In Proc. Int. Conf. on Augmented, Virtual Environments and Three-Dimensional Imaging, ICAV3D-2001, May 30 -June 1, 2001, Mykonos, Greece

POCS reconstruction of stereoscopic views

Ryszard Stasiński

Institute of Electronics
and Telecommunications
Technical University of Poznań
Piotrowo 3A
PL-60-965 Poznań, Poland
rstasins@et.put.poznan.pl

Janusz Konrad

Department of Electrical
and Computer Engineering
Boston University
8 Saint Mary's Street
Boston, MA 02215, USA
jkonrad@bu.edu

ABSTRACT

This paper presents an application of POCS (projection onto convex sets) methodology to the reconstruction of intermediate stereoscopic views. The basic problem in such a reconstruction, resulting from disparity compensation, is that of the recovery of a regularly-sampled image from its irregularly-spaced samples. This problem also arises in other image processing and coding applications. The results reported here improve our previous POCS-based reconstruction method by locally adapting the algorithm to the density of image samples. We also extend the method to color images by implementing the method in the luminance-chrominance (Y-U-V) space.

1. INTRODUCTION

In order to compute intermediate views in a stereoscopic or multiview representation of a 3-D scene, the usual problem is that of recovering regularly-spaced images samples (intensity and color) based on irregularly-spaced samples. This is due to disparity compensation and is similar to motion compensation used for temporal interpolation of image sequences.

The interpolation of either regularly- or irregularly-spaced samples based on the knowledge of a regularly-sampled image has been extensively treated in the literature and has found numerous practical applications in image processing and coding. The case of computation of regularly-spaced samples based on regular ones has been explored to a lesser degree. The primary reason for this are difficulties associated with the extension of Shannon's sampling theory to signals defined over irregular sampling grids. In such cases, Shannon's theory is not applicable and alternative methods must be found to reconstruct or approximate the original continuous signal.

Although some results on the reconstruction of band-limited functions from irregularly-spaced samples are available (e.g., [1]), their practical usefulness is limited; theoretical constraints on the maximum spacing of irregular samples under the perfect reconstruction condition cannot be satisfied by arbitrarily-distributed image samples after disparity or motion compensation. By relaxing the perfect reconstruction condition, alternative methods have been proposed such as the polynomial interpolation or iterative reconstruction [2].

In this paper, we extend a reconstruction method based on *projections onto convex sets* (POCS) [3] proposed by us earlier [4]. We improve the convergence of reconstruction by locally adapting the algorithm to the density of image samples. We also extend the method to color images by implementing it in the luminance-chrominance (Y-U-V) space.

2. PROPOSED APPROACH

Let $g = \{g(\mathbf{x}), \mathbf{x} = (x, y)^T \in R^2\}$ be a continuous 2-D projection of the 3-D world onto an image plane and let $g_\Lambda = \{g(\mathbf{x}), \mathbf{x} \in \Lambda\}$ be a discrete image obtained from g by sampling over a lattice Λ [5]. Let's assume that g is band-limited, i.e., $G(\mathbf{f}) = \mathcal{F}\{g\}$=0 for $\mathbf{f} \notin \Omega$ where $\mathcal{F}$ is the Fourier transform, $\mathbf{f} = (f_1, f_2)^T$ is a frequency vector and $\Omega \subset R^2$ is the spectral support of g. If the lattice Λ satisfies the multi-dimensional Nyquist criterion [5], the Shannon sampling theory allows to perfectly reconstruct g from g_Λ. However, in the case of irregular sam-

This work was carried out with partial support from the Natural Sciences and Engineering Research Council of Canada under Strategic Grant STR224122 while the second author was at INRS-Télécommunications.

pling the theory is not applicable. Therefore, the general goal is to develop a method for the reconstruction of g from an irregular set of samples $g_\Psi = \{g(\mathbf{x}_i), \mathbf{x}_i \in \Psi \subset R^2, i = 1, ..., K\}$, where Ψ is an irregular sampling grid.

