3-2014

A Robust Quasi-dense Matching Approach for Underwater Images

Hu Han
University of New Hampshire, Durham

Yuri Rzhanov
University of New Hampshire, Durham, Yuri.Rzhanov@unh.edu

Thierry Boyer
Parks Canada Agency

Recommended Citation

H. Hu, Rzhanov, Y., and Boyen, T., “Robust Quasi-dense Matching Approach for Underwater Images”, American Society for Photogrammetry and Remote Sensing (ASPRS) 2014. Louisville, KY, 2014.

A ROBUST QUASI-DENSE MATCHING APPROACH FOR UNDERWATER IMAGES

Han Hu[1], Graduate student
Yuri Rzhanov[1], Professor
Thierry Boyer[2]
[1]Center for Coastal and Ocean Mapping, University of New Hampshire, Durham, NH 03824
[2]Parks Canada Agency, 1800 Walkley Road, Ottawa, ON, K1A 0M5, Canada
hhu@ccom.unh.edu
yuri.rzhanov@unh.edu
thierry.boyer@pc.gc.ca

ABSTRACT

While different techniques for finding dense correspondences in images taken in air have achieved significant success, application of these techniques to underwater imagery still presents a serious challenge, especially in the case of "monocular stereo" when images constituting a stereo pair are acquired asynchronously. This is generally because of the poor image quality which is inherent to imaging in aquatic environments (blurriness, range-dependent brightness and color variations, time-varying water column disturbances, etc.). The goal of this research is to develop a technique resulting in maximal number of successful matches (conjugate points) in two overlapping images. We propose a quasi-dense matching approach which works reliably for underwater imagery. The proposed approach starts with a sparse set of highly robust matches (seeds) and expands pair-wise matches into their neighborhoods. The Adaptive Least Square Matching (ALSM) is used during the search process to establish new matches to increase the robustness of the solution and avoid mismatches. Experiments on a typical underwater image dataset demonstrate promising results.

KEYWORDS: Underwater image, dense matching, match propagation, affine transformation, 3D reconstruction

BACKGROUND

Optical underwater imagery is of great importance for scientists in hydrography and oceanography because it allows intuitive interpretation of scenery by human operators. In the area of image processing, dense matching has been one of the most widely investigated topics for a long time since it is an essential phase for many applications, such as 3D reconstruction (Ahmadabadian et al., 2013) and image enhancement (HaCohen et al., 2011). However, underwater image processing is more challenging than that in air due to following limitations. (1) The scene illumination is non-uniform. Typically underwater image is acquired with either ambient or artificial illumination. Both cases lead to time-varying illumination pattern. In the former case it is because of water column and surface perturbations, and in the latter case because of the motion of the light source. (Note that this paper considers imagery acquired by a single moving underwater camera.) (2) Light attenuation with range is significant due to medium absorption and scattering. Because of those different environmental factors, underwater images exhibit different properties compared with the images taken in air. Thus, matching of underwater images is more difficult in the following aspects: a) underwater images have much fewer salient points, like corners (even man-made objects are often covered with sediments and vegetation), b) the light-absorbing medium (water) leads to a violation of the brightness constancy constraint, c) wavelength-dependent absorption of light does not allow to rely on color constancy, and d) the presence of suspended particles introduces noise in optical measurements.

Since the brightness and color constancy does not hold for underwater imagery, the traditional intensity-based matching methods that compare the gray-scales of pixels from two small image windows (patches) cannot be used for underwater image matching as it usually can for the matching of in-air images. Straightforward application of this method leads to a high probability of identifying a wrong match as a good one (false positive). Thus, for underwater images, the correlation score obtained from the intensity-based matching cannot be accepted as confidently as that for air imaging. However, the correlation measurement can still be an important indicator for similarity estimation between patches as long as the intensity-based matching method can be combined with other measurements to identify the correct image correspondences. Adaptive Least Square Matching (ALSM) incorporates adaptive geometric properties between the patches that are matched. This method utilizes an iterative optimization of the local geometric warping parameters between image patches such that the sum of absolute gray-scale differences (Srikham et al., 2010) between corresponding pixels is minimized. In this paper, both the warping parameters and gray-scale similarity measurement contribute to the final decision about the goodness of the match. The "match-and-expand" procedure is used to implement the dense pixel-wise matching (Lhuillier and Quan, 2002). Our inspiration comes from (Kannala and Brandt, 2007), where the local geometric transformation is incorporated in a procedure for matching images with wide baseline. Starting with a sparse set of feature matches (seed matches), the search for the pixel-wise relationships is then iteratively expanded into the neighborhoods of seeds. In the process of searching for new matches, the ALSM technique is used instead of the traditionally used normalized cross-correlation. With ALSM, the correlation measurement is meaningful. Simultaneously the optimal warping parameters are obtained and then compared with the parameters of the neighboring seed. Experiments on a typical underwater dataset demonstrate high performance of the proposed approach under different imaging conditions.

