AN EXAMPLE OF SURFACE FAULT ESTIMATION ON THE BASIS OF SUPER-RESOLUTION APPROACH[1]

Wojciech JAMROZIK

Silesian University of Technology at Gliwice, Department of Fundamentals of Machinery Design
18A Konarskiego Str., 44-100 Gliwice, Poland, wojciech.jamrozik@polsl.pl

Summary
The paper presents an image processing (fusion) technique for resolution enhancement. The technique can be applied to detection of product and process faults. Obtaining images with desired resolution is one of the most important stages in a whole vision system applied to industrial applications. A group of techniques which allows to enhance the image spatial resolution is a super resolution. This group differs from simple interpolation approaches. The application of this technique to the surface fault estimation of concrete stones has been presented in the paper. A set of images was acquired on a fully automatic production line between stones forming and curing stages with the use of CCD camera and then processed. Several super resolution methods have been investigated. Afterwards simple analysis of synthetic images was carried out. Provided results of experiments has shown that super resolution approach is promising when image resolution enhancement is necessary for reliable product diagnostic.

Keywords: image fusion, super-resolution, image processing, paving stones.

PRZYKŁAD ESTYMACJI WAD POWIERZCHNIOWYCH Z ZASTOSOWANIEM NADROZDZIELCZOŚCI

Streszczenie
Artykuł prezentuje technikę przetwarzania (fuzji) obrazów pozwalającą na zwiększanie ich rozdzielczości. Technika ta może zostać zastosowana w procesie detekcji uszkodzeń produktów i procesów. Uzyskanie obrazów o żądanej rozdzielczości jest jednym z najważniejszych zadań w systemie wizyjnym zastosowanym w przemyśle. Grupa technik, które pozwalają na zwiększenie rozdzielczości obrazów, nazywana jest nadrozdzielczością i w znacznym stopniu różni się od podejść interpolacyjnych. Zastosowanie tej techniki do estymacji wad powierzchniowych betonowej kostki brukowej została zaprezentowana w artykule. Na w pełni zautomatyzowanej linii produkcyjnej, pomiędzy fazą formowania, a dojrzewania kostek, przy użyciu kamery CCD zarejestrowany został zbiór obrazów. W następnej kolejności zbiór ten został przetworzony. Proste metody analizy obrazów zostały zastosowane do otrzymanych obrazów syntetycznych. Przedstawione wyniki eksperymentów, dowodzą, że nadrozdzielczość jest obiecującą techniką, kiedy wymagane jest zwiększenie rozdzielczości obrazów w procesie diagnozowanie produktów.

Słowa kluczowe: fuzja obrazów, nadrozdzielczość, przetwarzanie obrazów, kostka brukowa.

1. INTRODUCTION

In recent years there has been an increasing interest in detection of product faults at an early stage of the production process. Because the complexity of continuous technical processes has also grown significantly, many advanced techniques have been worked out, including the application of a vision system to fault detection of products, processes as well as machines. Frequently, during simultaneous monitoring of the product and process a direct relationship between machinery wear or faults and properties of a final product can be revealed. This relationship could be identified only in a case when products are observed with the appropriate sensors. In some cases to obtain reliable data that is suitable for further analysis, applications of more sophisticated techniques are required. This problem is clearly noticeable in case of cameras (e.g. CCD) used to product and process observation.

When spatial dimensions of potential symptoms, representing faults on the product surface or in its

[1] All research data was collected at Betra (Racibórz) member of the Awbud Group

geometrical structure, are relatively small in comparison to the whole product dimensions, and moreover it is not certain that the failure appears at a fixed place, the area observed by the camera is often insufficient. In such cases resolution of acquired images is often too low. One of possible solutions is the use of expensive high performance cameras. However, there is also an alternative solution for these kinds of problems. To obtain images with a sharp and unequivocal representation of desired objects, instead of expensive acquisition devices, proper image processing methods can be applied to images of lower resolution.

An example of image processing methods useful in fault detection is a super resolution (SR), which enhance the image spatial resolution. The paper deals with the application of SR to detect some small surface faults, which could be symptoms of machine wear or incorrectnesses of production process parameters.

