Underwater Image Mosaics for AUV-Mounted Cameras

Erik Makino Bakken
Norwegian Defence Research Establishment (FFI)
Kjeller, Norway
erik-makino.bakken@ffi.no

Øivind Midtgaard
Norwegian Defence Research Establishment (FFI)
Kjeller, Norway
oivind.midtgaard@ffi.no

***Abstract*—This paper presents a novel method for automatically selecting the highest quality image sections for an underwater optical mosaic. We have constructed an image quality measure, which is used to determine the optimal seam between two overlapping images. We also present our preprocessing algorithm, which is a novel combination of existing methods. The full mosaic processing chain is demonstrated on images collected with a HUGIN autonomous underwater vehicle (AUV), using a camera system consisting of a high-sensitivity, still-image camera and a synchronized strobe with multiple light emitting diodes (LEDs).**

Keywords—mosaic, AUV, underwater optical imaging

I. Introduction

Autonomous Underwater Vehicles (AUVs) have proven to be efficient assets for detailed imaging and mapping of the seafloor. The most commonly applied sensor is sonar, which provides an excellent combination of resolution and range. Optical sensors, however, provide unique scene information that cannot be obtained from acoustical sensors, but imaging can be challenging due to strong absorption and scattering of light in water. In order to collect high-quality images the camera has to operate close to the scene, typically at only a few meters range. Each image then only covers a small area, and larger objects may not fit into a single image. To get a complete, high-resolution view of a larger scene one has to stitch multiple images together into a mosaic [1], [2], [3]. Important applications are for instance archaeology [4], [5] and habitat mappings [6], [7].

There are several challenges with underwater imaging. First, light attenuation is severe and depends on the wavelength of the light, causing certain colors to dominate. This also introduces an uneven color distribution because the distance to the seafloor varies within the image. Backscattering in the water volume causes reduced image contrast, while forward scattering of reflected light reduces image resolution. At lower depths, there is insufficient natural light, and imaging requires the use of artificial light, which typically becomes non-uniformly distributed across the scene. Due to the difficult light propagation conditions, image preprocessing is necessary to improve contrast, normalize illumination, and adjust color balance. Previous works have studied color correction of underwater images with natural lighting [8], [9], but also in the context of AUVs with artificial lighting [10]. To correct non-uniform illumination, we first considered homomorphic filtering [11], but we instead chose to make an empirical model for illumination and attenuation, by taking the pixel-wise average over the image set [12].

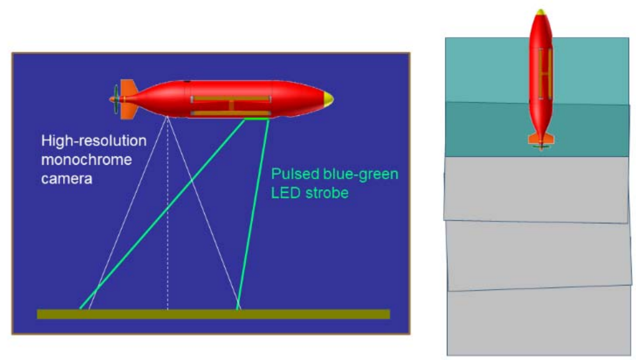

Fig. 1. Imaging geometry of the TileCam still-image camera and strobe.

In the cases of overlapping images, it is important to choose the image section of better quality. To obtain this, one could make a Voronoi tessellation based on the minimum distance to the image centers [3]. In our data the better part of the image is not confined to the center, but can sometimes be in the upper or lower half depending on the vehicle altitude and orientation. To determine which image section to use, we have developed a quality measure, which is used to find the relative quality of two image sections covering the same seafloor patch. To determine the seam between images, we have developed an optimization algorithm based on a seam algorithm [13] from the OpenCV library [14]. The algorithm avoids splitting objects in overlapping images if there are large, local differences between the co-registered images. Furthermore, it uses the proposed quality measure so that the higher quality image section is used.

