ACQUISITION PROCESSING CHAIN FOR DYNAMIC PANORAMIC IMAGE SEQUENCES

Markus Beermann and Eric Dubois

School of Information Technology and Engineering (SITE)
University of Ottawa, Canada

ABSTRACT

The acquisition of digital image data is not just a single step but actually consists of a whole chain of processing tasks. These have been extensively studied over many years. However, with different types of cameras different specific aspects of processing become important. For a panoramic camera setup, specific challenges as well as opportunities derive from the extremely wide angle of view. We present an integrated signal processing chain that mosaics 6 planar simultaneously acquired images into a cubic panorama. Challenges that are unique to omnidirectional imaging are addressed in order to maximize the obtained visual quality with respect to dynamic range, color reproduction, and spatial resolution.

Index Terms**—** Panoramic Image Acquisition, Image Acquisition Processing, Multi-Camera Acquisition

1. INTRODUCTION

The ability to capture dynamic panoramic image sequences can provide impressive viewing impressions (e.g. IMAX cinema) but can also be just the first step to new even more fascinating applications like virtual tourism, augmented virtual presence, etc. [1].

In a static scene, a single rotating camera can take a sequence of images which are merged to form a single panorama (see e.g. [2]). In addition, different exposure settings may subsequently be chosen and a high-dynamic-range representation of the scenery be generated [3]. However, not all scenes are static. Specialized panoramic cameras allow for simultaneous capture of dynamic video frames over almost a full sphere of view. The geometry and calibration challenges that derive from implementations of such cameras have recently been addressed and can be assumed mostly understood [4]. In the following, a signal processing chain [5] for generating high-quality panoramic images from the data supplied by a multi-CCD camera is presented. The focus is not on the calibration and rectification issues. Rather, we will discuss specific issues of a multi-CCD panoramic setup for the remaining signal processing chain that derive from the wide angle of view.

2. PANORAMIC IMAGE SEQUENCE ACQUISITION

The camera that we are working with consists of 6 CCD sensors, five of which are mounted in a circle pointing horizontally outwards, and the sixth is pointing upwards. Due to the use of wide-angle lenses, the total field of view almost covers a full sphere, except for the south pole area up to approximately the southern Arctic Circle. The sensors perform synchronized exposures, capturing 768×1024 pixels each, arranged in a Bayer pattern. Information regarding the calibration of this panoramic camera can be obtained from the camera manufacturer[1]. It is also discussed in [1] or [4].

The goal of the processing chain, starting from the six simultaneously captured sensor images, is the projection onto the six faces of a cube to form a cubic panorama [6] in a well defined color space, namely sRGB [7]. This representation allows for easy handling of the data for storage (based on rectangular faces of the cube) and rendering (e.g., using OpenGL) as well as further processing, e.g. [8, 9].

Along with the raw image data, side information is stored, e.g., the camera exposure settings, GPS information, etc.

Dynamic range There is alway the chance that the dynamic range of a scene exceeds the commonly available 8bit range of the camera. But this probability is significantly higher for panoramic cameras. For example, the top camera could be facing a bright sky or ceiling lamps, while some other camera faces a dark corner. One approach to capture the full dynamic range of a scenery is to perform multiple exposures with different shutter times. Thereby, different parts of the range are captured in different images, which can be merged later [1, 3]. Unfortunately, this merging can be very difficult if the scene is not static, and can lead to strange artifacts for dynamic panoramic sequences. Instead, we enable individual auto-exposure control of each of the 6 CCD sensors in the camera. A compromise has to be found for the visible area of each CCD, only, instead of the full view of the camera.

Resolution of the Panorama The wide field of view of the panoramic camera also spans a huge amount of detail in the scenery which requires a high sampling density. But the amount of managable pixels is limited. Therefore, special care is taken to extract and preserve the captured spatial bandwidth.