It should be noted that commercial software packages working reliably on image sequences acquired in air like PhotoScan of Agisoft LLC (Agisoft LLC, 2014), fail to produce a point cloud from imagery used in this paper.

The rest of this paper is organized as follows. Section 2 gives a brief review of the principles of the ALSM. Section 3 describes the proposed algorithm in detail. Section 4 shows some experiments on a typical underwater image dataset. The last section presents the conclusions.

ADAPTIVE LEAST SQUARE MATCHING

The Adaptive Least Square Matching (ALSM) (Gruen, 1985) technique has been widely used in Photogrammetry and Computer Vision since it was proposed in 1985. The process of the ideal pinhole camera imaging is in principle the projection from a Euclidean 3-dimensional coordinate system to a 2-dimensional plane. Due to the change of viewpoints and properties of the medium, the patches from different images that correspond to the same object in 3D scene differ both in shape and color (gray-scale levels of contributing pixels). Traditionally, the ALSM assumes that there is a geometric affine transformation and a radiometric linear gray-scale transformation between corresponding image pixels within the patches.

We denote the small corresponding image patches as discrete two-dimensional functions $f(x, y)$ and $g(x, y)$ respectively where $f(x, y)$ is the function that describes the patch in the reference image and $g(x, y)$ describes the patch in the target image. If the image location is represented by homogeneous vectors, the pixel mapping can be written in matrix form.

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \tag{1}$$

where $\mathrm{a}_0, \mathrm{a}_1, \mathrm{a}_2$ and $\mathrm{b}_0, \mathrm{b}_1, \mathrm{b}_2$ are the affine geometric parameters. $(\mathrm{x}_2, \mathrm{y}_2)$ is the image coordinate in the patch of target image $\mathrm{g}(\mathrm{x}, \mathrm{y})$ and $(\mathrm{x}_1, \mathrm{y}_1)$ is the image coordinate in the patch of reference image $\mathrm{f}(\mathrm{x}, \mathrm{y})$.

The relationship of gray-scales between corresponding pixels is assumed to be linear,

$$g(x_2, y_2) = h_0 + h_1 f(x_1, y_1) \tag{2}$$

where h_0 and h_1 are constants. To summarize, the pixel in the patch of the target image corresponds to the pixel in the patch of the reference image through an affine transformation. Also, the corresponding pixels are linearly related by their gray-scales.

Linearized version of (1) and (2) gives:

$$v = c_1 dh_0 + c_2 dh_1 + c_3 da_0 + c_4 da_1 + c_5 da_2 + c_6 db_1 + c_7 db_2 + c_8 db_3 - \Delta \tag{3}$$

where $c_1,\ldots,c_8$ are the partial derivatives of $g(x_2,y_2)$with respect to the geometric affine and radiometric linear gray-scale parameters; dh_0, dh_1, …, db_3 are the corrections of the parameters and Δ is the constant brightness difference.

Each pixel pair in the corresponding patches gives a linear equation in the form of equation (3) and from all the pixel pairs between the image patches we obtain a set of linear equations which can be written in a matrix form:

$$V = AX - L \tag{4}$$

Equation (4) is the observation equation where A is the design matrix and X is the correction vector of observables (geometric and radiometric parameters).

$$X = (dh_0\ \ dh_1\ \ da_1\ \ da_2\ \ da_3\ \ db_0\ \ db_1\ \ db_2)^T \tag{5}$$

The brightness difference between the corresponding pixels in the matching patches is minimized by iterative solution of the observation equation and updating of the correction vector X until the correction of each element in X becomes smaller than a predefined threshold value.

With a rough estimate of locations of the corresponding image patches, the optimal affine and brightness transformation parameters can be obtained using the ALSM. The center of the image patch is considered to be the final refined correspondence location. The ALSM is applied to every eligible putative match around the seed match. The best ones (i.e., with the highest similarity score) are chosen as the final matches and added to the seed list for further propagation. However, if the initial guess of patch location is far from the true location, the ALSM will not converge to a final solution within a certain number of iterations. In this case the putative match is rejected.

