

# Developing a Visual Stopping Criterion for Image Mosaicing Using Invariant Color Histograms

Armagan Elibol and Hyunjung Shim<sup>(✉)</sup>

School of Integrated Technology, Yonsei University, Incheon, Republic of Korea  
{aelibol,kateshim}@yonsei.ac.kr

**Abstract.** For over a decade, image mosaicing techniques have been widely used in various applications *e.g.*, generating a wide field-of-view image, 2D optical maps in remote sensing or medical imaging. In general, image mosaicing combines a sequence of images into a single image referred to as a mosaic image. Its process is roughly divided into the iterative image registration and blending. Unfortunately, the computational cost of iterative image registration increases exponentially given a large number of images. As a result, mosaicing for a large scale scene is often prohibitive for real-time applications. In this paper, we introduce an effective visual criterion to reduce the number of image mosaicing iterations while retaining the visual quality of the mosaic. We analyze the change in invariant color histograms of the mosaic image over iterations and use it to determine a termination condition. Based on various experimental evaluations using four different datasets, we significantly improve the computational efficiency of mosaicing algorithm.

**Keywords:** Image mosaicing · Visual quality · Optical mapping · Invariant color histogram

## 1 Introduction

Image mosaicing is a class of techniques that register overlapping images and combine them into a larger image [12]. Since mosaicing is effective to create a wide-field-of-view image (i.e., a mosaic image) from a set of images and/or video, the resultant mosaics have been very useful for different scientific studies such as geology [5,9], biology [10] or archaeology [1,11]. Especially with the rapid development of mobile platforms, it becomes possible to obtain optical data of areas beyond the human reach. Mosaics of these areas can help revealing locations of areas of interest or visualize temporal changes in the morphology of bio-diversity of the terrain. For that, mosaics are analyzed by a human expert and provide the global perspective on the area of interest.

In general, image mosaicing is composed of two main phases: iterative image registration for aligning image pairs and image blending for obtaining the final

mosaic. An image registration process is composed of a pairwise and global registration. While pairwise registration is to identify the transformation between two overlapping images in the sequence, global registration extracts the best possible transformation parameters of each image with respect to a common mosaic coordinate frame. Image blending imposes the smooth transition along the seam in a final mosaic image after global registration and this improves the final quality of the mosaic. The blending is necessary because photometric differences are the main source of seams and they can occur even under the perfect geometric alignment.

Image mosaicing is accomplished via iterating pairwise image registration and global registration (updating the estimate of camera trajectory) using possible overlapping image pairs. Considering time-consecutive images, they generally present significant overlaps. While registering them, their registration parameters can serve as an initial estimate of camera trajectory. However, this initial estimate suffers from error accumulation. This is because the absolute homography, a planar transformation between an input frame and global frame, is derived from multiple relative homographies, a planar transformation between two input frames. When computing each relative homography, we purely rely on correspondences, which vary upon the performance of feature descriptors and matching algorithm. Consequently, each relative homography potentially hides the error caused by incorrect correspondences. Since the absolute homography aggregates multiple relative homographies, the errors from each homography are accumulated in the absolute homography.

Non-consecutive overlapping image pairs can be predicted by this coarse estimate. Registering non-consecutive overlapping image pairs helps improve the trajectory and mosaic. Once overlapping image pairs are identified, global registration methods can be employed in order to find the best transformation parameters between image coordinate frame and a global frame. Note that we can choose an arbitrary image frame to fix the global coordinate system. In our implementation, we choose the first frame as the global frame. Global registration is done by minimizing an error defined by the distance of correspondences between image pairs. This step requires the non-linear optimization, which comes with high computational cost. This cost increases drastically if we are given a large number of input images to create a huge mosaic.

