Real Time Turbulent Video Super-Resolution Using MPEG 4

Barak Fishbain*, Leonid P. Yaroslavsky, Ianir A. Ideses
Department of Electrical Engineering - Physical Electronics, the Iby and Aladar Fleischman Faculty of Engineering, Tel-Aviv University, Tel Aviv 69978, Israel

ABSTRACT

It has been shown that one can make use of local instabilities in turbulent video frames to enhance image resolution beyond the limit defined by the image sampling rate. This paper outlines a real-time solution for the implementation of super-resolution algorithm on MPEG-4 platforms. The MPEG-4 video compression standard offer, in real-time, several features, such as motion extraction with quarter pixel accuracy, scene segmentation to video object planes, global motion compensation and de-blocking and de-ringing filters, which can be incorporated into the super-resolution process to produce enhanced visual output. Experimental verification on real-life videos is also provided.

Keywords: Turbulent video, Real-Time, Super-Resolution, MPEG-4

1. INTRODUCTION

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 [1] 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. Methods of generating stabilized videos from turbulent videos, including real time ones, were reported in Refs.[2,3,4,5]. In [6], the idea was advanced of making a profit from atmosphere turbulence-induced image geometrical spatial/temporal degradations to compensate image sampling artifacts and generate stabilized images of the stable scene with higher resolution than that defined by the camera sampling grid.

Recovering the high-frequency information is possible as multiple low-resolution (LR) observations may provide additional information about the high-frequency data. This information is introduced through sub-pixel displacements in the sampling grid, which enables the recovery of resolution. To this end, it is required that the LR observations contain different but related views of the scene. Super-resolution principles are detailed in the works of S. Srinivasan and R. Chellappa [7], Galatsanos, Wernick and Katsaggelos [8], R.R. Schultz [9] and T. J. Schultz [10]. Several researchers treat the problem of high resolution image recovery by designing an efficient multi-frame filtering algorithms, that account for both intra-frame (spatial) and inter-frame (temporal) correlations, for restoring image sequences that are degraded both by blur and noise [11, 12]. Others have formulated solutions to global motion problems, usually from an application perspective [13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. With that, all of the above methods take solely the raw sequences' frames as input, without taking into consideration the presence of moving objects and super-resolution is achieved through computational complex algorithms.

Nowadays most digital footage data is transmitted and stored using the International Telecommunication Union (ITU) and Moving Picture Experts Group (MPEG) coding standards [23,24]. Super-resolution techniques that have been designed for raw, uncompressed, video may not be effective when applied to compressed video because they do not incorporate the compression process into their models. This has raised the need for super-resolution techniques which utilize the standards' various features [25,26,27,28]. However, the suggested methods do not comply with real-time constrains. A real-time super-resolution method, over video encoder hardware, was presented in [29]. Yet, the suggested method utilized proprietary compression standard and exploited only the sequence's motion field. Using a common compression standard, such as ITU H.264 or MPEG-4 presents several benefits, such as real-time compatibility and broad availability of software and hardware implementations.
This paper describes a practical super-resolution scheme which utilizes MPEG-4 features for producing, in real-time, good quality higher-resolution videos from low-resolution turbulent degraded video streams with discrimination of turbulent from real motion which is caused by moving objects or global camera translations.

* barak@eng.tau.ac.il;phone 972-3-640-8014; fax 972-3-641-0189

Real-Time Image Processing 2008, edited by Nasser Kehtarnavaz, Matthias F. Carlsohn,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6811, 681106, © 2008 SPIE-IS&T · 0277-786X/08/$18

2. ALGORITHM DESCRIPTION

The core of the suggested SR method is elastic registration, with sub-pixel accuracy, of available video frames of stable scenes followed by resampling of the frames according to the registration results. The first step of the processing of each current frame is updating a reference frame used for elastic registration. The next step is computing, with sub-pixel accuracy, a map of pixel displacements in the current frame with respect to the reference frame. The pixel displacement map is then analyzed and segmented to separate pixels of the real moving objects from those that belong to the stable scene and are displaced solely due to the atmosphere turbulence or global camera movements.