2.1. POCS-based reconstruction algorithm

We use the POCS methodology [3] to reconstruct image g. This methodology involves a set theoretic formulation, i.e., finding a solution as an intersection of property sets rather than by a minimization of a cost function. We use the following sets [4]:

- A_0 - set of all images g such that at $\mathbf{x}_i \in \Psi$, $i = 1, ..., K$ (irregular sampling grid) $g(\mathbf{x}_i) = g_\Psi(\mathbf{x}_i)$,
- A_1 - set of all band-limited images g, i.e., such that $G(\mathbf{f}) = 0$ for $\mathbf{f} \notin \Omega$.

If the membership in A_0 can be assured by a sample replacement operator $\mathcal{R}$ (to enforce proper image values on Ψ), and the membership in A_1 – by suitable bandwidth limitation (low-pass filtering) $\mathcal{B}$, then the iterative reconstruction algorithm can be expressed as follows:

$$\begin{aligned} g^{k+1} &= \mathcal{BR}g^k = \mathcal{B}[g^k + \mathcal{S}_\Psi(g - g^k)], \quad (1) \\ &= \mathcal{RB}g^k = \mathcal{B}g^k + \mathcal{S}_\Psi(g - \mathcal{B}g^k), \end{aligned}$$

where g^k is the reconstructed image after k iterations and $\mathcal{S}_\Psi$ is a sampling operator that extracts image values (luminance/color) on the irregular grid Ψ. Note that equation (1), proposed in [6], results in an approximation rather than interpolation of g_Ψ; the last step is that of low-pass filtering. In order to implement equation (1) on a computer, a suitable discretization must to be applied. In [6], equation (1) was implemented as follows:

$$g_\Lambda^{k+1} = \mathcal{B}[g_\Lambda^k + \alpha \mathcal{I}_{\Psi/\Lambda}(g_\Psi - \tilde{g}_\Lambda^k)], \quad (2)$$

where the lowpass filtering $\mathcal{B}$ is implemented over Λ and α is a parameter that allows control of convergence and stability of the algorithm. The symbol $\tilde{g}_\Lambda^k$ denotes a bilinearly-interpolated image g_Λ^k needed to recover image samples on Ψ. Also, note that an interpolation function $\mathcal{I}_{\Psi/\Lambda}$ replaces the sampling operator $\mathcal{S}_\Psi$. This function interpolates image samples $(g_\Psi - \tilde{g}_\Lambda^k)$ defined on Ψ in order to recover samples on Λ. Sauer and Allebach have studied three interpolators $\mathcal{I}_{\Psi/\Lambda}$: one derived from bilinear interpolation and two based on triangulation with planar facets [6]. The implementation (2) of the reconstruction algorithm (1) suffers from two deficiencies. First, by processing all images on Λ there is little flexibility in shaping the spectrum of g_Λ^k; any practical lowpass filtering on Λ must suppress high frequencies since a slow roll-off transition band must be used to minimize ringing on sharp luminance/color transitions. Secondly, the interpolation operator $\mathcal{I}_{\Psi/\Lambda}$, especially the one based on triangulation (better performance), is involved computationally.

In our earlier work [4] an alternative implementation of the algorithm (1) has been proposed. Since our goal is the reconstruction of image samples obtained from motion or disparity compensation, a 1/2-, 1/4- or 1/8-pixel precision of motion or disparity vectors is usually sufficient. Therefore, it has been proposed to implement (1) on an oversampled grid matching that precision:

$$g_{\Lambda_P}^{k+1} = \mathcal{B}[g_{\Lambda_P}^k + \alpha \mathcal{S}_{\Psi/\Lambda_P}(g_{\Psi/\Lambda_P} - g_{\Lambda_P}^k)]. \quad (3)$$