In this paper, we aim to obtain a mosaic image using a reduced number of overlapping image pairs with retaining the visual quality as well as possible to the one using all image pairs. In this way, we can reduce the computational cost introduced by global registration as well as the cost of identifying and registering overlapping image pairs. In [4], the importance of overlapping image pairs have been evaluated by using a weighted shortest path algorithm. Although the importance of the overlapping image pairs were evaluated through their shortest alternative paths and final mosaics were nearly identical to their counterpart ones, the visual quality of image registration and intermediate mosaics were not analyzed. In this paper, we propose to use a deformation and viewpoint invariant color histogram [2] (referred to as an invariant histogram for the rest of paper.)

to measure the changes in visual quality of mosaic after each iteration of the image mosaicing process. The important property of the invariant histogram is that it is invariant under any mapping of the surface that is locally affine. This property is particularly beneficial to measure the image similarity under a wide class of viewpoint changes or deformations. Since images are warped with different transformation parameters to compose the mosaic, the change in the invariant histogram is caused by the misregistration between images in our application. Therefore, we find that the change in invariant histogram is an adequate measure to evaluate our mosaicing process. The proposed method can be integrated into various existing frameworks in image mosaicing to improve their computational efficiency.

## 2 Invariant Histogram Based Mosaic Image Quality Monitoring

Standard color histograms are sensitive to changes in the viewpoint. Domke and Aloimonos [2] proposed a new color histogram that is invariant to an arbitrary transformation of locally affine surface. They weight pixels using gradients of different color channel. In our context, individual image is warped in global frame to form a mosaic assuming the target surface being locally affine. If the alignment between images remains same, applying an arbitrary transformation does not change the invariant histogram [2]. Our proposal is to generate the intermediate mosaics and compare its invariant histogram with that of previous iteration. If the ratio of change is lower than a threshold, we terminate mosaicing iterations. Our method can be interpreted as adding constraint to image mosaicing framework by monitoring invariant histograms of the mosaics produced at each iteration. A standard image mosaicing pipeline combined with our method is illustrated in Fig. 1. To compare histograms of two images  $a$  and  $b$ , we employ the same metric in [2]. For an image  $a$  and  $b$ , computation of differences between their histograms is given in Eq. 1.

$$d(\mathbf{h}^a, \mathbf{h}^b) = \frac{\sum_c (\mathbf{h}_c^a - \mathbf{h}_c^b)^2}{\sum_c (\mathbf{h}_c^b)^2} \quad (1)$$

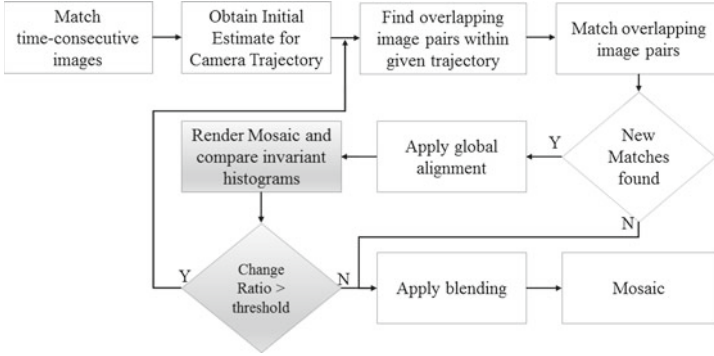
where  $\mathbf{h}_c$  denotes the histogram value for color channel  $c$  and computed as follows:

$$\mathbf{h}_c = \sum_{s, s_c=c} |f_x(s)g_y(s) - f_y(s)g_x(s)| \quad (2)$$

where  $f$  and  $g$  denote derivatives in two color channels [2].