The displacement map found for pixels of the stable scene serves, at the next processing step, for placing pixels of each current frame, according to their positions determined by the displacement map, into the reference frame, which is correspondingly up-sampled to match the sub-pixel accuracy of the displacement map. For up-sampling, different image interpolation methods can be used. Among them, discrete sinc-interpolation is the most appropriate as the one with the least interpolation error [30]. As a result, output stabilized and enhanced in its resolution frame is accumulated. In this accumulation process it may happen that several pixels of different frames are to be placed in the same position in the output enhanced frame. In order to make best use of all of them, these pixels must be averaged. For this averaging, we suggest, computing median of those pixels in order to avoid influence of outliers that may appear due to possible anomalous errors in the displacement map.

After all available input frames are used in this way, the enhanced and up-sampled output frame contains, in positions where there were substitutions from input frames, accumulated pixels of the input frames and, in positions where there were no substitutions, interpolated pixels of the reference frame. Substituted pixels introduce to the output frame high frequencies outside the base-band defined by the original sampling rate of the input frames. Those frequencies were lost in the input frames due to the sampling aliasing effects. Interpolated pixels that were not substituted do not contain frequencies outside the base-band. In order to finalize the processing and take full advantage of the super-resolution provided by the substituted pixels, an iterative re-interpolation algorithm was used [31, 32]. Once iterations are stopped, the output-stabilized and resolution-enhanced image obtained in the previous step is sub-sampled to the sampling rate determined by selected enhanced bandwidth and then subjected to additional processing aimed at camera aperture correction, denoising and reducing blocking (de-blocking) and ringing effects (de-ringing).

3. SUPER RESOLUTION WITHIN THE MPEG-4 FRAMEWORK

A digital image sequence coded at a low bit rate using a motion-compensated video compression standard should contain little data redundancy. However, the success of a particular super-resolution enhancement algorithm is predicated on sub-pixel-resolution overlap (i.e., redundancy) of moving objects from frame-to frame. If an MPEG bit stream is coded at a relatively high bit rate (e.g., a compression ratio of 15:1), enough data redundancy exists within the bit stream to successfully perform super-resolution enhancement within the decoder.

3.1. MPEG Picture Types

MPEG-encoded image sequences are divided into groups of pictures (GOPs) composed of primarily three different frame types: intra-coded frames (I-pictures), predictive-coded frames (P-pictures), and bi-directionally predictive-coded frames (B-pictures). I-pictures are coded independently without reference to other pictures, using block-DCT (discrete cosine transform) compression [33]. P-pictures obtain predictions from temporally preceding I- or P-pictures, while B-pictures are predicted from the nearest preceding and/or upcoming I- or P-pictures.

In the video observation model, the I-picture which begins a particular GOP is the reference frame. The SR algorithm requires motion vectors between the reference frame and all other frames. Therefore, only the neighboring frames up to and including the first P-picture in the current GOP and the frames down to and excluding the last P-picture in the previous GOP are integrable with the I-picture.

3.2. Computing the Reference Frame

In the super-resolution under turbulence conditions, the reference image is an estimate of the stable scene. Such an image has to be obtained from the input sequence itself [34,35]. In order to achieve optimal results, the reference image should have the following properties:

- The reference image should contain only the static background with no moving objects in it.
- It should contain no turbulent induced geometric distortion and global camera movement.

In order to generate a reference image, one has to, for each time window, (i) take the current frame as the reference; (ii) compute motion vectors for each frame in the time window with respect to this reference frame; (iii) find pixel-wise means for motion vectors; (iv) resample pixels of each frame to the positions defined be the means; this will generate a stabilized sequence. This process presents high computational load, and therefore an approximate solution is suggested.

The suggested solution takes the first I-frame since the following frames have predictions which are directly connected to macro-blocks within the I-picture. Integration of more than one GOP into the SR process is desired, therefore, for every new I-frame the translations with regards to the reference I-frame have to be computed.

3.3. Motion analysis

Estimating accurate sub-pixel-resolution motion vectors is a critically important component in super-resolution enhancement algorithms. The goal is to compute the motion fields as quickly and as accurately as possible, so that super-resolution enhancement becomes practical desktop computing application.