QUASI-DENSE MATCHING BASED ON PROPAGATION

1. Initial Matching

The aim of the first step is to obtain a sparse set of matches and to estimate the epipolar geometry to initialize the propagation process. This is achieved by the extraction and matching of a sparse set of feature keypoints – seed matches. Success at this step is crucial as it determines the quality of the final result. Seed matches should have no outliers, so the most reliable salient point detectors (SIFT (Lowe, 2004 and Lowe, 1999) and SURF (Bay et al., 2006)) have been chosen for this operation. Matching of the conjugate features requires reliable descriptors. It has been found that Daisy descriptor (Tola et al., 2010) often outperforms the ones proposed in SIFT and SURF. A fundamental matrix which essentially describes the epipolar geometry for an image pair is then found using RANSAC-based (Fischler and Bolles, 1981) robust method to guarantee the absence of outliers in the set of seed matches (see e.g., (Zhang et al., 1995; Hartley and Zisserman, 2000)). The recovered epipolar geometry is later used to prevent the propagation process from expanding into incorrect regions. This epipolar constraint on propagation is especially important for underwater imagery. Since the image correlation measurement is preferred in our algorithm and if we consider all the neighboring pixel pairs around the seed match as putative matches without imposing the epipolar constraint, the matches which are not on the epipolar line may potentially have higher similarity scores than the true match because of lesser trustworthiness of the correlation method for underwater imagery.

2. Quasi-Dense matching using ALSM

After the first stage of obtaining an initial sparse set of seed matches, we will use these seed matches to produce a quasi-dense correspondences between the two images. Quasi-dense matching was proposed by Lhuillier and Quan (Lhuillier and Quan, 2002; Lhuillier, 1998) and developed further to be applicable to wide baseline images (Megyesi and Chetverikov, 2004; Megyesi and Chetverikov, 2003) and multiple views (Koskenkorva et al., 2010). This method is considered to be the “golden mean” between sparse and dense approaches (Khropov and Konushin, 2010). The method is based on the propagation of point correspondences into their neighborhoods.

We start with a brief review of the match propagation algorithm. The incorporation of the ALSM technique will be described in the subsequent paragraphs. The putative matches are searched in the immediate spatial neighborhood of seed match (x, x'). At first, all the neighboring pixel correspondences (u, u') are checked for satisfying the epipolar constraint and 2D discrete disparity limit. The distance from u' in the target image to its corresponding epipolar line calculated by u should be less than a small predefined threshold. The 2D discrete disparity limit is defined as follows (see (Lhuillier and Quan, 2002) for more detail). Let $N(x)$ and $N(x')$ denote all the neighboring pixels of x and x' from the seed match. The possible matches (u, u') should be limited within the following area

$$N(x,x') = \{(u,u'), u \in N(x), u' \in N(x'), ||(u'-u)-(x'-x)||_{\infty} \leq \epsilon\} \tag{6}$$

After this all the similarities over eligible matches are measured by Zero-mean Normalized Cross-Correlation (ZNCC). The matches are next ranked by their ZNCC scores and added to a local list of matches. The best new matches in this local list are added to the seed list and the final dense match results if they have not been found by previous propagation. The propagation always starts with the most reliable seed (i.e., with the largest correlation score) which is then removed from the seed list. The entire matching process continues until no seed matches remain in the seed list.

In order to adjust the traditional quasi-dense matching for underwater imagery, our proposed approach incorporates the ALSM in the propagation process to minimize the number of mismatches. The modification is described below.

The first extension to the match propagation algorithm is applied at the first stage of the initial matching. The initial set of feature matches is matched using the ALSM in order to obtain the affine parameters for the transformation between the initial seed patches (image regions around x and x'). The affine parameters are adhered to the seed matches and are used later in the propagation step.

The original quasi-dense algorithm (Lhuillier and Quan, 2002) does not include the normalization of the local image patches and the search for new matches is performed in the non-altered neighborhood of x'. This limited the application of the original quasi-dense algorithm to wide-baseline images because in such images viewpoints differ substantially and hence the deformation between two corresponding patches can be significant. In underwater imagery, viewpoints are almost always far away from each other because it is difficult to control viewpoints due to currents and platform buoyancy. This paper proposes to search for the putative matches not in the original neighborhood, but in the image patch warped by the affine transformation with the parameters found for the seed match (x, x').