where $\mathcal{B}$ is implemented on Λ_P, that is a $P \times P$-times denser (oversampled) lattice than Λ, and P equals 2, 4, or 8 depending on motion/disparity vector precision. Clearly, Λ is a sub-grid of Λ_P, i.e., $\mathbf{x} \in \Lambda \Rightarrow \mathbf{x} \in \Lambda_P$. g_{Ψ/Λ_P} is the nearest-neighbor interpolation of g_Ψ on Λ_P, defined at each $\mathbf{x}_i \in \Psi$ as follows:

$$g_{\Psi/\Lambda_P}(\mathbf{y}) = \begin{cases} g_\Psi(\mathbf{x}_i) & \text{if } \|\mathbf{x}_i - \mathbf{y}\| \le \|\mathbf{x}_i - \mathbf{z}\|, \\ 0 & \text{otherwise.} \end{cases}$$

for all $\mathbf{y}, \mathbf{z} \in \Lambda_P$. Similarly, $\mathcal{S}_{\Psi/\Lambda_P}$ denotes the nearest-neighbor sampling, i.e., sampling on $\mathbf{y} \in \Lambda_P$ that is nearest to $\mathbf{x}_i \in \Psi$. In other words, the implementation (3) is performed on a denser lattice Λ_P and the positions of the irregular samples from Ψ are quantized to the nearest position on Λ_P. This allows us to avoid the cumbersome interpolation $\mathcal{I}_{\Psi/\Lambda}$ under the assumption that a suitable value of P is selected.

2.2. Adaptation of the relaxation coefficient

The choice of the relaxation coefficient α in equation (3) has a direct impact on the convergence properties of the algorithm; the greater the α, the faster the convergence, but only up to some α_{max} above which the algorithm becomes unstable. Experiments have shown that the value of α_{max} in (3) is closely related to the properties of the irregular sampling grid. Namely, the algorithm has been most prone to instability in image regions where irregular sampling grid is the densest. Clearly, when increasing α above α_{max}, the algorithm starts to diverge in those image regions where the number of irregular samples per area is the highest. That is why it is proposed

to introduce an additional α-correcting term in equation (3) as follows:

$$g^{k+1}_{\Lambda_P} = \mathcal{B}[g^k_{\Lambda_P} + (\alpha/d_\Psi)\mathcal{S}_{\Psi/\Lambda_P}(g_{\Psi/\Lambda_P} - g^k_{\Lambda_P})]. \quad (4)$$

where d_Ψ are samples of a function describing local density of irregular grid. We expect that algorithm implementations based on (4) will allow higher values of α_{max}, and therefore faster convergence than those based on formulation (3). As it will turn out, the values of α_{max} become only marginally dependent on the degree of variation in the local densities of irregular grids.

To be a good descriptor of local grid density, the function d should equal 1 where the grid is regular, should be greater than 1 in areas where there are more samples of irregular grid than those of regular one, and less than 1 when converse is true. Experiments show that the actual definition of the function is not critical; various functions d seem to work almost equally well. It has been decided that the d function is computed by counting occurrences of irregular samples in a 1×1 square neighborhood of each node of the regular grid, and then by filtering the results by a 5×5 separable smoothing filter. We obtain in this way a regularly-sampled function d; as samples on the irregular grid d_Ψ the nearest-neighbor samples of d (on the regular grid) are taken.

2.3. Implementation

The implementation of equation (4) would, in general, require more memory and be less efficient than that of equation (2), however we opt for an implementation in the frequency domain in order to reduce the computational complexity. Let

$$e^k = (g_{\Psi/\Lambda_P} - g^k_{\Lambda_P})/d_\Psi$$

be the reconstruction error defined on Λ_P weighted by the inverse of the irregular grid density function. Then, each iteration of the reconstruction algorithm (4) consists of 3 steps:

1. Fourier transform of the error e^k sampled on Ψ/Λ_P:

$$E^k(\mathbf{f}) = \mathcal{F}\{\sum_i e^k(\mathbf{x}_i)\delta(\mathbf{x} - \mathbf{x}_i)\}, \quad (5)$$
$$\mathbf{x}_i \in \Psi/\Lambda_P,$$

   where δ is the Kronecker delta.