## 3 Experimental Results

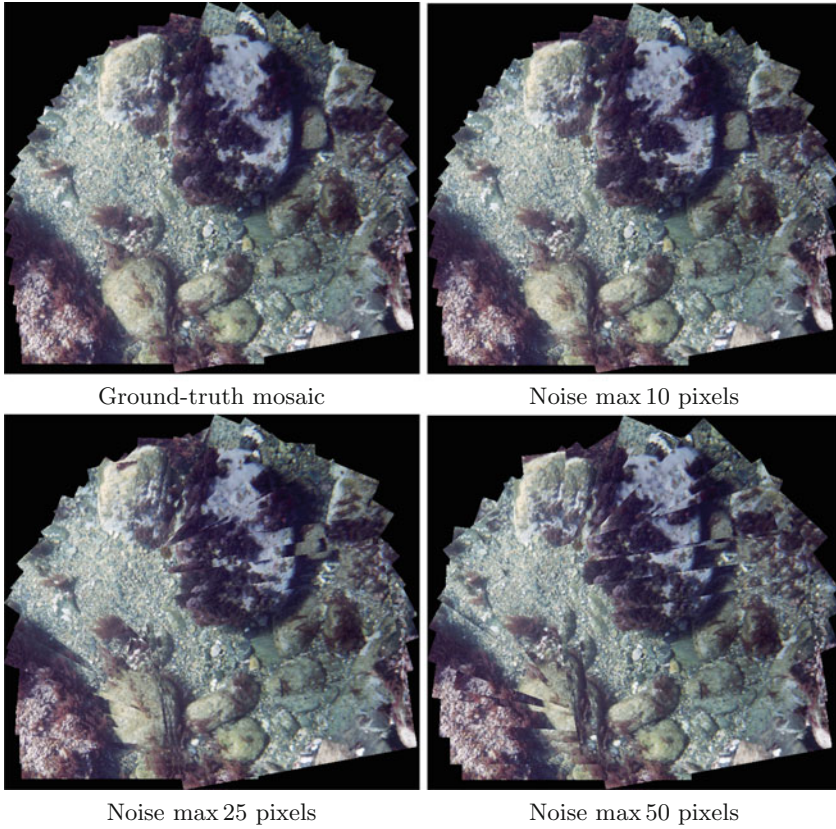
We have conducted various experiments on four different datasets. The first experiment is to measure how invariant histogram varies upon misregistration



**Fig. 1.** A mosaicing pipeline of proposed method.

in the mosaic and to monitor the value of the metric given in Eq. 1. For that, we use 33 images of  $384 \times 288$  pixels cropped from high resolution mosaic. We register images to the mosaic directly in order to obtain their image-to-mosaic planar transformations. Given these transformation parameters, the mosaic is generated by a bottom-up strategy. This mosaic serves a ground-truth as illustrated in Fig. 2. For image registration, we extract the Scale Invariant Feature Transform (SIFT) [8] features and apply Random Sample Consensus (RANSAC) eliminate outliers and estimate the planar transformation. To analyze the robustness of proposed method, we generate the misalignment in image pairs and report the effects of misalignment in the quality of mosaic. To simulate misalignments, we add a Gaussian random noise with zero mean and several levels of standard deviation to the translation parameters both  $x$  and  $y$  direction. Then, we obtain misaligned mosaics due to the erroneous parameters. The invariant histograms of misaligned mosaics were compared with the one of ground truth mosaic by using Eq. 1. For each variance level of noise, we randomly draw 1000 samples of noise. From this experiment, we observed how the value has changed and how the registration errors have evolved over the significance of noise. Furthermore, to quantify the errors in camera trajectory, we register images pairwise. A totally 528 image pairs were registered and the total number of correspondences over these pairs becomes 142,317. For each noisy transformation set, a symmetric transfer error [7] is computed.

We summarize our results in Table 1. Numbers given in the table are statistically computed over 1000 trials for each noise level. For higher level of noise, mosaics that have the maximum symmetric transfer error within trials are illustrated in Fig. 2. We find that starting from the noise level of 10 pixels, a visual disturbance on mosaic can be easily recognizable. This provides some insights for choosing a threshold. For the experiments with real image sequences, we terminate the iteration if the change between histograms is smaller than or equal  $10^{-4}$  in two consecutive iterations. Taking into account the mosaics in Fig. 2 and symmetric transfer errors in Table 1, it can be concluded that symmetric transfer error may not provide fully accurate information about the visual quality of the mosaics. However, the noise level of parameters is strongly correlated with



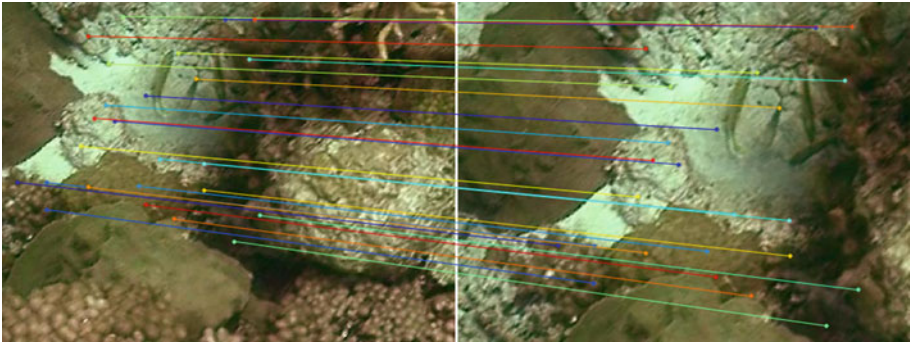
**Fig. 2.** Mosaics obtained with additive noise on the translation parameters of their planar image-to-mosaic transformations