The mapping of one such unstable image to a stable one can be obtained by registering a spatial neighborhood surrounding each pixel in the image in the reference image. In its simplest form, it is sufficient to find, for each pixel, the translations along the x and y axes. It has been suggested that such a registration can be implemented using optical flow methods [36,37] or correlation methods [38]. This paper suggests using motion vector data available in MPEG encoded videos.

MPEG Motion estimation techniques aims to find a 'match' to the current block or region that minimizes the energy in the motion compensated residual (the difference between the current block and the reference area). This usually involves evaluating the residual energy at a number of different offsets. The choice of measure for 'energy' affects computational complexity and the accuracy of the motion estimation process. Eq. (1), eq. (2) and eq. (3) describe three energy measures, Mean Squared Error (MSE), Mean Absolute Error (MAE) and Sum of Absolute Error (SAE), where the block size is NxN, C and R are the current and reference frames respectively and Ω is the corresponding sampled areas. SAE is probably the most widely-used measure of residual energy for reasons of computational simplicity.

$$MSE = \frac{1}{N^2} \sum_{i,j \in \Omega} \left(C_{i,j} - R_{i,j} \right)^2 \tag{1}$$

$$MAE = \frac{1}{N^2} \sum_{i,j \in \Omega} \left| C_{i,j} - R_{i,j} \right| \tag{2}$$

$$SAE = \sum_{i,j \in \Omega} \left| C_{i,j} - R_{i,j} \right| \tag{3}$$

Many fast search algorithms have been proposed, such as Logarithmic Search, Hierarchical Search, Cross Search, Unsymmetrical Cross Multi Hexagon-grid (UMHexagonS) and Enhanced Predictive Zonal Search (EPZS) [23, 39, 40]. In each case, the performance of the algorithm can be evaluated by comparison with Full Search. Suitable comparison criteria are compression performance and computational performance. Other criteria may be helpful; for example, some 'fast' algorithms such as Hierarchical Search are better-suited to hardware implementation than others.

3.3.1. Sub-pixel Resolution Motion Estimation

As described in Sect. 2 the super-resolution process requires fractional translations rather than just integer values. The MPEG-4 standard defines half-pixel vectors in MPEG-4 Simple Profile and quarter-pixel vectors in Advanced Simple profile and H.264. Sub-pixel motion estimation requires the encoder to interpolate between integer sample positions in the reference frame. Interpolation is computationally intensive. Calculating sub-pixel samples for the entire search window is not usually necessary. Instead, it is sufficient to find the best integer-pixel match (using one of the fast search algorithms discussed above) and then to search interpolated positions adjacent to this position. In the case of quarter-pixel motion estimation, first the best integer match is found; then the best half-pixel position match in the immediate neighborhood is calculated; finally the best quarter-pixel match around this half-pixel position is found.

3.3.2. Real Motion Extraction

The pixel displacement map is analyzed and segmented to separate pixels of the real moving objects from those that belong to the stable scene and are displaced solely due to the atmosphere turbulence. One of the key contributions of MPEG-4Visual is a move away from the ‘traditional’ view of a video sequence as being merely a collection of rectangular frames of video. Instead, MPEG-4 Visual treats a video sequence as a collection of one or more video objects (VO). MPEG-4 Visual defines a video object as a flexible ‘entity that a user is allowed to access (seek, browse) and manipulate (cut and paste) [41].An instance of a VO at a particular point in time is a *video object plane* (VOP). A video scene may be made up of a background object and a number of separate foreground objects. This approach is potentially much more flexible than the fixed, rectangular frame structure of earlier standards. The separate objects may be exploited for efficient background separation. One drawback, however, is the fact that the MPEG-4 standard was designed for natural videos, therefore the presence of turbulent motion might lead to a significant number of errors in the process of segmenting the scene into VO's. To this end, the absolute difference from the reference as well as statistical analysis of displacement magnitudes and angles [2,3] are used for segmenting pixels of the real moving objects from those that belong to the stable scene and are displaced solely due to the atmosphere turbulence.