2. SUPER-RESOLUTION

The super-resolution (SR) approach is a process of generating images characterized by higher resolution in comparison to source images. The source can consist of one or more images. As opposed to interpolation, the images which are generated by super-resolution techniques are characterized not only by a higher pixel number but also by more resolving power. The next advantage of SR methods is minimization of blur, which is a common effect of image interpolation.

There are two main steps all SR approaches. First is the image registration, which should be understood as the process of geometrical aligning of two or more images. The second step is the image reconstruction. It is the fusion of previously aligned images in order to generate new synthetic high resolution image. This process is shown in Fig. 1.

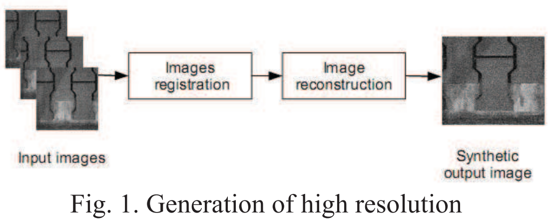


Fig. 1. Generation of high resolution image

2.1. Image registration

In the paper image registration was carried out using Fourier-Mellin transform [3, 13]. It was based on Fourier Shift Theorem [1], which was proposed for registration of translated images, because Fourier Transform (FT) was translation invariant. The Mellin transform is very close to Fourier transform but it uses polynomial kernel. It gives a transform-space image that is invariant to translation, rotation and scale. A similar effect can be obtained using FT when input data is converted to log-polar coordinates, thus scale and rotation are changed to vertical and horizontal of sets can be measured.

2.2. Image reconstruction

Image reconstruction is a phase of the SR method. Previously registrated images are aggregated in a new synthetic image. This image is characterized by higher spatial resolution than input images used in the reconstruction process. Several image reconstruction methods have been taken into consideration. Details of these methods are discussed below.

Iterated backprojection This SR algorithm was introduced by Irani and Peleg in [4]. In this method the key element solving the following equation is defined:

$$g_k(m,n)=\sigma_k\left(h\left(f(x,y)\right)\right)+\eta_k(x,y) \qquad (1)$$

where:

- g_k is an acquired kth image,
- f is a high resolution image after reconstruction,
- h is a blurring operator,
- η_k is an additive noise term,
- σ_k is a downsampling operator. It provides transformation between SR image dimensions and low resolution image dimensions.

The SR image estimation process starts with initial guess $f^{(0)}$ of the high resolution image. Set of low resolution images $g^{(0)}$ is obtained from HR guess. The difference between the original LR image and simulated ones is computed. If $f^{(0)}$ is the ideal guess of the HR image the difference equals to zero. In another case it is used to improve the HR image quality, by backprojecting each value of the difference image into its place on the HR image. The process is repeated iteratively to minimize an error function, given by Eq. 2:

$$e^{(n)}=\sqrt{\sum_k\sum_{(x,y)}\left(g_k(x,y)-g_k^{(n)}(x,y)\right)^2} \qquad (2)$$

Projection onto convex sets In POCS technique [9, 11] blurring caused by camera PSF (point spread function), as well as effects of undersampling are taken into account. The low resolution sequence is described for each frame $g(x, y, k)$. It is assumed that the estimated HR image for frame k is desired at time t_r. For each pixel, within the LR image set a convex set can be defined by Eq. 3:

$$C_{t_r}(x,y,t)=\left\{y(p,q,t_r):\left|r^{(d)}\right|(x,y,k)\le\delta_0\right|\right\} \qquad (3)$$

where:

$$r^{(d)}(x,y,k) \doteq g(x,y,k) - \sum_{(p,q)} d(p,q,t_r) h_{t_r}(p,q;x,y,k) \quad (4)$$

is the residuum, associated with d - member of the constrained set, h_{tr} is the blur PSF and the combination of relative motion of object and camera. The quantity δ_0 reflects statistical confidence of actual image membership to the C_{tr} set. An estimation of the high-resolution version of the reference image is determined iteratively starting from some arbitrary initialization. Successive iterations are obtained by projecting the previous estimated image into the consistency set with an amplitude constraint set that restricts the grey levels of the estimated image to the range [0, 255].