**Table 1.** Change on invariant histograms and computed symmetric transfer errors with different levels of noise. Change on histograms is computed by using Eq. 1

Noise Level (in pixels)	Change on histogram			Symmetric Transfer Error (in pixels)		
	Mean	Standard Deviation	Maximum	Mean	Standard Deviation	Maximum
2.50	1.18E-05	7.28E-06	5.66E-05	2.74	0.67	8.26
5.00	1.45E-04	9.46E-05	6.18E-04	5.29	1.25	11.18
10.00	1.02E-03	4.37E-04	2.89E-03	10.48	2.47	17.28
25.00	4.84E-03	2.12E-03	1.44E-02	26.09	6.16	36.63
50.00	1.62E-02	7.08E-03	4.79E-02	52.24	12.30	70.34

to the visual errors in mosaics. On the other hand, it should be noted that the noise in our experiments was only added to the translation parameters. Having small noise on the rotation and scale parameters can provoke more noticeable errors on the final mosaic.

Finally, we have evaluated the computational performance of our method on three datasets (referred as Underwater Dataset I (UWDI), Underwater Dataset

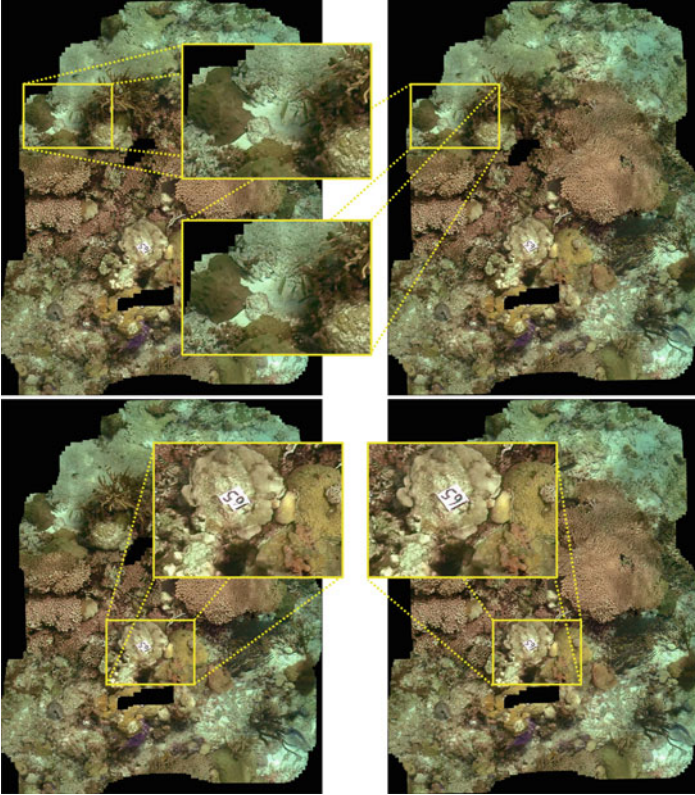


**Fig. 3.** Image pairs and established correspondences between them. This is the image pair where the symmetric transfer error had a maximum value in the UWDI.

II (UWDII), and aerial). They are extracted from a high-resolution image using real trajectory parameters of different Unmanned Vehicle (UVs). The UWDI is composed of 555 images of  $512 \times 384$  pixels. Total number of successfully registered (An image pair is considered successfully matched if it has a minimum of 20 inliers.) overlapping image pairs is 18,392 and total number of correspondences is 7,992,010. The UWDII consists of 460 images of  $572 \times 380$  pixels. This dataset is relatively sparse, having only 1,897 overlapping image pairs, and presents two non-overlapping time-consecutive image pairs. Such properties of dataset falls apart traditional methods, which requires overlap between time-consecutive images. The total number of correspondences is 828,947. The aerial