II. Image Data/Experiment

The images used in this study have been recorded in varied environments with the TileCam optical camera system which has been specially designed for the HUGIN AUVs through a collaborative development effort between Norsk Elektro Optikk (NEO), FFI, and Kongsberg Maritime. It consists of a high-resolution, high-sensitivity still image camera and a synchronized strobe with multiple light emitting diodes (LEDs). The camera is mounted near the vehicle rear end and oriented downwards. To reduce backscatter, the LEDs are

Global Oceans 2020: Singapore - U.S. Gulf Coast | 978-1-7281-5446-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/IEEECONF38699.2020.9389090

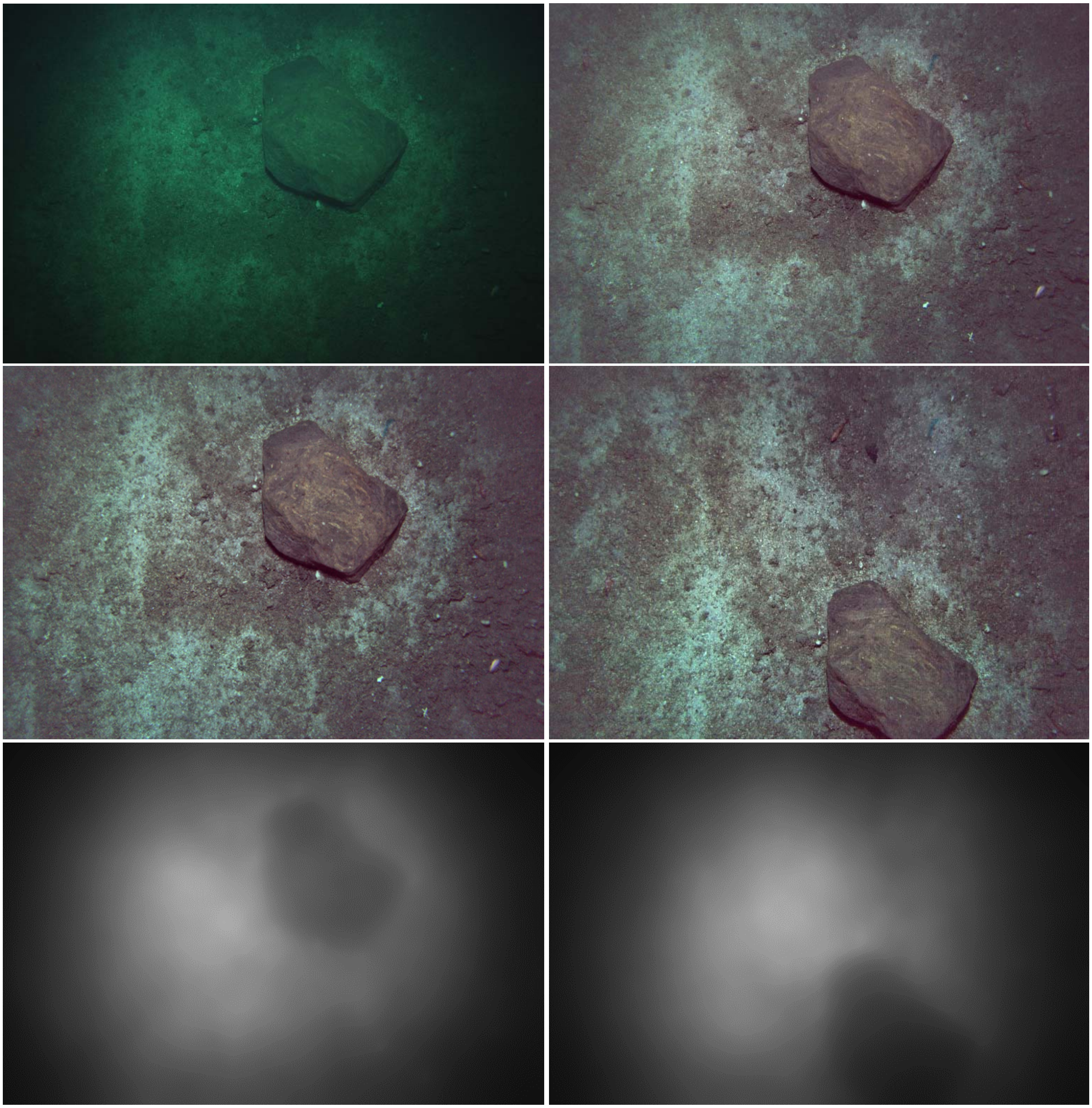

Fig. 2. TileCam example images (2004 x 1336 pixels) captured at 6.5 meters altitude. Top left: Raw image. Top right: Image corrected with frame average. Middle left: Image after preprocessing. Middle right: The next image after preprocessing. Bottom left: The quality of the first image. Bottom right: The quality of the next image.

separated from the camera and tilted backwards to illuminate the camera field of view (Fig. 1).

Two TileCam system variants exist. The standard version combines a monochrome camera and blue-green diodes to maximize the imaging range, while a special version applies a color camera and white diodes to provide color images in clear waters.

The image data presented in this paper was collected with FFI's HUGIN-HUS AUV in clear Norwegian coastal waters in October 2015 at around 200m depth. An example image is shown in the top left of Fig. 2

III. METHOD

A. Processing Chain