Each putative match (u, u') from the local normalized image patch is first checked for epipolar constraint and 2D disparity limit as it is normally done in the original propagation algorithm. The location of the match that satisfies the above constraints is given as a rough estimate for a putative match to the ALSM. The refined location for this putative match (u_r, u_r') is returned from the ALSM together with its geometric affine and radiometric gray-scale linear parameters. Actually, u is equivalent to u_r since the center of image patch in the reference image does not change during the ALSM computations. u_r', the correspondence of u_r in the target image, is obtained by iterating equation (4). Now the uniqueness of (u_r, u_r') should be checked to guarantee that the match has not been covered on the previous propagation steps. If the match (u_r, u_r') already exists in the final dense match list, it is rejected. The affine parameters of (u_r, u_r') are compared with those of the seed match (x, x'). The match (u_r, u_r') is accepted and added to the list of local candidates only if the differences between the parameters are sufficiently small. The underlying assumption for this comparison of affine parameters is that the local observed surface is smooth, which means that the local geometric warping parameters for the two neighboring matches should vary continuously. Likewise, the proposed algorithm proceeds until no seed match can be taken out from the seed list for propagation. Figure 1 illustrates the workflow of this proposed algorithm. Compared with the seeds in the original proposed algorithm, the seed matches in our approach have the geometric affine transformation parameters as an extra property which is used for normalization of local image patches and surface smoothness check.

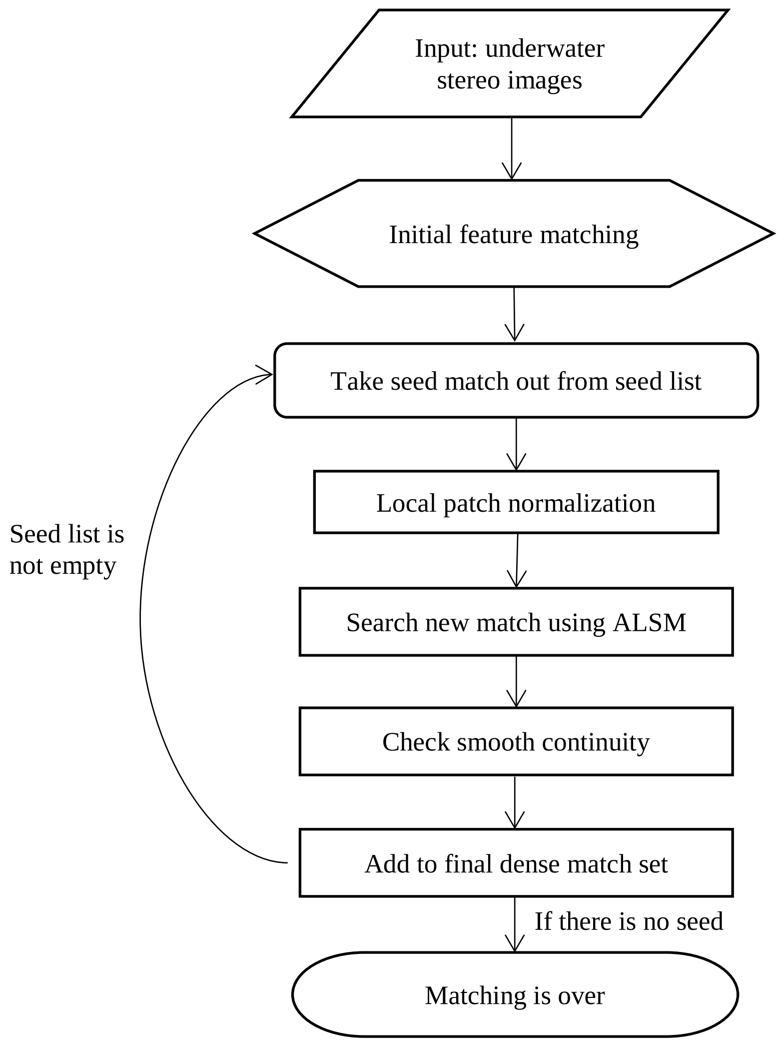

Figure 1. The workflow of the proposed algorithm

EXPERIMENTS

The approach described above was tested on a typical underwater image dataset. The dimensions of the images are 712 pixels in width and 1072 pixels in height. Two different image pairs from the dataset were selected for the test. The first image pair covers a smooth surface of a sunk airplane and does not have dark (poorly textured) area. The surface covered by the second image pair has depth discontinuities and large dark areas. SIFT detector was used for the initial feature matching step and produced 304 feature matches and 155 feature matches respectively (Figure 2).