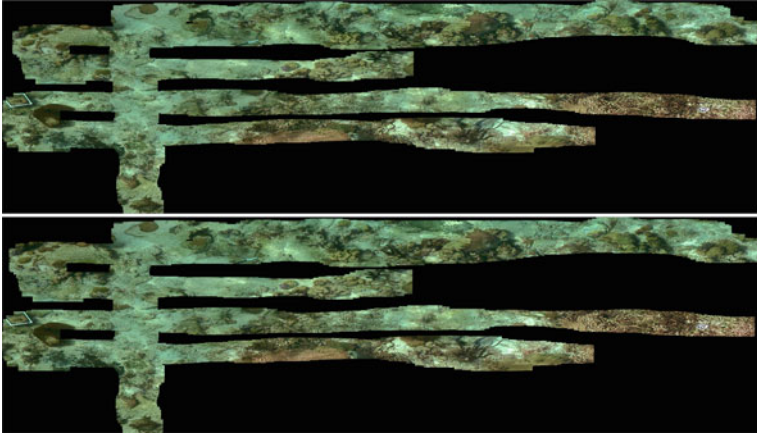
**Table 2.** Summary of results obtained using proposed method during the image mosaicing process. Strategy 'Without' represents the framework in the Fig. 1 without proposed steps.

Dataset	Strategy	Successful Pairs	Unsuccessful Pairs	Avg. Error in pixels	Std. Deviation Pairs in pixels	Maximum in pixels
UWDI	with proposed method	3,094	0	10.11	6.59	55.22
	without	18,320	2,902	7.34	3.06	31.14
	The approach in [6]	14,646	424	7.66	3.18	31.31
	AGA	18,392	135,343	7.33	3.06	31.14
UWDII	with proposed method	1,091	13	1.27	2.38	55.01
	without	1,877	976	0.83	0.33	4.70
	AGA	1,897	103,673	0.83	0.33	4.89
Aerial Dataset	with proposed method	980	0	1.12	0.43	4.80
	without	3,260	264	1.07	0.42	5.02
	The approach in [6]	3,792	202	1.06	0.42	4.94
	AGA	4,299	30,417	1.06	0.42	4.94



**Fig. 4.** Mosaics obtained with (left) and without using the proposal (right). (Top) Zoomed region where the symmetric transfer error is maximum. (Bottom) Small misalignment on final mosaics.

dataset comprises 264 images of  $387 \times 288$  pixels having 4,299 matched image pair and the total number of correspondences is 432,086. Our termination criterion is integrated into the image mosaicing method in [3] because this mosaicing algorithm allows to handle randomly ordered image sequence. In this way, we can manage the case when there are non-overlapping time-consecutive images like in the UWDI. Table 2 presents the summary of the results. The second column corresponds to the tested method. The third column shows the total number of successfully matched image pairs. The fourth column contains the total number of image pairs that were not successfully matched and we denote them as *unsuccessful pairs*. The last three columns correspond to the average symmetric transfer error, the standard deviation, and maximum error calculated using all the correspondences identified by All-against-all (AGA) matching strategy. For the UWDI, global registration is carried out using five points (four corners and the center of the image). Since the UWDI and aerial dataset provide an overlap between time-consecutive images, we make a comparison with the method in [6].



**Fig. 5.** Obtained mosaics of the UWDII with (top) and without using the proposal (bottom).



**Fig. 6.** Obtained mosaics of the aerial dataset with and without using the proposal.

Based on our experiments, we find that the maximum symmetric transfer error usually appears on overlapping image pairs with a big change on scale. Since their scale varies, one of them may not be visible in a final mosaic. Therefore, the visual quality of the final mosaic does not reflect the maximum symmetric transfer errors entirely as seen in Fig. 3. From the results presented in Table 2, mosaics can be obtained with a small number of image matching attempts without disturbing the final visual quality. Figs. 4, 5, and 6 show the obtained mosaics with and without using our proposal.



Although the computational times are not reported here, our method significantly reduces the total number of image mosaicing iterations and image matching attempts. The bottleneck of proposed method is the rendering phase, generating mosaic at each iteration and computing the invariant histogram. The time spent for rendering step can be reduced by applying the multiscale image analysis.

