DIGITAL COMPUTATIONAL IMAGING

Yaroslavsky L. P.

School of Electrical Engineering, Department of Physical Electronics,
Faculty of Engineering, Tel Aviv University, Israel

Introduction

Imaging has always been the primary goal of informational optics. The whole history of optics is, without any exaggeration, a history of creating and perfecting imaging devices. Starting more than 2000 years ago from ancient magnifying glasses, optics has been evolving with ever increasing speed through Galileo's telescope and van Leeuwenhoek's microscope, through mastering new types of radiations and sensors to modern wide variety of imaging methods and devices of which most significant are holography, methods of computed tomography, adaptive optics, synthetic aperture and coded aperture imaging, digital holography. The main characteristic feature of this latest stage of the evolution of optics is integrating of physical optics with digital computers. With this, informational optics is reaching its maturity. It is becoming digital and imaging is becoming computational. Present day main trends in digital computational imaging are:

- Development and implementation of new digital image acquisition, image formation and image display methods and devices.
- Transition from digital image processing to real-time digital video processing
- Widening front of research toward to 3D imaging and 3D video communication

We illustrate these trends on an example of three developments: (i) a new family of optics-less image sensors that base their image formation capability solely on numerical processing of radiation intensity measurements made by a set of simple radiation sensors with natural cosine-law spatial sensitivity; (ii) real-time digital video processing for perfecting visual quality and achieving super-resolution of video streams distorted by camera noise and atmospheric turbulence; (ii) 3D video communication paradigm based on the use of computer generated display holograms.

Optics-less image "smart" sensors

One can treat images as data that indicate location in space and intensities of sources of radiation. In conventional optical imaging systems, the task of determining positions of sources of light is solved by lenses and the task of measurement of light source intensities is solved by plane light sensitive sensors, such as photographic films or CCD/CMOS electronic sensor arrays. Lens directs light from each of light sources to a corresponding place of the sensor's plane, and sensor's output signal at this place provides an estimate of the light source intensity.

Lenses are wonderful processors of directional information carried by light rays. They work in parallel with all light sources in their field of view and with the speed of light. However their high perfection has its price. Perfect lenses are large, heavy and very costly. In addition, lenses are not available for many kinds of practically important kinds of radiation such as, for instance, X-rays and radioactive radiation. This motivates search for lens-less imaging devices.

Recently, H. J. Caulfield and the present author suggested a concept of a new family of lens-less "smart" radiation sensors ([1, 2]), which exploits the H. J. Caulfield's idea of combining the natural cosine-law directional sensitivity of radiation sensors with the computational power of modern computers and digital processors to achieve sensor's spatial selectivity not limited by the laws of diffraction. "Smart" sensors consist of arrays of small elementary flat light sub-sensors with the cosine-low angular selectivity supplemented with a signal processing unit. This unit, the sensor's "brain", collects output signals $\{s_n\}$ of all available N elementary sub-sensors and estimates on this base intensities $\{A_k\}$ and angles $\{\vec{\theta}_{n,k}\}$ between

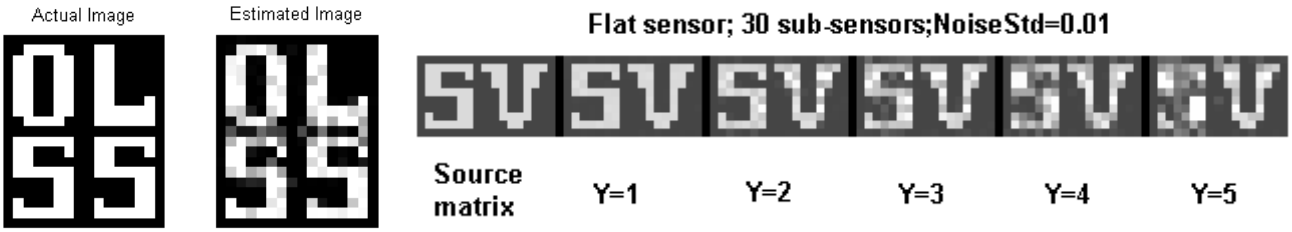
directions to the given number K of distant light sources and normals to the sub-sensor's surfaces. In view of the statistical nature of the random of sensor noise, statistically optimal estimates will be the maximum likelihood estimates. In an assumption that elementary sensor noise components are statistically independent and normally distributed with the same standard deviation, maximum likelihood estimates $\{\hat{A}_k, \hat{\theta}_k\}$ of intensities and directional angles of the sources can be found as solutions of the following equation:

$$\{\hat{A}_k, \hat{\theta}_k\} = \arg\min_{\{\hat{A}_k, \hat{\theta}_k\}} \left\{ \sum_{n=1}^{N} \left| s_n - \left[\sum_{k=1}^{K} \hat{A}_k \, \overline{\cos}(\vec{\theta}_{n,k}) \right] \right|^2 \right\}$$

where $\overline{\cos}(.) = \cos(.)$, if $\cos(.) \geq 0$ and $\overline{\cos}(.) = 0$ otherwise. For a single light source, an analytical solution of this equation is possible, which means that the computational complexity of estimation of intensity and directional angles of a single light source is of the order of the number of sensors and that the computations can be implemented in a quite simple hardware. For larger number K of light sources, solution of this equation requires optimization in K-dimensional space. The computational complexity of estimation of source parameters dramatically grows with the number of sources.

For locating multiple light sources, "smart" sensors can be used in two modes: (i) "localization" mode for localization and intensity estimation of light sources, when only a priori knowledge available is the number of light sources and (ii) "imaging" mode for estimation of intensities of the given number of light sources in the given locations, such as, for instance, in nodes of a regular grid on a certain distance from the sensor. In the latter case, the computational complexity and estimation errors, for a given number of elementary sensors and their noise level, are very substantially lower.

Two illustrative simulation results of localization and imaging of the given number of very distant and very proximal radiation sources by arrays of elementary sensors are presented in Fig. 1, a and b) correspondingly.



a) b)

Fig.1. (a) Imaging of an 19x16 array radiation sources by an optics-less spherical sensor with the number of sub-sensors 300; sub-sensor's SNR=100. (b) Imaging of an array 11x16 of radiation sources by a scanning optics-less flat sensor with 30 elementary sub-sensors for distances of sources from the sensor Y=1 to 5 (in units of inter-sub-sensor distance); sub-sensor's SNR=100.

In conclusion it is interesting to note that quite a number of examples of optics-free vision among live creatures can be found. Obviously, plants, such as sunflowers, that feature heliotropism must have a sort of "skin" vision to determine direction to sun in order to be able to direct their flowers or leaves accordingly. There are also many animals that have extra-ocular photo-reception. One can also find reports on the phenomenon of "skin" vision in humans. (see., for instance, [3]). The described "smart" sensors may cast a light on possible mechanisms of the skin vision phenomenon.

Real time digital processing for video denoising, deblurring and super-resolution

In long distance observation systems, images and video are frequently damaged by atmospheric turbulence, which causes spatially and temporally chaotic fluctuations in the index of refraction of the atmosphere ([4]) and results in chaotic, spatial and temporal geometrical distortions of neighborhoods of all pixels. This geometrical instability of image frames heavily worsens the quality of videos and hampers their visual analysis. To make visual analysis possible, it is required first of all to stabilize images of stable scenes while preserving real motion of moving objects that might be present in the scene.

In Refs. [5], methods of generating stabilized videos from turbulent videos were reported. It was also found that, along with image stabilization, image super-resolution on stable scenes can be achieved ([6]). The core of the stabilization and super-resolution method is elastic pixel wise registration, with sub-pixel accuracy, of available video frames of stable scenes followed by resampling of the frames according to the registration results. For achieving the required elastic registration of frames, a time window of several preceding and following frames of the video sequence is analyzed for each current video frame. For each pixel of the frame, its ***x***-***y*** displacements in all remaining frames of the window are found using the methods of block matching or optical flow methods. Then the displacement data are analyzed to derive their statistical parameters for distinguishing pixels that were displaced due to the atmospheric turbulence from those that belong to real moving objects. This distinction can be made on an assumption that turbulence-induced pixel displacements are relatively small and irregular in time while displacements caused by real movement are relatively large and, what is more important, contain a regular, in time, component.

On the base of these measurements, the output stabilized and super-resolved frame is generated on a sampling grid built by sub-divisions of the initial sampling grid. Nodes of the latter correspond, for each pixel, to mean values of found pixel displacements. Formation of the output frame consists of two steps.

On the first step, corresponding pixels from all time window frames are placed at the nodes of the sub-pixel grid according to their found displacements minus displacement mean values. Because pixel displacements are chaotic, it may happen in this process that two or more corresponding pixels from different frames have to be placed in the same position in the output frame. In these cases, robust to outliers average, such as median, can be taken as a replacement of those pixels.