Papoulis-Gerchberg Papoulis-Gerchberg algorithm [2, 7] is a special case of the POCS method. This method is based on two assumptions: some pixel values on a high resolution grid are known, the high frequency components in the high resolution image are equal to zero. To obtain the HR image the hi-res grid is formed. Then after position conversion known pixel values are placed on the high resolution grid. Next, the high frequency components are set to zero in the frequency domain. Then in spatial domain known pixel values are set on the high resolution grid. Setting the high frequency components equal to zero, this method interpolate the unknown values, thus corrects the aliasing for low frequency components. The entire procedure is repeated until all pixel values are found and corrected.

Structure adaptive normalized convolution The method is based on the framework of normalized convolution (NC), in which an image is approximated through projection onto a subspace [8]. The window function of adaptive NC is adapted to local linear structures. This leads to more samples of the same modality that are gathered for the analysis.

The normalized convolution is a method for image analysis that takes into account uncertainties in pixel values and at the same time allows spatial localization of possibly unlimited analysis functions. The projection into the subspace which is spanned by analysis functions is equivalent to a weighted least square problem, where the weights are induced from the certainty of the image and the desired localization of analysis functions. The result of Normalized Convolution is in each image pixel a set of expansion coefficients, one for each analysis function.

The normalized convolution can be considered as a local operator because it works in a finite neighborhood. When the area of application is relatively large the result is blurred. The use of an adaptive applicability function increases the quality of fusion from sparsely sampled data. The applicability function is an anisotropic Gaussian kernel with ability to adapt its shape and orientation to the local image structure. It ensures that only samples with similar intensity and gradient information are used for the local expansion. That avoids blurring along lines and edges.

3. ERROR MEASURES

Two error measures have been proposed in order to assess evaluated image reconstruction algorithms.

3.1. Peak signal to noise ratio

The peak signal to noise ratio (PSNR) is commonly used for image quality assessment [5, 10]. It is based on the mean squared error (MSE), which can be obtained using Eq. 5:

$$MSE = \frac{1}{mn} \sum_{i=1}^{n} \sum_{c=1}^{m} \left(x_{ic} - y_{ic} \right)^2 \quad (5)$$

where: x_{ic} and y_{ic} are values of a pixel i in a channel c of the original and compared image respectively, n is the number of pixels in each channel and is m number of channels. The PSNR can be obtained from MSE and the maximum signal (pixel) value s with the use of Eq. 6:

$$PSNR = 10 \lg \frac{s^2}{MSE} \quad (6)$$

The PSNR is expressed in dB. Higher values indicate that there are lower errors and according to that higher quality.

3.2. Structural similarity

The mean squared error measures have been widely used, but obtained results are characterized by insufficient correlation with the visual degradation [6, 10]. Wang et al. described in [10] the mean structural similarity (MSSIM) error measure. They stated that it has greater correlation with visual degradation than MSE and is not so computational demanding. The structural similarity (SSIM) error measure calculates the similarity in a local window. It combines differences in average, variation and correlation.

The input for SIMM are two intensity value sets, with n elements each, from corresponding windows in the original and compared image. The first averages μ_x and μ_y are calculated using Eq. 7:

$$\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i \quad (7)$$

Next, the calculation of variances σ_x and σ_y for two sets of the intensity values is conducted using Eq. 8:

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{i-1}^{n}\left(x_i - \mu_x\right)^2} \tag{8}$$

The last operation is correlation calculated between two sets of the intensity values using Eq. 9:

$$\sigma_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \mu_x\right)\left(y_i - \mu_y\right) \tag{9}$$

Previously calculated averages, variances and correlation are used to obtain the SSIM (Eq. 10):

$$SSIM_{xy} = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{10}$$

Constants c_1 and c_2 (small values) are used to prevent dividing by zero. SSIM values calculated for all windows are than averaged to obtain the mean structural similarity measure.