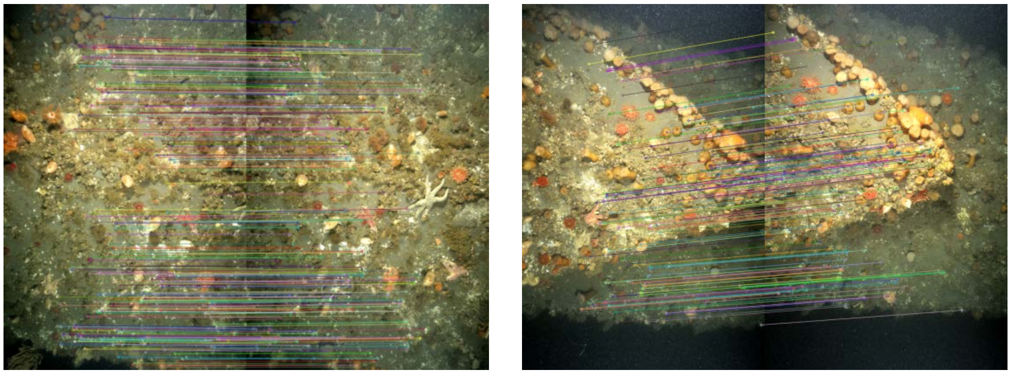

Figure 2. SIFT-detected matches for two underwater stereo image pairs.

After application of the match propagation using the ALSM, 391016 quasi-dense matches were obtained from the first pair of images and 284338 from the second pair. The ratio of successfully matched pixels to all pixels is 0.51 and 0.37 respectively. For the second image pair, the pixel matching ratio is lower because of the existence of low contrast areas and significant depth discontinuities. (Note that the ratio of 1 cannot be achieved in principle because the images overlap only partially.) However, we will demonstrate that this is still sufficient for the successful 3D reconstruction. The epipolar geometry of these image pairs has been recovered as described above. The quasi-dense matches obtained by the proposed approach were triangulated to build point clouds and then surfaces were reconstructed from point clouds to test the performance of the approach. The results are shown in Figure 3 and Figure 4.

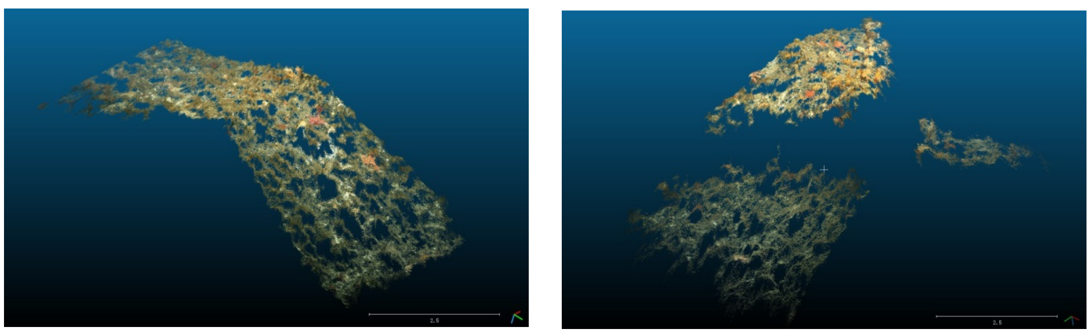

Figure 3. Visualization of triangulated point clouds.

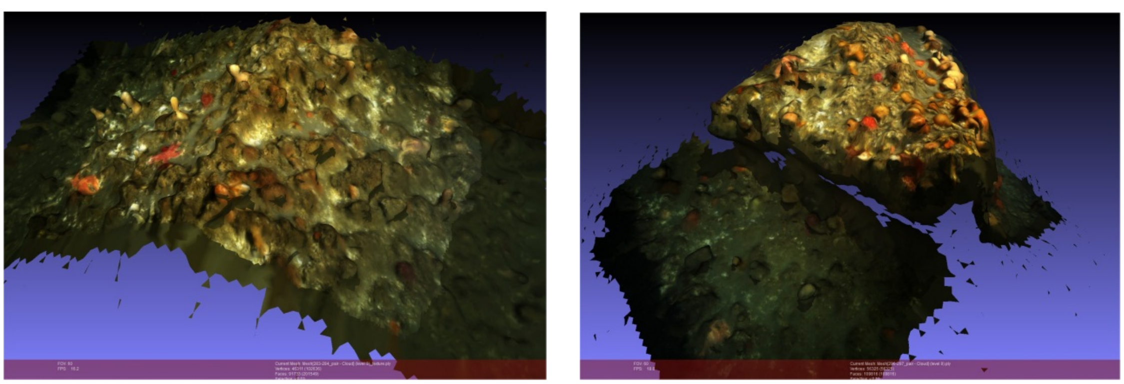

Figure 4. Reconstructed surfaces from the point clouds with image texture draped over the 3D surface. (The surface is constructed using Poisson surface reconstruction algorithm (Kazhdan et al., 2006) provided by software MeshLab (Visual Computing Lab, 2014).)