2. Update and bandwidth limitation:

$$G^{k+1}(\mathbf{f}) = \begin{cases} G^k(\mathbf{f}) + \alpha E^k(\mathbf{f}) & \mathbf{f} \in \Omega, \\ 0 & \mathbf{f} \notin \Omega, \end{cases} \quad (6)$$

   with $G^0 = 0$. Note that the smaller the Ω, the fewer the samples of $E^k(\mathbf{f})$ that need to be computed. This allows significant reduction of memory requirements and computational complexity by means of a pruned FFT [7] (see [4] for details on memory management in the proposed algorithm). A simple bandwidth limitation by zeroing parts of the spectrum leads in spatial domain to oscillations at sharp luminance/color transitions. In [4], a detailed discussion of the design of lowpass filters that minimize these effects can be found.

3. Computation of g^{k+1} on Ψ/Λ_P:

$$g^{k+1}(\mathbf{x}_i) = \mathcal{F}^{-1}\{G^{k+1}(\mathbf{f})\}|_{\mathbf{x}=\mathbf{x}_i \in \Psi/\Lambda_P}. \quad (7)$$

   This operation can be very efficiently implemented by computing the pruned inverse Fourier transform [7].

3. EXPERIMENTAL RESULTS

The proposed reconstruction algorithm has been tested experimentally on images with various irregular sampling grids Ψ. In our previous work [4], we presented results for both synthetic and natural disparity fields. Since the local sample density is less predictable in the case of natural disparity fields here we are comparing the new algorithm with the previous, non-adaptive one on natural data only. The disparities were computed from an ITU-R 601 stereopair *Flowerpot* using an optical flow-type algorithm [8], and subsequently used in disparity compensation to obtain the irregular grid Ψ. Then, the luminance and color of g_Ψ were computed using bicubic interpolation [9]. Using g_Ψ, g_Λ was reconstructed and compared with the original image. We tested the algorithm for P=4, 8 and 16 and various α's. We used lowpass filters proposed earlier in [4] since they give a good compromise between detail loss and aliasing.

Fig. 1(a) shows the PSNR evolution for luminance reconstruction error with fixed and adaptive α. Note that the experimentally optimized fixed α (highest stable value) for the luminance component was 0.4 for P=4, 0.3 for P=8 and 0.2 for P=16, while it was 0.7 for all P's in the adaptive case. Clearly, the convergence in the adaptive case is faster and the steady-state PSNR is higher; the benefits of α adaptation are evident.

Figs. 1(b-c) show the PSNR evolution for chrominance errors. Note that similarly to the luminance case the higher the P, the better the

performance of the algorithm, although the higher the computational complexity due to the higher oversampling rate. While the increase of oversampling from P=4 to P=8 shows up to a 1dB PSNR gain, a similar increase from P=8 to P=16 is much less evident.

Subjectively, the reconstructed images were of very high quality. This suggests the viability of the proposed algorithm for various high-quality reconstructions in image processing and coding.

4. REFERENCES

[1] H. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in *Wavelet: Mathematics and Applications*, J. Benedetto and M. Frazier, Eds., chapter 8, pp. 305–363. CRC Press, Boca Raton FL, 1994.

[2] A. Sharaf and F. Marvasti, "Motion compensation using spatial transformations with forward mapping," *Signal Process., Image Commun.*, vol. 14, pp. 209–227, 1999.

[3] P. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.

[4] R. Stasiński and J. Konrad, "POCS-based image reconstruction from irregularly-spaced samples," in *Proc. IEEE Int. Conf. Image Processing*, Sept. 2000, vol. II, pp. 315–318.

