Maeder, "An automatic image quality assessment technique incorporating high level perceptual factors," in Proc. IEEE Int. Conf. Image Process., 1998, pp. 414-418.

[65] Z. Wang and A. C. Bovik, "Embedded foveation image coding," *IEEE Trans. Image Process.*, vol. 10, no.10, pp. 1397-1410, 2001.

[66] S. Lee, M. S. Pattichis, and A. C. Bovik, "Foveated video quality assessment," *IEEE Trans. Multimedia*, vol. 4, no. 3, pp. 129-132, Mar. 2002.

[67] Z. Wang, L. Lu, and A. C. Bovik, "Foveation scalable video coding with automatic fixation selection," *IEEE Trans. Image Process.*, vol. 12, no. 2, pp. 243-254, Feb. 2003.