MSSIM takes into account the image luminance, contrast and structure comparison measures.

4. CONCRETE STONE PRODUCTION PROCESS

Nowadays, concrete stones are produced by fully automatic machines. Stones are formed on a special reinforced wooden product plate. On the plate a hollow mould is lowered. The whole forming set is laying on a vibrating table. Because each stone consists of two layers of concrete, the more robust in the lower part and more even (smooth) at the top, the mould is filled in two steps. Each of these steps is followed by vibration. That ensures proper mould fulfillment. The concrete in mould is tighten from above with the tightening stamp and the main vibration takes place. Fresh, wet stones are transported to curing racks. The curing racks can be enclosed as well as insulated and equipped with heating systems. Once the curing is completed stones are transported to the dry site of the line and packed. Simplified scheme of the process is presented in Fig. 2.

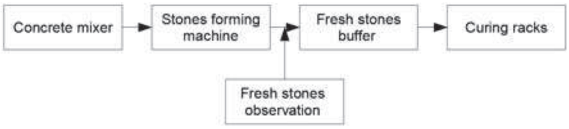


Fig. 2. Paving stones production process scheme

Because constant monitoring of the inner side of mould holes and the lower stamp side is difficult due to limited space and large amount of dust, the crucial place, in which the most of defects of stones (and indirectly of machine) can be identified, is located between the forming machine and the buffer before curing racks. There, the newly formed stones can be measured and their surfaces can be examinated. When some defects, mostly small-sized ones, are identified at this stage, decisions about further correction of process parameters, either presence of machine faults or excessive wear can be taken.

Among many different techniques proper for shape measurements and surface monitoring the choice of the vision system for fresh concrete stones observation is justified. The hardware side of this kind of system is relatively uncomplicated because the camera can be directly connected to the computer using USB or FireWire protocol. The main disadvantage of this approach is, in most cases, the small resolution of obtained images. Because of that some methods for image resolution enhancement are necessary to be applied.

5. IMAGE ACQUSISTION AND PROCESSING

Image acquisition has been performed with the use of a CCD camera. A test stand near the fresh stones conveyor is presented in Fig. 3. The acquisition speed was 30 fps, and the exposure time was set to 1/4000 s. The resolution of acquired images were 1024x768 or 640x480 pixels and it depended on a type of camera used. During the research about 3 millions images, from the side and from above the production plate, were acquired, but only 10000 most valuable ones were processed. To obtain the reference images required for comparison with estimated synthetic images, only a fragment of the production plate (with fresh stones) was observed. Images were downsampled to generate input data for the resolution enhancement process. These downsampled images resolutions were 256x192 and 160x120 pixels, thus the upasampling factor was 4.

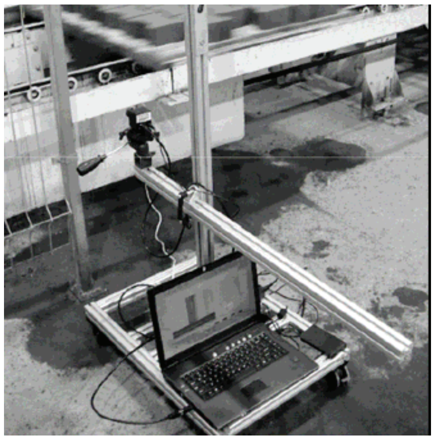
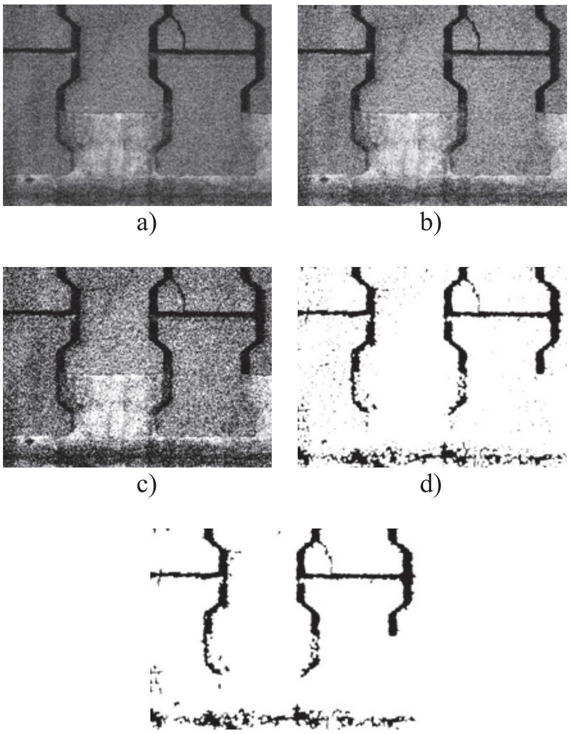