3.3.3. Global Motion Extraction

Macro blocks within the same video object may experience similar motion. For example, camera pan will produce apparent linear movement of the entire scene, camera zoom or rotation will produce a more complex apparent motion and macro blocks within a large object may all move in the same direction. Global Motion Compensation (GMC) enables an encoder to transmit a small number of motion (warping) parameters that describe a default ‘global’ motion. Additionally, the GMC can provide improved motion analysis and real motion extraction.

3.4. Generation of super-resolved stable frames

Once the data accumulation and interpolation stages are done, the output-stabilized and resolution-enhanced image obtained might be subjected to additional processing aimed at camera aperture correction, denoising and reducing blocking (de-blocking) and ringing effects (de-ringing). Many such filters designs have been proposed and implemented, ranging from relatively simple algorithms to iterative algorithms that are many times more complex than the encoding and decoding algorithms themselves [42, 43, 44, 45]. The goal of a de-blocking or de-ringing filter is to minimize the effect of blocking or ringing distortion whilst preserving important features of the image. MPEG-4 Visual describes a deblocking filter and a deringing filter: these are ‘informative’ parts of the standard and are therefore optional. Both filters are designed to be placed at the output of the decoder. With this type of post-filter, unfiltered decoded frames are used as the reference for motion-compensated reconstruction of further frames. This means that the filters improve visual quality at the decoder but have no effect on the encoding and decoding processes.

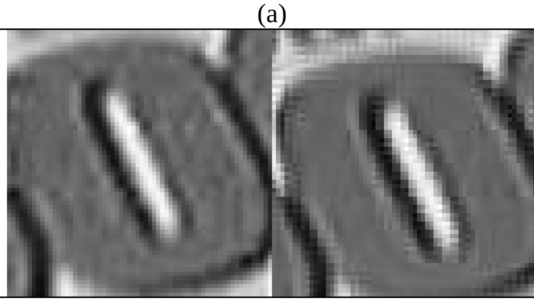
4. RESULTS

Experimental results are presented, in which decoded pictures from MPEG-4 bit streams containing translations due to global (Figure 1) or turbulent (Figure 2) motion. Figure 1 depicts the super-resolution process, through MPEG-4, applied on a real-life sequence containing global motion. The sequence was acquired with 320x240 pixels webcam. Figure 1(a) presents the first frame extracted from the sequence interpolated to twice, on each axis, of its original size. Figure 1(b) is the super-resolved image after accumulation of 50 frames. The translation of each frame with regards to the first frame, which is the first I-frame broadcasted, was computed using the MPEG-4 diamond search with quarter pixel accuracy. Figure 1(c) and (d) present fragments extracted from figures (a) and (b) respectively. The left-hand side parts of (c) and (d) are interpolated (Figure 1 (a)), while the right-hand parts are the corresponding super-resolved fragments.

A super-resolved frame computed using 150 frames of a turbulent degraded sequence is presented in Figure 2(a). Figure (b) depicts the absolute difference of the super-resolved image and the reference frame interpolated to four times of its original size. As one can see, most of the difference's energy is located in the vicinity of edges. The right-hand side of Figure 2(c) shows a fragment extracted from the interpolated reference frame, while the left-hand side is the corresponding super-resolved one. As one can see the later contains finer details of the scene. Finally, Figure 2(d) presents both spectra of the interpolated (on the left-hand side of Figure 2(d)) and of the super-resolved frames. It can be seen that the super-resolution process enhance data in higher frequencies.


(a)


(b)


(c)