The processing chain is shown in Fig. 3. First, we read the vehicle navigation and Doppler data files. Then we use these data together with the individual image data to calculate a local scene plane for each image, which later will be used to map the image onto the mosaic plane. In addition, the total area coverage of the mosaic is computed. An illumination model is constructed as the pixel-wise average of all the images in the data set.

After this initial setup, the images are processed in sequence. A quality measure is calculated for each pixel in the image, which will later be used for determining which part of the image should be used. Each image goes through a pre-processing step before being mapped into the mosaic. The image is first enhanced with the illumination model. Then, we use histogram equalization and color correction to further enhance the image. The final pre-processing step is to undistort the image.

After pre-processing, each image is mapped into the mosaic plane via its local plane calculated in the initial step. A seam between the image and the previous image is found based on a dynamic programming algorithm from OpenCV that we have modified to prefer the parts of the images with better quality. The image is then inserted into the mosaic, again using the quality measure. Finally, after processing of all the images, the mosaic is written to file.

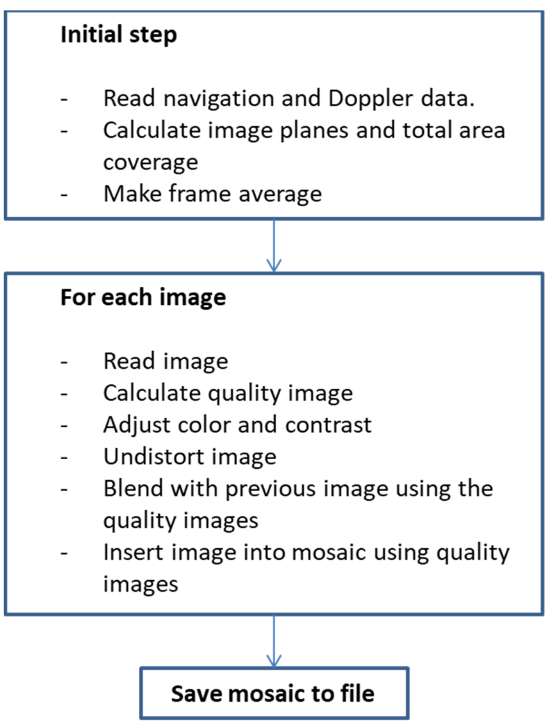

Fig. 3. Proposed mosaic processing chain.

B. Image Enhancement

Underwater images have reduced contrast because of backscattered light from the water volume. We have tested different methods to improve the contrast and found that a combined approach led to better results.