Fig. 3. The image acquisition stand

For the test purposes the normalization of images was performed. It consisted in the equalization of the histogram and gamma correction.

To detect faults on the concrete stone surface a simple processing of the interpolated and super resoluted images was applied. The image before processing is presented in Fig. 4(a). At first, a lookup table was used for histogram equalization (Fig. 4(b)). Then image binarization using adaptive thresholding was accomplished and the result of the operation is shown in Fig. 4(c). Next, the binary image was eroded (Fig. 4(d)). At last small objects were removed, to reduce the noise in the image. The image which was result of the whole processing procedure is presented in Fig. 4(e).



a) b) c) d) e)

Fig. 4. Results of processing of synthetic images at particaular procedure stages

6. RESULTS

In this subsection, the results of four SR reconstruction algorithms described in Sec. 2.2 were compared. The comparison was based on the error metrics described in Sec. 3. In Table 1 the obtained values of PSNR measure are gathered. It was stated that a typical PSNR value for the reconstructed image is greater than 20 dB [12]. PSNR values obtained for two and six frames were nearly equal, and it could be stated that during the reconstruction process the synthetic images of sufficient quality were possible to be generated. The lower value of PSNR obtained when eight frames were taken into consideration could be caused by too high translation between particular frames.

Tab. 1. PSNR values for proposed reconstruction methods

		Frames		
		2	6	8
SR method	Backprojection	17,63	17,72	17,91
	Normalized convolution	19,63	18,63	18,45
	Papoulis-Gerchberg	18,72	19,09	11,91
	POCS	19,63	19,45	19,36

The mean structural similarity values were presented in Table 2. Based on the results in Table 2, it could be stated that the MSSIM values analysis led to similar conclusions as it was in the case of the PSNR. Also it could be perceived that the image normalization led to better reconstruction results, through minimization of the influence of lightening conditions. Owing obtained result one stated that among all tested methods POCS was the most promising. The PSNR and especially MSSIM values confirmed that images reconstructed by the projection onto convex sets have the lowest visual degradation and noise. That made them a sufficient input for further image processing and analysis.

Exemplary results of synthetic image processing are presented in Fig. 5. The main difference between the superresoluted and interpolated images is in the presence of noise.

Tab. 2. MSSIM values for proposed reconstruction methods

			Frames		
			2	6	8
SR method	Real image	Backprojection	0,348	0,348	0,348
		Normalized convolution	0,408	0,400	0,400
		Papoulis-Gerchberg	0,392	0,372	0,101
		POCS	0,380	0,372	0,380
	Normalized image	Backprojection	0,832	0,836	0,836
		Normalized convolution	0,888	0,888	0,888
		Papoulis-Gerchberg	0,876	0,880	0,464
		POCS	0,896	0,892	0,893