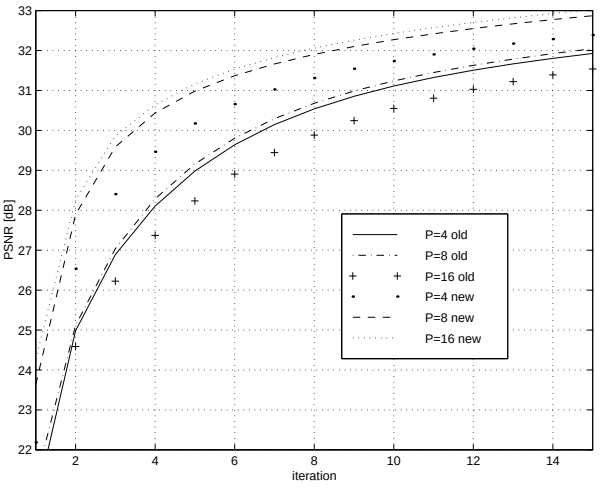
[5] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, no. 4, pp. 502–522, Apr. 1985.

[6] K.D. Sauer and J.P. Allebach, "Iterative reconstruction of band-limited images from nonuniformly spaced samples," *IEEE Trans. Circuits Syst.*, vol. 34, no. 12, pp. 1497–1506, Dec. 1987.

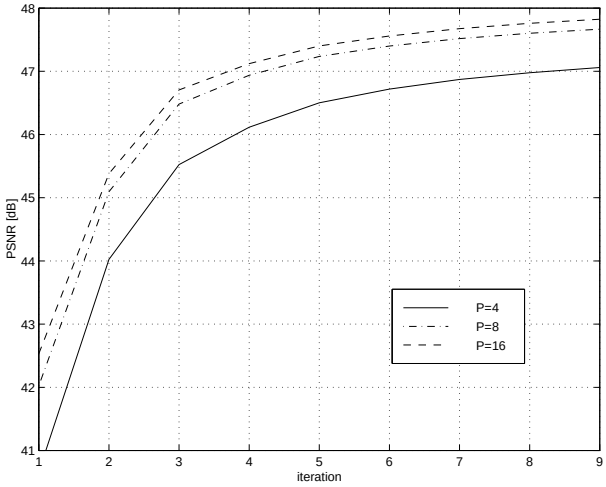
[7] R. Stasiński, "FFT pruning - a new approach," in *Signal Process. III: Theories and Applications (Proc. Third European Signal Process. Conf.)*, 1986, pp. 267–270.

[8] R. March, "Computation of stereo disparity using regularization," *Pattern Recognit. Lett.*, vol. 8, pp. 181–187, Oct. 1988.

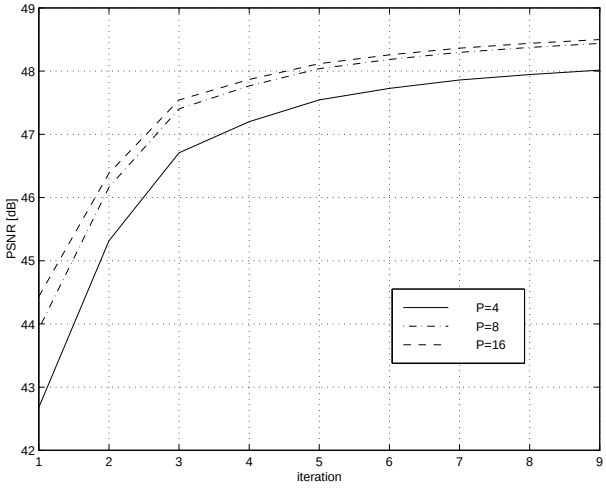
[9] R.G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.



(a) Luminance Y and its comparison with non-adaptive method ("old")



(b) Chrominance U



(c) Chrominance V

Figure 1: Evolution of PSNR of the reconstruction error.