3. RAW IMAGE DATA PROCESSING

At each instant of a sequence of captures ($j = 1, ..., J$), raw image data from each of the 6 CCDs ($i = 1, ..., 6$) is stored in a $M \times N = 1024 \times 768$ array $\breve{\mathbf{C}}^B_{i,j}$. The individual values $\breve{c}^B_{i,j}[m, n]$ have 8 bit precision and form a Bayer pattern, where the RGB components are stored in $\begin{smallmatrix}\mathrm{B}&\mathrm{G}\\\mathrm{G}&\mathrm{R}\end{smallmatrix}$ layout. The 6 CCD images are processed to result as the 6 faces of a cube by performing the following steps:

- Lens falloff correction
- De-modulation of Bayer pattern
- Edge-preserving noise filter
- Color transform to (s)RGB
- White Balancing
- Exposure normalization

---

[1] `http://www.ptgrey.com/products/ladybug2/index.asp`

- Retinex
- Up-sampling
- Mapping to cube faces

The individual steps are going to be presented in the following, not necessarily in the exact order in which they are applied.

While the raw CCD data as well as sRGB color space and likely display device are going to be using 8 bit precision values, all intermediate values discussed in the following are assumed to be at appropriately higher precision.

Lens Fall-Off Correction Starting from the raw CCD values, the effect of lens vignetting is to be corrected first. This is part of the camera calibration and distortion correction and has been discussed e.g. in [1, 4]. We include it here in a simplified manner for completeness:

$$\forall_{i,j} : c^B_{i,j}[m,n] = \frac{(m-m^c_i)^2+(n-n^c_i)^2+K^2}{K^2} \cdot \breve{c}^B_{i,j}[m,n], \qquad (1)$$

where (m^c_i, n^c_i) denotes the optical image center of the ith CCD.

Color De-Modulation Color image signals represented in a Bayer pattern are usually viewed as a spatial multiplexing of RGB values. However, the situation can also be interpreted as a frequency multiplex of Luma and Chroma components $\mathrm{LC_1C_2}$. Locally adaptive de-modulation of the $\mathrm{LC_1C_2}$ components followed by linear color transform results in state-of-the-art Bayer de-mosaicing performance at low complexity [10]. Introducing two chrominance carrier signals $P_1[m,n] = (-1)^{m+n}$ and $P_2[m,n] = ((-1)^m + (-1)^n)$, the multiplexed color signal $\mathbf{C}^B_{i,j}$ can be expressed as

$$\begin{aligned} c^B_{i,j}[m,n] = {} & c^{LCC}_{i,j,1}[m,n] + \\ & P_1[m,n] \cdot c^{LCC}_{i,j,2}[m,n] + P_2[m,n] \cdot c^{LCC}_{i,j,3}[m,n]. \end{aligned} \qquad (2)$$

Here $\mathbf{C}^{LCC}_{i,j,k}$, with $k = 1,2,3$, denote the three $\mathrm{LC_1C_2}$ components that we want to extract from $\mathbf{C}^B_{i,j}$. They can be recovered by a set of band-pass filters and de-modulation to base-band. See [10] for details and an adaptive choice of filters. After recovery of the $\mathrm{LC_1C_2}$ components, they can be transformed to camera RGB signal $\mathbf{C}^{RGB}_{i,j,k}$ by a linear color transform. While this approach is computationally efficient, it is capable of extracting and preserving almost the full bandwidth of the captured signal as well.

Color Transform The RGB components of the image signal are a result of the spectral properties of the camera's optical system and CCDs. In order to enable a realistic reproduction of the color of the panoramic imagery, the image signal contained in camera RGB is converted into sRGB color space [7]. The conversion into sRGB requires the image data to be available in CIE XYZ space. Therefore, the mapping from camera RGB to XYZ has to be determined first. Let $\mathbf{s}_X$, $\mathbf{s}_Y$, $\mathbf{s}_Z$ denote XYZ-responses to a random set of spectral densities, forming a uniform distribution on the xy-plane, and let $\mathbf{s}_R$, $\mathbf{s}_G$, $\mathbf{s}_B$ denote the vectors of corresponding RGB responses of the camera. Neglecting a spectral influence of the wide-angle lenses, $\mathbf{s}_R, \mathbf{s}_G, \mathbf{s}_B$ derive from the random densities in combination with the sensitivities of the CCDs and the attenuation of the IR cut-off filter. With $\mathbf{S}_{XYZ} = [\mathbf{s}_X, \mathbf{s}_Y, \mathbf{s}_Z]$ and $\mathbf{S}_{RGB} = [\mathbf{s}_R, \mathbf{s}_G, \mathbf{s}_B]$, a mapping from camera RGB to XYZ can be derived as [11]

$$\mathbf{T}_M = \left(\mathbf{S}_{RGB}{}^{\mathsf{T}} \mathbf{S}_{RGB}\right)^{-1} \mathbf{S}_{RGB}{}^{\mathsf{T}} \mathbf{S}_{XYZ}. \qquad (3)$$