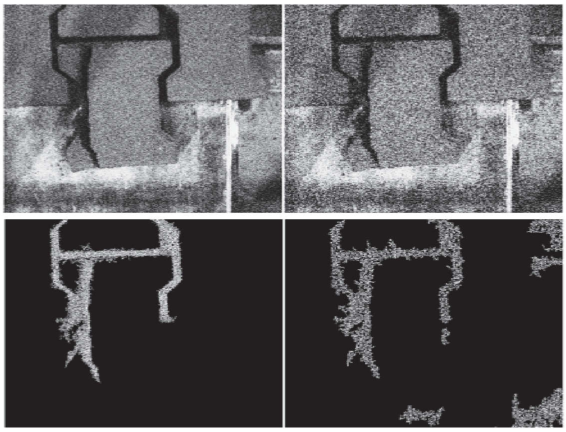


Fig. 5. Results of processing of synthetic images: SR (left), interpolated (right)

When analyzing the superresoluted images the degree of noise is considerably lower as in comparison to the corresponding interpolated one. The noise reduction can lead to better selectivity in fault detection tasks. When dealing with relatively small faults in interpolated images they can be reduced to the noise level and removed. In this manner some valuable information can be lost.

7. CONCULSIONS

As the result of the performed investigations one concluded that SR could be a promising method when sharp and clear small details are required to be found in a low resolution image. In the case of estimation of faults appearing during a production process SR methods provide images with minimized blur, which is a common effect of image interpolation.
The main disadvantage of presented methods is their computational complexity. High reconstruction time (exceeding 30 sec) makes the application in a real production process very difficult. The solution could be implementation of high efficiency and specialized DSP processors or elaboration of dedicated platforms based on FPGA chips.

REFERENCES

[1] Bracewell R. N., editor. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1965.
[2] Gerchberg R. W. *Super-resolution through error energy reduction*. Journal of Modern Optics, 21(9), pp. 709-20, 1974.
[3] Xiaoxin Guo, Zhiwen Xu, Yinan Lu, and Yunjie Pang: *An application of Fourier-Mellin transform in image registration*. In CIT '05: Proceedings of the The Fifth International Conference on Computer and Information Technology, pp. 619-623, Washington, DC, USA, 2005. IEEE Computer Society.
[4] Irani M., Peleg S.: *Improving resolution by image registration.* CVGIP: Graph. Models Image Process., 53(3), pp. 231-239, 1991.
[5] Mrak M., Grgic S., Grgic M.: *Picture quality measures in image compression systems*. In EUROCON 2003. Computer as a Tool. The IEEE Region 8, vol. 1, pp. 233-236, September 2003.
[6] Muresan D. D., Parks T. W.: *Image interpolation using adaptive linear functions and domains.* IEEE 2001 Western New York Image Processing Workshop.
[7] Papoulis A.: *A new algorithm in spectral analysis and band-limited extrapolation*. IEEE Trans. Circuits Syst, vol. 22, pp. 735-742, 1975.
[8] Pham T. Q., van Vliet L. J., Schutte K.: *Robust fusion of irregularly sampled data using adaptive normalized convolution*. EURASIP J. Appl. Signal Process., 2006(1), pp. 236-236, 2006.
[9] Stark H., Oskoui P. *High-resolution image recovery from image-plane arrays, using convex projections.* J. Opt. Soc. Am. A, 6(11), 1715-1726, 1989.
[10] Zhou Wang, Bovik A. C., Sheikh H. R., Simoncelli E. P.: *Image quality assessment: From error measurement to structural similarity*. IEEE Trans. Image Processing, 13, pp.600-612, 2004.
[11] Wheeler F. W., Hoctor R. T., Barrett E. B.: *Super-resolution image synthesis using projections onto convex sets in the frequency domain.* In IS&T/SPIE Symposium on Electronic Imaging, volume 5674 of Conference on Computational Imaging, pages 479_490, January 2005.
[12] Wu H. R. and Rao K. R.: *Digital Video Image Quality and Perceptual Coding (Signal Processing and Communications).* CRC Press, Inc., Boca Raton, FL, USA, 2005.
[13] Zieliński T. P.: *Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań.* W.T, Warszawa, 2007.

Wojciech JAMROZIK is, since 2007, PhD student in the Department of Fundamentals of Machinery Design at Silesian University of Technology at Gliwice. The man areas of his interest are applications of vision system in diagnostic systems, image processing, analysis and recognition. He deals also with the fusion of images and information.