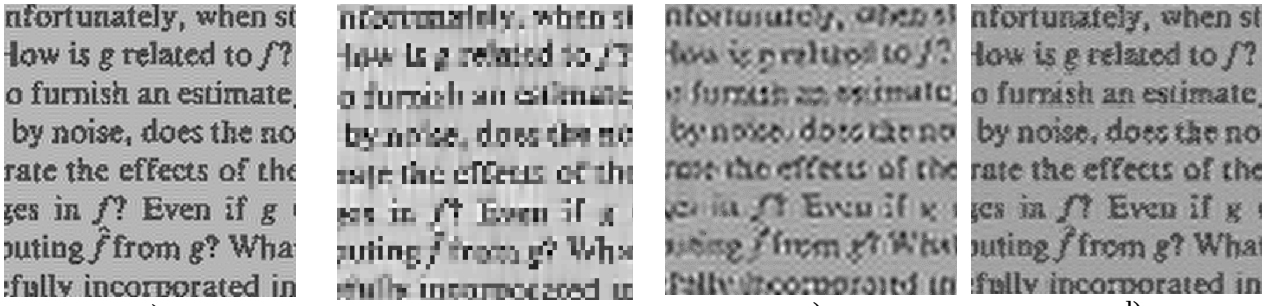
On the second step, sub-pixels that remain empty because of possible shortage of data in the selected time window, should be interpolated from available data. For the interpolation, different available methods for interpolation of sparse data can be used.

Pixels retrieved from the set of the time window frames contain, in form of aliasing components, high frequencies outside the image base-band defined by the original sampling rate of the input frames. Placing them into proper sub-pixel positions results in de-aliasing these frequencies and, therefore, in widening image bandwidth beyond the base band. The more pixels in different sub-sampling positions are available the higher degree of super-resolution achieved. The degree of achievable super-resolution depends on camera optics and fill factor, the number of frames used for data fusion and magnitude of turbulence induced pixel displacements ([7]).

The efficiency of the entire processing is illustrated in Fig. 2 by the results of computer simulations. As one can see from Fig. 8, even from as small number of the frames as 15, a substantial resolution enhancement is potentially possible.

3-D visualization and communication and computer generated display holograms

There are no doubts that the ultimate solution for 3-D visualization is holographic imaging. This is the only method capable of reproducing, in the most natural viewing conditions, 3-D images that have all the visual properties of the original objects including full parallax, and are visually separated from the display device. 3-D visual communication and display can be achieved through generating, at the viewer side, of holograms out of data that contain all relevant information on the scene to be viewed . Digital computers are ideal means for converting data on 3-D scenes into optical holograms for visual perception ([8, 9]).



a) b) c) d)

Fig. 2. Illustrative simulation results of resolution enhancement of turbulent video frames. a)- initial high resolution frame; b) an examples of a low resolution frame distorted by simulated random local displacements with standard deviation 0.5 inter-pixel distance; c) - resolution enhanced frame obtained by the described fusion process from 15 low resolution frames; d) final output frame obtained by interpolation of samples that are missing in frame e).

The core of the 3D digital holographic visual communication paradigm is the understanding that, for 3D visual communication, one does not need to record, at the scene site, a physical hologram of the scene and that the hologram has to be generated at the viewer side. To this goal, one needs to collect, at the scene side, and to transmit to the viewer site a set of data that will be sufficient to generate, at the viewer site, a synthetic hologram of the scene for viewing.

The major requirement to computer-generated display holograms is that they should provide natural viewing conditions for the human visual system and, in particular, separation of reconstructed images from the display device.

A crucial issue in transmitting data needed for the synthesis, at the viewer site, of display holograms is the volume of data to be collected and transmitted and the computational complexity of the hologram synthesis. The upper bound of the amount of data needed to be collected at the scene side and transmitted to the viewer site is, in principle, the full volumetric description of the scene geometry and optical properties. However, a realistic estimation of the amount of data needed for generating a display hologram of the scene is by orders of magnitude lower then the upper bound due to the limitations of the human visual system. This also has a direct impact on the computational complexity of the synthesis of holograms.