Combining $\mathbf{T}_M$ with the linear transform defined for sRGB, we obtain a mapping from camera RGB to linear sRGB. This is applied after the adaptive filter described in the next paragraph.

The target sRGB color space is non-linear and requires another processing step on each component of the linear sRGB signal, namely a gamma, shift, and normalization (see [7] for details). This step is postponed until the end of the processing chain, so that all further processing is performed in linear space.

Edge-Preserving Noise Filter Before spatial sampling is performed in CCD sensors, ideally optical low-pass filtering is performed matching the spatial bandwidth to the sampling density. In the panoramic camera, no or little optical pre-filtering is applied, resulting in some aliasing and cross-talk in and between $\mathrm{LC_1C_2}$ components.

Although the missing optical pre-filter in general allows the application of super-resolution methods [12, 13] on sequences of images, it can lead to undesirable staircasing, wrong colors, and dot patterns when images are considered one at a time. In our case, these effects especially occur in the Luma component after color demodulation. Here, we apply an adaptive filter motivated from [14],

$$\forall_{i,j,k} : \hat{\mathbf{C}}^{LCC}_{i,j,k} = F^e_k\{\mathbf{C}^{LCC}_{i,j,k}\}, \qquad (4)$$

where $\hat{\mathbf{C}} = F^e_k\{\mathbf{C}\}$ is defined as

$$\hat{c}[m,n] = \frac{\sum\limits_{m',n'} c[m',n'] \cdot w_k\left(c[m,n] - c[m',n']\right)}{\sum\limits_{m',n'} w_k\left(c[m,n] - c[m',n']\right)}. \qquad (5)$$

Here, indices (m', n') denote the 4-neighborhood of (m, n), i.e. $|m - m'| + |n - n'| \leq 1$. The weight function is

$$w_k(d) = \begin{cases} s & \text{if } |d| < \max\{\sigma, \mu \cdot \mathrm{MAD}(m,n)\} \\ -\kappa_k & \text{otherwise} \end{cases} \qquad (6)$$

using the median absolute deviation [15],

$$\mathrm{MAD}(m,n) = \underset{m',n'}{\mathrm{median}} \left\{ \left| c^{LCC}_{i,j,1}[m,n] - c^{LCC}_{i,j,1}[m',n'] \right| \right\}, \qquad (7)$$

which is again evaluated on the local 4-neighborhood. Robustness and a noise smoothing effect is achieved by introducing a noise level $\sigma =$ const, which prevents the enhancement of small signal variations. The enhancing feature controlled by $\kappa_k \geq 0$ is only used for the L component. The Chroma components are not enhanced, and $\kappa_k = 0$ for $k = \{2, 3\}$.

White Balance Automatic white balancing is not provided by the camera. In case the manual balancing is not accurately set, or lighting conditions vary over the sequence, balancing is performed after the acquisition. Application of the simple but efficient gray world assumption [5] has provided good results. Indeed, by the panoramic camera setup, a more complete observation of the world than by a single camera is available leading to a higher stability. For application of the gray world assumption, we compute the necessary averages over the 6 CCD images, $B^{cub}_{j,k} = \frac{1}{6MN} \sum\limits_{i=1}^{6} \sum\limits_{m=1}^{M} \sum\limits_{n=1}^{N} \hat{c}^{RGB}_{i,j,k}[m,n]$. To extend the horizon even further, averaging is additionally performed across capture instances, $B^{seq}_{j,k} = \frac{1}{1+\rho}\left(B^{cub}_{j,k} + \rho \cdot B^{seq}_{j-1,k}\right)$. Here, a recursive method has been chosen for complexity reasons. Thereby, the conversion from raw data into cubes can be completed one cube at a time. Adaptation speed is controlled by ρ. The gray world assumption is applied through normalization by the average green ($k = 2$),

$$\forall_{i,j,k} : \mathbf{C}^{RGB}_{i,j,k} = \hat{\mathbf{C}}^{RGB}_{i,j,k} \cdot \frac{1}{\lambda + 1} \left(\lambda + \frac{B^{seq}_{j,2}}{B^{seq}_{j,k}} \right). \tag{8}$$

It may be visually beneficial to allow for only a partial application of the gray world assumpion. The parameter λ is used to vary the extent of the offline color balancing.