## 4 Conclusion and Future Work

Lately, great advancements in the mobile robotic platforms make it possible to obtain optical data from areas unreachable by humans. In most of the cases, a single image is not sufficient to provide an overview of the area of interest. To this end, Image mosaicing has been an indispensable tool for creating a large-area optical map from the images collected by mobile platforms. Without any prior on camera trajectory, a common mosaicing strategy is to apply the AGA image matching and then to perform global registration. This approach is exhaustive as it also attempts to register images that do not overlap. Therefore, its algorithmic complexity grows quadratically with the total number of images, which limits its usage in a small scale dataset.

Our experiments showed that invariant color histograms can be used as a visual stopping criterion during image mosaicing process. Also, we find that symmetric transfer error may not be an accurate indicator of visual quality of final mosaic, especially when camera trajectory provides scale changes and high overlapping area between both consecutive and non-consecutive images. Another important point can be stressed that identifying all overlapping image pairs may not be necessarily improving the visual quality of mosaic although it improves the camera trajectory estimate. In the future, we plan to extend invariant histograms based stopping criterion for mosaicing with low-overlapping image pairs.

**Acknowledgments..** Authors would like to thank Underwater Vision Laboratory of Computer Vision and Robotics Institute of University of Girona for providing high-resolution test images and real trajectory parameters. This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Republic of Korea, under the IT Consilience Creative Program (NIPA-2014-H0201-14-1002) supervised by the NIPA (National IT Industry Promotion Agency). Aerial High-resolution image was retrieved from <https://unsplash.com/stevenlewis> on the 27th of April, 2015.

## References

1. Bingham, B., Foley, B., Singh, H., Camilli, R., Delaporta, K., Eustice, R., Mallios, A., Mindell, D., Roman, C., Sakellariou, D.: Robotic tools for deep water archaeology: Surveying an ancient shipwreck with an autonomous underwater vehicle. *J. Field Rob.* **27**(6), 702–717 (2010)
2. Domke, J., Aloimonos, Y.: Deformation and viewpoint invariant color histograms. In: *BMVC*, pp. 509–518 (2006)

3. Elibol, A., Gracias, N., Garcia, R.: Fast topology estimation for image mosaicing using adaptive information thresholding. *Rob. Auton. Syst.* **61**(2), 125–136 (2013)
4. Elibol, A., Gracias, N., Garcia, R., Kim, J.: Graph theory approach for match reduction in image mosaicing. *J. Opt. Soc. Am. A.* **31**(4), 773–782 (2014). <http://josaa.osa.org/abstract.cfm?URI=josaa-31-4-773>
5. Escartin, J., Garcia, R., Delaunoy, O., Ferrer, J., Gracias, N., Elibol, A., Cufi, X., Neumann, L., Fornari, D.J., Humpris, S.E., Renard, J.: Globally aligned photomosaic of the lucky strike hydrothermal vent field (Mid-Atlantic Ridge, 3718.5°N): Release of georeferenced data, mosaic construction, and viewing software. *Geochem. Geophys. Geosyst.* **9**(12), Q12009 (2008)
6. Gracias, N., Zwaan, S., Bernardino, A., Santos-Victor, J.: Mosaic based navigation for autonomous underwater vehicles. *IEEE J. Oceanic Eng.* **28**(4), 609–624 (2003)
7. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press, Harlow (2004)
8. Lowe, D.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**(2), 91–110 (2004)
9. Park, J.Y., Choi, J.Y., Jeong, E.Y.: Applying an underwater photography technique to nearshore benthic mapping: A case study in a rocky shore environment. *J. Coastal Res. SI* **64**, 1764–1768 (2011)
10. Pizarro, O., Williams, S.B., Jakuba, M.V., Johnson-Roberson, M., Mahon, I., Bryson, M., Steinberg, D., Friedman, A., Dansereau, D., Nourani-Vatani, N., Bongiorno, D., Bewley, M., Bender, A., Ashan, N., Douillard, B.: Benthic monitoring with robotic platforms - the experience of Australia. In: *IEEE International Underwater Technology Symposium (UT)*, pp. 1–10 (2013)
11. Scaradozzi, D., Sorbi, L., Zoppini, F., Gambogi, P.: Tools and techniques for underwater archaeological sites documentation. In: *Oceans - San Diego 2013*, pp. 1–6 (2013)
12. Szeliski, R.: Image alignment and stitching: A tutorial. *Found. Trends® Comput. Graph. Vis.* **2**(1), 1–104 (2006)