Several computationally inexpensive and at the same time quite sufficient solutions for creating 3D visual sensation with synthetic display holograms have been suggested ([8,9]):

- Multiple view compound macro-holograms. In this method, the scene to be viewed is described by means of multiple view images taken from different directions in the required view angle, and, for each image, a hologram is synthesized separately with an account of its position in the viewing angle. Ref. 10 describes an example of synthesis of a compound multiple view macro-hologram. Fig. 3 illustrates viewing such a synthetic compound hologram and an example of the reconstructed images for one particular observation view angle.
- Composite stereo-holograms. This is a special case of multiple view compound macro holograms. Composite stereo holograms are synthetic Fourier holograms that reproduce only horizontal parallax ([11, 12]). When viewed with two eyes, they are capable of creating 3D sensation thanks to stereoscopic vision. With such holograms arranged in a circular composite hologram, full 360 degrees view of the scene can be achieved. An example of such a 360 degrees is stored at MIT museum of holography ([14])
- _"Programmed diffuser" holograms ([8, 13]). This method assumes that objects to be viewed are specified in the object coordinate system by their "macro" shape, by the magnitude of the object reflectivity distribution in the object plane tangent to the object surface and by the directivity pattern of the diffuse component of its surface. The diffuse light scattering from the object surface is simulated by adding to the object wave front phase defined by the object macro-shape a pseudo-random phase component (a "programmable diffuser"), whose correlation function corresponds to the given directivity

pattern of the object diffuse surface. This pseudo-random phase component is combined with the deterministic phase component defined by the object shape to form the distribution of the phase of the object wavefront.

Fig. 3 Viewing a compound computer-generated hologram (left) and one of the views reconstructed from the hologram (right).

LITERATURE

1. H. J. Caulfield, L. P. Yaroslavsky, Flat Accurate Nonimaging Point Locator, Digital Holography and Three-Dimensional Imaging 2007 Conference, Vancouver, BC, Canada, June 18-20, 2007
2. Leonid Yaroslavsky and H. John Caulfield, Computational Vision in Nature and Technology, http://arxiv.org/abs/physics/0703103
3. Bongard M.M. and Smirnov M. S. , On "skin vision" of r. Kuleshova, Biofizika. 1965;10:148-54.
4. M. C. Roggermann and B. Welsh," Imaging Through Turbulence", CRC Press, Inc, 1996.
5. B. Fishbain, L. P. Yaroslavsky, I. A. Ideses, Real Time Stabilization of Long Range Observation System Turbulent Video, Journal of Real Time Image Processing, v. 2, Number 1 / October, 2007.
6. L. Yaroslavsky, B. Fishbain, G. Shabat, I. Ideses, Super-resolution in turbulent videos: making profit from damage, Optics Letters, Nov. 1, 2007/Vol.32, No. 21
7. L. Yaroslavsky, B. Fishbain, G. Shabat, I. Ideses, Super-Resolution Of Turbulent Video: Potentials and Limitations, SPIE Conf. Electronic Imaging, Image Processing: Algorithms and Systems VII, Proceedings SPIE, v# 6812, San Jose, Jan. 2008,
8. L. Yaroslavkii, N. Merzlyakov, Methods of Digital Holography, Consultants Bureau, N. Y., 1980
9. L. Yaroslavsky, Digital Holography and Digital Image Processing, Kluwer Scientific Publishers, Boston, 2004
10. V.N. Karnaukhov, N.S. Merzlyakov, M.G. Mozerov, L.P. Yaroslavsky, L.I Dimitrov, E. Wenger, Digital Display Macro-holograms. In: Computer Optics, ISSN 0134-2452, Ed. Ye. P. Velikhov and A.M. Prokhorov, Issue 14-15, part 2, Moscow, 1995
11. V.N. Karnaukhov, N.S. Merzlyakov, L.P. Yaroslavsky, Three dimensional computer generated holographic movie, Sov.Tech. Phys. Lett., v.2, iss.4, Febr. 26, 1976, p.169-172.
12. Yatagai T. Stereoscopic approach to 3-D display using computer-generated holograms, Appl. Opt., 1976, v. 15, No 11, p. 2722-2729
13. N.S. Merzlyakov, L.P. Yaroslavsky, Simulation Of Light Spots on Diffuse Surfaces by Means of a "Programmed Diffuser", Sov. Phys. Tech. Phys., v. 47, iss.6, 1977, p. 1263-1269
14. Yaroslavskii, L.P., 360-degree Computer Generated Display Hologram, http://museumcollections.mit.edu/FMPro?-lay=Web&wok=107443&-format=detail.html&-db=ObjectsDB.fp5&-find