Exposure Normalization By allowing each CCD its own, time varying, automatically controlled exposure settings, the resulting signals $\mathbf{C}^{RGB}_{i,j,k}$ are not aligned with regard to their amplitudes. Pixel values from each CCD depend on the respective shutter time $t_{i,j}$ and gain setting $g_{i,j}$ in a linear manner. Alignment of all CCD images into a common illumination space is achieved by multiplying each value by $a_{i,j} = \frac{\bar{a}}{t_{i,j} \cdot g_{i,j}}$. The dynamic range of the resulting illumination space can be significantly higher than the 8bit that most display devices usually are capable of representing. After multiplication by factors $a_{i,j}$, the illumination space can still be alligned to the target dynamic range of a display by choosing $\bar{a}$. A practical normalization for small exposure variation is by the maximum amplification $\bar{a} = \max_{i,j}\{a_{i,j}\}$. For a dynamic sequence of panoramas experiencing different lighting conditions, we perform a smooth illumination adaptation. As a compromise for the six images from one capture, we first determine

$$\forall_j : \log(\tilde{a}_j) = \frac{1}{6} \sum_{i=1}^{6} \log(a_{i,j}). \tag{9}$$

Instead of a single fixed normalization $\bar{a}$, we allow $\bar{a}_j$ to vary over the sequence of captures, and ensure smooth variation by convolution of $\tilde{a}_j$ as a function of j with a smoothing kernel $\mathbf{h}$,

$$\log(\bar{\mathbf{a}}) = \log(\tilde{\mathbf{a}}) * \mathbf{h}. \tag{10}$$

Note that normalization using $\bar{\mathbf{a}}$ is applied as part of the following step.

Retinex Retinex algorithms aim at the reduction of dynamic range of image signals, mimicking the functionality of the human visual system and thus leading to natural looking results. In our case, Retinex is performed using the bilateral filter as proposed in [16]. This approach is based on the concept of image decomposition [17] into detail layers, and individual compression of each layer. The choice of the bilateral filter avoids so-called halo-effects which occur if linear convolutions are used for the decomposition instead. In accordance to [16], a single decomposition stage provides sufficient results. The decomposition is applied to the luminance of the image signal. Therefore, the white-balanced RGB signal $\mathbf{C}^{RGB}_{i,j,k}$ is first transformed to XYZ color space, resulting in $\mathbf{C}^{XYZ}_{i,j,k}$, of which only the Luma component is used. Although Retinex is meant to be applied to the image signals in the common illumination space, the decomposition is performed based on the Luma, $\mathbf{C}^{XYZ}_{i,j,1}$, before exposure normalization is performed. Thereby, the separation of base content and detail content in each image is based on the amplitude scale chosen by the auto-exposure control. A compressed signal is obtained by local scaling of each pixel,

$$\forall_{i,j,k} : \grave{c}^{RGB}_{i,j,k}[m,n] = r_{i,j}[m,n] \cdot c^{RGB}_{i,j,k}[m,n], \tag{11}$$

where the scaling factors are determined as

$$\mathbf{R}_{i,j} = \frac{255}{F\{\mathbf{C}^{XYZ}_{i,j,1}\}} \left(\frac{a_{i,j}}{\bar{a}_j} \frac{F\{\mathbf{C}^{XYZ}_{i,j,1}\}}{255} \right)^{\gamma}. \tag{12}$$

Here, $\hat{\mathbf{C}} = F\{\mathbf{C}\}$ denotes the bilateral filter,

$$\begin{aligned} \hat{c}[m,n] &= \frac{1}{K_{m,n}} \sum_{m',n'} c[m,n] \cdot K_{m,n} \\ \text{where } K_{m,n} &= \sum_{m',n'} w_A\left(c[m,n] - c[m',n']\right) \cdot \\ &\qquad w_S\left(\sqrt{(m-m')^2 + (n-n')^2}\right). \end{aligned} \tag{13}$$