1) Frame Average

The first method for increasing the contrast is to normalize with a frame average [12]. The idea is to calculate a pixel-wise average of the images, which gives a model for the uneven light distribution, and then divide each image by this average image to compensate. Since the light distribution depends on the sensor height above the seafloor, one has to make a separate light model for each height interval and interpolate. We typically use height intervals between 0.5m and 1.0m. If the total number of images is small, scene features might be visible in the light model. To avoid that, we calculate the median of five images at a time, and take the average of the medians. The method works well for a flat seabed, but inaccuracies occur when the seabed has height variations. The result after applying this method can be seen in the top right of Fig. 2.

2) CLAHE

To further improve the image contrast we apply Contrast Limited Adaptive Histogram Equalization (CLAHE [15]). The illumination in our underwater images varies throughout the image, and in particular, the corners of the images are darker with lower signal to noise ratio. This makes CLAHE well suited for contrast enhancement. For color images, we first transform the image to the HSI (Hue, Saturation, Intensity) color space, apply CLAHE to the intensity image, and then transform back to the BGR (Blue, Green, Red) space. As seen in the middle left image in Fig. 2, CLAHE sharpens the image and works well together with the frame average processing. Since a frame average correction also improves the contrast, we use other parameters for CLAHE when combining the two methods.

3) Color Correction

A challenge with underwater color imaging is that light attenuation depends on the photon wavelength, causing red, green, and blue light to attenuate differently. Moreover, the attenuation depends on water quality. In our test data, red light is reduced much more than green and blue, and the raw images look green with a hint of blue, as evident from the top left image in Fig. 2.

An approach for color intensity correction is to assume that an image should contain similar amounts of red, blue, and green. This is the gray world assumption [8]. We correct the image by scaling the two colors with the lowest average intensity up to the color with the highest intensity.

For color images, the frame average is calculated and applied for each channel separately. This method improves the contrast, but it also effectively balances the colors. With color correction by scaling only, the color balance is uneven, and some parts have a green tone, while other parts appear red. Correction with the frame average is done first, and after that, we scale the colors to have the same average intensity. One could think that the frame average is enough to even out the colors, but we found the result to be improved by an additional color scaling.

C. Image Quality

To ensure good quality imagery of the full survey area, the camera frame rate is set sufficiently high to provide a certain degree of overlap between successive images. There may also be some overlap in the swaths of different survey lines. This introduces a need to select the image sections of highest relative quality for use in the mosaic. One common approach for this is to make a Voronoi tessellation based on the minimum distance to the image centers [3], thus assuming that the quality decreases radially from the center.

In murky underwater environments, the best-illuminated image section typically has the higher quality. As evident from the imaging geometry in Fig. 1, the illumination pattern in our data shifts along the vertical image axis as vehicle altitude and orientation vary. The better part of the image is then not confined to the center and we seek a through-the-sensor approach for quality estimation. The raw image intensity is not a robust quality measure as water volume backscattering also increases this quantity. We thus constructed a measure of small-scale image texture, by first applying a gradient operator. For two image sections covering the same scene with comparable sensing geometries, the differences in filter output can be attributed to differences in image quality, i.e. contrast and resolution. We then apply a low pass filter to the gradient image as we expect the quality to vary smoothly. The bottom row in Fig. 2 shows the results for the two example images. As expected, dark shades indicate poor quality regions along the image edges. Perhaps more surprisingly, the boulder also attains low quality values in both images. This is because the boulder has less small-scale texture than the surrounding seafloor and we apply a relative quality measure that can only be used to rank corresponding image sections after accurate co-registration.

D. Geometry

From earlier calibration experiments we have estimated the camera matrix and distortion parameters for our TileCam system. The images have radial distortion, which is rectified by using OpenCV library functions.

The main assumption for the mosaic processing is that the scene is planar for each image. In most cases, this is a good approximation of the scene. However, if objects raise significantly above the seafloor, the assumption will not hold. This will cause an error that depends on the object height compared to the AUV altitude. With this assumption, each image is projected into a local plane before it is projected orthographically into the mosaic plane.