From the constructed point clouds and textured surfaces of these two underwater image pairs it follows that our proposed algorithm is robust enough to provide a sufficient number of image correspondences for the successful reconstruction of the 3D surface.

CONCLUSION

In this paper, a quasi-dense matching algorithm which is sufficiently robust for the application to underwater imagery has been reported. The proposed algorithm works differently from existing methods in several ways to be adapted to underwater imaging environment. The ALSM technique is incorporated in the process of the match propagation. In addition, the local geometric warping parameters for each putative match are also compared with those of the neighboring seed match to choose the best dense matches. The seed match is augmented by an array of affine transformation parameters which is returned from the ALSM and the match expansion procedure is carried out in the normalized (warped) neighborhood of the seed.

The experiments on a typical dataset demonstrated that this algorithm is robust to the change of environment factors (in comparison with air) and it is able to provide a sufficient number of successful correspondences between the overlapping underwater images. The algorithm described in this paper can be useful in a wide range of applications such as underwater image modeling and rendering.

ACKNOWLEDGEMENTS

This work was funded by NOAA grant NA10NOS4000073.

REFERENCES

Agisoft LLC, 2014. PhotoScan. http://www.agisoft.ru/products/photoscan/. (last date accessed: 15 February 2014)

Ahmadabadian, A.H., et al., 2013. A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 78: 157-167.

Bay, H., T. Tuytelaars and L. Van Gool, 2006. Surf: Speeded up robust features. *European Conference on Computer Vision*, May 7-13, 2006.

Fischler, M.A. and R.C. Bolles, 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6): 381-395.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3): 175-187.

HaCohen, Y., et al., 2011. Non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics*, 30, 4, 70: 1-70: 9.

Hartley, R. and A. Zisserman, 2000. Multiple view geometry in computer vision, Cambridge University, Cambridge.

Kannala, J. and S.S. Brandt, 2007. Quasi-dense wide baseline matching using match propagation. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 18-23, 2007.

Kazhdan, M., M. Bolitho and H. Hoppe, 2006. Poisson surface reconstruction. *Proceedings of the fourth Eurographics symposium on Geometry processin*, June 26 to 28, 2006.

Khropov, A. and A. Konushin, 2006. Guided Quasi-Dense Tracking for 3D Reconstruction. *International Conference Graphicon*, July 1-5, 2006.

Koskenkorva, P., J. Kannala and S.S. Brandt, 2010. Quasi-dense wide baseline matching for three views. *20th International Conference on Pattern Recognition*, August 23-26, 2010.

Lhuillier, M., 1998. Efficient dense matching for textured scenes using region growing. *The Ninth British Machine Vision Conference, September 14-17*, 1998.

Lhuillier, M. and L. Quan, 2002. Match propagation for image-based modeling and rendering. *Pattern Analysis and Machine Intelligence*, 24(8): 1140-1146.

Lowe, D.G., 1999. Object recognition from local scale-invariant features. *The proceedings of the seventh IEEE international conference on Computer Vision*, September 20-27: 1150-1157.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2): 91-110.

Megyesi, Z. and D. Chetverikov, 2004. Affine propagation for surface reconstruction in wide baseline stereo. *Proceedings of the 17th International Conference on Pattern Recognition*, August 23-26, 2004.

Megyesi, Z. and D. Csetverikov, 2004. Affine dense matching for wide baseline stereo. *Proceedings of 2rd International Symposium on 3D Data Processing, Visualization and Transmission*, September 6-9, 2004.

Srikham, M., C. Pluempitiwiriyawej and T. Chanwimaluang, 2010. Comparison of dense matching algorithms in noisy image. *Second International Conference on Digital Image Processing*, February 26, 2010.

Tola, E., V. Lepetit and P. Fua, 2010. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 32(5): 815-830.

Visual Computing Lab, ISTI, CNR, 2014. MeshLab. http://meshlab.sourceforge.net/. (last date accessed: 14 February 2014)

Zhang, Z., et al., 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial intelligence*, 78(1): 87-119.