The weight functions are $w_{S/A}(d) = \exp\{-\frac{d^2}{\sigma^2_{S/A}}\}$. The weighted average in (13) is evaluated over the neighborhood (m', n') around (m, n) where $w_S(d)$ is significant. For Retinex, a large spatial extent is necessary, and this operation can be very time consuming. Several means to accelerate the bilateral filter operation have been propsed, see e.g. [16, 18]. Among available possibilities, simple sub-sampling of the evaluated neighborhood provides very satisfacory results. Namely, m' (and n' similarly) is chosen to be outof $\{m, m \pm s, m \pm 2s, m \pm 3s, ...\}$, where s is a positive integer.

The compression parameter γ from (12) may be determined by visual inspection of the results of the retinex procedure. As a rule of thumb, it should be chosen so that the dynamic range of the illumination space fits into the dynamic display range, which is typically 8bit. With the input images captured with 8bit precision, the illumination space has a dynamic range of $\log_2 \max_{i,j}\{a_{i,j}\} - \log_2 \min_{i,j}\{a_{i,j}\} +$ 8 bit, so that $\gamma \approx 8/\left(\log_2 \max_{i,j}\{a_{i,j}\} - \log_2 \min_{i,j}\{a_{i,j}\} + 8\right)$.

Up-sampling When mapping the camera images onto the cube faces, resampling becomes necessary. Target pixel positions on the cube faces will not in general map onto integer pixel positions in the processed CCD image matrices. A common way to determine fractional pixel image values is to apply bilinear interpolation between the closest available values. The frequency response of the bilinear interpolator depends in the distance from available pixels and tends to attenuate high frequency components if the target image value is located in the middle between available pixels. In order to avoid the discrimination of these positions, we implement a two-step interpolation method as described in [19]. In the first step, the image resolution is increased by a factor of two (per dimension) using a separable higher-order interpolation filter. Separability and the fixed resampling ratio allow for an implementation of low complexity. Fractional pixel positions are then computed using bilinear interpolation on the high-resolution grid. The high-resolution signal does not have components in the upper half of its frequency band, and high frequency attenuation through bilinear filtering takes significantly less effect.

Geometric Texture Mapping Since the 6 sensors of the panoramic camera are not co-located, and in order to account for manufacturing variations as well as the lens distortion, geometric texture mapping is based on calibration provided by the camera manufacturer. Thereby, a calibrated mapping of coordinates from an imaginary sphere around the camera center onto the camera images is available (see also [1]). Since the images are slightly overlapping, coordinates from more than one sensor may be valid for a target on the sphere. Due to parallax of the camera setup, the coordinate mapping depends on the diameter of the assumed sphere. For the acquisition processing, we can choose between two sphere radii, a smaller for indoor and a larger one for outdoor scenes. In addition, more advanced methods may be applied to account for parallax effects and merge neighboring CCD images so that a consistent panorama is obtained [20, 21, 1], but this is out-of the scope of this paper.

For display purposes, storage, and further processing, it is convenient to map the panorama on a cube instead of a sphere [9]. By including this additional coordinate transform, we obtain a direct mapping from target cube face coordinates to calibration corrected CCD image coordinates. A single resampling step is performed based on this mapping. Overlapping images are blended using alpha masks.

4. SUMMARY

We have presented our implementation of an acquisition processing chain designed to match the specific needs of a dynamic multi-camera acquisition system. These needs derive firstly from the wide angle of view and secondly from viewing applications that want to offer a high image quality when viewing one panorama but also a high density of panoramas when striding through a scene. Due to the almost spheric field of view, the probability of the dynamic range exceeding the capabilities of the A/D converters in the camera is significantly higher than for standard digital cameras. This issue has been approached by individual adjustment of each if the camera's CCDs for capture and amplitude alignment and compression using a Retinex algorithm at a later stage. The wide field of view also covers a high amount of content in the scenery which requires a large number of pixels in order to result in a satisfactory viewing impression when one region of interest is explored. For a video camera, a resolution of 1024×768 pixels still is considered high resolution, but for a single view it is not. So particular care is taken in obtaining and preserving the full spatial bandwidth of the image signal, especially in the stages of Bayer de-mosaicing and re-sampling. Some small artifacts remaining from the de-mosaicing due to color crosstalk as well as a sensor noise floor are addressed by an edge enhancing locally adaptive filter. Another aspect of the wide angle of view is the more complete observation of the *world* resulting in a higher stability when performing white balancing using the *gray world assumption*.