From the DVL we obtain four height measurements. By using least squares, a plane is computed, which is a local approximation of the scene. An alternative is to use an altimeter to measure the height above the seafloor, and then assume that the plane is parallel to the forward/starboard plane of the AUV. We chose to use the DVL heights since they give more information than a single measurement. The local plane is given in global coordinates, and its position and orientation are computed from the vehicle navigation solution. After the image is mapped onto the local plane, it is projected onto the mosaic plane by an orthographic projection.

E. Seam

For consecutive images, we use a modified version of the OpenCV class *DpSeamEstimator*, which is based on [13]. In this class a proposed seam path is assigned a higher cost if it crosses regions with large differences in color or intensity between the images. This will make the seam avoid splitting objects if there is a difference in the images due to changes in the scene or if the images are taken from different angles. We modify these costs to include the quality of the pixel. Given these parameters, a dynamic programming algorithm searches for the optimal (i.e. minimal cost) seam path. The seam will usually be located in an area where the quality of the images is similar, as can be seen in Fig. 4. The seam crosses the rock, but in an area with small differences between the images.

F. Blending

In the transition between the images, there will be a small ramp where the images are blended, and here we use alpha-blending. It is a simple way to blend by linearly transitioning from one image to the other. If I_1 and I_2 are the intensities of a pixel in two different images, then the intensity of a pixel in the ramp is given by

$$I_{\text{Blended}} = \alpha I_1 + (1-\alpha) I_2, \qquad (1)$$

where α goes from 1 to 0 linearly with distance as one transitions from the first to the second image.

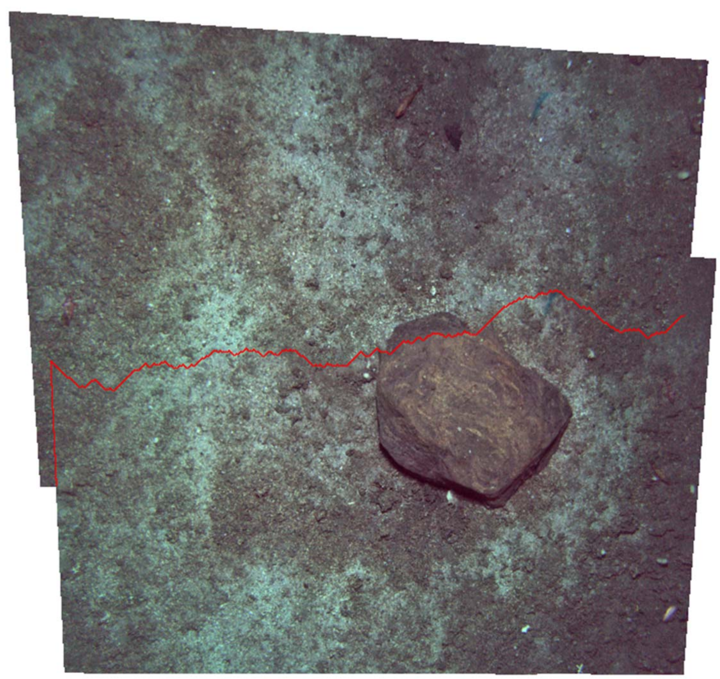

Fig. 4. The estimated seam (red curve) between the two images in the middle row of Fig. 2.

IV. RESULTS

We have successfully tested the proposed processing chain on multiple data sets with color and graytone TileCam images collected in different environments. Fig. 6 displays an example mosaic constructed from 18 individual images captured at around six meters altitude in relatively clear water. We observe that the resulting color intensities are consistent within each image and between images. The image transitions are smooth, because the seam is confined to regions with small differences in contrast and color between the images. Using the image parts with higher contrast, increases the level of visible seafloor details as can be appreciated in Fig. 5, which zooms in on the top part of the mosaic.