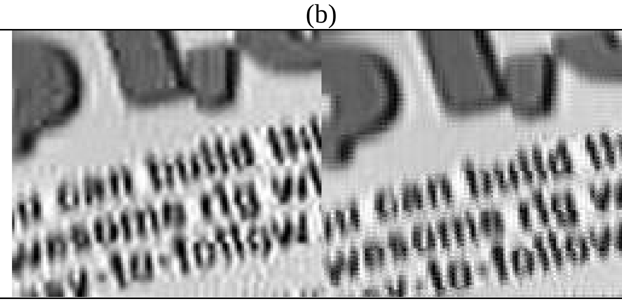
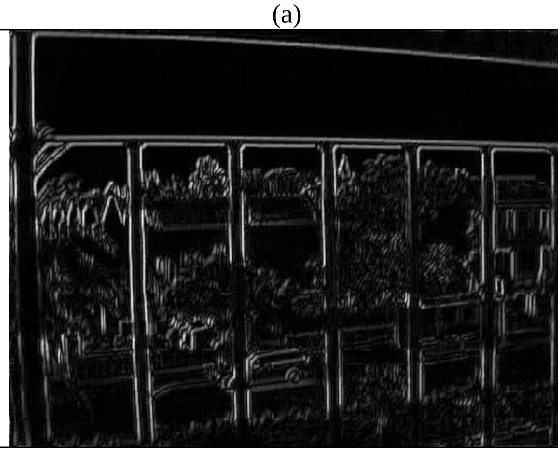
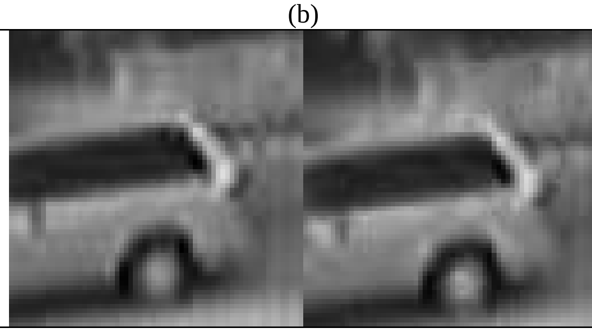

(d)

Figure 1 – Super Resolution from Global Motions through MPEG-4

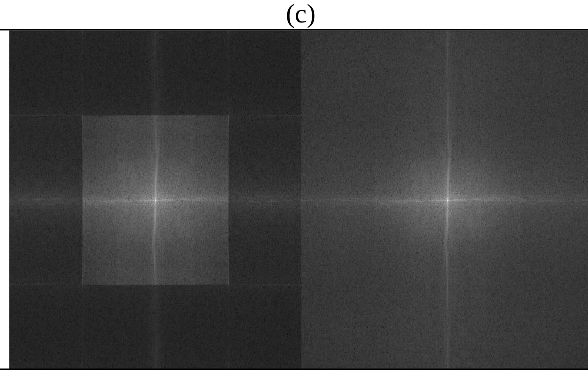
(a)



(b)



(c)



(d)

Figure 2 - Super Resolution of turbulent degraded videos MPEG-4.

5. CONCLUSION

This paper shows the utilization of an incremental SR algorithm within the MPEG-4 video encoder. The paper suggests the use of inbuilt features of the MPEG-4 compression standard in order to reduce the computational complexity and for enhancing the visual output. To this end, the use of MPEG-4's motion extraction, video object planes, global motion compensation and de-blocking and de-ringing filters is suggested. The experiments carried out exhibit a clear quality increase of the super-resolved image.

ACKNOWLEDGEMENTS

The research was partially funded by the Israeli Ministries of Transportation and of Science Culture and Sport.

REFERENCES

[1] M. C. Roggermann and B. Welsh, *Imaging Through Turbulence*, CRC Press, Inc, 1996.

[2] B. Fishbain, L. P. Yaroslavsky and I. A. Ideses, "Real-time stabilization of long range observation system turbulent video", Journal of Real-Time Image Processing Vol.2(1), pp. 11-22, Oct. 2007.

[3] B. Fishbain, L. P. Yaroslavsky, I. A. Ideses, A. Shtern, O. Ben-Zvi, "Real-time stabilization of long-range observation system turbulent video", Proc. of Real-Time Image Processing/Electronic Imaging 2007, SPIE Vol. 6496, San-Jose, CA, USA, 28 January – 1 February 2007.

[4] Sh. Gepshtein, A. Shtainman, B. Fishbain, L. Yaroslavsky, "Restoration of atmospheric turbulent video containing real motion using elastic image registration", The 2004 European Signal Processing Conference (EUSIPCO-2004), John Platt, pp. 477-480, Vienna, Austria, 2004.