5. REFERENCES

[1] Matthew Uyttendaele, Antonio Criminisi, Sing Bing Kang, Simon Winder, Richard Szeliski, and Richard Hartley, "Image-based interactive exploration of real-world environments," *IEEE Transactions on Computer Graphics and Applications*, vol. 24, no. 3, pp. 52–63, May 2004.

[2] Aljoscha Smolić and Thomas Wiegand, "High-resolution video mosaicing," in *Proc. IEEE Int. Conference on Image Processing ICIP '01*, Oct. 2001, vol. 3, pp. 872–875.

[3] Paul Debevec and Jitendra Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. of ACM SIGGRAPH 1997*, 1997.

[4] Sei Ikeda, Tomokazu Sato, and Naokazu Yokya, "High-resolution panoramic movie generation from video streams acquired by an omnidirectional multi-camera system," in *IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems 2003*, 2003.

[5] Rajeev Ramanath, Wesley E. Snyder, Youngjun Yoo, and Mark S. Drew, "Color image processing pipeline," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 34–43, Jan. 2005.

[6] Derek Bradley, Alan Brunton, Mark Fiala, and Gerhard Roth, "Image-based navigation in real environments using panoramas," in *HAVE 2005 – IEEE International Workshop on Haptic Audio Visual Environments and their Application*, Ottawa, ON, Canada, 2005.

[7] Michael Stokes, Matthew Anderson, Srinivasan Chandrasekar, and Ricardo Motta, "A standard default color space for the internet – sRGB," http://www.w3.org/Graphics/Color/sRGB, 1996.

[8] Chris Warrington, Gerhard Roth, and Eric Dubois, "Markerless augmented reality for cubic panorama sequences," in *5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, Oct. 2006.

[9] Mark Fiala and Gerhard Roth, "Automatic alignment and graph map building of panoramas," in *HAVE 2005 – IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, Oct. 2005.

[10] Eric Dubois, "Filter design for adaptive frequency-domain bayer demosaicking," in *Proc. IEEE Int. Conference on Image Processing ICIP '06*, Oct. 2006.

[11] Gaurav Sharma, Ed., *Digital Color Imaging Handbook*, CRC Press, 2003.

[12] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.

[13] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Dynamic demosaicing and color super-resolution of video sequences," in *Proceedings of the SPIE conference on image reconstruction from incomplete data III*, Oct. 2004, vol. 5562.

[14] Markus Beermann and Jens-Rainer Ohm, "Non-linear up-sampling for image coding in a spatial pyramid," in *Proc. SPIE Visual Communications and Image Processing VCIP '07*, San Jose, USA, Jan. 2007.

[15] Michael J. Black, Guillermo Sapiro, David H. Marimont, and David Heeger, "Robust anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421–432, Mar. 1998.

[16] Frédo Durand and Julie Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 2002, pp. 257–266, ACM Press.

[17] Jack Tumblin and Greg Turk, "LCIS: A boundary hierarchy for detail-preserving contrast reduction," in *Proceedings of SIGGRAPH 99, ACM SIGGRAPH*, 1999, pp. 83–90.

[18] Michael Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1141–1151, Oct. 2002.

[19] Thomas Wedi, "Adaptive interpolation filters and high-resolution displacements for video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 4, 2006.

[20] Xiaoyong Sun and Eric Dubois, "A novel algorithm to stitch multiple views in image mosaics," in *Proc. 2004 IEEE Int. Conf. Acoustics Speech Signal Processing*, May 2004, pp. III–481–III–484.

[21] Alan Brunton and Chang Shu, "Belief propagation for panorama generation," in *Proc. Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT 2006)*, June 2006.