Figs. 4-6 indicate that the vehicle's aided inertial navigation solution is sufficiently accurate for successful co-registration of succeeding images. For longer time intervals, though, navigation drift introduces mosaic distortions. When the current image overlaps imagery from a previous survey line, this becomes evident as image feature mismatches.

A limitation with our proposed method is the assumption of planar scenes. Violation of this assumption introduces geometrical distortions, such as the small discontinuities of the boulder edges across the image seam in Fig. 4.

Fig. 5. A more detailed view of the upper part of Fig. 6.

Fig. 6. Mosaic consisting of 18 images.

V. CONCLUSION

The quality of underwater optical images is significantly affected by difficult lighting conditions, severe light absorption, and scattering. We have developed a novel combination of methods for improving the contrast, intensity distribution, and color balance of the images. Furthermore, we have demonstrated an approach to select the best sections of overlapping images based on a relative quality measure. The mosaic processing chain has been successfully tested on TileCam color and graytone images captured by HUGIN AUVs in varying underwater conditions. In future work, we aim to implement image based matching to improve the co-registration for large mosaics. We will also seek to improve the range-dependent illumination model by taking the inclination of the local scene planes into account.

[1] H. S. Sawhney, S. Hsu, and R. Kumar, “Robust video mosaicing through topology inference and local to global alignment,” Proc. Eur. Conf. Computer Vision, 1998.

[2] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, “Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," Journal of Field Robotics, vol. 27, no. 1, 2010.

[3] O. Pizarro and H. Singh, “Toward large-area mosaicing for underwater scientic applications," IEEE Journal of Oceanic Engineering, vol. 28, no. 4, 2003.

[4] R. Eustice, H. Singh, J. Leonard, and M. Walter, “Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters,” The International Journal of Robotics Research, 25(12), 2006

[5] M. Ludvigsen, B. Sortland, and H. Singh, “Applications of geo-referenced underwater photo mosaics in marine biology and archaeology," Oceanography, vol. 20, no. 4, 2007.

[6] D. Lirman, N. R. Gracias, B. E. Gintert, A. C. R. Gleason, P. R. Reid, S. Negahdaripour, and P. Kramer, “Development and application of a video-mosaic surveytechnology to document the status of coral reef communities,” Environ. Monit. Assess., 125:59–73, 2007

[7] D. Corrigan, K. Sooknanan, J. Doyle, C. Lordan, and A. Kokaram, “A Low-Complexity Mosaicing Algorithm for Stock Assessment of Seabed-Burrowing Species,” IEEE Journal of Oceanic Engineering, vol. 44, no. 2, 2019.

[8] G. Bianco, M. Muzzupappa, F. Bruno, R. Garcia, and L. Neumann, “A new color correction method for underwater imaging,” The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W5, 2015.

[9] K. Iqbal, R. A. Salam, A. Osman, and A. Z. Talib, “Underwater image enhancement using an integrated colour model,” IAENG International Journal of Computer Science, 34(2), 2007.

[10] J. W. Kaeli, H. Singh, C. Murphy, and C. Kunz, “Improving color correction for underwater image surveys,” OCEANS'11 MTS/IEEE KONA, 2011

[11] S. Bazeille, I. Quidu, L. Jaulin, and J.-P. Malkasse, “Automatic underwater image pre-processing,” CMM’06 - Caractérisation du Milieu Marin, 2006.

[12] R. P. Gutiérrez, “Image blending techniques and their application in underwater mosaicing,” Ph.D. dissertation, Universitat de Girona, Spain, 2013.

[13] V. Kwatra, A. Schödl, E. Irfan, G. Turk, and A. Bobick, “Graphcut textures: Image and video synthesis using graph cuts," ACM Transactions on Graphics, 2003.

[14] OpenCV, viewed 3 July 2020, https://opencv.org.

[15] K. Zuiderveld, “Contrast limited adaptive histogram equalization," in Graphics Gems IV (P. S. Heckbert, ed.), ch. VIII.5, pp. 474-485, Academic Press, 1994.