[5] L. P. Yaroslavsky, B. Fishbain, I. A. Ideses, D. Slasky and Z. Hadas, "Simple Methods for Real-time Stabilization of Turbulent Video", proc. of ICO Topical Meeting on Optoinformatics/Information Photonics, ITMO, pp. 138-140, St Petersburg Russia, 2006.

[6] L.P. Yaroslavsky, B. Fishbain, G. Shabat and I. Ideses, "Super-resolution in turbulent videos: making profit from damage", Optics Letters, Vol. 32(20), pp. 3038-3040, 2007.

[7] S. Srinivasan and R. Chappella, "Image Sequence Stabilization, Mosaicking and Superresolution", *Video and Image Processing Handbook* (edited by A.C. Bovic), Dallas: 2000, chapter 3.13, pp. 259-268.

[8] N. P. Galatsanos, M. N. Wernick and A. K. Katsaggelos, "Multi-channel Image Recovery", *Video and Image Processing Handbook* (edited by A.C. Bovic), Dallas: 2000, chapter 3.7, pp. 161-174.

[9] R.R. Schultz, Extraction of High-Resolution Frames from Video Sequences, IEEE, Trans. On Image Proc. Vol. 5(6), pp 996-1011, 1996.

[10] T. J. Schultz, "Multiframe Image Restoration", *Video and Image Processing Handbook* (edited by A.C. Bovic), Dallas: 2000, chapter 3.8 , pp. 175-189.

[11] M. Ozkan, M. Sezan, "Efficient Multiframe Wiener Restoration of Blurred and Noisy Image Sequences", IEEE, Transactions on Image Processing Vol. 1(10), pp 453-475, 1994.

[12] L. Guan, Restoration fof randomly blurred Images by the Wiener Filter, IEEE, Transactions on Acoustics, Speech and signal Processing Vol. 37(4), pp 589-595, 1989.

[13] G. Adiv, "Determining 3-D motion and structure from optical flow generated by several moving objects", IEEE Transactions on Pattern Analysis and Machine Intelligence 7, pp. 384-401, 1985.

[14] D. Capel and A. Zisserman, "Automated mosaicing with super-resolution zoom", in IEEE Computer Vision and Pattern Recognition, pp. 885-891, 1998.

[15] M. Irani and S. Peleg, "Improving resolution by image registration", Graph. Models Image Process. Vol. 53, pp. 231-239, 1991.

[16] R.C. Hardie, K. J. Barnard and E. E. Armstrong, “Joint map registration and high resolution image estimation using a sequence of undersampled images”, IEEE Transactions on Image Processing 6, pp. 1261-1633 (1997).

[17] M. Elad and A. Feuer, “Recursive Optical Flow Estimation – Adaptive Filtering Approach”, Journal of Visual Communication and Image Representation, Vol. 9(2), pp. 119-138, 1998.

[18] N. Goldberg, A. Feuer and G.C.Goodwin, " Super - Resolution Reconstruction Using Spatio - Temporal Filtering" Journal of Visual Communications and Image Representation, Vol. 14 (2003). pp. 508--525.

[19] D. Capel and A. Zisserman, "Computer vision applied to superresolution", IEEE Signal Processing Magazine 20(3), pp. 75-86, 2003.

[20] S. Baker and T. Kanade, "Limits on super-resolution and how to break them", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24(9), pp. 1167–1183, 2002.

[21] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, “Fast and Robust Multi-Frame Super-resolution”, IEEE Transactions on Image Processing, Vol. 13(10), pp. 1327-1344, 2004.

[22] S. Farsiu, M. Elad, and P. Milanfar, “Multi-Frame Demosaicing and Super-Resolution of Color Images”, IEEE Transactions on Image Processing, Vol. 15, pp. 141-159, 2006.

[23] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression Video Coding for Next-generation Multimedia*, John Wiley, England 2003.

[24] V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards*, Norwell, MA: Kluwer, 1995.

[25] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian Resolution Enhancement of Compressed Video", IEEE Transactions on Image Processing", Vol. 13(7), 2004.

[26] E. Shor, L. Yaroslavsky, Resolution improvement by sub-pixel image registration, in: Visualization, Imaging, and Image Processing, VIIP 2004, 9/6/2004 - 9/8/2004, Marbella, Spain, Editor(s): Prof. Juan J. Villanieva

[27] B. K. Gunturk, Y. Altunbasak and R. M. Mersereau, "Super-Resolution Reconstruction of Compressed Video Using Transform-Domain Statistics", IEEE Transactions on Image Processing, Vol. 13(1), 2004.

[28] D. Chen and R. R. Schultz, "Extraction of High-Resolution Video Stills from MPEG Image Sequences", Proceedings of International Conference on Image Processing (ICIP), vol.2, pp. 465-469 Chicago, IL, USA, 4-7 Oct. 1998.

[29] G.M. Callico, A. Nunez, R. P. Llopis, R. Sethuraman, "Low-cost and real-time super-resolution over a video encoder IP", Proceedings. Fourth International Symposium on Quality Electronic Design, pp. 79-84, San-Jose CA-USA, 24-26 March 2003.

[30] L. Yaroslavsky, “Fast Discrete Sinc-Interpolation: A Gold Standard for Image Resampling”, In: *Advances in Signal transforms*: Theory and Applications, J. Astola, L. Yaroslavsky, Eds., EURASIP Book Series on Signal Processing and Communications, Hindawi, 2007

[31] P.J.S.G. Ferreira, “Interpolation and the Discrete Papoulis-Gerchberg Algorithm”, IEEE Transactions on Signal Processing, Vol. 42(10), pp. 2596-2606, 1994.

[32] A. Papoulis, “A New Algorithm in Spectral Analysis and Band-Limited Extrapolation”, IEEE Transactions on Circuits and Systems Vol. 22(9), pp. 735-742, 1975.

[33] J. L. Mitchell, W. B. Pennebaker, C. E. Fogg, and D. J. LeGall, MPEG Video Compression Standard.New York, NY: Chapman and Hall, 1996.

[34] B. Cohen, V. Avrin, M. Belitsky, and I. Dinstein, "Generation of a restored image from a video sequence recorded under turbulence effects", Optical Engineering, vol. 36 no. 12, pp. 3312-3317, 1997.

[35] C. Bondeau, E. Bourennane, Restoration of images degraded by the atmospheric turbulence, Proceedings of the 4th International Conference on Signal Processing (ICSP), pp. 1056-1059, vol.2, Beijing China, 1998

[36] T. Brox, A. Bruhn, N. Papenberg & J. Weickert, High Accuracy Optical Flow Estimation based on Theory for Wrapping, Proc. 8th European Conference on Computer Vision, Prague, Czech Republic, 2004, 4:25-36.

[37] L. J. Barron, D. J. Fleet and S. S. Beachemin, Performance of Optical Flow Techniques, International Journal of Computer Vision 12, 1994, 43-77.

[38] L. Yaroslavsky, Digital Holography and Digital Image Processing, Kluwer Scientific Publishers, Boston, 2004

[39] Z. Chen, J. Xu, Y. He and J. Zheng, "Fast integer-pel and fractional-pel motion estimation for H.264/AVC", Journal of Visual Communication and Image Representation, Vol.17(2), Pages 264-290, April 2006.

[40] A. M. Tourapis, "Enhanced predictive zonal search for single and multiple frame motion estimation," in Proc. of Visual Communications and Image Processing (VCIP), pp. 1069-1079, San-Jose CA., Jan. 2002.

[41] ISO/IEC 14496-2, Amendment 1, Information technology – coding of audio-visual objects – Part 2:Visual, 2001.

[42] J. Chou, M. Crouse and K. Ramchandran, A simple algorithm for removing blocking artifacts in block transform coded images, IEEE Signal Process. Lett., 5, February 1998.

[43] S. Hong, Y. Chan and W. Siu, A practical real-time post-processing technique for block effect elimination, Proc. IEEE ICIP96, Lausanne, September 1996.

[44] T. Meier, K. Ngan and G. Crebbin, Reduction of coding artifacts at low bit rates, Proc. SPIE Visual Communications and Image Processing, San Jose, January 1998.

[45] Y. Yang and N. Galatsanos, Removal of compression artifacts using projections onto convex sets and line modeling, IEEE Trans. Image Processing, 6, October 1997.